Andre Martinho de Almeida
David Eckersall
Ingrid Miller *Editors*

Proteomics in Domestic Animals: from Farm to Systems Biology

Proteomics in Domestic Animals: from Farm to Systems Biology

Andre Martinho de Almeida • David Eckersall •
Ingrid Miller
Editors

Proteomics in Domestic Animals: from Farm to Systems Biology

Editors
Andre Martinho de Almeida
Instituto Superior de Agronomia
Universidade de Lisboa
Tapada da Ajuda, Lisbon, Portugal

David Eckersall
ERA Chair Laboratory, Faculty of Veterinary Medicine
University of Zagreb
Zagreb, Croatia

Ingrid Miller
Institute of Medical Biochemistry
University of Veterinary Medicine
Vienna, Austria

ISBN 978-3-319-69681-2   ISBN 978-3-319-69682-9 (eBook)
https://doi.org/10.1007/978-3-319-69682-9

Library of Congress Control Number: 2017962424

© Springer International Publishing AG 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Acknowledgments

We could not complete this book without a word of appreciation to all the authors who have actively contributed to this truly international project that included research teams from Argentina, Australia, Austria, Belgium, Canada, Denmark, France, Germany, Italy, the Netherlands, New Zealand, Norway, Portugal, Spain, Switzerland, Thailand, United Kingdom, and the United States. Their expertise and dedication to this project has been a key feature in successfully completing this book. Most of the authors of the different chapters of this book were members of COST action FA1002—Proteomics in Farm Animals, funded by the European Science Foundation (http://www.cost.eu/COST_Actions/fa/FA1002). The networking funds and the activities engendered by this COST action have also been a determinant factor in the publication of this book. We finally thank Springer's Editorial staff who have first encouraged us to embrace this project and have helped us to edit this book in all the necessary steps. This project could not have been completed without all these important contributions. To all, thank you very much.

About the Book

In the last 20 years, the development in the so-called post-genomic tools, such as proteomics, transcriptomics, or metabolomics, has been immense. They have allowed important scientific advances in different areas of life sciences, including animal and veterinary sciences. This book focuses on proteomics and its use in farm animal research.

The first part of this book introduces and explains in detail the major principles and techniques associated with proteomics including the differences between gel-free and gel-based proteomics, as well as the major challenges facing mass spectrometry and bioinformatics. Furthermore, we provide an insight into how specific but common problems associated with farm animal samples may be solved.

The second part of the book has a more illustrative nature, highlighting important achievements in different areas of research within the animal science. These include for instance proteomics in blood and associated fluids, skeletal muscle, wool, mammary gland, colostrum research, and dairy products or proteomics in adipose and hepatic tissues to name but a few. The overall objective of this part of the book is to illustrate the importance of using proteomics in farm animal science.

Finally, the book concerns also the uses of other post-genomic tools, namely, transcriptomics and metabolomics, and how they interact with proteomics in a systems biology approach.

The entire book has been built around a philosophy that aims to bring novel users of proteomics into the field and to consider or continue using proteomics in their research projects dedicated to farm animal and veterinary sciences. Therefore, concepts and technologies herein described are easy to access and easy to follow and understand, rendering this book particularly suitable for all members of the farm animal research community: students, professors, researchers, technicians, and regulatory bodies.

Contents

Proteomics in Domestic Animals on a Farm to Systems Biology Perspective: Introductory Note 1
André M. de Almeida, Ingrid Miller, and P. David Eckersall

Considerations for Farm Animal Proteomic Experiments: An Introductory View Gel-Based Versus Non-gel-Based Approaches 7
John D. Lippolis and Jarlath E. Nally

Quantitative Gel Electrophoresis 17
Victoria J. Mansour and Jens R. Coorssen

Sample Preparation for 2DE Using Samples of Animal Origin 37
Hugo Osório, André M. de Almeida, and Alexandre Campos

Gel-Free Proteomics 55
Dominique Baiwir, Paolo Nanni, Susanne Müller, Nicolas Smargiasso, Denis Morsa, Edwin De Pauw, and Gabriel Mazzucchelli

Proteomic Research in Farm Animal Serum and Plasma 103
Laura Soler and Ingrid Miller

Proteomic Research in Urine and Other Fluids 121
Gloria Isani, Enea Ferlizza, and Jarlath E. Nally

Colostrum Proteomics Research: A Complex Fluid with Multiple Physiological Functions 149
Lorenzo E. Hernández-Castellano, Craig R. Baumrucker, Josef Gross, Olga Wellnitz, and Rupert M. Bruckmaier

Proteomics in Milk and Dairy Products 169
Alessio Soggiu, Paola Roncada, and Cristian Piras

Proteomics in Skeletal Muscle Research 195
Eva Veiseth-Kent, André M. de Almeida, Brigitte Picard, and Kristin Hollung

Proteomics and the Characterization of Fatty Liver Metabolism in Early Lactation Dairy Cows 219
Björn Kuhla and Klaus L. Ingvartsen

Proteomics Research in the Adipose Tissue 233
Fabrizio Ceciliani, Cristina Lecchi, Jeanne Bazile, and Muriel Bonnet

Proteomics and Mammary Gland Research in Dairy Species 255
André M. de Almeida and Peter David Eckersall

Proteomics in Wool and Fibre Research 281
Jeffrey E. Plowman

Proteomics of Meat Products 297
Gianluca Paredi, Federica Mori, and Andrea Mozzarelli

Proteomics in Fish and Aquaculture Research 311
Pedro M. Rodrigues, Samuel A.M. Martin, Tomé S. Silva, Surintorn Boonanuntanasarn, Denise Schrama, Márcio Moreira, and Cláudia Raposo

The Use of Proteomics to Study Biomarkers of Stress and Welfare in Farm Animals 339
Anna Marco-Ramell, Ana M. Gutiérrez, Antonio Velarde, Jose J. Cerón, and Anna Bassols

Bioinformatics Support for Farm Animal Proteomics 361
Aivett Bilbao and Frédérique Lisacek

Peptidomics on Farm Animal Research 387
Silvina Fadda, Enrique Sentandreu, and Miguel Angel Sentandreu

Studying the Animal Transcriptome: State of the Art and Challenges in the Context of Animal and Veterinary Sciences 421
José Ricardo Parreira and Susana de Sousa Araújo

NMR Metabolomics *pari passu* with Proteomics: Two Relevant Tools for Animal Sciences Combined 447
Mariana Palma, André M. de Almeida, and Manolis Matzapetakis

Omics and Systems Biology: Integration of Production and Omics Data in Systems Biology 463
Kasper Hettinga and Lina Zhang

Proteomics in Domestic Animals on a Farm to Systems Biology Perspective: Final Remarks and Future Prospects 487
André M. de Almeida, Ingrid Miller, and P. David Eckersall

Proteomics in Domestic Animals on a Farm to Systems Biology Perspective: Introductory Note

André M. de Almeida, Ingrid Miller, and P. David Eckersall

In the post-World War II era and with the objective to increase food production, particularly in high-value products such as dairy, meat, aquaculture products or eggs, animal production has become increasingly intensified. Indeed, aiming to obtain optimal economic results and improvements in genetics/selection, nutrition, mechanization, management and housing has led animal production and animal science to what they are today, a highly efficient industry that operates in a global market where animal products are traded across the globe with major transport and environmental costs associated.

Over seven decades later, the challenges faced by animal production are enormous. Perhaps the most significant is the increase in world population. From an estimated 7.6 billion when this book is being written (July 2017), the number of human beings on planet earth is expected to rise to nearly 10 billion in 2050 and over 11 billion in 2100 (United Nations World Population Prospects). The main part of such growth will be in developing and emerging countries, particularly in Asia, Africa and to a lesser extent Latin America. This implies first of all the need for an increase in food production such as cereals and other staple foods that, with economic growth, will be followed by an increasing demand in animal products. This will in turn cause a demand in feedstuffs (such as cereals and legumes) for

A.M. de Almeida (✉)
Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, Lisbon, Portugal
e-mail: aalmeida@isa.ulisboa.pt

I. Miller
University of Veterinary Medicine Vienna, Vienna, Austria
e-mail: Ingrid.Miller@vetmeduni.ac.at

P.D. Eckersall
School of Veterinary Medicine, University of Glasgow, Glasgow, Scotland, UK

Faculty of Veterinary Medicine, University of Zagreb, Zagreb, Croatia
e-mail: david.eckersall@glasgow.ac.uk

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_1

animal feeding and increased intensification, similarly to that which has happened in industrialized nations over the last 70 years. Interestingly, in industrialized countries, the trend will be the opposite, with decreasing populations that will in turn cause a likely decrease in the food demand, particularly staple foods, while certain sectors such as organic or certified (PDO, protected designation of origin, or PGI—protected geographical indication) will increase their proportional importance. Another equally important challenge is the sustainability of animal production. Indeed and in both industrialized and developing countries, the environmental footprint associated with animal production activities is too big to be sustainable. These occur in the form of water pollution from large-scale swine, dairy or feedlots, soybean and corn farming-associated deforestation in tropical countries, high levels of water consumption, the production of greenhouse gases by cattle and the road/boat/airfreight transport of animal products (e.g. wool, beef, mutton) or production factors (soybean, corn, etc.). A third major challenge is related to animal welfare and the existence and maintenance of current animal production systems. Indeed, most systems were first developed in the 1950s when animal welfare was not a major concern of the public opinion. Since then, numerous changes have occurred, and perhaps the most striking are the production conditions of laying hens that, upon public opinion and negative press, were banned in the European Union in 2012, are already banned in some American states and will likely be banned in the majority of the industrialized world. Finally, global warming and the advance of desertification in certain areas of the globe, such as in California, the Mediterranean basin, Australia, Chile, Argentina and Southern Africa, will necessarily lead to a regression in historically important areas for crop and animal production. Furthermore, global warming will likely lead to longer summers and temperature increase in presently temperate regions, hence creating the conditions for tropical diseases to move towards the poles. This problem is worsened by the constant flow of people, animals and animal products across the globe that makes biosecurity law enforcement particularly difficult.

All such challenges will require immediate responses from farmers, consumers, retail commerce, legislators, animal production technicians and, of course, the animal and veterinary sciences academic community. Solutions for these problems or most likely for several combinations of these challenges will have to be proposed, tested and implemented. In the last 20 years, animal production has moved to a completely different context. Indeed, aspects such as animal selection, reproduction and physiology have moved from a compartmental perspective to one more focused on changes at the molecular level that would in turn be related to classical disciplines. The sequencing of entire genomes (e.g. cattle, chicken, pig or sheep) in recent years is an impressive landmark for animal science that has opened important avenues in the study of animal biology and production. Therefore, in the last two decades, farm animal science has moved far beyond simple genome sequencing and gene function. Indeed, the use of post-genomics tools such as the omic sciences has allowed an unparalleled approach to gene function and consequently an overall view on how organisms (and in this case farm animal species) work and how they react to particular conditions. Omics include several subdisciplines. These include

transcriptomics, proteomics, lipidomics and metabolomics. Each of these subdisciplines is dedicated to a particular aspect. For instance, proteomics may be defined as the science that studies the proteome or the proteins present in a given organelle, cell, tissue, fluid, organ, organism or population, as well as the post-translational modifications (PTMs) to which these proteins are subjected to. By analogy other omics are dedicated to particular classes of compounds: transcriptomics for the transcriptome or all RNAs in a given organelle, cell, tissue, fluid, organ, organism or population, lipidomics for the lipids and metabolomics for the metabolites. Life science studies frequently involve a multi-omics approach, frequently defined as the systems biology approach.

Being in their vast majority proteinaceous material of great value in human nutrition and other uses, farm animal products such as dairy, meat, eggs, wool and honey have an enormous potential to be analysed using proteomics in order to provide greater insight and understanding of their highly complex roles in vivo and as they are processed postharvest. This discipline is, in the opinion of the authors of this book, the most significant of the post-genomics tools, particularly in the context of animal science. Indeed, the protein synthetic process underlies all of the important aspects of animal growth and production, and therefore the potential for proteomics is enormous. However, the value of proteomics in animal sciences is not restricted to product and production issues. Proteins are the key to understanding biological processes in cells and tissues, being vital to most cellular functions but also having essential roles in the extracellular compartments. Using proteomics to identify and quantify protein changes, interactions and modifications will become an ever more essential tool to use in the study of the physiology and pathophysiology of farm animals.

Proteomics itself has also made impressive advances in the last two decades. Initial proteomics approaches were conducted essentially using two-dimensional electrophoresis, in the beginning, and also lacking methods for protein identification, until mass spectrometry gained importance through increased sensitivity and protein validation methods like Western blot were developed. Today's proteomics is more often based on the mass spectrometric approaches for both protein identification and quantitation, in many cases combined with previous liquid chromatographic or other sample pre-fractionation steps. In recent years validation has also moved from the classical molecular methods to mass spectrometry-based methods, provided that protein databases are extensive and of good quality. Gel-free methods, once set up, are easy to implement and tend to work very well for the majority of biological samples arising from studies with farm animals and have the benefit of high degree of automation and throughput. However, the instrumentation for such studies is extremely expensive and difficult to implement and requires ever-growing costs, normally not compatible with traditionally small budgets such as those of animal and veterinary sciences. Gel-based proteomics is less cost-intensive, though with lower throughput and often needs a higher input of practical skills and method optimization. It is still one of the most important tools in farm

animal proteomics. In practice, results of both approaches have proven to be complementary, leading to different sets of target proteins, and their combination is frequently suggested as leading to the best results, particularly in poorly studied organisms. Finally, the combination of proteomics with other post-genomic tools, particularly transcriptomics and metabolomics, in the so-called systems biology approach is taking animal and veterinary sciences to a completely new level.

The importance of proteomics and other post-genomics tools in animal sciences has been recognized by the academic community, funding bodies and specialized scientific publications. The use of proteomics has numerous benefits in this field of research and in certain areas such as dairy science, meat science, lactation biology and wool and fibre production, to name but a few; such benefits are particularly important. They include, for instance, animal adaptation to nutritional changes, product characterization and certification and animal welfare monitoring.

The use of proteomics in animal sciences is the subject of this book. We have edited it with two major goals: (1) to initiate scientists that have never or have rarely used proteomics to the field and (2) to illustrate the benefits and potential for proteomics approaches to specific areas of animal science. A third goal, about interaction of proteomics with other post-genomics tools, specifically transcriptomics and metabolomics, has also been included as the combination of omic technologies and is likely to have major contributions in the not too distant future. Consequently, the initial chapters of this book are focused on methodology aspects and on the different approaches in farm animal proteomics studies. We provide an emphasis on the gel-based and gel-free dichotomy, sample preparation and basics on the use of mass spectrometry for proteomics studies. These technical chapters were written by experts in protein science and mass spectrometry, not necessarily working specifically in animal science. They have however an important connection to the field, being aware of specificities of animal science that make it very different from other sciences where proteomics is widely used, like plant sciences or cancer research. These more technical chapters are followed by a set of chapters that address the importance of proteomics in specific fields. Chapters on adipose tissue proteomics, aquaculture proteomics, mammary gland proteomics, muscle and meat proteomics, liver proteomics, blood and derived fluids proteomics or colostrum proteomics are included in this part of the book. These chapters were written by animal and veterinary science researchers that have for long used proteomics and have become experts in these specific areas. After this section, we continue the book with two chapters on proteomics-associated post-genomics tools, specifically transcriptomics and metabolomics (NMR—nucleic magnetic resonance based) and how they interact with proteomics. Finally, though before the conclusions, we end the book with a specific chapter on the interaction of the different omics in a systems biology perspective.

We hope that this book will become an important international source of proteomics literature and inspiration for animal scientists in the years to come. Above all we trust that this book will encourage our fellow animal and veterinary

scientists to adopt proteomics as a research discipline and will finally use it in the framework of their research projects.

Lisbon, Vienna and Zagreb, 1st of August 2017

André M. de Almeida

Ingrid Miller

David Eckersall

Considerations for Farm Animal Proteomic Experiments: An Introductory View Gel-Based Versus Non-gel-Based Approaches

John D. Lippolis and Jarlath E. Nally

Abstract Preparing for a proteomic experiment will require a number of important decisions. Because of the complexity of most samples, one of the first important decisions is how to separate proteins prior to analysis by the mass spectrometer. There are two basic approaches; the first approach is gel-based electrophoresis that typically separate proteins based on molecular weight and/or isoelectric point. The second approach is non-gel-based or liquid chromatography that typically separates peptides based on hydrophobicity. We discuss some of the pros and cons of each separation method to allow the proper alignment of research objectives and scientific methodologies.

1 Introduction

Proteomics includes the identification and characterization of the protein content from complex biological samples. Although the bovine and human genomes contain more than 20,000 genes, they likely encode as many as a million distinct proteins due to alternative gene splicing and posttranslational modifications (Elsik et al. 2009; Kelemen et al. 2013; Yang et al. 2016). Such biological complexity illustrates why there is no single standard analytical method for preparing protein samples for a proteomic experiment. The choice of analytical methods will depend on the specific experimental question, the type of sample for analysis, and methodologies available. Analytical methods can be used independently or as complementary approaches used in tandem. There are three basic steps to a proteomic

J.D. Lippolis (✉)
Ruminant Diseases and Immunology Research Unit, National Animal Disease Center, Agricultural Research Service (ARS), United States Department of Agriculture (USDA), Ames, IA, USA
e-mail: john.lippolis@ars.usda.gov

J.E. Nally
Bacterial Diseases of Livestock Research Unit, National Animal Disease Center, Agricultural Research Service (ARS), United States Department of Agriculture (USDA), Ames, IA, USA

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_2

experiment; the first is sample preparation, then sample separation prior to mass spectrometry, and, finally, protein detection and identification. The objective of this chapter is to provide researchers who are planning proteomic experiments insight into the basic considerations necessary to achieve their desired scientific outcomes. We will discuss the reasons why sample preparation is critical to a successful proteomic experiment. In addition, we will explain when protein separation protocols are necessary and help readers understand the pros and cons of gel-based versus gel-free separation as they determine the best application for their unique circumstances.

1.1 Sample Preparation

In proteomics, sample preparation is critical. The number of proteins identified, and confidence in the data, depends on the quality and complexity of starting material (Bodzon-Kulakowska et al. 2007). Careful attention to possible sources of contamination is important prior to a successful proteomic experiment. For example, an experiment to define the proteome of a bacterium such as *Escherichia coli* grown in a simple culture media might comprise preparation as straightforward as multiple rounds of centrifugation and washing to remove media components. In contrast, analyses of protein expression by *E. coli* when grown in a complex biological fluid, such as milk during bovine mastitis, will require protocols to first enrich for proteins of *E. coli* and eliminate as much contaminating milk protein as possible (Lippolis et al. 2009). Failure to eliminate contaminants will result in an experiment in which the primary goal of identifying bacterial proteins is compromised by the identification of milk proteins.

Proteomic experiments that use tissues are more complex compared to those that use single cells. If a sample is a tissue, size and location of the samples are important considerations to allow for a representative and consistent mixture of cell types. Careful attention to possible sources of variation is important to identify and control.

Protein sample preparation may also include various specific isolation or enrichment techniques that utilize physical characteristics of the protein or peptides. For example, the isolation and identification of major histocompatibility complex (MHC)-associated peptides can illustrate this idea (Hunt et al. 1992; Lippolis et al. 2002). MHC molecules are cell surface proteins that combine with peptides derived from intracellular protein degradation. The combination of the MHC and peptide fragments allows the immune system to survey intracellular proteins and detect foreign proteins. To isolate MHC-associated peptides, cell lysates are incubated with an MHC-specific antibody bound to protein A sepharose beads. The antibody bound to beads allows for the MHC/peptide complex to be separated from other proteins in the cell lysate by centrifugation and multiple washing steps. The peptides are then separated from the antibody and MHC complex by a simple size exclusion filter. The result is a clean sample of a

complex mixture of peptides (thought to be as many as 10,000 unique peptides) that were bound to MHC molecules.

Another important tool in sample preparation is the depletion of highly abundant proteins. In plasma, ten proteins make up approximately 90% of the total protein (Cho 2007). If nothing is done to reduce the levels of those abundant proteins, the number of proteins identified in that sample will be relatively few. Techniques such as N-linked glycopeptide enrichment, cysteinyl-peptide enrichment, size fractionation, combinatorial hexapeptide libraries, and immunoaffinity column depletion of abundant serum proteins can be used to enrich for proteins of interest (Whiteaker et al. 2007; Bandow 2010).

Finally, meaningful proteomic experiments must be based on consistent sample preparation techniques. Therefore, preliminary experiments to demonstrate reproducible protein isolation and preparation are critical to subsequent experimental success. For example, analyses of sample preparations by SDS-PAGE to ensure an absence of protein degradation due to experimental artifact or protease contamination can be an important step to ensure a successful experiment.

1.2 Why Sample Separation Prior to Mass Spectrometry Is Necessary

A proteomic study typically requires that complex biological samples be subjected to some type of protein or peptide separation, which also serves to improve mass spectrometry function. Mass spectrometers can efficiently obtain sequence information from peptides up to ~20 amino acids long (Steen and Mann 2004). For this reason, most proteomic experiments use a “bottom-up” approach whereby the protein mixture is first digested with a protease (e.g., trypsin) to generate such peptides. However, the digestion of thousands of proteins results in potentially millions of peptides. The first step of peptide identification begins when ionized peptides enter the mass spectrometer, and the mass-to-charge ratio of all the peptides is determined in what is referred to as the MS scan. The mass spectrometer will sequentially isolate the most abundant ions and individually fragment them. The mass-to-charge ratio of the resulting peptide fragments is then obtained (MSMS scan). It is from the daughter ions that computer algorithms predict the peptide sequence and identify the parent protein. After multiple MSMS scans, an MS scan reexamines the input sample to determine the next set of most abundant ions. An ion exclusion list ensures that the mass spectrometer does not isolate the most abundant ion repeatedly. This cycle is repeated throughout the mass spectrometer analytical run.

If the protein sample was obtained from a single spot excised from a two-dimensional (2-D) polyacrylamide gel, the number of proteins will likely be small and relatively few peptides generated. In this case, the mass spectrometer can sequence all the abundant ions and many of the less abundant ions, for not all

peptides will ionize equally. The end result will be a large number of peptide sequences identified for each protein in the sample preparation. The greater the coverage of the protein identified, the greater the confidence in the identification of the protein. In contrast, if a more complex proteome sample is digested and peptides are minimally fractionated before ionization and entry into the mass spectrometer, then thousands of ions would be identified in each MS scan. In this case, the cycle of MS and MSMS scans would only analyze a small fraction of the total ions and only the most abundant. Thus, complexity and dynamic range of a sample will limit results to the most abundant and fail to identify low abundance proteins of interest (Huber et al. 2003; Brunet et al. 2003; Reinhardt and Lippolis 2006; Zolotarjova et al. 2008). Hence, sample separation prior to MS allows the user to control the number of ions entering the mass spectrometer at any given time and provides the instrument the necessary time to analyze all potential peptide ions.

Proteomic experiments can range from the quantification of hundreds of specific proteins to the identification of thousands of unknown proteins. The most basic is to determine the identity of all proteins that exist in a sample. This experimental approach is typically referred to as shotgun proteomics, whereby proteins are identified in a random manner and often without any preexisting knowledge of protein content. As discussed above, significant limitations in analyzing a complex proteome include the number of proteins in the sample and the dynamic range of the proteins in the mixture. Some proteomes can be dominated by a small number of very abundant proteins. For example, nearly half of all protein in plasma is albumin, and nearly 90% of the complete protein content comprises only ten proteins (Cho 2007). Differences in protein amounts in a complex proteome between the highly abundant and lesser abundant proteins can be greater than seven orders of magnitude (Stasyk and Huber 2004). Without separation of protein samples into fractions, the mass spectrometry will primarily identify the abundant proteins and few, if any, of the less abundant proteins.

The power of shotgun proteomics is exemplified by experiments that aim to identify the complete proteome of an organism. In 2014, two groups of researchers presented their work to define the human proteome (Kim et al. 2014; Wilhelm et al. 2014). One group reported detection of 84% of the approximately 21,000 protein-coding genes in humans (Kim et al. 2014). This was achieved by extracting proteins from 30 different human tissue or primary cell types. Protein samples were initially separated in SDS-PAGE gels, and each lane of the gel was cut into multiple gel slices. Proteins contained in gel slices were digested with trypsin and peptides extracted from the each gel slice for separation on a reverse-phase HPLC column directly connected to a mass spectrometer (LC-MS/MS). The result of this work was approximately 25 million high-resolution mass spectra acquired from more than 2000 LC-MS/MS runs. The run time for a single LC-MS/MS would be approximately 2 h; thus, the total run time of this experiment is nearly 6 months of uninterrupted mass spectrometer time. Although 84% of predicted proteins were identified, it was not reported how many of these proteins were identified by the presence of a single peptide. Proteins identified, or quantified, using a single peptide are typically done so with much less confidence than multiple peptides. Perhaps,

more importantly, it also makes it difficult to determine the biological significance of such identifications.

1.3 Which Sample Separation Is Best?

The choice of method of sample separation is based on the experimental goals, the expectations of the data, and the resources available (Monteoliva and Albar 2004; Baggerman et al. 2005; Jafari et al. 2012; Abdallah et al. 2012). There are two major types of sample separation techniques used in proteomic experiments. The first is gel-based separation, which is typically two-dimensional gel electrophoresis. The second type of sample separation is non-gel based or various forms of liquid chromatography. Details about gel-based and non-gel-based separation techniques will be detailed in subsequent chapters in this book. However, we will briefly detail the pros and cons of each type of sample separation technique (Fig. 1).

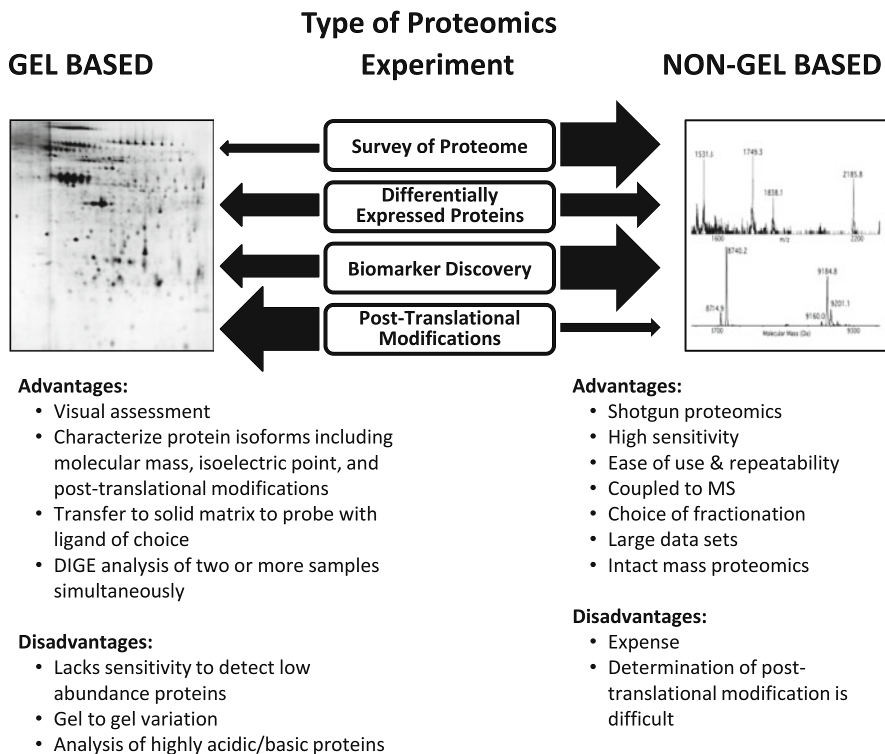

Fig. 1 Advantages and disadvantages of gel based and non-gel based proteomic experiments

1.4 Gel-Based Proteomics

Since individual proteins can be differentiated by their respective isoelectric point (p*I*) and molecular mass, the use of gel-based approaches provides a powerful method to separate complex protein samples. In 2-D gel electrophoresis, protein samples are solubilized prior to separation on immobilized pH gradients (IPG) in the first dimension, according to p*I*, and in the second dimension by sodium dodecyl sulfate polyacrylamide gels (SDS-PAGE) according to molecular mass. Highly resolved proteins can then be detected on protein gels with an array of protein stains. Proteins of interest can then be excised, digested, and identified by mass spectrometry.

Advantages of gel-based proteomics include the ability to obtain highly resolved individual protein spots from complex samples since IPG strips now come in lengths of up to 24 cm with increasingly more defined pH gradients. Mass spectrometry data of excised protein spots can then be correlated to the identity of specific protein isoforms, e.g., the identification of posttranslational modifications (Witchell et al. 2014). Methodologies exist to allow direct gel-to-gel comparisons to quantitate differential protein expression between samples, e.g., 2-D DIGE (Schuller et al. 2015). The separation of proteins of a gel-based platform also allows for their transfer to membranes for more defined experimental questions, e.g., to identify protein antigens that react with convalescent sera (Nally et al. 2005; Monahan et al. 2008) or to identify proteins reactive with a defined ligand. Conversely, the separation of proteins is often considered laborious and more of an art form. Finally, membrane-associated proteins can be difficult to solubilize prior to separation, so an appropriate zwitterionic detergent should be used (Nally et al. 2005).

1.5 Non-gel-Based Proteomics

Liquid chromatography has a long history of use to separate proteins and peptides based on various physical properties. With the advent of electrospray ionization (ESI), the ability to directly link chromatography to mass spectrometry has made this method especially convenient, reproducible, and powerful. The advantage of a non-gel approach for proteomics is the ease of sample separation techniques. Once proteins are solubilized and digested with trypsin, the separation techniques are straightforward and very reproducible. However, the more complex the proteome, the greater separation that will be needed to get the depth of coverage desired. One option is the use of very long HPLC columns. A typical HPLC column connected to a mass spectrometer is 10 cm long. Some investigators have used 50 cm columns with shallow gradients to obtain single runs that achieve near complete coverage of a yeast proteome (Nagaraj et al. 2012). Alternatively, biphasic columns can be used that combine two HPLC packing materials into one column. The most common

HPLC is the use of strong-cation exchange (SCX) beads packed in front of a typical reverse-phase (RP) packing material. A sample is injected onto the SCX, and subsequent injections of increasingly concentrated salt solutions move a fraction of the peptides from the SCX matrix to the RP matrix. In between each salt injection, a normal RP HPLC gradient is run to separate and analyze the peptides bound to the RP matrix. This multidimensional protein identification technology approach is called MudPit (Washburn et al. 2001). Very long columns or MudPit proteomic experiments can identify thousands of proteins in a sample. In addition, liquid chromatography is compatible with methodologies that give quantitative information [e.g., amine-reactive isobaric tagging reagents (Ross et al. 2004)].

Subsequent to the shotgun approach described above is the possibility of specific examination of protein quantification by mass spectrometry. Once proteins are identified, specific peptides from each protein can be used as markers for that peptide. A mass spectrometer can then be programmed to identify only those peptides specified using the peptide's mass-to-charge ratio (m/z). In this technique, referred to as selected reaction monitoring (SRM) or multiple reaction monitoring (MRM), specific peptides can be easily isolated, identified, and quantified. This technique can examine and quantify approximately 500 peptides, or 125 proteins, per MRM experiment (Shi et al. 2016). MRM experiments require significant assay development time.

As previously mentioned, the process to digest proteins with a specific protease for analysis by mass spectrometry is referred to as bottom-up proteomics. In contrast, top-down proteomics is the analysis of whole proteins or protein complexes by mass spectrometry (Han et al. 2008; Toby et al. 2016). Top-down proteomics requires high-resolution instruments to accurately evaluate and distinguish whole proteins. Constant innovation and improvement to mass spectrometers is making top-down proteomics more accessible. Top-down techniques have been used to determine the molecular weight of intact proteins (Nally et al. 2005). However, significant work is needed in the field before quantitation of whole proteomes can be analyzed by top-down methods.

The disadvantage of the non-gel-based approach is that when a proteome is digested, specific peptides of interest can get lost in the complex mixture of thousands of peptides. A proteome of potentially tens of thousands of proteins digested into millions of peptides with expression differences that can be seven orders of magnitude makes identification of this complex mixture difficult.

1.6 Protein Detection and Identification by Mass Spectrometry

Ionization of peptides is the first step in mass spectrometry. The two most frequently used ionization methods are electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) (Yates 1998). ESI is most common

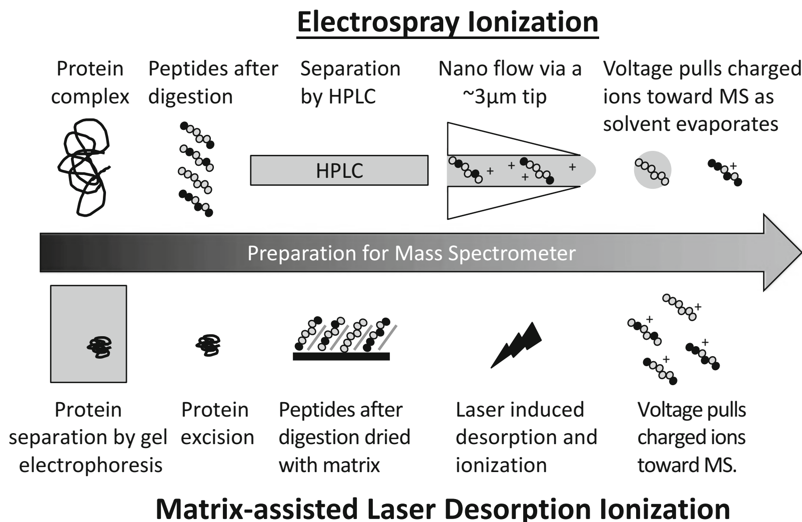

Fig. 2 Peptide ionization methods used in proteomic experiments

because it can be linked directly to a nanoliter per minute HPLC system. Peptides in volatile HPLC solutions are protonated with an acid and sprayed toward the mass spectrometer entry through a small emitter (~3 µm inner diameter). Charged peptides can enter and be manipulated by the mass spectrometer. MALDI requires the mixing of the peptide with a UV-absorbing molecule and the formation of crystals. A laser is used to strike the crystalline structures resulting in the sublimation of the matrix and the ionization and release of the associated peptides. At this point, the charged peptides can enter and be manipulated by the mass spectrometer. The key difference between the two ionization methods is that MALDI used a static sample. Therefore, the sample to be analyzed by MALDI-MS needs to contain only a few proteins. MALDI-MS is typically associated with the analysis of protein spots from a 2-D gel. The advantage of MALDI-MS is a relatively easy and fast analysis of sample. Conversely, ESI-MS can be used to identify thousands of proteins in a single sample. The quality and quantity of data obtained in an ESI-MS experiment will depend on the ability to separate peptides in the HPLC to allow the mass spectrometer the time necessary to isolate and fragment each peptide in the mixture (Fig. 2).

1.7 Conclusions

The type of proteomic research question will often determine the type of protein separation suited for the experiment. In addition, the choice between gel and gel-free protein separation in a proteomic experiment may simply be a matter of

preference, convenience, or the possession of specific equipment or mass spectrometer. One laboratory may possess and have the necessary skills with HPLC equipment, whereas another laboratory may prefer working with 2-D gel equipment. Each type of protein separation can be used independently or in combination. In most cases, some sort of protein or peptide separation is necessary to achieve the goals of a proteomic experiment. Careful consideration of the strengths of each type of separation and how they align with project goals prior to the experiment will help to ensure a successful outcome.

References

Abdallah C, Dumas-Gaudot E, Renaut J et al (2012) Gel-based and gel-free quantitative proteomics approaches at a glance. Int J Plant Genomics 2012:1–17. https://doi.org/10.1155/2012/494572

Baggerman G, Vierstraete E, De Loof A, Schoofs L (2005) Gel-based versus gel-free proteomics: a review. Comb Chem High Throughput Screen 8:669–677

Bandow JE (2010) Comparison of protein enrichment strategies for proteome analysis of plasma. Proteomics 10:1416–1425. https://doi.org/10.1002/pmic.200900431

Bodzon-Kulakowska A, Bierczynska-Krzysik A, Dylag T et al (2007) Methods for samples preparation in proteomic research. J Chromatogr B 849:1–31. https://doi.org/10.1016/j.jchromb.2006.10.040

Brunet S, Thibault P, Gagnon E et al (2003) Organelle proteomics: looking at less to see more. Trends Cell Biol 13:629–638

Cho WCS (2007) Proteomics technologies and challenges. Genomics Proteomics Bioinformatics 5:77–85. https://doi.org/10.1016/S1672-0229(07)60018-7

Elsik CG, Tellam RL, Worley KC et al (2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science 324:522–528. https://doi.org/10.1126/science.1169588

Han X, Aslanian A, Yates JR (2008) Mass spectrometry for proteomics. Curr Opin Chem Biol 12:483–490. https://doi.org/10.1016/j.cbpa.2008.07.024

Huber LA, Pfaller K, Vietor I (2003) Organelle proteomics: implications for subcellular fractionation in proteomics. Circ Res 92:962–968. https://doi.org/10.1161/01.RES.0000071748.48338.25

Hunt DF, Henderson RA, Shabanowitz J et al (1992) Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. Science 255:1261–1263

Jafari M, Primo V, Smejkal GB et al (2012) Comparison of in-gel protein separation techniques commonly used for fractionation in mass spectrometry-based proteomic profiling. Electrophoresis 33:2516–2526. https://doi.org/10.1002/elps.201200031

Kelemen O, Convertini P, Zhang Z et al (2013) Function of alternative splicing. Gene 514:1–30. https://doi.org/10.1016/j.gene.2012.07.083

Kim M-S, Pinto SM, Getnet D et al (2014) A draft map of the human proteome. Nature 509:575–581. https://doi.org/10.1038/nature13302

Lippolis JD, Bayles DO, Reinhardt TA (2009) Proteomic changes in Escherichia coli when grown in fresh milk versus laboratory media. J Proteome 8:149–158. https://doi.org/10.1021/pr800458v

Lippolis JD, White FM, Marto JA et al (2002) Analysis of MHC class II antigen processing by quantitation of peptides that constitute nested sets. J Immunol 169:5089–5097

Monahan AM, Callanan JJ, Nally JE (2008) Proteomic analysis of Leptospira interrogans shed in urine of chronically infected hosts. Infect Immun 76:4952–4958. https://doi.org/10.1128/IAI.00511-08

Monteoliva L, Albar JP (2004) Differential proteomics: an overview of gel and non-gel based approaches. Brief Funct Genomic Proteomic 3:220–239. https://doi.org/10.1093/bfgp/3.3.220
Nagaraj N, Kulak NA, Cox J et al (2012) System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. Mol Cell Proteomics 11:M111.013722. https://doi.org/10.1074/mcp.M111.013722
Nally JE, Whitelegge JP, Aguilera R et al (2005) Purification and proteomic analysis of outer membrane vesicles from a clinical isolate of Leptospira interrogans serovar Copenhageni. Proteomics 5:144–152. https://doi.org/10.1002/pmic.200400880
Reinhardt TA, Lippolis JD (2006) Bovine milk fat globule membrane proteome. J Dairy Res 73:406–416. https://doi.org/10.1017/S0022029906001889
Ross PL, Huang YN, Marchese JN et al (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169. https://doi.org/10.1074/mcp.M400129-MCP200
Schuller S, Sergeant K, Renaut J et al (2015) Comparative proteomic analysis of lung tissue from guinea pigs with leptospiral pulmonary haemorrhage syndrome (LPHS) reveals a decrease in abundance of host proteins involved in cytoskeletal and cellular organization. J Proteome 122:55–72. https://doi.org/10.1016/j.jprot.2015.03.021
Shi T, Song E, Nie S et al (2016) Advances in targeted proteomics and applications to biomedical research. Proteomics 16:2160–2182. https://doi.org/10.1002/pmic.201500449
Stasyk T, Huber LA (2004) Zooming in: fractionation strategies in proteomics. Proteomics 4:3704–3716. https://doi.org/10.1002/pmic.200401048
Steen H, Mann M (2004) The ABC"s (and XYZ"s) of peptide sequencing. Nat Rev Mol Cell Biol 5:699–711. https://doi.org/10.1038/nrm1468
Toby TK, Fornelli L, Kelleher NL (2016) Progress in top-down proteomics and the analysis of proteoforms. Annu Rev Anal Chem 9:499–519. https://doi.org/10.1146/annurev-anchem-071015-041550
Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242–247. https://doi.org/10.1038/85686
Whiteaker JR, Zhang H, Eng JK et al (2007) Head-to-head comparison of serum fractionation techniques. J Proteome Res 6:828–836. https://doi.org/10.1021/pr0604920
Wilhelm M, Schlegl J, Hahne H et al (2014) Mass-spectrometry-based draft of the human proteome. Nature 509:582–587. https://doi.org/10.1038/nature13319
Witchell TD, Eshghi A, Nally JE et al (2014) Post-translational modification of LipL32 during Leptospira interrogans infection. PLoS Negl Trop Dis 8:e3280. https://doi.org/10.1371/journal.pntd.0003280
Yang X, Coulombe-Huntington J, Kang S et al (2016) Widespread expansion of protein interaction capabilities by alternative splicing. Cell 164:805–817. https://doi.org/10.1016/j.cell.2016.01.029
Yates JR (1998) Mass spectrometry and the age of the proteome. J Mass Spectrom 33:1–19. https://doi.org/10.1002/(SICI)1096-9888(199801)33:1<1::AID-JMS624>3.0.CO;2-9
Zolotarjova N, Mrozinski P, Chen H, Martosella J (2008) Combination of affinity depletion of abundant proteins and reversed-phase fractionation in proteomic analysis of human plasma/serum. J Chromatogr A 1189:332–338. https://doi.org/10.1016/j.chroma.2007.11.082

Quantitative Gel Electrophoresis

Victoria J. Mansour and Jens R. Coorssen

Abstract Gel electrophoresis is the most widely used technique for the analysis of protein samples, and there are a variety of methods that can be used to investigate single proteins as well as highly complex protein mixtures. To ensure reproducible and reliable separations of such samples and the resolution of distinct protein species, there has been a substantial amount of research dedicated to optimising methods to the refined techniques available today. There are thus a number of factors that have a marked influence on the practice of quantitative proteomics. The procedures and reagents involved in preparing a protein sample can have a significant effect on the composition of the proteome and/or its resolution by electrophoresis. Furthermore, since most proteins are colourless, a protein stain is required to detect the resolved proteome; thus, it is essential that the characteristics of the stain enable optimal detection regardless of protein type and/or concentration. Notably, to obtain reliable quantitative data, the approach by which images are acquired is equally important.

Keywords Gel electrophoresis • Protein stain • Fluorescence • Sample preparation • Quantitative analysis • Protein detection • Deep imaging • 1D/2D/3D gel electrophoresis

Abbreviations

1DE	One-dimensional electrophoresis
2DE	Two-dimensional electrophoresis
cCBB	Colloidal Coomassie brilliant blue

[†]V.J. Mansour née Gauci.

V.J. Mansour
School of Medicine, Western Sydney University, Penrith, NSW, Australia

J.R. Coorssen (✉)
Faculty of Applied Health Sciences and Faculty of Mathematics and Science, Brock University, St Catharines, ON, Canada
e-mail: jcoorssen@brocku.ca

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_3

CBB	Coomassie brilliant blue
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propane sulfonate
DIGE	Difference gel electrophoresis
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
LLD	Lowest limit of detection
LDR	Linear dynamic range
IEF	Isoelectric focussing
IPV	Inter-protein variability
IRF	Infrared fluorescence
MS	Mass spectrometry
MW	Molecular weight
NCCB	Neuhoff colloidal Coomassie brilliant blue
PAGE	Polyacrylamide gel electrophoresis
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
SR	SYPRO Ruby
TBP	Tributyl phosphine

1 Introduction

The cell is the basic structural, functional and biological unit for all known organisms. They are composed of a variety of biomolecules such as deoxyribonucleic acid (DNA), ribonucleic acid (RNA), proteins, carbohydrates, lipids and their metabolites. DNA carries the primary genetic code, and when transcribed into RNA, this code can then be translated to produce proteins; thus, changes to DNA and/or RNA can result in protein variations. As proteins are fundamental functional units, they are involved in a myriad of biological processes ranging from the transport of molecules across cell membranes to metabolism and cell signalling. This vast array of functions is further modulated by a large variety of post-translational modifications that further fine-tune the functions and localisations of these proteins. Thus, proteoforms or protein species (i.e. splice and post-translational variants, isoforms, and mutants) largely account for all the physiological (dys)functions that underlie healthy and disease phenotypes.

The entire complement of proteins expressed by a genome in a given biological sample, whether it be a whole organism, tissue, fluid, cell or organelle, was originally referred to as the proteome, a term introduced by Marc Wilkins in 1994 at the 2D electrophoresis meeting held in Siena, Italy. Changes to the proteome can be elicited in response to a vast range of external and internal stimuli and can include increased and/or decreased abundance of existing proteins, expression of additional proteins and/or cessation of specific proteins, as well as the introduction of post-translational modifications. Thus, proteomes are orders of magnitude more complex than

genomes, and most proteoforms cannot be 'predicted' from DNA or RNA sequence information. Thus, the term proteome has come to more realistically mean the full complement of proteoforms present in a given sample. Thus, to comprehensively study proteomes, a variety of techniques are employed—gel electrophoresis, mass spectrometry, liquid chromatography, X-ray crystallography, confocal microscopy and protein/antibody arrays—to understand the breadth of protein structure and function which is now more widely referred to as proteomics. Here we focus on sample preparation, gel-based separation and in-gel detection methods that facilitate the delivery of reliable quantifiable data.

2 Protein Sample Preparation

Before a proteome can be most effectively investigated, appropriate preparation techniques must be applied. These include sampling, sample handling, extraction and potential fractionation methods, and their use is highly dependent on the sample type and/or the research objective and may thus include approaches to simplify overall sample complexity as well as enhance detection of extremely low-abundance proteoforms (Coorssen and Yergey 2015). Since even a single cell consists of a complex assortment of proteoforms, it is critical that the buffers/solutions used ensure complete solubilisation of all species so as to ensure as complete an assessment of a proteome as possible (Mansour née Gauci et al. 2016). The most important features for an efficient solubilisation buffer are:

- A chaotropic agent to disrupt the non-covalent forces promoting protein secondary, tertiary and quaternary structures
- A detergent to promote solubilisation of the hydrophobic regions of proteins
- A reducing agent to prevent covalent bond formation between cysteine residues

If all these features are utilised, there is a much improved likelihood of analysing all the protein species within a given sample.

For gel-based proteomics, the Laemmli buffer system is commonly used for one-dimensional electrophoresis (1DE), utilising sodium dodecyl sulphate (SDS) as the detergent which, when used in combination with heat, generally ensures complete protein denaturation (Laemmli 1970; Cannon-Carlson and Tang 1997). For two-dimensional electrophoresis (2DE), the original buffer system introduced in 1975 (O'Farrel 1975) has undergone a number of significant improvements, all resulting in better protein solubilisation. These have mainly included substituting Nonidet P-40 with the zwitterionic detergent 3-[(3-cholamidopropyl) dimethylammonio]-1-propane sulfonate (CHAPS) and the inclusion of tributyl phosphine (TBP) and thiourea to complement β-mercaptoethanol/dithiothreitol (DTT) and urea (Perdew et al. 1983; Rabilloud 1998; Rabilloud et al. 1997; Herbert et al. 1998).

The method used for protein extraction can also have a considerable impact on the efficacy of solubilisation and thus proteome assessment. Conventional manual methods of homogenisation have been shown to be somewhat heterogeneous in

effect (Butt and Coorssen 2005b). This was overcome using automated frozen disruption, yielding more uniform homogenisation, with the frozen powdered sample providing greater surface area for protein solubilisation and thus improving overall protein yield (Butt and Coorssen 2005b). The additional benefits to this homogenisation method are that proteoform integrity is not compromised as the sample is kept at colder temperatures and its application isn't limited by the sample type (Butt and Coorssen 2005b; Gauci et al. 2013; Partridge et al. 2016).

As mentioned above, a single cell presents a complex variety of proteoforms of widely varying abundance; thus, complexity is markedly increased when working with samples from multicellular organisms. Although the means by which proteome accessibility and solubilisation are achieved is important, this complexity poses further complications for gel electrophoresis, or indeed any available analytical method. The most common approach to overcome this is to pre-fractionate the protein sample prior to analysis by 1DE, 2DE and/or mass spectrometry (MS), which can be done by either physical or chemical means. Methods used for pre-fractionation include:

- Centrifugation
  - Ultracentrifugation to produce total membrane and soluble fractions (Pasquali et al. 1997; Molloy et al. 2000; Butt and Coorssen 2005b; Butt et al. 2006)
  - Differential detergent or density gradient fractionation to isolate subcellular fractions (Hurkman and Tanaka 1986; Fialka et al. 1997; Ramsby et al. 1994; Molloy et al. 1998; Bernocco et al. 2008; Völkl et al. 1997)
  - ProteoMiner™ or equaliser beads to selectively bind proteins that are then collected by centrifugation (Thulasiraman et al. 2005; Boschetti and Righetti 2008; Castagna et al. 2005)
- Solvent extraction of hydrophobic proteins (Molloy et al. 1999; Ferro et al. 2000; Schröder and Hasilik 2006)
  - Isolating subgroups of membrane proteins (Pasquali et al. 1997; Taylor et al. 1997; Molloy et al. 2000; Pedersen et al. 2003)
- Chromatography to enrich or deplete specific types of proteins (Fountoulakis et al. 1999a,b; Wissing et al. 2000; Ghosh et al. 2004; Smith et al. 2004; Larsen et al. 2005; Wang et al. 2003; Wasinger et al. 2005; Shen et al. 2005)
- Electrophoretic separation of proteins (Hannig 1978; Corthals et al. 1997; Völkl et al. 1997; Herbert and Righetti 2000; Görg et al. 2002)

Determining the most appropriate protein sample preparation methods to use will not only depend on the research objective but the benefits and/or disadvantages of the method(s) selected (Coorssen and Yergey 2015).

3 Gel-Based Resolving Technologies

Most simply, electrophoresis involves the separation (i.e. resolution) of a mixture of molecules in an electric field. The means by which these molecules migrate depends on their charge and molecular weight (MW). Thus, those proteins with a negative charge will move away from the cathode (−) and towards the anode (+), and those of high MW will migrate more slowly than lower MW molecules. However, depending on the protein and gel preparation and the electrophoresis conditions, the migration patterns can be 'tuned' with some selectivity.

A variety of mediums (e.g. cellulose acetate and filter strips, agarose, agar and starch gels) were initially used to resolve protein samples by electrophoresis, but it wasn't until the late 1950s that polyacrylamide gels became the preferred medium (Raymond and Weintraub 1959; Raymond and Wang 1960; Chrambach and Rodbard 1971, 1972). This gel matrix has a number of superior qualities such as reproducible pore formation; pore size can be small or large depending on acrylamide concentration, and once polymerised, it is transparent, stable, flexible and insoluble (and non-toxic). After widespread acceptance of polyacrylamide as the standard electrophoresis medium came the development of SDS-polyacrylamide gel electrophoresis (PAGE) or more commonly referred to as 1DE (Laemmli 1970; Cannon-Carlson and Tang 1997). By incorporating SDS into the protein preparation buffer, electrophoresis buffer and gel matrix polypeptides were maintained as single linear molecules. Also, as SDS conveys an overall negative charge, it promotes unidirectional movement towards the anode, and the MW for any polypeptide can be estimated with good reproducibility (Weber and Osborn 1969; Poduslo and Rodbard 1980). The Laemmli buffer system is highly efficient for the resolution of most proteins, except those with a molecular weight below ~10 kilodaltons (Schägger and von Jagow 1987). This is because it uses glycine as the trailing ion which migrates slowly during stacking and doesn't allow for small polypeptides to be resolved from the SDS zone (Schägger and von Jagow 1987). To overcome this, glycine was replaced by tricine and the sieving properties of the acrylamide gel adjusted enabling proteins of 1–100 kilodaltons to be highly resolved from the SDS zone (Schägger and von Jagow 1987).

While 1DE can be efficient at resolving proteins across a range of MWs, these are denaturing strategies that disrupt native protein conformations and interactions. However, it is also feasible to use 1DE in a native PAGE approach. Blue-native PAGE uses Coomassie brilliant blue (CBB) during protein sample preparation and in the cathode buffer to ensure complete saturation of the system since CBB-protein binding is not uniform (De St. Groth et al. 1963; Compton and Jones 1985; Tal et al. 1985; de Moreno et al. 1986). The benefits of this approach include unidirectional protein migration due to the negative charge imparted by CBB, real-time direct visualisation of protein separations, approximation of MW for resolved proteins/ complexes and retention of functional activity (Schägger and von Jagow 1991; Schägger et al. 1994). Alternatively, clear-native PAGE (i.e. omitting CBB from the sample and cathode buffer) can be used to prepare highly pure native protein

complexes; resolution is dependent on the native charge of a complex rather than the charge conveyed by CBB, and proteins retain their functional activity (Schägger et al. 1994; Wittig and Schägger 2005), enabling studies relating stoichiometry and function.

Using 1DE can be useful in simplifying the complexity of the original protein sample and investigating simple protein mixtures. However, since this resolution technique solely relies on MW as the criterion for separation, it cannot be assumed that any band represents only a single type of polypeptide as it is highly likely that more than one protein or proteoform will have the same approximate MW. Thus, it has proven useful as a first dimension of separation in some bottom-up proteomic approaches that yield lists of proteins potentially present based only on 'shotgun' assessments of the peptides found in a given gel region (Coorssen and Yergey 2015; Oliveira et al. 2014). However, this does not provide information on native proteoforms, which are the basis for molecular mechanisms and thus biological processes. In considering this, it was realised that separations based on two physico-chemical characteristics of proteins would increase the likelihood that a single resolved band/spot will represent a single proteoform. This is the foundation for 2DE (O'Farrel 1975; Klose 1975; Scheele 1975). This technique initially resolves proteins within a pH gradient relying on their isoelectric point (*pI*); polypeptides migrate through the gel to a pH region where their net charge is zero (i.e. isoelectric focusing—IEF). This gel is then positioned on top of a 1D SDS-PAGE gel (Laemmli method) to then resolve these proteins based on their size (i.e. MW). The resolving power of this method was recognised immediately. As with all newly developed methods, problems were identified over time, but significant refinements and optimisations (i.e. sample preparation, protein solubilisation buffer, immobilised pH gradients, zoom and ultra-zoom IEF gels) saw 2DE become the gold standard for highly reproducible and reliable separations of complex protein samples (Bjellqvist et al. 1982; Rabilloud et al. 1994, 1997; Rabilloud 1998; Sanchez et al. 1997). With all of these efforts, significant improvements were seen for the neutral pH region of resolution, which meant that some proteins were stacked at the pH extremes of the gels and remained unresolved. It is now recognised that a third electrophoretic separation of these stacked regions and other areas of the gel can effectively resolve these protein species (Butt and Coorssen 2005a; Wright et al. 2014a).

4 In-Gel Detection of Proteins

Once a protein sample has been resolved, the in-gel resolution pattern is usually revealed with the use of protein stains. There is a vast array of protein stains available, but their use depends on their detection sensitivity as well as the sample characteristics and research objectives (Miller et al. 2006). Thus, it is desirable to have a stain that can detect proteins of all types across a wide dynamic range. Since

quantitative top-down proteomics is dependent on the protein stain, it is essential that a stain:

- Provides reproducible detection of low amounts of protein (lowest limit of detection—LLD)
- Delivers a linear relationship between the amount of protein and staining intensity (linear dynamic range—LDR)
- Has low staining variability between different proteins (inter-protein variability—IPV)
- Delivers a high signal-to-noise ratio
- Is compatible with downstream microchemical characterisation techniques such as MS
- Has a relatively straightforward protocol for use
- Is of low cost (if possible)

These issues have been the subject of a recent detailed review (Gauci et al. 2011), and thus this section will only focus on the most common and competitive protein stains used for quantitative proteomics.

Densitometric protein stains absorb light energy when illuminated revealing the stained proteins as black or coloured spots or bands. A benefit of densitometric stains is real-time spot/band visualisation during staining, enabling qualitative assessment of protein resolution. This can be very useful for recognising the presence and/or absence of protein, general observational increases and/or decreases in concentrations, determination of MW and whether there were issues with the resolving capabilities of the gel technology employed (e.g. diffuse spots/bands, irregular dye fronts). It had however long been assumed that densitometric stains had a very limited dynamic range and therefore their sensitivity was insufficient for quantitative proteomics. However, like much dogma that has surrounded gel-based, top-down approaches to proteomics, this too may now need to be reconsidered with the advent of better imaging platforms (Noaman et al. 2017).

In the 1960s, the first quantitative assessment of gel-resolved proteins was achieved using Coomassie brilliant blue (CBB) (De St. Groth et al. 1963). This initiated a plethora of studies attempting to improve the limit of detection of CBB beyond ~30 ng of gel-resolved protein, but these were of limited success (Reisner et al. 1975; Kahn and Rubin 1975; Blakesley and Boezi 1977; Choi et al. 1996). Due to the properties of densitometric stains, any dye remaining in the gel matrix contributes to background absorbance, thus concealing low-abundance proteins and limiting coverage of a resolved proteome. Attempts to increase gel washing (i.e. destaining) to reduce background noise came at the price of also removing stain from proteins hence affecting quantitative evaluation. Again, detection sensitivity was limited. It wasn't until 1985 that significant improvements to the CBB stain formulation, with the introduction of colloidal CBB (cCBB), dramatically reduced background staining, thereby improving detection sensitivity by four- to fivefold for a variety of proteins (4–8 ng of protein) (Neuhoff et al. 1985, 1988). Other advantages of cCBB included low cost, a simple staining protocol and full compatibility with downstream protein identification technologies (e.g. MS).

However, an increasing recognition of proteome complexity (i.e. myriad proteoforms) fuelled the search for still more sensitive protein stains. Within 20 years of the introduction of cCBB came the development of densitometric silver stain, which was declared to be ~100-fold more sensitive than cCBB (Switzer et al. 1979). Although sub-nanogram detection sensitivity with silver stain was attainable, proteins stained in this way were incompatible with subsequent MS analysis. This was later rectified with changes to the staining method, but this sacrificed detection sensitivity (Yan et al. 2000; Chevallet et al. 2006, 2008). A number of other factors must also be considered when using silver stain:

- It is a tedious and laborious method.
- There is no definitive endpoint to the staining which thus potentially increases inter-experiment variability.
- There is high staining variability among proteins (i.e. IPV).
- There is only a small window in which protein concentration and staining are linear (i.e. LDR).
- Staining of other biomolecules also occurs.

While a number of other densitometric dyes have been investigated as alternatives to cCBB and silver stain, none have surpassed the quantitative detection sensitivity of these two protein stains.

Based on the detection characteristics of densitometric stains and the era in which they were being developed, the limiting factor was always the level of background noise. For this reason, fluorescent stains became favoured for in-gel protein detection since they are largely detectable only when bound to proteins, thus enhancing detection sensitivity by minimising background and enabling signal integration. The current commercial gold standard is SYPRO Ruby (SR) having a detection sensitivity of ~1–2 ng of protein, with an apparent wide LDR and minimal staining variability between proteins (i.e. IPV) (Berggren et al. 2000; Cong et al. 2008). A limitation to the use of SR is the excessive cost. Even though attempts to extend the use of SR by dilution or reuse were investigated, the limited quantitative characterisation of these strategies raised doubt in terms of practical application (Kreig et al. 2003; Ahnert et al. 2004).

Since the chemical structure of SR is proprietary, optimisation of the staining method proved challenging. Based on the fact that SR contains ruthenium, alternative ruthenium stains were developed, but their main advantage over SR was substantially lower cost (Rabilloud et al. 2000; Berggren et al. 2002; Berger et al. 2004). In the search for other competitive fluorescent stains, naturally fluorescent compounds were investigated and led to the development of deep purple (Bell and Karuso 2003). This stain was purported to deliver picogram detection sensitivity, but its quantitative performance relative to SR was found to be quite poor upon more extensive testing on 2DE gels (Mackintosh et al. 2003; Chevalier et al. 2004; Harris et al. 2007). Some reasons for this limited performance could be related to the high background noise, poor staining of specific types of proteins and limited photostability (Smejkal et al. 2004). Like densitometric stains, a large number of studies were dedicated to discovering new fluorescent stains to compete with SR;

but it wasn't until relatively recently that cCBB would again be used in a different context.

Close to 40 years after its initial characterisation as an in-gel protein stain, cCBB was also shown to possess infrared fluorescent properties when bound to protein in-gel (Luo et al. 2006). Soon after this revelation, preliminary assessment of its application for 2DE-based top-down proteomics revealed that it performed quite similar to SR but with an added advantage of reduced background (Harris et al. 2007). This prompted a more detailed quantitative assessment of cCBB infrared detection (cCBB-IRFD), in which 14 CBB formulations (including commercial formulations) were investigated using a single standardised staining protocol (Butt and Coorssen 2013). This study revealed two top performers when examining quantitative criteria using 1DE (i.e. LLD, LDR, IPV)—the original cCBB formulation developed by Neuhoff in 1985 (NCCB) and Bio-Safe™ Coomassie Stain developed by Bio-Rad (also a colloidal formulation) (Neuhoff et al. 1985, 1988; Butt and Coorssen 2013). In 2DE analyses, cCBB-IRFD had a slight advantage over SR, exhibiting an almost 1% increase in overall total spot detection (Butt and Coorssen 2013). However, when tested using an optimised protocol for NCBB, cCBB-IRFD provided pronounced improvements in 1DE and 2DE analyses (Gauci et al. 2013). In 1DE analyses using a variety of protein standards, cCBB-IRFD delivered sensitivity similar to SR, minimal IPV and reduced background noise when NaCl was used in the destaining protocol (Gauci et al. 2013). When assessed in parallel with SR, the LDR, lowest practical sensitivity and MS compatibility were shown to be significantly better with the new cCBB formulation and wash protocol. Perhaps even more importantly, cCBB-IRFD outperformed SR in terms of the total spot numbers detected, regardless of sample type (i.e. membrane and soluble proteomes from both plant and mammalian tissues; Fig. 1). However, the most important finding was that cCBB-IRFD delivered 18–36% more coverage of the proteomes than SR (with only 3–9% selective detection) (Gauci et al. 2013). Additional advantages to this staining and detection protocol are high reproducibility (even after gel storage), low cost, safe use and no need for additional measures (e.g. protection from light). With all these features, coupled cCBB-IRFD is a low-cost, effective protein staining and detection protocol for high-sensitivity quantitative proteomic analyses.

The stains mentioned above are 'general' and thus used to detect the entire proteome. However, it is known that post-translational modifications (i.e. one critical route to generating proteoforms) influence protein activity, localisation within a cell and/or interactions with other biomolecules. One of the most pervasive and important modifications is phosphorylation, and this subset of proteoforms is commonly referred to as the phosphoproteome. A wide variety of staining strategies have been devised for in-gel detection of the phosphoproteome (Cutting and Roth 1973; Hegenauer et al. 1977; Kaufmann et al. 2001; Larsen et al. 2001; Yamada et al. 2007), but the most successful stain thus far is Pro-Q Diamond (Steinberg et al. 2003; Schulenberg et al. 2004). Since detection relies heavily on the presence of phosphorylated residues, sensitivity does differ (i.e. β-casein (five residues) and pepsin (one residue) are detected at 1–2 ng and 8 ng, respectively

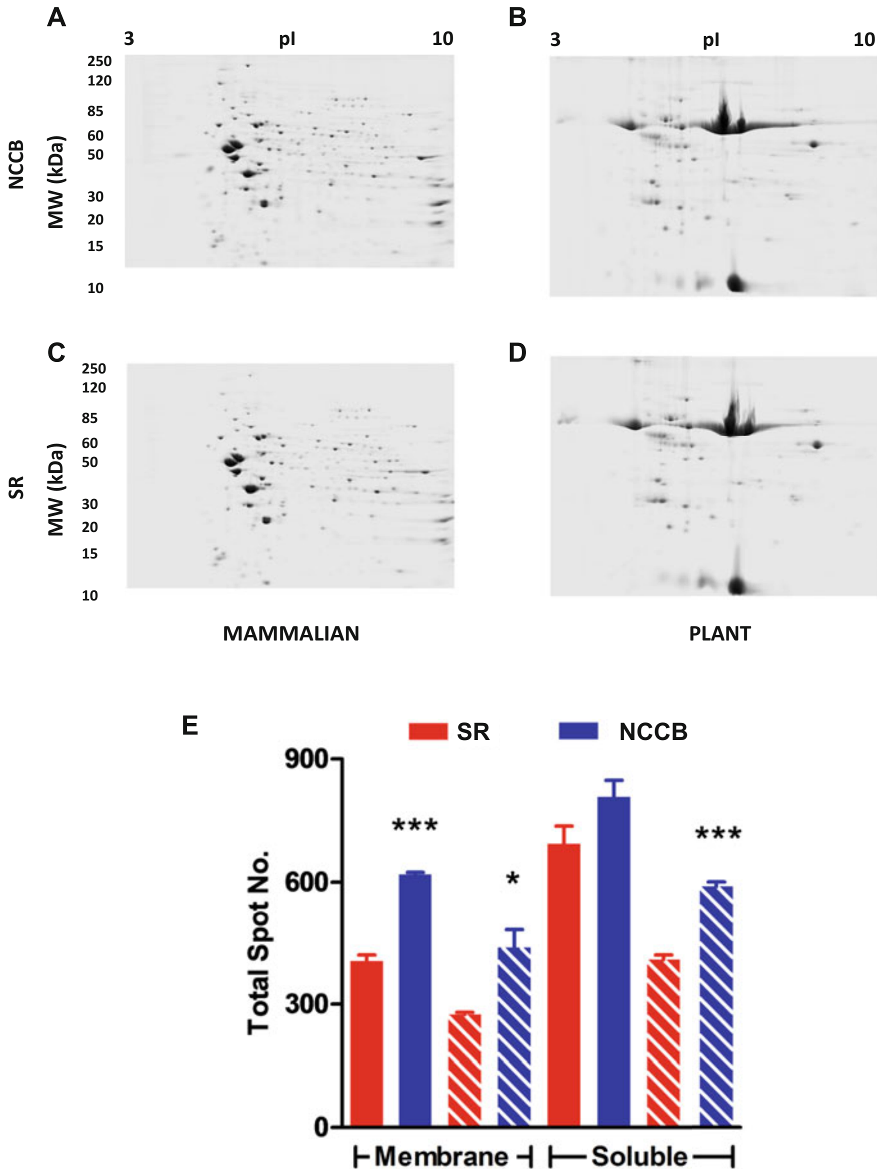

Fig. 1 Total spot detection of mammalian and plant proteomes: NCCB and SR. (**a**–**d**) Representative 2DE gel images for membrane fractions from mammalian (mouse brain—**a** and **b**) and plant (*Arabidopsis thaliana*—**c** and **d**) proteomes. (**e**) Total spot number analysis of mouse brain (*solid bars*) and *Arabidopsis thaliana* leaf (*hatched bars*) proteomes. Error bars represent SEM (n = 3). Statistically significant differences in comparison to SR are denoted *, where * $p < 0.05$, *** $p < 0.0001$; Student's t-test. Modified from Journal of Proteomics, Vol. 2, Gauci *et al*., 'Coomassie Blue Staining for High Sensitivity Gel-Based Proteomics', pages 96–106, Copyright 2013, with permission from Elsevier https://doi.org/10.1016/j.jprot.2013.01.027

(Steinberg et al. 2003)). Another modification that can easily be detected for gel-resolved proteins is glycosylation (Keyser 1964; Wardi and Michos 1972; Matthieu and Quarles 1973; Furlan et al. 1979; Muñoz et al. 1988), with Pro-Q Emerald 488 being the preferred detection stain (Hart et al. 2003). To date Pro-Q Emerald 488 is the best-performing glycoprotein stain available (Steinberg et al. 2001). However, like Pro-Q Diamond, detection sensitivity of Pro-Q Emerald 488 is determined by the carbohydrate content of a proteoform (Hart et al. 2003). A significant advantage to the use of these phospho- and glycoprotein stains is that a single gel-resolved proteome can be assessed for both of these sub-proteomes as well as the total resolved proteome, the only caveat being that staining must be sequential so as to avoid fluorophore quenching (Wu et al. 2005; Steinberg et al. 2003). Using such coupled staining strategies a single gel can provide a deeper understanding of the proteome in terms of changing proteoform concentrations and their modifications. A variety of other staining techniques further enhance this ability to delve more deeply into gel-resolved proteomes, including the detection of calcium-binding proteins, penicillin-binding proteins and oxidised proteins (Hill et al. 1994; Zhao et al. 1999; Yan et al. 1998; Hurd et al. 2007).

For the above densitometric and fluorescent stains, all are used post-electrophoresis to detect resolved proteins in-gel via non-covalent interactions. It is also possible to stain the proteins via covalent labelling prior to electrophoresis; this provides the basis for difference gel electrophoresis (DIGE) (Ünlü et al. 1997). These reactive CyDyes (e.g. Cy2, Cy3, Cy5), each fluorescing at a different wavelength, are usually similar in molecular weight and charge, so in theory they do not substantially alter the native resolving pattern of the proteome (Tonge et al. 2001). Thus, up to three samples can be differentially labelled and simultaneously resolved on the same gel (i.e. multiplexing), minimising the need for large numbers of gels and thus purportedly limiting inter-gel variability. Other 2D-DIGE experimental setups can include the use of an internal standard, a sample representing all samples to be tested, which can assist in quantitative analysis (Friedman et al. 2004; Alban et al. 2003). While this approach is cost-effective, is somewhat less labour intensive and provides a format for reliable comparisons, the labelling of proteins is limited to either lysine (minimal labelling) or cysteine (saturation labelling) residues (Ünlü et al. 1997; Shaw et al. 2003; Viswanathan et al. 2006). These CyDyes are also available with a variety of reactive groups (http://www.lumiprobe.com/fluorophore-chart), but, as noted, there has been little quantitative evaluation of the primary staining characteristics. Since the number of cysteine or lysine residues differs across proteoforms and their post-translational modifications (or other modifications near these residues) may influence labelling in different ways, potentially large variations in staining between species may be expected (IPV). This becomes even more of an issue in terms of cysteine labelling since many proteins do not contain this amino acid (or only have a single residue) and thus will go undetected. For 2DE, there is a limit to how much protein can be loaded onto a single gel, and if using DIGE the amount of protein loaded per sample is reduced, i.e. for a 100 μg total protein load, only ~33 μg from each sample would be loaded if comparing three samples. Although DIGE is useful for the detection and identification of

proteins (Timms and Cramer 2008), it must be noted that this load limitation can contribute to the difficulties known to be associated with the detection of low-abundance protein species, assuming labelling is even optimal. In this case, it may be necessary to spike the sample mix with unlabelled extra material or to run an additional (semi)preparative gel in parallel, but bear in mind that the sensitivity of the subsequent MS system for protein identification must be considered. Additionally, if no robotic spot picker is available, post-staining with visible dyes (silver, CBB) will be needed for manual spot picking, which bears the risk of slight differences in the pattern, due to different stainability of proteins (Gauci et al. 2011) or minor spot shifts between labelled and unlabelled material (Tonge et al. 2001; Shaw et al. 2003).

5 Other Considerations

While ensuring that gel electrophoresis and staining methods are optimal for reliable and reproducible quantitative detection of proteomes, the computational software and imaging instrumentation also have a significant impact on the quantitative information obtained (Mansour née Gauci et al. 2016; Gauci et al. 2013; Shaw et al. 2003; Miura 2001; Noaman et al. 2017). As mentioned previously, most stains either absorb light (densitometry) or emit light (fluorescence), and since the imaging instrumentation is the first step in obtaining useful quantitative data, it must provide a means to both produce and capture this signal. Most imaging instruments are comprised of a white light and/or UV lamp (i.e. standard 'light box')—or laser system—and a detector (i.e. usually a charge-coupled device (CCD) or a photomultiplier tube (PMT)) (Miura 2003). While CCD cameras are generally more cost-effective, laser systems currently appear to offer higher detection sensitivity (Miura 2003). The manner in which images are created is also important as it can have a significant effect on the quantitative information obtained; the best formats for analysis are TIFF, PNG and GIF files, with most analysis software accommodating 8–16 bit TIFF files (Marengo et al. 2005; Mansour née Gauci et al. 2016).

Even though the imaging instrumentation and image format are important considerations, there is also the choice of which computational software to use—Melanie, PDQuest, Z3, Progenesis, Delta2D and Proteomeweaver, as well as a number of freeware packages (Marengo et al. 2005). With most software programs, there is no limit to the extent of manual intervention; thus, any manual user changes extending outside the algorithms for spot detection and analysis must be noted in publications so as to facilitate an open forum for more comparative studies. The protein stain to be used can be a factor that determines which instrumentation is needed for acquiring high-quality images (or alternatively an existing imager may pose a limitation to which protein stains can be used). This circumstance in combination with the unknown composition of the sample in question can have a dramatic impact on the quantitative outcome since the concentration of each protein species cannot be evaluated until after staining and gel image acquisition are

complete. Thus, when hyper-abundant spots are imaged, they usually have large intensities that, with some stains, can result in self-quenching, which will affect the reliability of the quantifiable data. However, by excising these spots and re-imaging the gel, 'deep imaging' can markedly improve the extent of information obtained from any given gel (Wright et al. 2014a, b). This further highlights how all aspects of the top-down proteomic approach—sample, stain, imaging instrumentation and analysis software—can significantly alter the quantitative outcomes and thus require thorough consideration during the research design phase.

After the completion of gel electrophoresis, staining, imaging and quantification, the end goal is to characterise and quantify those proteoforms of interest (i.e. new, absent or otherwise changed in pattern) and then to identify them by MS. The most common ionisation methods for MS are electrospray ionisation (ESI) and matrix-assisted laser desorption ionisation (MALDI). To ensure that MS provides reliable protein identifications, the protein stain used must not interfere with the preparation and processing methods used for MS (Gauci et al. 2013; Lauber et al. 2001). The major limitations that hinder the confidence in a protein identification search are contamination and whether the protein of interest is present in the database. Differences in experimentally observed pI and MW following 2DE vs. theoretical values in protein databases generally provide the first indication that one is dealing with a proteoform rather than the 'parent' protein species described in the database.

6 Conclusions

Since the establishment of reliable and reproducible gel electrophoretic and staining methods, a variety of research designs are available to promote farm animal proteomics. In addition to the obvious potential impacts on husbandry, health and production, many 'farm' animals are also gaining popularity as models for disease (Bendixen et al. 2011). Gel electrophoresis provides a robust and well-established avenue with which to explore proteome changes in organs, fluids (e.g. blood, semen, venom), subcellular organelles and/or other samples. The strategies to explore these sample proteomes by gel electrophoresis are not limited to just the overall global map profile, but any sub-proteome (i.e. membrane, cytosolic, phosphorylated, glycosylated) which can reveal additional protein alterations to further increase our understanding of how they may influence different physiological states.

References

Ahnert N, Patton WF, Schulenberg B (2004) Optimized conditions for diluting and reusing a fluorescent protein gel stain. Electrophoresis 25(15):2506–2510

Alban A, David SO, Bjorkesten L, Andersson C, Sloge E, Lewis S, Currie I (2003) A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 3(1):36–44

Bell PJL, Karuso P (2003) Epicocconone, a novel fluorescent compound from the fungus *Epicoccum nigrum*. J Am Chem Soc 125(31):9304–9305

Bendixen E, Danielsen M, Hollung K, Gianazza E, Miller I (2011) Farm animal proteomics — a review. J Proteomics 74(3):282–293

Berger K, Wissmann D, Ihling C, Kalkhof S, Beck-Sickinger A, Sinz A, Paschke R, Führer D (2004) Quantitative proteome analysis in benign thyroid nodular disease using the fluorescent ruthenium II tris(bathophenanthroline disulfonate) stain. Mol Cell Endocrinol 227(1–2):21–30

Berggren K, Chernokalskaya E, Steinberg TH, Kemper C, Lopez MF, Diwu Z, Haugland RP, Patton WF (2000) Background-free, high sensitivity staining of proteins in one- and two-dimensional sodium dodecyl sulfate-polyacrylamide gels using a luminescent ruthenium complex. Electrophoresis 21(12):2509–2521

Berggren KN, Schulenberg B, Lopez MF, Steinberg TH, Bogdanova A, Smejkal G, Wang A, Patton WF (2002) An improved formulation of SYPRO Ruby protein gel stain: comparison with the original formulation and with a ruthenium II tris (bathophenanthroline disulfonate) formulation. Proteomics 2(5):486–498

Bernocco S, Fondelli C, Matteoni S, Magnoni L, Gotta S, Terstappen GC, Raggiaschi R (2008) Sequential detergent fractionation of primary neurons for proteomics studies. Proteomics 8 (5):930–938

Bjellqvist B, Ek K, Giorgio Righetti P, Gianazza E, Görg A, Westermeier R, Postel W (1982) Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. J Biochem Biophys Methods 6(4):317–339

Blakesley RW, Boezi JA (1977) A new staining technique for proteins in polyacrylamide gels using coomassie brilliant blue G250. Anal Biochem 82(2):580–582

Boschetti E, Righetti PG (2008) The ProteoMiner in the proteomic arena: a non-depleting tool for discovering low-abundance species. J Proteomics 71(3):255–264

Butt RH, Coorssen JR (2005a) Postfractionation for enhanced proteomic analyses: routine electrophoretic methods increase the resolution of standard 2D-PAGE. J Proteome Res 4:982–991

Butt RH, Coorssen JR (2005b) Pre-extraction sample handling by automated frozen disruption significantly improves subsequent proteomic analyses. J Proteome Res 5(2):437–448

Butt RH, Coorssen JR (2013) Coomassie blue as a near-infrared fluorescent stain: a systematic comparison with sypro ruby for in-gel protein detection. Mol Cell Proteomics 12(12):3834–3850

Butt RH, Lee MWY, Pirshahid SA, Backlund PS, Wood S, Coorssen JR (2006) An initial proteomic analysis of human preterm labor: placental membranes. J Proteome Res 5(11):3161–3172

Cannon-Carlson S, Tang J (1997) Modification of the Laemmli sodium dodecyl sulfate-polyacrylamide gel electrophoresis procedure to eliminate artifacts on reducing and nonreducing gels. Anal Biochem 246:146–148

Castagna A, Cecconi D, Sennels L, Rappsilber J, Guerrier L, Fortis F, Boschetti E, Lomas L, Righetti PG (2005) Exploring the hidden human urinary proteome via ligand library beads. J Proteome Res 4(6):1917–1930

Chevalier F, Rofidal V, Vanova P, Bergoin A, Rossignol M (2004) Proteomic capacity of recent fluorescent dyes for protein staining. Phytochemistry 65(11):1499–1506

Chevallet M, Diemer H, Luche S, Dorsselaer AV, Rabilloud T, Leize-Wagner E (2006) Improved mass spectrometry compatibility is afforded by ammoniacal silver staining. Proteomics 6(8):2350–2354

Chevallet M, Luche S, Diemer H, Strub J-M, Dorsselaer AV, Rabilloud T (2008) Sweet silver: a formaldehyde-free silver staining using aldoses as developing agents, with enhanced compatibility with mass spectrometry. Proteomics 8(23-24):4853–4861

Choi J-K, Yoon S-H, Hong H-Y, Choi D-K, Yoo G-S (1996) A modified coomassie blue staining of proteins in polyacrylamide gels with bismark brown R. Anal Biochem 236(1):82–84

Chrambach A, Rodbard D (1971) Polyacrylamide gel electrophoresis. Science 172(3982):440–451

Chrambach A, Rodbard D (1972) Polymerization of polyacrylamide gels: efficiency and reproducibility as a function of catalyst concentrations. Sep Sci Technol 7(6):663–703

Compton SJ, Jones CG (1985) Mechanism of dye response and interference in the Bradford protein assay. Anal Biochem 151(2):369–374

Cong W-T, Hwang S-Y, Jin L-T, Choi J-K (2008) Sensitive fluorescent staining for proteomic analysis of proteins in 1-D and 2-D SDS-PAGE and its comparison with SYPRO Ruby by PMF. Electrophoresis 29(21):4304–4315

Coorssen J, Yergey A (2015) Proteomics is analytical chemistry: fitness-for-purpose in the application of top-down and bottom-up analyses. Proteomes 3(4):440

Corthals GL, Molloy MP, Herbert BR, Williams KL, Gooley AA (1997) Prefractionation of protein samples prior to two-dimensional electrophoresis. Electrophoresis 18(3–4):317–323

Cutting JA, Roth TF (1973) Staining of Phospho-proteins on acrylamide gel electropherograms. Anal Biochem 54:386–394

de Moreno MR, Smith JF, Smith RV (1986) Mechanism studies of coomassie blue and silver staining of proteins. J Pharm Sci 75(9):907–911

De St. Groth SF, Webster RG, Datyner A (1963) Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. Biochimica et Biophysica Acta 71:377–391

Ferro M, Seigneurin-Berny D, Rolland N, Chapel A, Salvi D, Garin J, Joyard J (2000) Organic solvent extraction as a versatile procedure to identify hydrophobic chloroplast membrane proteins. Electrophoresis 21(16):3517–3526

Fialka I, Pasquali C, Lottspeich F, Ahorn H, Huber LA (1997) Subcellular fractionation of polarized epithelial cells and identification of organelle-specific proteins by two-dimensional gel electrophoresis. Electrophoresis 18(14):2582–2590

Fountoulakis M, Takács M-F, Berndt P, Langen H, Takács B (1999a) Enrichment of low abundance proteins of Escherichia coli by hydroxyapatite chromatography. Electrophoresis 20(11):2181–2195

Fountoulakis M, Takács M-F, Takács B (1999b) Enrichment of low-copy-number gene products by hydrophobic interaction chromatography. J Chromatogr A 833(2):157–168

Friedman DB, Hill S, Keller JW, Merchant NB, Levy SE, Coffey RJ, Caprioli RM (2004) Proteome analysis of human colon cancer by two-dimensional difference gel electrophoresis and mass spectrometry. Proteomics 4(3):793–811

Furlan M, Perret BA, Beck EA (1979) Staining of glycoproteins in polyacrylamide and agarose gels with fluorescent lectins. Anal Biochem 96(1):208–214

Gauci V, Wright E, Coorssen J (2011) Quantitative proteomics: assessing the spectrum of in-gel protein detection methods. J Chem Biol 4(1):3–29

Gauci VJ, Padula MP, Coorssen JR (2013) Coomassie blue staining for high sensitivity gel-based proteomics. J Proteomics 90:96–106

Ghosh D, Krokhin O, Antonovici M, Ens W, Standing KG, Beavis RC, Wilkins JA (2004) Lectin affinity as an approach to the proteomic analysis of membrane glycoproteins. J Proteome Res 3(4):841–850

Görg A, Boguth G, Köpf A, Reil G, Parlar H, Weiss W (2002) Sample prefractionation with Sephadex isoelectric focusing prior to narrow pH range two-dimensional gels. Proteomics 2(12):1652–1657

Hannig K (1978) Continuous free-flow electrophoresis as an analytical and preparative method in biology. J Chromatogr A 159(1):183–191

Harris LR, Churchwood MA, Butt RH, Coorssen JR (2007) Assessing detection methods for gel-based proteomic analyses. J Proteome Res 6:1418–1425

Hart C, Schulenberg B, Steinberg TH, Leung W-Y, Patton WF (2003) Detection of glycoproteins in polyacrylamide gels and on electroblots using Pro-Q Emerald 488 dye, a fluorescent periodate schiff-base stain. Electrophoresis 24:588–598

Hegenauer J, Ripley L, Nace G (1977) Staining acidic phosphoproteins (phosvitin) in electrophoretic gels. Anal Biochem 78:308–311

Herbert B, Righetti PG (2000) A turning point in proteome analysis: sample prefractionation via multicompartment electrolyzers with isoelectric membranes. Electrophoresis 21(17):3639–3648
Herbert BR, Molloy MP, Gooley AA, Walsh BJ, Bryson WG, Williams KL (1998) Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reducing agent. Electrophoresis 19(5):845–851
Hill IE, Hogue CWV, Clark ID, Macmanus JP, Szabo AG (1994) Detection of calcium binding proteins on polyacrylamide gels using time-resolved lanthanide luminescence photography. Anal Biochem 216(2):439–443
Hurd TR, Prime TA, Harbour ME, Lilley KS, Murphy MP (2007) Detection of reactive oxygen species-sensitive thiol proteins by redox difference gel electrophoresis: implications for mitochondrial redox signaling. J Biol Chem 282(30):22040–22051
Hurkman WJ, Tanaka CK (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. Plant Physiol 81:802–806
Kahn R, Rubin RW (1975) Quantitation of submicrogram amounts of protein using coomassie brilliant blue R on sodium dodecyl sulfate-polyacrylamide slab-gels. Anal Biochem 67(1):347–352
Kaufmann H, Bailey JE, Fussenegger M (2001) Use of antibodies for detection of phosphorylated proteins separated by two-dimensional gel electrophoresis. Proteomics 1(2):194–199
Keyser JW (1964) Staining of serum glycoproteins after electrophoretic separation in acrylamide gels. Anal Biochem 9:249–252
Klose J (1975) Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissue. Humangenetik 26:231–243
Kreig RC, Paweletz CP, Liotta LA, Petricon EF III (2003) Dilution of protein gel stain results in retention of staining capacity. Biotechniques 35:376–378
Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227(5259):680–685
Larsen MR, Sørensen GL, Fey SJ, Larsen PM, Roepstorff P (2001) Phospho-proteomics: evaluation of the use of enzymatic de-phosphorylation and differential mass spectrometric peptide mass mapping for site specific phosphorylation assignment in proteins separated by gel electrophoresis. Proteomics 1(2):223–238
Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jorgensen TJD (2005) Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. Mol Cell Proteomics 4:873–886
Lauber WM, Carroll JA, Dufield DR, Kiesel JR, Radabaugh MR, Malone JP (2001) Mass spectrometry compatibility of two-dimensional gel protein stains. Electrophoresis 22(5):906–918
Luo S, Wehr NB, Levine RL (2006) Quantitation of protein on gels and blots by infrared fluorescence of coomassie blue and fast green. Anal Biochem 350:233–238
Mackintosh JA, Choi H-Y, Bae S-H, Veal DA, Bell PJ, Ferrari BC, Van Dyk DD, Verrills NM, Paik Y-K, Karuso P (2003) A fluorescent natural product for ultra sensitive detection of proteins in one-dimensional and two-dimensional gel electrophoresis. Proteomics 3(12):2273–2288
Mansour née Gauci VJ, Noaman N, Coorssen JR (2016) Gel-staining techniques – dyeing to know it all. In: eLS. Wiley, Chichester. http://www.els.net. https://doi.org/10.1002/9780470015902.a0002686.pub3
Marengo E, Robotti E, Antonucci F, Cecconi D, Campostrini N, Righetti PG (2005) Numerical approaches for quantitative analysis of two-dimensional maps: a review of commercial software and home-made systems. Proteomics 5(3):654–666
Matthieu J-M, Quarles RH (1973) Quantitative scanning of glycoproteins on polyacrylamide gels stained with periodic acid-Schiff reagent (PAS). Anal Biochem 55(1):313–316
Miller I, Crawford J, Gianazza E (2006) Protein stains for proteomic applications: *which, when, why?* Proteomics 6(20):5385–5408
Miura K (2001) Imaging and detection technologies for image analysis in electrophoresis. Electrophoresis 22(5):801–813

Miura K (2003) Imaging technologies for the detection of multiple stains in proteomics. Proteomics 3(7):1097–1108

Molloy MP, Herbert BR, Walsh BJ, Tyler MI, Traini M, Sanchez J-C, Hochstrasser DF, Williams KL, Gooley AA (1998) Extraction of membrane proteins by differential solubilization for separation using two-dimensional gel electrophoresis. Electrophoresis 19(5):837–844

Molloy MP, Herbert BR, Williams KL, Gooley AA (1999) Extraction of Escherichia coli proteins with organic solvents prior to two-dimensional electrophoresis. Electrophoresis 20(4–5):701–704

Molloy MP, Herbert BR, Slade MB, Rabilloud T, Nouwens AS, Williams KL, Gooley AA (2000) Proteomic analysis of the Escherichia coli outer membrane. Eur J Biochem 267(10):2871–2881

Muñoz G, Marshall S, Cabrera M, Horvat A (1988) Enhanced detection of glycoproteins in polyacrylamide gels. Anal Biochem 170(2):491–494

Neuhoff V, Stamm R, Eibl H (1985) Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: a systematic analysis. Electrophoresis 6(9):427–448

Neuhoff V, Arold N, Taube D, Ehrhardt W (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using coomassie brilliant blue G-250 and R-250. Electrophoresis 9(6):255–262

Noaman N, Abbineni PS, Withers M, Coorssen JR (2017, September) Coomassie staining provides routine (sub)femtomole in-gel detection of intact proteoforms: expanding opportunities for genuine Top-down Proteomics. Electrophoresis. https://doi.org/10.1002/elps.201700190

O'Farrel PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021

Oliveira BM, Coorssen JR, Martins-de-Souza D (2014) 2DE: the phoenix of proteomics. J Proteomics 104:140–150

Partridge MA, Gopinath S, Myers SJ, Coorssen JR (2016) An initial top-down proteomic analysis of the standard cuprizone mouse model of multiple sclerosis. J Chem Biol 9(1):9–18

Pasquali C, Fialka I, Huber LA (1997) Preparative two-dimensional gel electrophoresis of membrane proteins. Electrophoresis 18(14):2573–2581

Pedersen SK, Harry JL, Sebastian L, Baker J, Traini MD, McCarthy JT, Manoharan A, Wilkins MR, Gooley AA, Righetti PG, Packer NH, Williams KL, Herbert BR (2003) Unseen proteome: mining below the tip of the iceberg to find low abundance and membrane proteins. J Proteome Res 2(3):303–311

Perdew GH, Schaup HW, Selivonchick DP (1983) The use of a zwitterionic detergent in two-dimensional gel electrophoresis of trout liver microsomes. Anal Biochem 135(2):453–455

Poduslo JF, Rodbard D (1980) Molecular weight estimation using sodium dodecyl sulfate-pore gradient electrophoresis. Anal Biochem 101(2):394–406

Rabilloud T (1998) Use of thiourea to increase the solubility of membrane proteins in two-dimensional electrophoresis. Electrophoresis 19(5):758–760

Rabilloud T, Valette C, Lawrence JJ (1994) Sample application by in-gel rehydration improves the resolution of two-dimensional electrophoresis with immobilized pH gradients in the first dimension. Electrophoresis 15(1):1552–1558

Rabilloud T, Adessi C, Giraudel A, Lunardi J (1997) Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis 18(3–4):307–316

Rabilloud T, Strub J-M, Luche S, Girardet JL, van Dorsselaer A, Lunardi J (2000) Ruthenium II tris (bathophenanthroline disulfonate), a powerful fluorescent stain for detection of proteins in gel with minimal interference in subsequent mass spectrometry analysis. Proteome 1(1):1–14

Ramsby ML, Makowski GS, Khairallah EA (1994) Differential detergent fractionation of isolated hepatocytes: biochemical, immunochemical and two-dimensional gel electrophoresis characterization of cytoskeletal and noncytoskeletal compartments. Electrophoresis 15(1):265–277

Raymond S, Wang Y-J (1960) Preparation and properties of acrylamide gel for use in electrophoresis. Anal Biochem 1(4–5):391–396

Raymond S, Weintraub L (1959) Acrylamide gel as a supporting medium for zone electrophoresis. Science 130:711

Reisner AH, Nemes P, Bucholtz C (1975) The use of coomassie brilliant blue G250 perchloric acid solution for staining in electrophoresis and isoelectric focusing on polyacrylamide gels. Anal Biochem 64(2):509–516

Sanchez J-C, Rouge V, Pisteur M, Ravier F, Tonella L, Moosmayer M, Wilkins MR, Hochstrasser DF (1997) Improved and simplified in-gel sample application using reswelling of dry immobilized pH gradients. Electrophoresis 18(3–4):324–327

Schägger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem 166(2):368–379

Schägger H, von Jagow G (1991) Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Anal Biochem 199(2):223–231

Schägger H, Cramer WA, Vonjagow G (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. Anal Biochem 217(2):220–230

Scheele GA (1975) Two-dimensional gel analysis of soluble proteins: characterization of guinea pig exocrine pancreatic proteins. J Biol Chem 250:5375–5385

Schröder B, Hasilik A (2006) A protocol for combined delipidation and subfractionation of membrane proteins using organic solvents. Anal Biochem 357(1):144–146

Schulenberg B, Goodman TN, Aggeler R, Capaldi RA, Patton WF (2004) Characterization of dynamic and steady-state protein phosphorylation using a fluorescent phosphoprotein gel stain and mass spectrometry. Electrophoresis 25(15):2526–2532

Shaw J, Rowlinson R, Nickson J, Stone T, Sweet A, Williams K, Tonge R (2003) Evaluation of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes. Proteomics 3(7):1181–1195

Shen Y, Kim J, Strittmatter EF, Jacobs JM, Camp DG, Fang R, Tolié N, Moore RJ, Smith RD (2005) Characterization of the human blood plasma proteome. Proteomics 5(15):4034–4045

Smejkal GB, Robinson MH, Lazarev A (2004) Comparison of fluorescent stains: relative photostability and differential staining of proteins in two-dimensional gels. Electrophoresis 25(15):2511–2519

Smith SD, She Y-M, Roberts EA, Sarkar B (2004) Using immobilized metal affinity chromatography, two-dimensional electrophoresis and mass spectrometry to identify hepatocellular proteins with copper-binding ability. J Proteome Res 3(4):834–840

Steinberg TH, Top KPO, Berggren KN, Kemper C, Jones L, Diwu Z, Haugland RP, Patton WF (2001) Rapid and simple single nanogram detection of glycoproteins in polyacrylamide gels and on electroblots. Proteomics 1(7):841–855

Steinberg TH, Agnew BJ, Gee KR, Leung W-Y, Goodman T, Schulenberg B, Hendrickson J, Beechem JM, Haugland RP, Patton WF (2003) Global quantitative phosphoprotein analysis using multiplexed proteomics technology. Proteomics 3(7):1128–1144

Switzer RC, Merril CR, Shifrin S (1979) A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. Anal Biochem 98(1):231–237

Tal M, Silberstein A, Nusser E (1985) Why does Coomassie Brilliant Blue R interact differently with different proteins? A partial answer. J Biol Chem 260(18):9976–9980

Taylor RS, Fialka I, Jones SM, Huber LA, Howell KE (1997) Two-dimensional mapping of the endogenous proteins of the rat hepatocyte Golgi complex cleared of proteins in transit. Electrophoresis 18(14):2601–2612

Thulasiraman V, Lin S, Gheorghiu L, Lathrop J, Lomas L, Hammond D, Boschetti E (2005) Reduction of the concentration difference of proteins in biological liquids using a library of combinatorial ligands. Electrophoresis 26(18):3561–3571

Timms JF, Cramer R (2008) Difference gel electrophoresis. Proteomics 8(23–24):4886–4897

Tonge R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Currie I, Davison M (2001) Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics 1(3):377–396

Ünlü M, Morgan ME, Minden JS (1997) Difference gel electrophoresis. A single gel method for detecting changes in protein extracts. Electrophoresis 18(11):2071–2077

Viswanathan S, Unlu M, Minden JS (2006) Two-dimensional difference gel electrophoresis. Nat Protoc 1(3):1351–1358

Völkl A, Mohr H, Weber G, Fahimi HD (1997) Isolation of rat hepatic peroxisomes by means of immune free flow electrophoresis. Electrophoresis 18(5):774–780

Wang YY, Cheng P, Chan DW (2003) A simple affinity spin tube filter method for removing high-abundant common proteins or enriching low-abundant biomarkers for serum proteomic analysis. Proteomics 3(3):243–248

Wardi AH, Michos GA (1972) Alcian blue staining of glycoproteins in acrylamide disc electrophoresis. Anal Biochem 49(2):607–609

Wasinger VC, Locke VL, Raftery MJ, Larance M, Rothemund D, Liew A, Bate I, Guilhaus M (2005) Two-dimensional liquid chromatography/tandem mass spectrometry analysis of Gradiflow™ fractionated native human plasma. Proteomics 5(13):3397–3401

Weber K, Osborn M (1969) The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. J Biol Chem 244(16):4406–4412

Wissing J, Heim S, Flohé L, Bilitewski U, Frank R (2000) Enrichment of hydrophobic proteins via Triton X-114 phase partitioning and hydroxyapatite column chromatography for mass spectrometry. Electrophoresis 21(13):2589–2593

Wittig I, Schägger H (2005) Advantages and limitations of clear-native PAGE. Proteomics 5(17):4338–4346

Wright EP, Partridge MA, Padula MP, Gauci VJ, Malladi CS, Coorssen JR (2014a) Top-down proteomics: enhancing 2D gel electrophoresis from tissue processing to high-sensitivity protein detection. Proteomics 14(7–8):872–889

Wright EP, Prasad KAG, Padula MP, Coorssen JR (2014b) Deep imaging: how much of the proteome does current top-down technology already resolve? PLoS One 9(1):e86058

Wu J, Lenchik NJ, Pabst MJ, Solomon SS, Shull J, Gerling IC (2005) Functional characterization of two-dimensional gel-separated proteins using sequential staining. Electrophoresis 26(1):225–237

Yamada S, Nakamura H, Kinoshita E, Kinoshita-Kikuta E, Koike T, Shiro Y (2007) Separation of a phosphorylated histidine protein using phosphate affinity polyacrylamide gel electrophoresis. Anal Biochem 360(1):160–162

Yan L-J, Orr WC, Sohal RS (1998) Identification of oxidized proteins based on sodium dodecyl sulfate–polyacrylamide gel electrophoresis, immunochemical detection, isoelectric focusing, and microsequencing. Anal Biochem 263(1):67–71

Yan JX, Wait R, Berkelman T, Harry RA, Westbrook JA, Wheeler CH, Dunn MJ (2000) A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization- mass spectrometry. Electrophoresis 21(17):3666–3672

Zhao G, Meier TI, Kahl SD, Gee KR, Blaszczak LC (1999) BOCILLIN FL, a sensitive and commercially available reagent for detection of penicillin-binding proteins. Antimicrob Agents Chemother 43(5):1124–1128

Sample Preparation for 2DE Using Samples of Animal Origin

Hugo Osório, André M. de Almeida, and Alexandre Campos

Abstract Protein separation by two-dimensional gel electrophoresis (2DE) is a well-established technique for proteome analysis. Despite the enormous potential of this methodology for resolving proteins in complex samples, the fact is that its use is limited and often avoided due to inadequate sample preparation. It is understood as being especially critical for 2DE-based proteomics research that samples need to be of good quality, and this goal could be accomplished by optimizing or adapting the sample preparation protocols to the specificities (i.e., chemical composition) of biological samples. Sample preparation protocols should also comply with the objectives of the research. For selection of the adequate protocol and for its optimization, it is therefore important first to know the distinct steps that constitute a protocol and the main objectives underlying each step. Secondly, it is important to understand the potential limitations of the biological samples hence to perform the necessary adjustments/alterations to the protocol. This chapter aims to review the major steps of sample preparation protocols for 2DE with particular focus on optimization, taking into account specificities of the biological samples. The main technical aspects concerning sample tissue disruption and cell lysis, protein extraction, fractionation, enrichment, and solubilization prior to 2DE will be presented and discussed. In this chapter, the most common biological samples employed in proteomics and 2DE research and corresponding sample

H. Osório
i3s-Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal

IPATIMUP-Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal

Faculty of Medicine, Department of Pathology and Oncology, University of Porto, Porto, Portugal

A.M. de Almeida
Instituto Superior de Agronomia, Universidade de Lisboa, Lisboa, Portugal

A. Campos (✉)
CIIMAR – Interdisciplinary Centre of Marine and Environmental Sciences, University of Porto, Matosinhos, Portugal
e-mail: amoclclix@gmail.com

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_4

preparation methods will be reviewed. This follows a brief discussion on 2DE and MALDI-TOF approaches. A selection of methods for 2DE sample preparation is included based on the authors' experience.

Keywords Protein extraction • Protein separation • Sample cleanup • Fractionation • Concentration • Depletion

List of Abbreviations

2DE	Two-dimensional gel electrophoresis
CHAPS	3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate
CPLL	Combinatorial peptide ligand libraries
DDM	*n*-Dodecyl β-D-maltoside
DTT	Dithiothreitol
HPLC	High-performance liquid chromatography
IAA	Iodoacetamide
IEF	Isoelectric focusing
LC-MS	Liquid chromatography coupled with mass spectrometry
MALDI-TOF	Matrix-assisted laser desorption/ionization—time of flight
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

1 Introduction

Proteomics is a field of growing importance in the context of animal and veterinary sciences (Almeida et al. 2015; Eckersall et al. 2012). In fact, and as the majority of the animal-origin products can be considered as proteinaceous products, proteomics plays a pivotal role in animal research, being highly complementary to other omics methods together with classical animal experimental approaches. Accordingly, the importance of proteomics has been described in several publications, including, for instance, areas as diverse as meat science (Paredi et al. 2012, 2013) and dairy science (Roncada et al. 2012) or even as a tool in the creation of vaccines against tick-borne disease (Marcelino et al. 2012b). Proteomics may be divided into two major types of approaches: gel-based and gel-free. More details about the differences in the two approaches may be found in other chapters of this book and in other publications (Almeida et al. 2015). Briefly, in the gel-based approach, proteins are extracted from tissues, fluids, or organs, denatured and separated in two-dimensional gels (loaded with the same amount of protein), and may be further stained using different techniques such as colloidal Coomassie, silver nitrate, fluorescent dyes, etc. Gels are then scanned using a densitometer or photographic device and subsequently analyzed using an appropriate software that will compare

protein abundance (relative spot intensity) in the different experimental groups. Protein spots in the gel showing differential intensities are finally selected for excision, digested with a protease such as trypsin, and identified using mass spectrometry techniques such as matrix-assisted laser desorption/ionization—time of flight (MALDI-TOF) (Soares et al. 2012).

The success of a gel-based proteomic experiment is dependent on two major factors. The first is the existence of reproducible gels where the same amount of protein is loaded to achieve similar distribution of the different protein spots across the gel with a distinct shape so that they may be recognized by the gel analysis software. The second factor, in common with gel-free proteomic approaches, has to do with the representation of that particular species in public protein databases in which mass spectrometry results are searched for. Generally speaking and when considering farm animals, only cattle (Ferreira et al. 2013), pig (de Almeida and Bendixen 2012), or rabbit (Miller et al. 2014) have very good public databases, with all the other species requiring homology searches in these handful of species to adequately identify their differentially abundant proteins. Nevertheless, this is a field that has been advancing very quickly, and the establishment of mass spectrometry (MS) databases such as the Bovine PeptideAtlas (Bislev et al. 2012) has been contributing significantly to improve the quality and accurateness of protein identifications. For further details, please refer to the mammary gland proteome chapter in this book.

In spite of these limitations, 2DE is still by far the most accessible technique for proteomics approaches in animal and veterinary sciences. Its usefulness is however dependent on having reproducible 2DE gels that in turn are dependent on a complex process of sample preparation. Sample preparation aims essentially at protein enrichment and purification and minimizing the effect of putative contaminants (e.g., lipids, nucleic acids, minerals, etc.) that may interfere with protein migration. When contaminants are in high concentrations in the sample, they will cause disruptive patterns rendering the gel useless. In fact, as animal tissues and fluids are highly complex in their composition, sample preparation could be the Achilles heel of any gel-based proteomic experiment. Strategies have to be devised in order to conduct a good, reproducible protein extraction protocol that will reduce the number of contaminants to a minimum and at the same time will minimize loss of protein content. Such strategies can be complex and while adequate to one type of tissue may prove to be useless in a different tissue or even in similar tissues from different areas of the body that may differ, for instance, in fat content.

In this chapter, we focus on the different approaches regarding sample preparation and handling that can be of use in a gel-based proteomic experiment. It is based on the experience of the authors who have run 2DE gels on numerous types of different samples that required different approaches on protein extraction and sample handling, from sheep and rabbit skeletal muscle (Almeida et al. 2010; Almeida et al. 2016) to cat urine (Ferlizza et al. 2015); from bacterial cell cultures (Correia et al. 2013; Marcelino et al. 2012a) to ruminant milk and colostrum (Hernández-Castellano et al. 2016), parasites (Ferreirinha et al. 2016), yeast extracts (Barbosa et al. 2016), and human serum (Gomes et al. 2013); or from

mussel organs (Puerto et al. 2011) to plant material (Almeida et al. 2012; Cruz De Carvalho et al. 2014; Sousa et al. 2014) or the goat mammary gland (Cugno et al. 2016) and lamb plasma (Hernández-Castellano et al. 2014). We hope this chapter will serve as a reference for sample preparation, particularly suitable to (recalcitrant) animal tissues.

2 Biological Samples

2DE has shown to be a versatile technique being applied to characterize the proteomes of many different biological tissues and fluids from humans and other mammals, insects, and lower invertebrates. The major accomplishments have been however in the field of human health with the proteomes of a variety of human tissues and cell lines being mapped (Finehout et al. 2004; Pixton et al. 2004; Quaranta et al. 2001; Anderson et al. 2004; Seow et al. 2000) as well as proteomes from other mammalian model species (Gazzana and Borlak 2007). The interest for uncovering the cell metabolism and the functions of the distinct cell organelles has led to the fractionation of the biological materials and to the analysis of the proteomes of mitochondria, microsomes, chloroplasts, lysosomes, other cell structures, and protein complexes (Lescuyer et al. 2003; Kanaeva et al. 2005; Jan Van Wijk 2000; Journet et al. 2002). In addition, 2DE/Western blot combined approaches provide an accurate protein selection for targeted proteomics experiments (Ferreirinha et al. 2016; Sousa et al. 2014).

Investigations carried out in the 1980s and 1990s in not-so-ready-to-solubilize samples, particularly from plant origin (Damerval et al. 1986; Burstin et al. 1993; Damerval et al. 1987, 1994), have enabled to establish new protocols of sample preparation, which contributed to the wider use of 2DE in many different types of biological materials. Subsequent proteomic investigations have been covering so diverse biological materials as the underwater mussel adhesive byssal threads, the bivalve shell matrix proteins, the hemolymph and silk gland proteomes, venom toxins, etc. (Gantayet et al. 2013; Immel et al. 2016; Cai et al. 2008; Valente et al. 2009; Hou et al. 2007). The growing number of tissues and species investigated by proteomics also reflects the extended dissemination of this methodology among the various fields of biology and animal and environmental sciences. On the other hand, the recent genomic and transcriptomic projects carried out in several non-model species like the Atlantic salmon, Pacific oyster, and whiteleg shrimp, as to mention a few in the field of aquaculture, have opened new possibilities for comprehensive proteomic studies and thereby for the discovery of new molecular and physiological functions in these species. 2DE has also contributed to the analysis of animal microbiomes (Wallace et al. 2017; Snelling and Wallace 2017) and with this to the development of the recent research field of metaproteomics.

3 Sample Preparation

3.1 Tissue Homogenization

Sample preparation is usually achieved following a sequence of procedures summarized in Fig. 1. For homogenization of animal soft tissues and to break animal cells, freeze-thawing, sonication, and tissue blending methods have been mostly applied (Cugno et al. 2016; Campos et al. 2013; Almeida et al. 2010). On the other hand, to disrupt animal and plant hard tissues, liquid nitrogen is frequently used (Almeida et al. 2010; Damerval et al. 1986). The rapid freezing of the biological material allowed by liquid nitrogen also helps to preserve the DNA, RNA, and most of the proteins. Moreover, lysis of bacterial cells, with thick and resistant cell walls, as well as yeast cells can be achieved by applying high pressure with a French pressure cell device (Moore et al. 2016; Nally et al. 2005). Tissue homogenization should be conducted at low temperature to minimize activity of endogenous proteases. Furthermore, buffer solutions containing detergents can be used in tissue homogenization. Detergents will act on membrane lipids and contribute to cell lysis and solubilization of membrane proteins. Commonly used detergents are *n*-dodecyl β-D-maltoside (DDM), a nonionic detergent that does not interfere with the native state of proteins and thus its main application is for solubilization of membrane protein complexes. Sodium dodecyl sulfate (SDS) is the most widely used detergent

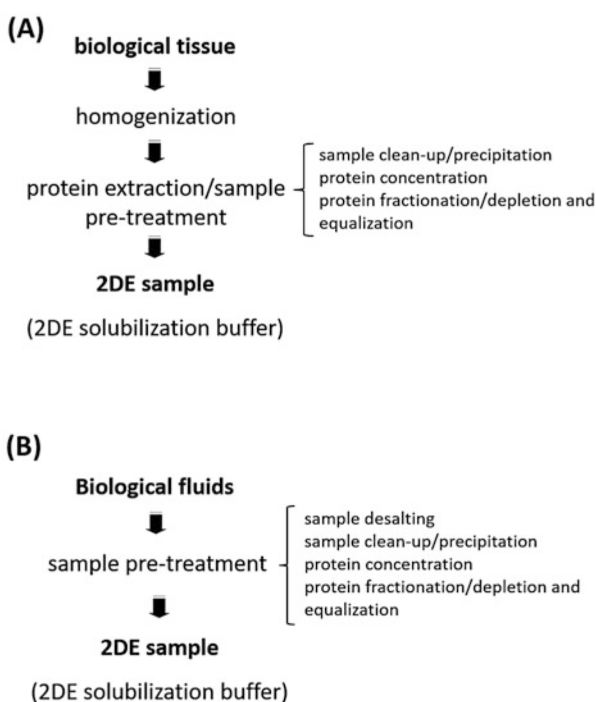

Fig. 1 General procedures of sample preparation of biological tissues (**a**) and fluids (**b**) applied in proteomics. A description of each step is presented in the main text of the article

to solubilize proteins; however, the ionic properties of this detergent are not compatible with 2DE, and its concentration should not exceed 0.25% (w/v) in protein samples. Neutral detergents such as 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS) are preferred in sample preparations for 2DE. Tissue homogenization solutions must further include protease inhibitors to reduce or stop any proteolytic activity in the samples. The samples for 2DE are typically prepared under denaturing conditions, in solutions containing a neutral detergent (i.e., CHAPS) and denaturing agents such as urea and thiourea (Görg et al. 2000).

3.2 Protein Extraction

Protein extraction follows tissue disruption and homogenization. In its simplest form, it consists in the direct solubilization of the proteins from the tissue homogenates with buffer solutions that are compatible with the isoelectric focusing (IEF). Such a solution typically contains strong protein denaturants such as urea, at concentrations between 6 M and 9 M, and thiourea (2M). Urea is used because it is a strong disruptor of protein noncovalent bonds and a neutral molecule compatible with the low ionic strength gel environments of IEF. These solutions still contain at least one nonionic detergent such as CHAPS with concentrations up to 4% (w/v) and a reducing agent to disrupt protein intra- and intercysteine disulfide bonds, such as dithiothreitol (DTT) (20–100 mM) (Görg et al. 2000). The standard 2DE solution used in our laboratory in different biological tissues consists of urea (7M), thiourea (2M), CHAPS (4% w/v), and DTT (65 mM) (Cugno et al. 2016; Campos et al. 2013; Puerto et al. 2011).

Many animal tissues are rich in compounds that interfere particularly with IEF. Among the interfering compounds are salts, lipids, polysaccharides, and nucleic acids. In order to improve IEF and the overall resolution of 2DE, these compounds need to be removed from protein samples prior to electrophoresis. Interfering compounds can be removed by selective precipitation of the proteins, for instance, with cold acetone and trichloroacetic acid (TCA) (Damerval et al. 1986; Campos et al. 2013). Several protein extraction protocols based on the principle of protein precipitation and solvent partition have been employed, for instance, to the liver, brain, ovarian tissues, longissimus dorsi muscle, and biological fluids to improve 2DE protein detection and resolution from those tissues (Jia et al. 2014; Shan et al. 2012; Hao et al. 2015; Fic et al. 2010).These protocols have been also employed to membrane protein solubilization and sample cleanup for optimal 2DE (Devraj et al. 2009). Two protocols generally applied for sample cleanup and concentration are schematically presented in Fig. 2. These have been initially developed for plant tissues but are nowadays generally used in many types of animal tissues.

Optimization of sample preparation is thus necessary and strongly recommended in proteomic studies. Figure 3 illustrates an investigation on sample preparation

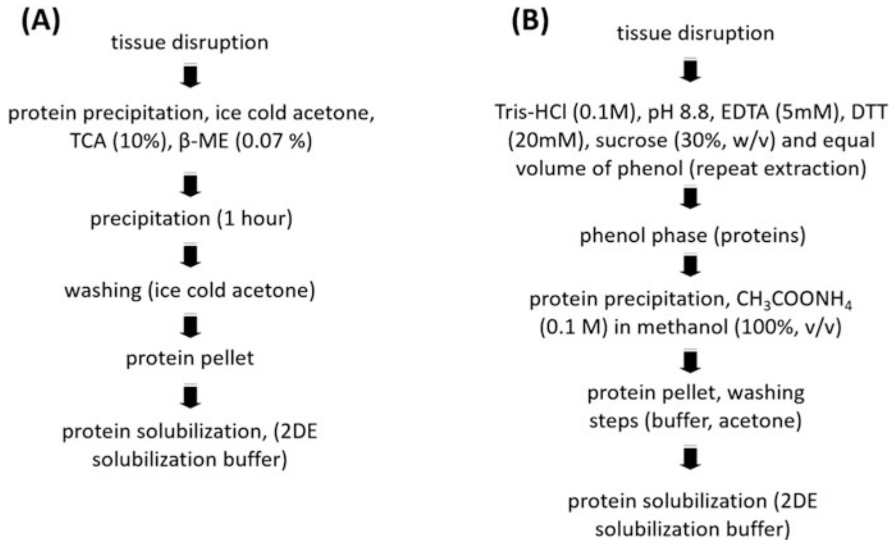

Fig. 2 Suggested sample preparation procedures for recalcitrant tissues, to allow the removal of salts, lipids, and other IEF interfering compounds. (**a**) Acetone/TCA precipitation described by Damerval et al. (1986); (**b**) phenol extraction described by Hurkman and Tanaka (1986). β-mercaptoethanol (β-ME) is a protein reducing agent, similar to DTT.

from tissues of the marine sentinel species *Mytilus galloprovincialis* performed in our laboratory (Campos et al. 2013). Our results demonstrated that precipitation methods can decrease gel background and protein spot streaking, allowing overall improvement of gel spot detection (Campos et al. 2013).

To improve the analysis of biological fluids like plasma, urine, cerebrospinal fluid, hemolymph, and others, specific sample pretreatment steps can be employed (Fig. 1). The sample pretreatment strategy can be complemented with more sensitive detection systems such as the differential gel electrophoresis (DIGE) to optimize protein detection in 2DE gels (Sun et al. 2011; Clement et al. 2013). The characterization of the protein composition of body fluids requires in most cases a sample concentration step (Ferlizza et al. 2015). Furthermore, proteins such as albumin, immunoglobulins, transferrin, haptoglobin, fibrinogen, and others, which are highly abundant in the biological fluids, can be depleted or equalized to both reduce sample complexity and protein concentration range, leading to an improvement in the detection of less abundant proteins (Kaur et al. 2010; Cao et al. 2013). Moreover, body fluids rich in salts such as urine are desalted prior to IEF by protein precipitation, liquid chromatography, or dialysis. Some of the most used protein fractionation, depletion, and concentration methods are described in the following sections.

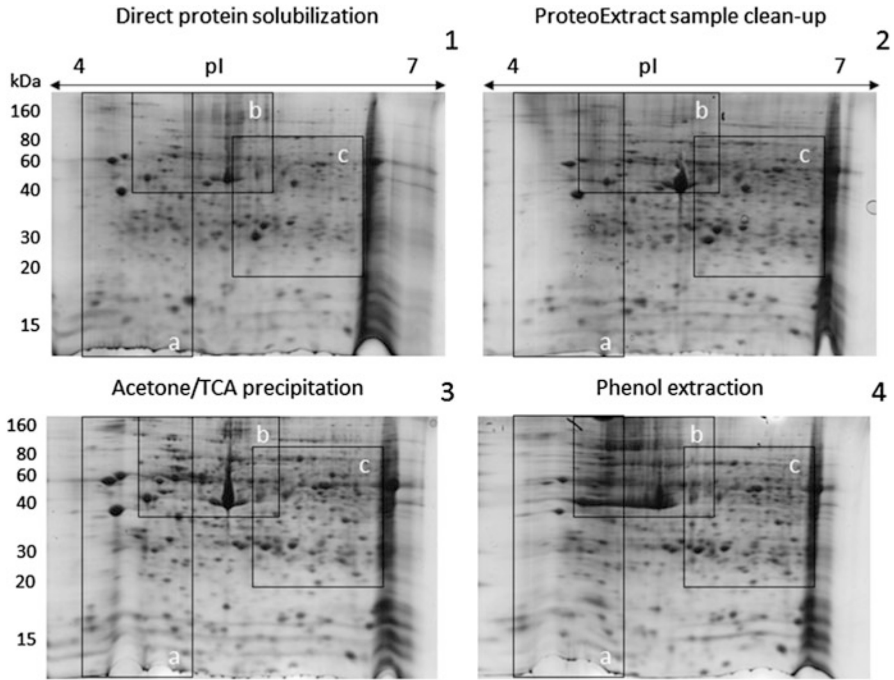

Fig. 3 Representative two-dimensional gel electrophoresis of proteins from the digestive gland of *Mytilus galloprovincialis* obtained with different sample preparation methods: 1, direct protein solubilization; 2, ProteoExtract sample cleanup kit (Calbiochem, San Diego, CA, USA); 3, cold acetone and TCA precipitation; and 4, phenol extraction (adapted from Campos et al. 2013). Protein samples (150 μg) were loaded on 7 cm IEF gel strips. Gels were stained with colloidal coomassie blue. Boxed gel sections show variations in protein resolution with different extraction methods. Section a, variations in protein resolution and horizontal streaking at the acidic end of 2DE gels; section b, protein spot streaking of high-abundance proteins with method 4; section c, variations in protein resolution with methods 3 and 4 allowing the detection of more protein spots

3.3 Sample Pretreatment Methods

3.3.1 Sample Fractionation

A typical proteome has a high dynamic range of protein concentrations where highly abundant proteins coexist together with trace proteins (Zubarev 2013). In addition, those proteins may be organelle specific such as mitochondria or nuclei and/or may have a precise localization such as the cytosol or membrane. So, after protein extraction, the crude protein samples might need to be fractionated (Fig. 1) in order to improve the overall proteome coverage by reducing its dynamic concentration range. Here we summarize some of the most relevant protein fractionation methodologies.

Density gradient centrifugation has been a key procedure to separate different cell organelles and protein complexes. It can be a rate-zonal separation, where proteins are separated by their size or mass, or an isopycnic separation, where proteins are separated based on their density (Lee et al. 2010).

Another relevant procedure for protein fractionation is liquid chromatography where proteins can be separated based on their size, charge, or hydrophobicity (Walls and Loughran 2011). Size exclusion chromatography is a technique where proteins migrate through a porous matrix starting with the elution of high molecular weight proteins and a higher retention of low molecular weight proteins. In ion exchange chromatography, proteins are separated on their charge using anion exchange resins for separation of positively charged proteins or cation exchange resins for separation of negatively charged proteins. Protein separation by reversed-phase chromatography is an additional possibility. This approach is one of the most popular separation procedures in proteomics workflows. After enzymatic digestion of protein extracts, peptides are separated by C18 reversed-phase chromatography online coupled to a mass spectrometer (LC-MS).

Body fluids, such as human blood, exhibit an additional challenge for analysis since levels of different proteins vary by ten orders of magnitude (Service 2008). Abundant protein depletion strategies combined with protein equalization and/or combinatorial peptide ligand libraries (CPLL) (Santucci et al. 2012) are employed to normalize overall protein content. Immunoaffinity protein removal is the most employed protein depletion strategy. Commercial kits were optimized to deplete albumin and antibody components or the top abundant proteins in human plasma/ serum or spinal fluid samples in preparation for mass spectrometry or 1D and 2D gel electrophoresis (Liu et al. 2011). Sample depletion can be performed in batch mode with antibody-coated microbeads, using spin filters or with multiple affinity removal columns for high-performance liquid chromatography (HPLC). On the other hand, in CPLL, antibodies are replaced by peptide ligands that show selective affinity to proteins. At the CPLL approach, an equal number of ligands is available for both high- and low-abundance proteins. When samples are incubated with CPLL, high-abundance proteins will saturate their high-affinity ligands, and surplus unbound proteins are washed away. As a result, low-abundance proteins are concentrated on their specific affinity ligands. Unlike immunodepletion that is limited to the number of antibodies available for capturing proteins, CPLL can be developed for any kind of sample or protein bait (Boschetti and Righetti 2008).

3.3.2 Protein Concentration

Protein concentration procedures are mostly employed in diluted biological samples such as fluids. Depending on sample type, different methodological procedures may be followed for protein concentration. Acetone precipitation is a simple method for protein concentration and removal of undesirable compounds that interfere with IEF, as described above. Importantly, this approach may lead to protein denaturation hampering protein resolubilization (Simpson and Beynon

2010). Trichloroacetic acid-induced protein precipitation (Rajalingam et al. 2009) is a standard method for proteomic analysis. It was observed that a concentration range of 5–45% (w/v) of TCA precipitates the maximum amount of proteins; however, it has low efficiency for precipitating disordered/unfolded proteins. Ammonium sulfate precipitates proteins by reducing its solubility (Burgess 2009). After precipitation and resolubilization, proteins can be separated by size exclusion chromatography allowing buffer replacement. Companies commercialize optimized kits for protein precipitation for 2D gels, such as 2D Cleanup Kit (GE Healthcare), ReadyPrep™ 2D Cleanup Kit (Bio-Rad), or ProteoExtract Protein Precipitation Kit (Calbiochem, Merck Millipore).

Protein concentration using centrifugal concentrators is a typically non-denaturing methodology useful for protein concentration, desalting, and purification (Andrew et al. 2001). A 3 kDa molecular weight cutoff (MWCO) membrane is appropriated for proteome concentration.

Vacuum concentration equipment has a vacuum pump tube attached to a centrifuge with a cold trap, allowing the reduction of the protein sample volume in a fast process. This procedure should be performed at room temperature to avoid sample degradation. Freezing of aqueous samples associated with the evaporation process is frequently observed. Importantly, most of other molecules will also concentrate, including metabolites, lipids, or salts. So, it is important to use complementary protein cleanup procedures. In addition, protein evaporation to dryness should be avoided due to posterior inefficient resolubilization. Sample lyophilization with a freeze dryer is another possibility; however it has similar drawbacks.

3.4 Total Protein Estimation

The estimation of total protein concentration in the biological samples before performing the 2DE is fundamental, in order to ensure the reproducibility of the methodology and the quality of the results. With the estimation of the total protein, one can standardize the 2DE analysis, based on a fixed amount of protein to be loaded on each 2DE gel. This procedure validates the quantitative comparison of individual proteins between gels. Total protein estimation is mostly carried out by conventional colorimetric methods of Folin-Lowry and Bradford and the bicinchoninic acid (BCA) method, which requires the use of a protein standard, usually bovine serum albumin (BSA). Total protein can be further estimated by the absorption of tyrosine and tryptophan at 280 nm. However, this method is unspecific and therefore less reliable since many other compounds such as alcohols, certain buffer ions, and nucleic acids all absorb at 280 nm. These methods have been extensively described and protocols optimized for a variety of biological samples. More detailed information can be found, for instance, in recent publications (Sapan and Lundblad 2015; Okutucu et al. 2007; Kao et al. 2008). Likely the main bottleneck in the use of these methods is that all, with greater or lesser extent,

are sensitive to many of the chemicals used in the preparation of the extraction and solubilization buffers, namely, detergents such as SDS and CHAPS, reducing agents, and metal ions, which introduce bias in protein measurements. More information on interfering chemicals can be found in the following reviews (Aschermann et al. 2008; Noble and Bailey 2009). This bias is usually circumvented by diluting the sample and the chemicals that make up the buffer to concentrations that are tolerable by the quantification methods and by including the solubilizing buffer in the same concentrations/dilutions in blank and calibration curve. Alternatively, the interfering chemicals may be removed by cleaning the sample, for instance, with protein precipitation as described above.

4 IEF and SDS-PAGE

Two-dimensional gel electrophoresis of proteins starts with IEF that separates proteins according to their isoelectric point and subsequently with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) that will separate proteins in a second dimension according to their molecular mass. Both electrophoretic processes are carried out in protein-denaturing conditions. Since the first work done by O'Farrell (1975), some critical technological developments took place in IEF allowing to greatly improve the overall reproducibility of 2DE. Among those are the precast IEF gels with immobilized pH gradients, commercialized in different lengths and pH gradients, and sophisticated instruments that allow different electrophoretic programs to be tested simultaneously. IEF gels of broad pH range (pH 4–7 or 3–10) and increased length (17–24 cm long) are mostly used to increase the number of detected proteins from proteomes and hence are most suited for discovery-based proteomics (Valério et al. 2016), while narrow-pH-range IEF gels have been used in specific cases to improve the resolution and separation of proteins with similar isoelectric points and proteoforms (Martins et al. 2014). A typical IEF protocol starts with the sample loading that takes about 12 h and allows simultaneously the rehydration of IEF gel strips. The electrophoretic program consists of stepwise voltage increments controlled by time. IEF protocols thus differ according to the characteristics of the IEF gels, i.e., their length and pH gradient, but also with sample composition. For instance, IEF of samples with high salt content can be improved by increasing the time of the separation and voltage and by applying and replacing periodically paper wicks to absorb the salts accumulated in the electrodes along the run (Wu et al. 2010). Other procedures have been proposed, for instance, to reduce the acidic region streaking due to the presence of nucleic acids in the samples (Roy et al. 2014; Brown and Norris 2016). Before the second dimension (SDS-PAGE), proteins are reduced and alkylated, by incubating IEF strips with DTT and iodoacetamide, for example. IEF strips are then assembled on top of the SDS-PAGE gels, and electrophoresis is applied to transfer proteins from the IEF to the SDS-PAGE gels. Standard SDS-PAGE is carried out with homogenous gels with 10–12 % polyacrylamide which allows separation of proteins in the range of

10–250 kDa. SDS-PAGE can be performed in many different protein gel electrophoresis systems. However, for comparative proteome studies, the use of multi-gel electrophoretic systems is recommended, to achieve reproducible separations between gels which will increase the consistency of results in comparative proteomic investigations.

5 Protein Identification with MALDI-TOF/TOF MS

Mass spectrometry (MS) approaches, such as MALDI-TOF/TOF, allow protein identification of the selected 2DE gel spots. Coomassie blue-based stains are recommended for MS analysis as this dye does not bind covalently to proteins and can be easily washed from the gel prior to MALDI-TOF/TOF analysis. In the case of silver staining, it is important to use a protocol compatible with MS without formaldehyde or glutaraldehyde. In the fixation and sensitization steps, these molecules may cause cross-linkage of internal lysine residues hindering protein digestion with enzymes that cut at this amino acid such as trypsin.

Excision of the selected 2D gel spots may be performed with a gel spot picker (e.g., OneTouch, Gel Company). Afterward, gel spots are washed with acetonitrile/ammonium bicarbonate solution, followed by protein reduction with DTT (Osorio and Reis 2013). Next, gel spots are treated with iodoacetamide (IAA) in order to alkylate the reduced cysteine residues. It is worth noting that this leads to carbamidomethylation of cysteine, an irreversible modification that has to be considered in the MS protein identification procedure. Next, proteins are digested with an appropriate protease. Trypsin is the most popular enzyme used in proteomics, a serine protease that cuts at the C-terminus side of lysine and arginine not followed by proline (Osorio and Reis 2013). However, and especially for low molecular weight proteins, which may have a reduced number of these amino acids, it is important to consider alternative proteases such as chymotrypsin or Asp-N. Then, peptides are extracted from the gel with trifluoroacetic acid (TFA) and dried. After resuspension with 0.1% TFA, the peptide mixture is spotted on a MALDI sample plate using reversed-phase C18 tips. The most usual matrix is α-cyano-4-hydroxycinnamic acid (CHCA) in 50% acetonitrile/0.1% TFA at a concentration of 8 mg/mL. After plating, the peptide mass spectrum is usually acquired in reflector positive mode. Some peptides may be further analyzed by MS/MS sequencing allowing a more detailed protein characterization. Protein identification is performed by peptide mass fingerprint—PMF—or by PMF+MS/MS (Osorio and Reis 2013). A search algorithm (e.g., Mascot, Matrix Science) performs this approach after selecting a proteome from a protein sequence database, such as UniProt or NCBI, together with a taxonomic selection.

6 Conclusions

Together with good experimentation procedures that include a thorough sample selection and collection, sample preparation is one of the key aspects in 2DE, perhaps even of higher importance than other subsequent steps of the process: protein quantification, electrophoresis, and protein identification using mass spectrometry. In fact, sample preparation will define the quality of the whole gel and the success of the following steps. In this chapter we have focused on the different aspects of the 2DE process but paying a particular attention to those related to sample preparation, particularly when handling recalcitrant samples. Several approaches are suggested in order to obtain good-quality and reproducible gels that in turn will lead to good-quality proteomic experiments. They may include, for instance, steps such as depletion and fractionation or even a step as simple as protein precipitation that should be considered during experimental design.

Acknowledgments A. Campos' work is supported by postdoctoral grant (SFRH/BPD/103683/2014) from FCT. IPATIMUP integrates the i3S Research Unit, which is partially supported by FCT, the Portuguese Foundation for Science and Technology. This work was financed by FEDER (Fundo Europeu de Desenvolvimento Regional) funds through the COMPETE 2020, Operacional Programme for Competitiveness and Internationalization (POCI), Portugal 2020, and Portuguese funds through FCT (Foundation for Science and Technology/Ministério da Ciência, Tecnologia e Inovação) in the framework of the projects "Institute for Research and Innovation in Health Sciences" (POCI-01-0145-FEDER-007274), "PEst-C/SAU/LA0003/2013" and "UID/multi/04423/2013".

References

de Almeida AM, Bendixen E (2012) Pig proteomics: a review of a species in the crossroad between biomedical and food sciences. J Proteome 75:4296–4314

Almeida AM, Campos A, Francisco R et al (2010) Proteomic investigation of the effects of weight loss in the gastrocnemius muscle of wild and NZW rabbits via 2D-electrophoresis and MALDI-TOF MS. Anim Genet 41:260–272

Almeida AM, Parreira JR, Santos R et al (2012) A proteomics study of the induction of somatic embryogenesis in Medicago truncatula using 2DE and MALDI-TOF/TOF. Physiol Plant 146:236–249

Almeida AM, Bassols A, Bendixen E et al (2015) Animal board invited review: advances in proteomics for animal and food sciences. Animal 9:1–17

Almeida AM, Palhinhas RG, Kilminster T et al (2016) The effect of weight loss on the muscle proteome in the Damara, Dorper and Australian Merino ovine breeds. PLoS One 11:e0146367

Anderson NL, Polanski M, Pieper R et al (2004) The human plasma proteome: a nonredundant list developed by combination of four separate sources. Mol Cell Proteomics 3:311–326

Andrew SM, Titus JA, Zumstein L (2001) Dialysis and concentration of protein solutions. Curr Protoc Immunol Appendix 3:A.3H.1-5

Aschermann, K., Lutter P, Wattenberg A (2008) Current status of protein quantification technologies. Precision and statistical significance Bioprocess Int 44–53

Barbosa AD, Pereira C, Osório H et al (2016) The ceramide-activated protein phosphatase Sit4p controls lifespan, mitochondrial function and cell cycle progression by regulating hexokinase 2 phosphorylation. Cell Cycle 15:1620–1630

Bislev SL, Deutsch EW, Sun Z et al (2012) A Bovine PeptideAtlas of milk and mammary gland proteomes. Proteomics 12:2895–2899

Boschetti E, Righetti PG (2008) The ProteoMiner in the proteomic arena: a non-depleting tool for discovering low-abundance species. J Proteome 71:255–264

Brown S, Norris G (2016) Improved consistency in 2D gel electrophoresis: sheep plasma as a test case. Electrophoresis 38:906–913

Burgess RR (2009) Protein precipitation techniques. Methods Enzymol 463:331–342

Burstin J, Zivy M, Devienne D, Damerval C (1993) Analysis of scaling methods to minimize experimental variations in 2-dimensional electrophoresis quantitative data: application to the comparison of maize inbred lines. Electrophoresis 14:1067–1073

Cai K, Chen K, Liu X et al (2008) Differential expression of haemolymph proteome of resistant strain and susceptible strain for BmNPV in Bombyx mori L. Sheng Wu Gong Cheng Xue Bao 24:285–290

Campos A, Puerto M, Prieto A et al (2013) Protein extraction and two-dimensional gel electrophoresis of proteins in the marine mussel Mytilus galloprovincialis: an important tool for protein expression studies, food quality and safety assessment. J Sci Food Agric 93:1779–1787

Cao Z, Yende S, Kellum JA, Robinson RAS (2013) Additions to the Human plasma proteome via a tandem MARS depletion iTRAQ-based workflow. Int J Proteomics 2013:654356

Clement CC, Aphkhazava D, Nieves E et al (2013) Protein expression profiles of human lymph and plasma mapped by 2D-DIGE and 1D SDS-PAGE coupled with nanoLC-ESI-MS/MS bottom-up proteomics. J Proteome 78:172–187

Correia M, Michel V, Osorio H et al (2013) Crosstalk between Helicobacter pylori and gastric epithelial cells is impaired by docosahexaenoic acid. PLoS One 8:e60657

Cruz De Carvalho R, Bernardes Da Silva A, Soares R et al (2014) Differential proteomics of dehydration and rehydration in bryophytes: Evidence towards a common desiccation tolerance mechanism. Plant Cell Environ 37:1499–1515

Cugno G, Parreira JR, Ferlizza E et al (2016) The goat (Capra hircus) mammary gland mitochondrial proteome: a study on the effect of weight loss using blue-native page and two-dimensional gel electrophoresis. PLoS One 11:e0151599

Damerval C, De Vienne D, Zivy M, Thiellement H (1986) Technical improvements in two-dimensional electrophoresis increase the level of genetic variation detected in wheat-seedling proteins. Electrophoresis 7:52–54

Damerval C, Le GM, Blaisonneau J et al (1987) A simplification of Heukeshoven and Dernick's silver staining of proteins. Electrophoresis 8:158–159

Damerval C, Maurice A, Josse JM, De Vienne D (1994) Quantitative trait loci underlying gene product variation: A novel perspective for analyzing regulation of genome expression. Genetics 137:289–301

Devraj K, Geguchadze R, Klinger ME et al (2009) Improved membrane protein solubilization and clean-up for optimum two-dimensional electrophoresis utilizing GLUT-1 as a classic integral membrane protein. J Neurosci Methods 184:119–123

Eckersall PD, Miller I, de Almeida AM (2012) Proteomics, a new tool for farm animal science. J Proteome 75:4187–4189

Ferlizza E, Campos A, Neagu A et al (2015) The effect of chronic kidney disease on the urine proteome in the domestic cat (Felis catus). Vet J 204:73–81

Ferreira AM, Bislev SL, Bendixen E, Almeida AM (2013) The mammary gland in domestic ruminants: a systems biology perspective. J Proteome 94:110–123

Ferreirinha P, Correia A, Teixeira-Coelho M et al (2016) Mucosal immunization confers long-term protection against intragastrically established Neospora caninum infection. Vaccine 34:6250–6258

Fic E, Kedracka-Krok S, Jankowska U et al (2010) Comparison of protein precipitation methods for various rat brain structures prior to proteomic analysis. Electrophoresis 31:3573–3579

Finehout EJ, Franck Z, Lee KH (2004) Towards two-dimensional electrophoresis mapping of the cerebrospinal fluid proteome from a single individual. Electrophoresis 25:2564–2575

Gantayet A, Ohana L, Sone ED (2013) Byssal proteins of the freshwater zebra mussel, Dreissena polymorpha. Biofouling 29:77–85

Gazzana G, Borlak J (2007) Improved method for proteome mapping of the Liver by 2-DE MALDI-TOF MS. J Proteome Res 6:3143–3151

Gomes C, Almeida A, Ferreira JA et al (2013) Glycoproteomic analysis of serum from patients with gastric precancerous lesions. J Proteome Res 12:1454–1466

Görg a OC, Boguth G et al (2000) The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis 21:1037–1053

Hao R, Adoligbe C, Jiang B et al (2015) An optimized trichloroacetic acid/acetone precipitation method for two-dimensional gel electrophoresis analysis of Qinchuan cattle longissimus dorsi muscle containing high proportion of marbling. PLoS One 10:e0124723

Hernández-Castellano LE, Almeida AM, Ventosa M et al (2014) The effect of colostrum intake on blood plasma proteome profile in newborn lambs: low abundance proteins. BMC Vet Res 10:85

Hernández-Castellano LE, Almeida AM, Renaut J et al (2016) A proteomics study of colostrum and milk from the two major small ruminant dairy breeds from the Canary Islands: a bovine milk comparison perspective. J Dairy Res 83:366–374

Hou Y, Xia Q, Zhao P et al (2007) Studies on middle and posterior silk glands of silkworm (Bombyx mori) using two-dimensional electrophoresis and mass spectrometry. Insect Biochem Mol Biol 37:486–496

Hurkman WJ, Tanaka CK (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. Plant Physiol 81:802–806

Immel F, Broussard C, Catherinet B et al (2016) The shell of the invasive bivalve species Dreissena polymorpha: biochemical, elemental and textural Investigations. PLoS One 11:1–28

Jan Van Wijk K (2000) Proteomics of the chloroplast: experimentation and prediction. Trends Plant Sci 5:420–425

Jia JL, Zhang LP, Wu JP et al (2014) Establishment of the optimum two-dimensional electrophoresis system of ovine ovarian tissue. Genet Mol Res 13:6528–6538

Journet A, Chapel A, Kieffer S et al (2002) Proteomic analysis of human lysosomes: application to monocytic and breast cancer cells. Proteomics 2:1026–1040

Kanaeva IP, Petushkova NA, Lisitsa AV et al (2005) Proteomic and biochemical analysis of the mouse liver microsomes. Toxicol In Vitro 19:805–812

Kao SH, Wong HK, Chiang CY, Chen HM (2008) Evaluating the compatibility of three colorimetric protein assays for two-dimensional electrophoresis experiments. Proteomics 8:2178–2184

Kaur P, Reis MD, Couchman GR et al (2010) SERPINE 1 links obesity and diabetes: a pilot study. J Proteomics Bioinform 3:191–199. https://doi.org/10.4172/jpb.1000139

Lee YH, Tan HT, Chung MCM (2010) Subcellular fractionation methods and strategies for proteomics. Proteomics 10:3935–3956

Lescuyer P, Strub J-M, Luche S et al (2003) Progress in the definition of a reference human mitochondrial proteome. Proteomics 3:157–167

Liu B, Qiu F, Voss C et al (2011) Evaluation of three high abundance protein depletion kits for umbilical cord serum proteomics. Proteome Sci 9:24

Marcelino I, de Almeida AM, Brito C et al (2012a) Proteomic analyses of Ehrlichia ruminantium highlight differential expression of MAP1-family proteins. Vet Microbiol 156:305–314

Marcelino I, de Almeida AM, Ventosa M et al (2012b) Tick-borne diseases in cattle: applications of proteomics to develop new generation vaccines. J Proteome 75:4232–4250

Martins JC, Campos A, Osório H et al (2014) Proteomic profiling of cytosolic glutathione transferases from three bivalve species: Corbicula fluminea, Mytilus galloprovincialis and Anodonta cygnea. Int J Mol Sci 15:1887–1900
Miller I, Rogel-Gaillard C, Spina D et al (2014) The Rabbit as an experimental and production animal: from genomics to proteomics. Curr Protein Pept Sci 15:134–145
Moore SM, Hess SM, Jorgenson JW (2016) Extraction, enrichment, solubilization, and digestion techniques for membrane proteomics. J Proteome Res 15:1243–1252
Nally JE, Whitelegge JP, Aguilera R et al (2005) Purification and proteomic analysis of outer membrane vesicles from a clinical isolate of Leptospira interrogans serovar Copenhageni. Proteomics 5:144–152
Noble JE, Bailey MJA (2009) Quantitation of Protein. Methods Enzymol 463:73–95
O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. J Biol Chem 250:4007–4021
Okutucu B, Dınçer A, Habib Ö, Zıhnıoglu F (2007) Comparison of five methods for determination of total plasma protein concentration. J Biochem Biophys Methods 70:709–711
Osorio H, Reis CA (2013) Mass spectrometry methods for studying glycosylation in cancer. Methods Mol Biol 1007:301–316
Paredi G, Raboni S, Bendixen E et al (2012) "Muscle to meat" molecular events and technological transformations: The proteomics insight. J Proteome 75:4275–4289
Paredi G, Sentandreu MA, Mozzarelli A et al (2013) Muscle and meat: New horizons and applications for proteomics on a farm to fork perspective. J Proteome 88:58–82
Pixton KL, Deeks ED, Flesch FM et al (2004) Sperm proteome mapping of a patient who experienced failed fertilization at IVF reveals altered expression of at least 20 proteins compared with fertile donors: case report. Hum Reprod 19:1438–1447
Puerto M, Campos A, Prieto A et al (2011) Differential protein expression in two bivalve species; Mytilus galloprovincialis and Corbicula fluminea; exposed to Cylindrospermopsis raciborskii cells. Aquat Toxicol 101:109–116
Quaranta S, Giuffrida MG, Cavaletto M et al (2001) Human proteome enhancement: high-recovery method and improved two-dimensional map of colostral fat globule membrane proteins. Electrophoresis 22:1810–1818
Rajalingam D, Loftis C, JJ X, Kumar TKS (2009) Trichloroacetic acid-induced protein precipitation involves the reversible association of a stable partially structured intermediate. Protein Sci 18:980–993
Roncada P, Piras C, Soggiu A et al (2012) Farm animal milk proteomics. J Proteome 75:4259–4274
Roy A, Varshney U, Pal D (2014) Avoiding acidic region streaking in two-dimensional gel electrophoresis: case study with two bacterial whole cell protein extracts. J Biosci 39:631–642
Santucci L, Candiano G, Bruschi M et al (2012) Combinatorial peptide ligand libraries for the analysis of low-expression proteins: validation for normal urine and definition of a first protein MAP. Proteomics 12:509–515
Sapan CV, Lundblad RL (2015) Review of methods for determination of total protein and peptide concentration in biological samples. Proteomics Clin Appl 9:268–276
Seow TK, Ong SE, Liang RCMY et al (2000) Two-dimensional electrophoresis map of the human hepatocellular carcinoma cell line, HCC-M, and identification of the separated proteins by mass spectrometry. Electrophoresis 21:1787–1813
Service RF (2008) Proteomics. Proteomics ponders prime time. Science 321:1758–1761
Shan LP, Li H, Fan Y et al (2012) Comparison of different methods of rat amniotic fluid sample preparation for 2-D electrophoresis. Chinese. J Med Genet 29:669–672
Simpson DM, Beynon RJ (2010) Acetone precipitation of proteins and the modification of peptides. J Proteome Res 9:444–450
Snelling TJ, Wallace RJ (2017) The rumen microbial metaproteome as revealed by SDS-PAGE. BMC Microbiol 17:9

Soares R, Franco C, Pires E et al (2012) Mass spectrometry and animal science: protein identification strategies and particularities of farm animal species. J Proteome 75:4190–4206
Sousa R, Osorio H, Duque L et al (2014) Identification of Plantago lanceolata pollen allergens using an immunoproteomic approach. J Investig Allergol Clin Immunol 24:177–183
Sun W, Jiang Y, He F (2011) Extraction and proteome analysis of liver tissue interstitial fluid. Methods Mol Biol 728:247–257
Valente RH, Guimarães PR, Junqueira M et al (2009) Bothrops insularis venomics: a proteomic analysis supported by transcriptomic-generated sequence data. J Proteome 72:241–255
Valério E, Campos A, Osório H, Vasconcelos V (2016) Proteomic and Real-Time PCR analyses of Saccharomyces cerevisiae VL3 exposed to microcystin-LR reveals a set of protein alterations transversal to several eukaryotic models. Toxicon 112:22–28
Wallace RJ, Snelling TJ, McCartney CA et al (2017) Application of meta-omics techniques to understand greenhouse gas emissions originating from ruminal metabolism. Genet Sel Evol 49:9
Walls D, Loughran ST (eds) (2011) Protein chromatography: methods and protocols. Humana Press, New York
Wu HC, Chen TN, Kao SH et al (2010) Isoelectric focusing management: an investigation for salt interference and an algorithm for optimization. J Proteome Res 9:5542–5556
Zubarev RA (2013) The challenge of the proteome dynamic range and its implications for in-depth proteomics. Proteomics 13:723–726

Gel-Free Proteomics

Dominique Baiwir, Paolo Nanni, Susanne Müller, Nicolas Smargiasso, Denis Morsa, Edwin De Pauw, and Gabriel Mazzucchelli

Abstract Mass spectrometry (MS)-based proteomics is nowadays considered as a mainstream analytical tool for life sciences. This success is mainly to be attributed to the tremendous improvements in the core technology of mass spectrometry coupled with recent developments in bioinformatics and separation sciences. The level of information that is henceforth accessible through MS techniques is broad and encompasses the protein sequence as well as its quantity, modification(s), structure, and implication in macromolecular assemblies. The current chapter focuses on the existing gel-free MS-based quantitative approaches that are applicable in the field of animal sciences and summarizes their principal features together with their respective advantages and drawbacks. In this context, we will discuss both the label-free and label-based relative quantitation strategies as well as the existing alternatives for absolute quantitation. The specific case of posttranslational modifications, whose extensive characterization is becoming a crucial topic in MS-based proteome studies, is next addressed under the quantitative scope of phosphorylation and glycosylation. Lastly, other promising nonmainstream techniques are presented, emphasizing their potent use to solve major proteomics challenges of tomorrow.

D. Baiwir • D. Morsa
Laboratory of Mass Spectrometry-GIGA Proteomics Facility, University of Liège, Liège, Belgium
e-mail: D.Baiwir@ulg.ac.be; d.morsa@ulg.ac.be

P. Nanni
Functional Genomics Center Zurich, University/ETH Zurich, Zurich, Switzerland
e-mail: paolo.nanni@fgcz.uzh.ch

S. Müller
Biognosys AG, Schlieren, Switzerland
e-mail: susanne.mueller@biognosys.com

N. Smargiasso • E. De Pauw • G. Mazzucchelli (✉)
Laboratory of Mass Spectrometry-MolSys, Department of Chemistry, University of Liege, Liege, Belgium
e-mail: nsmargiasso@ulg.ac.be; E.Depauw@ulg.ac.be; gabriel.mazzucchelli@ulg.ac.be

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_5

Keywords Quantitative proteomics • Label-free quantitative proteomics • Label-based proteomics • Experimental design • PTMs • Isotope labeling

List of Abbreviations

AAA	Amino acid analysis
APEX	Absolute protein expression index
AQUA	Absolute quantification
cICAT	Cleavable isotope-coded affinity tag
CID	Collisional energy dissociation
DDA	Data-dependent acquisition
DESI	Desorption electrospray ionization
DIA	Data-independent acquisition
emPAI	Exponentially modified protein abundance index
ECD	Electron capture dissociation
ETD	Electron transfer dissociation
FFPE	Formalin fixation combined with paraffin embedment
FlexiQUANT	Full-length expressed stable isotope-labeled proteins for quantification
HCD	Higher-energy collision dissociation
H/D exchange	Hydrogen/deuterium exchange
HDX	Hydrogen deuterium exchange
HILIC	Hydrophilic interaction liquid chromatography
HRM	Hyper reaction monitoring
ICAT	Isotope-coded affinity tag
ICPL	Isotope-coded protein label
ICP-MS	Inductively coupled plasma mass spectrometry
ID	Isotopic dilution
IMS	Mass spectrometry imaging
iTRAQ	Isobaric tag for relative and absolute quantitation
LC	Liquid chromatography
LC-MS	Liquid chromatography coupled to mass spectrometry
LC-MS/MS	Liquid chromatography coupled to tandem mass spectrometry
LMD	Laser microdissection
MALDI	Matrix-assisted laser desorption ionization
MALDI-ISD	Matrix-assisted laser desorption ionization in-source decay
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS2	Tandem mass spectrometry
MS/MS	Tandem mass spectrometry
m/*z*	Mass-to-charge ratio
NGS	Next-generation sequencing
NHS	*N*-hydroxysuccinimide

PCS	Peptide-concatenated standard
PEO	Polyethylene oxide
PRM	Parallel reaction monitoring
PSAQ	Protein standard absolute quantification
PTM	Posttranslational modification
QC	Quality control sample
QconCAT	concatenation of *Q* (for quantification) peptides
QqQ	Triple quadrupole
RISQ	Recombinant isotope labeled and selenium quantified
ROI	Region of interest
RP-HPLC	Reverse-phase high-performance liquid chromatography
SILAC	Stable isotope labeling by amino acids in cell culture
SILAM	Stable isotope labeling of mammals
SWATH	Sequential windowed acquisition of all theoretical fragment ion mass spectra
TIC	Total ion current
TIMS	Trapped ion mobility spectrometry
TMT	Tandem mass tag
UPLC	Ultrahigh-performance liquid chromatography
XIC	Extracted ion chromatograms
XL	Cross-linked
ZIC-HILIC	Zwitterionic hydrophilic interaction liquid chromatography

1 Introduction

Depending of the biological question and the number of samples to be compared, an appropriate experimental design must first be carefully setup. The key points that should be accounted for by the experimentalist to succeed in this critical step are accordingly highlighted in the Sect. 2.

In the particular context of animal sciences, the two most widespread quantitative MS strategies, e.g., label-based and label-free methods, are in principle applicable. Nevertheless, a limitation arising from the restricted availability of relevant protein sequence databases may occasionally be prejudicial. The two forementioned methods are typically employed in the framework of differential quantitative analysis targeting common proteins identified across a series of samples. Absolute quantification by mass spectrometry is also possible and requires specific acquisition methods involving internal standards. Several approaches relying on these quantification strategies are described in the Sects. 3 and 4 with the objective to effectively guide the scientists within their research topics.

Specific issues related to the quantification of protein isoforms that commonly arise with glycoproteins and phosphorylated modified proteins are next discussed. Additionally, alternative examples of MS-based quantitative proteomics studies on

posttranslational modifications are highlighted. These aspects are gathered within the Sect. 5.

Finally, some alternative strategies that would help in understanding complex biological systems based on quantitative MS studies are described in the Sect. 6. These include applications in mass spectrometry imaging, top-down and middle-down proteomics, as well as interactions mapping through H/D exchange and cross-linking experiments.

2 Design of Experiments

The combination of experimental factors that impact proteomic studies is of primary importance and should be carefully considered at the early stage of the research. The use of appropriate controls and statistics is part of the experimental design. This will directly condition the false positive (type 1 errors) and false negative (type 2 errors) rates which may affect the final quality of the study and, in the worst cases, completely invalidate the results (Maes et al. 2015).

Fundamental principles of statistical experimental design were first described in 1926 by Ronald A. Fisher and further completed in the book of the same author in 1935 (Fisher 1926, 1971). R. A. Fisher introduced the concept of the *null hypothesis* in the significance testing approach which is still applied in modern design of experiments. In inferential statistics, the null hypothesis states that the effect of two treatments is comparable. In the significance testing approach, a null hypothesis is rejected when the observed data are significantly unlikely to have occurred if the null hypothesis was true. In other words, to validate a hypothesis, one should disprove its alternative. If the null hypothesis is incorrectly rejected, it will lead to a type 1 error. On the opposite, a failed rejection of the null hypothesis when the alternative is true will generate a type 2 error. The significance level is then the probability of obtaining a type 1 error when the null hypothesis is true.

These fundamental principles are discussed in the context of quantitative mass spectrometry-based proteomics in several reviews (Oberg and Vitek 2009; Maes et al. 2016; Ruderman 2017; Karp and Lilley 2007).

The current subchapter summarizes the key points that should be carefully addressed before starting every new research in which a proteomic approach is envisaged.

2.1 Defining the Research Question, the Related Hypothesis and the Appropriate MS-Based Proteomic Approach

Every research starts by clearly stating the scientific question of interest as well as identifying the sample types and groups of interest that will allow one to answer the

question. The question is then translated into specific hypotheses to be addressed. This process allows identification of the response variables to measure, to select the appropriate methodology accordingly, and to identify experimental factors that should be controlled to minimize their negative impact. Domon et al. described the options and considerations when selecting a quantitative proteomics strategy (Domon and Aebersold 2010). They considered three main approaches: shotgun (or discovery), directed, and targeted strategies. They proposed a rational framework for selecting optimal strategies to address the specific hypotheses under consideration. The different methodologies available in MS-based quantitative proteomics are further explained in the current chapter.

Whatever are the technologies selected and the global workflow applied, one should include the use of controls and consider the concepts of randomization, replication, and blocking. The number of samples included in a study will directly impact its statistical power. The power is defined as the probability of rejecting a null hypothesis when there is a significant difference between the considered groups, thus avoiding type 2 errors. The calculation of the power implies some knowledge the expected effects amplitudes we are looking for, defined as the signal, compared to the variation of the measured responses between samples from the same group, defined as the noise.

2.2 Controls

In addition to the biological positive and negative control samples which, when available, are part of the groups to be compared, two other kinds of control samples can be distinguished: the *experimental controls* and the *quality control samples* (QCs).

The experimental controls should be selected according the response variables to measure. Then, when an absolute quantitative study is conducted, the exact amount and purity of the controls should be known (Mallick and Kuster 2010; Russell and Lilley 2012; Geiger et al. 2011). In case of an enrichment procedure, i.e., for a peptide containing a posttranslational modification (PTM), the controls should contain the same PTM of interest. They should be incorporated in the sample at the early stage of the process to undergo all steps of the pre-analytical and analytical procedures. They are generally spiked into the biological samples and enable testing the ability of the global method to detect true variations.

The quality control samples are required to check the performances of the processes and equipments; they also allow to detect deviations by using control charts. They are typically run independently of the biological samples and on a periodical basis.

2.3 Replication, Randomization, Blocking

Concepts from experimental design such as replication, randomization, and blocking should be applied in all proteomic studies to avoid biased responses and to improve precision of the experiment's results.

As mentioned by Oberg et al., replication allows one to assess whether the observed difference in a measurement is likely to occur by a random chance. In addition, it also ensures the reliability of the conclusions deduced from the observed data (Oberg and Vitek 2009). Biological replication is performed to increase the measurement accuracy of the mean response values obtained for a population of interest using a subset of samples. Technical replication allows to reduce the impact of the process variability by analyzing multiple times the same biological sample. Scheerlinck et al. highlighted some considerations to minimize technical variations during sample preparation prior to label-free quantitative mass spectrometry analysis (Scheerlinck et al. 2015). When a method is reproducible and under control, the technical variability is mostly negligible compared to the biological one. In such circumstances, it is recommended to increase the number of biological replicates. This effect was highlighted in the work of C. Gan et al. where technical, experimental, and biological variations were estimated in a label-based quantitative proteomic study (Gan et al. 2007). Depending on the heterogeneity in the population of interest, up to ten biological replicates may be required to reach a sufficient statistical power. A. Paulovich et al. published an interesting work where they simulated different experimental scenarios for verification purposes including all statistical power calculation and false negative and positive rates based on a previous population study (Paulovich et al. 2008). Additional insights about reproducibility and interlaboratory ring trial studies making use of label-free proteomic methods are available in several recent publications (Choi et al. 2016; Al Shweiki et al. 2017).

Randomization is essential as it counteracts bias due to uncontrolled and unknown impacting variables. It is therefore mandatory to randomly select individuals from the underlying population and randomize the order of sample processing and analyses. This is even more crucial when a longitudinal collection of samples is conducted since randomization will eliminate any procedural temporal deviations.

By opposition to randomization, blocking helps to reduce the bias and variance due to predictable and known sources of experimental variations. For each source of variation, a balanced experiment is performed by alternating, among others, the experimental date, user, and reagent batch. By designing the experiment accordingly, each nuisance factor is expected to affect almost equitably all the groups to be compared (Ruderman 2017). When all the factors to consider cannot be balanced, randomization should be used instead, and a block-randomized design can be applied. J. A. Westbrook et al. discussed this strategy especially for label-based quantitative proteomics (Evans et al. 2012; Westbrook et al. 2015).

Sample pooling is generally considered when there is an incompatibility with the analytical resources, for budget consideration or for material availability purposes. One should therefore evaluate its impact properly from a statistical point of view. By pooling samples, the biological variability is reduced as all individual results are averaged. In addition, bias can be introduced due to the presence of outliers or nonlinear effects. Creating multiple sub-pools is generally mandatory to reach a sufficient study power (Paulovich et al. 2008; Oberg and Vitek 2009; Maes et al. 2016).

3 Differential Proteomics

Differential proteomics, or formerly termed expression proteomics, refers to the determination of quantitative maps of protein expression from different biological extracts such as organisms, tissues, organs, or cells (Smith and Veenstra 2003). It consists of quantitatively assessing the protein complement discrepancies between samples of interest, for instance, between a normal, or control, sample and its diseased counterparts. Historically developed with two-dimensional gel electrophoresis (Mustafa Ünlü et al. 1997; May et al. 2012), it is nowadays commonly implemented using online chromatography–mass spectrometry approaches. In this context, distinct strategies coexist and are principally subdivided in two categories referring to the label-based and the label-free methodologies. These are detailed in the following sections of the chapter.

3.1 Label-Based Proteomics

Label-based proteomics implies the chemical modification of proteins or peptides with stable isotope labels. Since these modifications are performed in the early phase of sample preparation, label-based quantification strategies are not affected by variability during sample handling and mass measurements. Therefore, samples can be fractionated leading to higher numbers of identified and quantified proteins, while maintaining low quantification errors. Although the current labeling options permit accurate comparison of several conditions, the experimental protocols are complex and, using a standard experimental design, a maximum of ten samples can be analyzed in parallel (Westbrook et al. 2015).

The labeling reaction may either be performed on postharvested samples, and thus qualified as in vitro labeling, or on metabolically active organisms, therefore referred to as in vivo labeling. The founding principles and objectives underlying both methods share similar key points consisting in the incorporation of stable isotope labels into a series of protein samples for the sake of relative quantitation (Nikolov et al. 2012). Their respective experimental implementations, features, advantages, and drawbacks are discussed hereafter.

3.1.1 In Vitro: Chemical Labeling

The in vitro labeling strategy is set as a postharvest methodology implemented on purified protein bulks. Because it does not target metabolically active samples, this method enables the study of a broader diversity of sample types with accurate control of the labeling step. Opposite to the metabolic labeling discussed in the next section, in vitro labeling protocols rely on a direct chemical linkage of isotopically labeled small molecules, referred to as tags, with reactive functions present on the proteins or peptides of interest. These functions usually correspond to either a free lysine or N-terminal amine groups or to a free cysteine sulfhydryl groups (Chahrour et al. 2015). Labeling reagents are commercially available as families of equivalent molecules sharing identical chemical skeleton with differential isotopic enrichments that enable their distinguishability using mass spectrometry.

The typical workflow consists of the differential tagging of the biological samples to analyze, for instance, samples collected at different timepoints along an experimental process, with a set of multiplexed stable isotope reagents. Depending on the labeling kit, from two (2-plex) to eight (8-plex) (Xie et al. 2011) and even ten (10-plex) (Werner et al. 2014), samples may be treated in parallel. The tagging step occurs either at the whole protein level, and is followed by the samples pooling and a unique one-pot enzymatic digestion, or at the proteolyzed peptides level consecutively to independent enzymatic digestions and prior to pooling. An alternative workflow consists of incorporating the labels during the enzymatic digestion itself by performing this one in heavy water ($H_2{}^{18}O$) (Busch et al. 2007) to substitute two ^{16}O for two ^{18}O at the peptides C-terminal extremity (Ye et al. 2009). Following the labeling and digestions steps, the workflow is usually pursued with the separation of peptides in liquid chromatography (LC) with subsequent identification by tandem mass spectrometry (MS/MS). The relative abundance of the corresponding differentially labeled peptides is established based on their respective peak intensities in the mass spectrum (Moritz and Meyer 2003). Depending on the reagent nature, the quantitation is either performed from the first stage MS1 spectrum (isotopic reagent) or from the second stage MS2 spectrum (isobaric reagent).

Isotopic Chemical Labeling

Isotopic labels consist of families of equivalent molecules with increasing nominal masses obtained through the gradual substitution of light atoms, i.e., ^{1}H, ^{12}C, and/or ^{14}N, for their heavy isotope counterparts, i.e., ^{2}H, ^{13}C, and/or ^{15}N. The particular substitution of ^{1}H for ^{2}H is however commonly avoided nowadays because the differentially tagged species do not exactly co-elute in reverse-phase liquid chromatographic separation RP-HPLC (Zhang et al. 2001). Once labeled, the resulting mass of the targeted proteins or peptides is offset to an extent agreeing with their respective tag mass differences in such a way that they appear at distinct mass-to-charge *(m/z)* ratios in the MS1 spectrum. In such circumstances, their relative abundance may be evaluated from the ratios of their respective area in the total ion current (TIC) chromatogram. Several isotopic tag families are available and discussed below.

Isotope-Coded Affinity Tag (ICAT)

The ICAT label family was introduced by Gygi and coworkers in 1999 and corresponds to biotinylated iodoacetamide or acrylamide derivatives that specifically react with the sulfhydryl functions of cysteine side chains (Gygi et al. 1999). The biotin moiety, dedicated to the affinity capture of labeled peptides onto a streptavidin-based column for purification purposes, and the reactive iodoacetamide moiety are interconnected by a polyethylene oxide (PEO) linker that may be isotopically enriched. In this context, the reagent exists both in a light or normal form and in a heavy form, in which eight ^{1}H were formerly substituted with ^{2}H before the deuterium coding gets later replaced with a ^{13}C coding (Zhang et al. 2006). The derivatization of cysteines typically occurs at the whole protein level, after denaturation and reduction, before the samples get combined and digested. In 2003, a cleavable cICAT was introduced to allow acidic removal of the biotin moiety prior to LC analysis as its hydrophobicity causes a narrowing of the elution time windows for tagged peptides (Li et al. 2003). While this technique has been successfully employed for quantitative analysis of animal origin samples (Guevel et al. 2011), it is often associated with poor sequence coverage as the method is only applicable to cysteine-bearing proteins (Schmidt et al. 2005). Hence, a high number of proteins are only quantified based on a single peptide, which drastically reduces quantitation accuracy. Additionally, the presence of the ICAT tag on the peptide influences its fragmentation spectra, making the identification process more difficult.

Isotope-Coded Protein Label (ICPL)

The ICPL tags were developed by Schmidt and coworkers in 2005 to overcome previously discussed limitations of ICAT (Schmidt et al. 2005). Historically intended to be applied at the whole protein level before being implemented post-digestion, this methodology aims at labeling primary amine groups localized on lysine side chain and N-terminal extremities (Lottspeich and Kellermann 2011). To this end, the ICPL reagents are organized around an isotopically coded nicotinic acid ester core that is functionalized with an *N*-hydroxysuccinimide (NHS) nucleofuge group. The combination of differential ^{2}H and ^{13}C enrichment enables the synthesis and commercialization of duplex as well as more recent triplex and quadruplex ICPL labeling kits.

Even though the sequence coverage gets improved over ICAT, studies revealed that a substantial proportion of identified peptides in trypsin digested samples lack a lysine and, as a consequence, are not quantifiable when the labeling is performed at the whole protein level (Mastroleo et al. 2009; Paradela et al. 2010). Typically, in such circumstances, ICPL allows for the limited quantification of approximately 70% of the identified proteins. Besides, the presence of tags on lysine residues decreases the trypsin digestion efficiency as this one becomes only effective after arginine residues. This effect results in the production of longer peptides that prevents efficient fragmentation by collisional energy dissociation (CID) and higher-energy collision dissociation (HCD). A solution is to perform the labeling reaction on the proteolyzed peptides, after the enzymatic digestion: as resultant N-terminal extremities become amenable for reaction (Leroy et al. 2010).

The ICPL methodology has already been successfully employed in several studies targeting samples from mammalian origin (Biniossek et al. 2013; Rainczuk et al. 2013; Maccarrone et al. 2014).

Dimethyl Labeling

Derivatization of lysine and N-terminal primary amines with isotopomeric dimethyl labels is an alternative methodology to ICPL labeling. This one relies on the reaction with formaldehyde to generate a Schiff base which is further reduced by the addition of cyanoborohydride (Sap and Demmers 2012). This addition results in the formation of a dimethylamine with a nominal mass shift of +28 Da. Incorporation of heavy labels is achieved by using ^{2}H and/or ^{13}C enriched formaldehyde reactants, possibly in combination with cyanoborodeuteride. In this context, studies exploiting from 3-plex up to 5-plexdimethyl labeling strategies have been reported (Boersema et al. 2008; Wu et al. 2014).

The labeling step is typically performed at the peptide level and has the advantage to use inexpensive reagents and to be applicable to virtually any sample. In-solution, online and on-column protocols have furthermore been developed for labeling of sample amounts ranging from submicrograms to milligram and implemented in various studies (Synowsky et al. 2009; Boersema et al. 2009). For instance, the method has been applied to highlight composition discrepancies in the bovine proteasome complexes extracted from the liver and from the spleen (Raijmakers et al. 2008).

Isobaric Chemical Labeling

As stated above, the quantitation approach for the isotopic chemical labeling considers the MS1 extracted ion chromatograms of the differentially tagged peptides. This strategy was historically developed with duplex kit and displays limited compatibility with higher-order multiplexing as this leads to substantial complexity of MS spectra and requires the use of high-resolution mass spectrometers to distinguish between highly charged species. For instance, triply charged peptides differentially tagged with triplex ICPL are only separated by 0.66 *m*/*z*. These limitations are alleviated using alternative labeling technologies based on isobaric kits which offer higher multiplexing capability. These reagents encompass isobaric tag for relative and absolute quantitation (iTRAQ) (Ross et al. 2004) and tandem mass tag (TMT) (Thompson et al. 2003) which allow parallel monitoring of two (TMT 2-plex), four (iTRAQ 4-plex), six (TMT 6-plex), eight (iTRAQ 8-plex), or ten (TMT 10-plex) experimental conditions at the same time. Recently, a new generation of reactants called combinatorial isobaric mass tags (CMTs) with wider multiplexing capability from 16-plex to 28-plex was introduced (Braun et al. 2015).

TMT and iTRAQ labeling is usually performed post-digestion at the peptide level. The tags share a common structure based on a nucleofuge NHS ester group targeting lysine and N-terminal amines, a balance or mass normalizer group and a reporter group, which dissociate from each other upon CID or HCD fragmentations.

Either TMT or iTRAQ labels have respective identical overall mass, and are therefore referred to as isobaric, but differ by the isotopic distribution on their constitutive atoms. At the peptide level, differentially tagged species are chromatographically equivalent and isobaric and give rise to a single peak in the MS1 spectrum. Upon fragmentation, the balance group is lost as a neutral, and the isolated reporter groups give rise to small ions whose respective masses are shifted from a fixed increment, typically 1 Da, in the MS2 spectrum. Their respective intensities are correlated with these of the initially differently tagged peptides, which enables further relative quantitation (Christoforou and Lilley 2012). The methodology is illustrated in Fig. 1.

With the exception of the 10-plex TMT kit detailed hereafter, the 1-Da increment on the reporter ion masses facilitates their discrimination even within low-resolution datasets. However, the masses of the reporter ions themselves are comprised between 113 and 121 Da (8-plex iTRAQ) or 126 and 131 Da (TMT 6-/10-plex) and require further method adjustments as this range is typically not acquired in common proteomics analysis. The isobaric method yet displays a disadvantage compared to the previously discussed isotopic strategy. Indeed, the isobaric tagging prevents straight application of targeted methods toward the preferential monitoring of peptides with differential abundances as this information is not accessible at the entire peptide level. Additionally, an issue popularly referred to as "ratio compression" may undermine the quantitation accuracy using isobaric tags (Ow et al. 2009; Martinez-Val et al. 2016). This phenomenon arises from the co-isolation of isobaric precursor peptides of diverse abundance with other steady peptides within the fragmentation window. These latter ions constitute a constant background that induces a shift in the measured reporter ion ratio toward unity, therefore not reflecting the true value awaited for the isolated peptide of interest. Two strategies involving either proton transfer reactions dedicated to a charge reduction of the co-isolated ions (Wenger et al. 2011) or additional fragmentation stages (MS3) (Ting et al. 2011) have been developed to reduce interferences and subsequent ratio distortion.

Isobaric Tag for Relative and Absolute Quantitation (iTRAQ)

iTRAQ kits are available as 4-plex and 8-plex kits and are built around a *N*-methylpiperazine reporter group, a carbonyl balance group and a NHS ester amine reactive group. The isotopic substitution is differentially performed using ^{13}C, ^{15}N, and ^{18}O atoms so that the mass of the reporter group varies between 114.1 and 117.1 Da for the 4-plex kit and 113.1 and 121.1 Da for the 8-plex kit. The mass of the balance group varies accordingly to keep the overall mass constant. Since its creation, iTRAQ never stops gaining in popularity and is now considered as the most employed labeling strategy nowadays with more than 400 entries found in PubMed for the single year 2016, well above TMT (238), stable isotope labeling by amino acids in cell culture (SILAC) (196) and ICAT (36). It has been used in numerous quantitative studies involving mammalian samples, for instance, to highlight the effect of gender and diet on pig skeletal muscle proteome (Hakimov et al. 2009) or to assess the dissimilarities in milk fat globule proteome between human and bovine (Yang et al. 2015b, 2016).

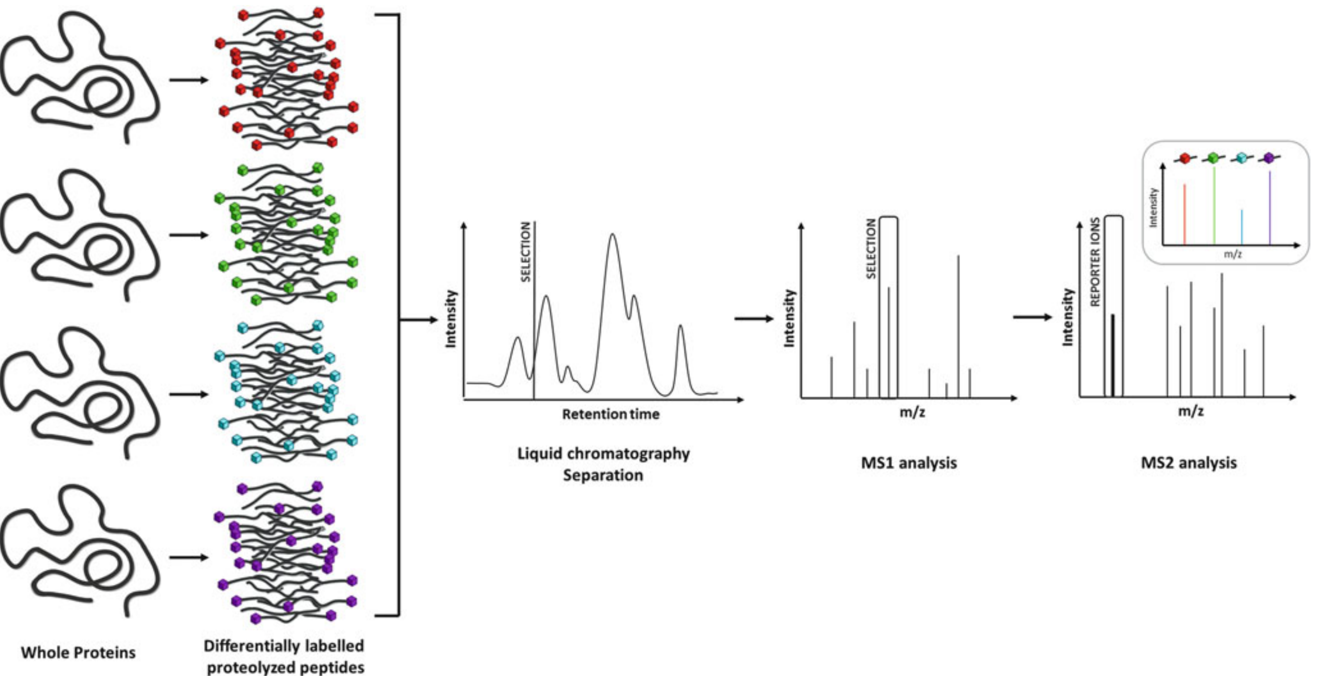

Fig. 1 Isobaric quantitative proteomic workflow illustrated for the 4-plex iTRAQ method. The labeling step is performed independently post-digestion before the samples are pooled. The peptides mix is subsequently separated using LC and analyzed using MS. The peptide of interest is selected (MS1 spectrum) and further fragmented for identification (MS2 spectrum). The relative quantitation is based on the reporter ions ratios in the MS2 spectrum

Tandem Mass Tag (TMT)

TMT reagents are available as duplex, 6-plex, and 10-plex kits and share a similar construction as iTRAQ molecules except that reporter groups now correspond to differentially isotopically enriched 1,5-dimethylpiperidines. The 6-plex kit corresponds to a differential incorporation of ^{13}C and ^{15}N isotopes in such a way that the mass of the reporter group is gradually incremented by 1 Da steps over the series. The 10-plex kit is set as an expansion of the 6-plex kit with four isotope variants relying on the small 6.32 mDa difference in $^{12}C/^{13}C$ and $^{14}N/^{15}N$ isotopic pairs (McAlister et al. 2012). Although marginal, this difference is sufficient to achieve baseline separation of reporter ions using high-resolution analyzers (Werner et al. 2012). While TMT offers higher multiplexing capability than iTRAQ, recent studies reveal that it is also less sensitive with labeling achievement on the order of 1.5-fold higher using 4-plex iTRAQ compared to 6-/10-plex TMT (Svinkina et al. 2015). The TMT labeling strategy has been exploited in many studies, namely, for quantification of posttranslational modification (PTM) (Liang et al. 2015), quantitative analysis of complex protein mixtures (Thompson et al. 2003), or to monitor the effect of specific inputs on plant samples (Lv et al. 2016).

3.1.2 In Vivo: Metabolic Labeling

Metabolic labeling refers to methods in which the machinery synthesis of living cells is used to incorporate stable isotope labels. This approach was first described in a proteomic context by Oda et al. who labeled proteins using nitrogen (^{15}N) in bacterial cell culture (Oda et al. 1999). In 2002, Matthias Mann and coworkers introduced the so-called stable isotope labeling by amino acids in cell culture (SILAC), which is still one of the most widely used methods exploiting metabolic labeling (Ong 2002). This method relies on the incorporation of amino acids with substituted stable isotopic nuclei (e.g., ^{2}H, ^{13}C, ^{15}N). Since mammalian cells cannot synthesize essential amino acids, external and controlled supply of such amino acids to cells in culture will lead to their incorporation into all newly synthesized proteins. By supplying a "light" or a "heavy" form of a particular amino acid (e.g., ^{12}C- and ^{13}C-labeled L-lysine, respectively) in the media formulation, light and heavy cell population can be generated. These cells can be directly mixed after the cell culture experiment as their related proteins remain distinguishable and then quantifiable by MS (Ong and Mann 2006; Van Hoof et al. 2007). In addition to cell culture, several approaches were followed to monitor protein dynamics in living animals. Nowadays, almost all model organisms from bacteria to rodents can be fully labeled with SILAC (Mann 2014).

"Stable isotope labeling of mammals" (SILAM) is an alternative method to SILAC intended to generate stable isotope-labeled animals. Rauniyar et al. detailed a method of metabolic labeling on rats with heavy nitrogen using ^{15}N-enriched spirulina diet (Rauniyar et al. 2013).

Due to the tight requirements on medium properties coupled with time-consuming and expensive processes, in vivo labeling methods tend to display limited applicability for quantitative proteomics studies (Sap and Demmers 2012). In particular, in the case of protocols targeting culture cell proteins, the request for dialyzed serum used to prevent natural amino acid contaminations may interfere with optimal cell line growth. Additionally, the downstream conversion of labeled arginines to prolines has been reported in the literature (Van Hoof et al. 2007). Considering higher organisms, studies relying on SILAC were reported with specimen such as mice (Krüger et al. 2008), flies (Sury et al. 2010), and plants (Gruhler et al. 2005), but the method is yet to be applied on human or bigger mammalian or human tissues due to technical, financial, and ethical considerations.

3.2 Label-Free Quantification

In contrast to label-based approaches, label-free quantification methods are cost-effective and simple and can be applied to dozens or even hundreds of samples, providing that sample preparation and LC-MS analysis are performed under strictly controlled conditions (Bantscheff et al. 2007; Neilson et al. 2011). In order to minimize the experimental variability, label-free proteomics workflows do not normally include protein fractionation steps and therefore may lead to lower numbers of identified and quantified proteins compared to label-based methods. Rapid development of new MS technologies and sample preparation workflows help to close this gap. In the early 2000s, the identification of a few hundred proteins in biological samples required days of sample pre-fractionation and MS analysis (Washburn et al. 2001), while current proteomics workflows and technologies allow for the identification of the complete yeast proteome in a single hour (Hebert et al. 2014). This impressive progress is the result of reproducible and multiplexed gel-free and bottom-up sample preparation protocols being developed which require minimal sample material (e.g., FASP, PCT, SP3, etc.) (Zougman et al. 2009; Hughes et al. 2014; Shao et al. 2015) and of improvements in the sensitivity and speed of MS systems (Richards et al. 2015). In the near future, the same technical advances may lead to the analysis of samples containing complete expressed mammalian proteomes with an estimated 10,000–12,000 proteins within a few hours (Mann et al. 2013). Unfortunately, for animals other than the most common model organisms, the same proteomics workflows are impaired by the limited availability of genomic data and/or well-annotated protein databases. This often leads to a lower number of identified proteins (Soares et al. 2012). Hopefully, the rise of next-generation sequencing (NGS) technologies will lead to greater availability of genomics resources for animal species.

Independently of the organism under study, the current proteomics label-free experiments go beyond producing long list of protein identifications. High-quality proteome-wide quantitative information can be produced even for a large set of samples. Figure 2 shows the main steps of the three most relevant bottom-up label-free

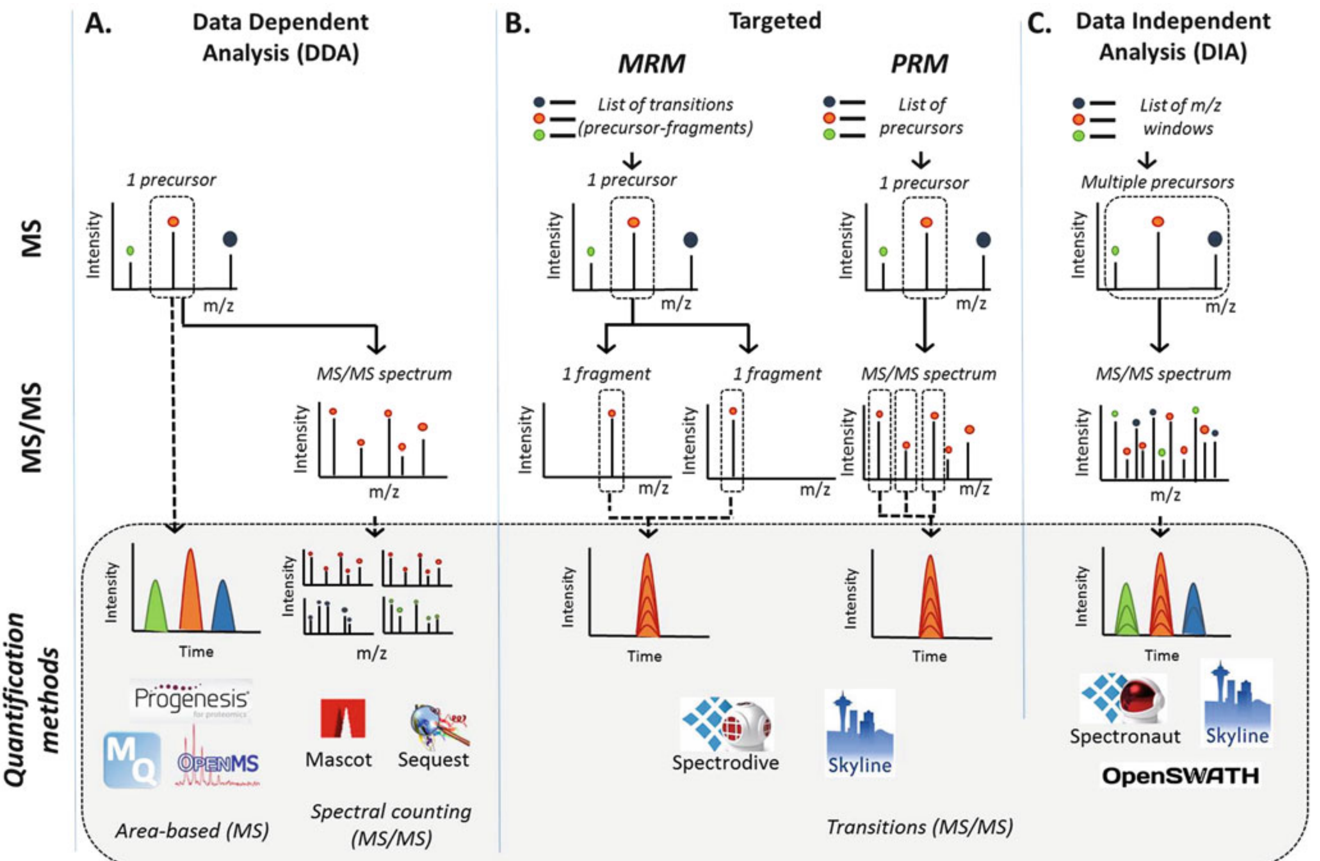

Fig. 2 Overview of the three main label-free quantification workflows. (**a**) Data-dependent acquisition (DDA), (**b**) multiple reaction monitoring (MRM) and parallel reaction monitoring (PRM), (**c**) data-independent acquisition (DIA)

quantification approaches: (A) data-dependent acquisition (DDA), with both MS and MS^2 quantification which is used for discovery studies; (B) targeted quantification [multiple and parallel reaction monitoring (MRM, PRM)]; and (C) data-independent acquisition (DIA) (Aebersold and Mann 2016). Each one of these workflows will be described in the following sections.

3.2.1 Data-Dependent Acquisition (DDA): Shotgun Proteomics

A typical protein identification workflow begins with extracting proteins from the samples, followed by their enzymatic digestion into peptides and analysis via reversed-phase liquid chromatography coupled to mass spectrometry (LC-MS). The mass spectrometer is normally operated in an alternating mode where both MS and MS/MS spectra are acquired, using an approach called data-dependent acquisition (DDA, Fig. 2a). The MS spectra contain the mass-to-charge ratio (m/z) of the peptide precursors and are normally acquired in a range between m/z 300 and 1500. The most intense peptide signals for each MS spectrum are selected for fragmentation, and the resulting fragments are acquired in MS^2 spectra (Elias and Gygi 2007; Nesvizhskii et al. 2007). The combination of MS and MS/MS spectra and the use of database search engines make it possible to identify peptide and protein sequences.

MS spectra of complex biological samples contain dozens or hundreds of peptide signals, but the current LC-MS platforms are not fast enough to isolate and fragment all these ions and instead only sequence the more intense precursors (hence the name shotgun proteomics). Consequently, in DDA experiments many of the least abundant analytes are not identified (Domon and Aebersold 2010). Nonetheless, shotgun proteomics datasets are rich in quantitative information that can be extracted in multiple ways from MS and MS^2 spectra.

The easiest way to perform label-free protein quantification is by spectral counting, where the relative amount of protein is estimated by comparing how many MS^2 spectra were acquired for the same protein in two or more LC-MS runs, or for different proteins within a single sample (Wang et al. 2006; Zhu et al. 2010). Most search engines provide spectral counting data, making this approach easily applicable for proteomics studies (Fig. 2a, lower panel). For absolute quantitation, Ishihama et al. introduced the exponentially modified protein abundance index (emPAI), which is proportional to protein content in a protein mixture (Ishihama et al. 2005). Since the detection of peptides in MS is severely affected by their physicochemical properties, the absolute protein expression index (APEX) (Lu et al. 2007) was introduced to account for variable peptide detection by MS techniques and improve the accuracy of the relative and absolute protein quantification.

Spectral counting MS^2 approaches have been frequently used in various animal proteomics studies since they are powerful and easy to perform [e.g., (Huan et al. 2015; Neely et al. 2015)], but they lack accuracy and linearity (Wang et al. 2006). The number of detectable peptide signals increases with larger protein abundance,

but given the way shotgun proteomics works, the number and type of identified peptides is biased toward abundant peptides and does not always correlate linearly with their actual abundance (Bantscheff et al. 2007). For this reason, alternative label-free protein quantification strategies that focus on MS instead of MS^2 spectra have been implemented. These approaches are capable of measuring changes in protein abundance with higher precision and accuracy (Fig. 2a, lower panel). MS spectrum-based methods rely on the correlation between protein concentration and the peak areas its peptides produce. More precisely, the peak area of one or more peptides in the MS spectra is integrated and used to compare the expression levels of a protein among samples.

Area-based quantification methods can be used for proteome-wide studies or for monitoring a few specific peptides. Large-scale proteomics studies require special software to combine data from several LC-MS analyses and to perform relative protein quantification. Such software (e.g., MaxQuant (Tyanova et al. 2016), Progenesis (NonLinear Dynamics 2017), or OpenMS (Sturm et al. 2008)) extract the retention time, m/z, and area of each peptide signal in each run and align them to correct for the natural chromatographic variability among LC-MS runs. The aligned and identified peptides are then quantified, and the data are processed to provide statistically evaluated quantification results at the protein level. Recently, this workflow has been applied to investigate mastitis in bovine milk (Mudaliar et al. 2016), the effects of castration on pig meat (Shi et al. 2016), farming practice on wild gilthead sea bream (Piovesana et al. 2016), growth path on beef production (Almeida et al. 2017), and weight loss on the mammary gland secretory tissue proteome of goat (Hernández-Castellano et al. 2016).

Examples of label-free quantification studies measuring only a few peptides are the evaluation of protein degradation throughout the ham dry-curing process (Gallego et al. 2016a, b) and the quantification of proteins of the parasite *Anisakis simplex* in fish (Fæste et al. 2016). In these studies quantification is performed without any of the abovementioned software, but by analyzing the extracted ion chromatograms (XICs) of each target peptide in each sample. This is feasible if the number of proteins of interest is small. However, if only a handful of proteins are of interest but these need to be quantified across many samples, targeted mass spectrometry monitoring methods are usually better suited than DDA. The following section describes the two most widely used targeted MS techniques: multiple reaction monitoring (MRM) and PRM.

3.2.2 Targeted MS Techniques: Multiple/Parallel Reaction Monitoring (MRM/PRM)

Multiple reaction monitoring (MRM) is a targeted MS approach which has been applied to a wide range of proteomics studies (Picotti et al. 2009). In a MRM experiment, only a list of preselected precursor (peptide) and fragment pairs, so-called transitions, are monitored (Fig. 2b). MRM analyses are performed on triple quadrupole (QqQ) instruments: during the MS analysis, peptide mixtures

enter the first quadrupole (Q1), which lets only a selected precursor *m*/*z* transit to the second quadrupole (Q2) for fragmentation (Lange et al. 2008). The fragments are transferred to the third quadrupole (Q3) which only lets fragments with a selected *m*/*z* pass to the detector (Fig. 3a). Within a few seconds, the mass spectrometer cycles through all transitions on the list and acquires the corresponding precursor and fragment intensities with much higher sensitivity than DDA. This allows quantification of medium- to low-abundant peptides down to 50 copies per cell (Picotti et al. 2009). Due to the use of a transition list, MRM ensures that the same set of proteins and peptides is measured across many samples without the missing data points typical for DDA experiments making it especially suitable for quantification studies (Picotti and Aebersold 2012).

A drawback of this type of targeted MRM methods is that the user needs to decide a priori which peptides and which fragments he wants to monitor in the experiment. The selected target peptides should be unique for the protein of interest (proteotypic), and both precursor and fragment should ionize well in order to be detectable in a mass spectrometer. This depends on peptide length, charge, post-translational modifications, the occurrence of missed cleavages, fragment charge, and other characteristics (Lange et al. 2008). Therefore, MRM transitions need to be optimized a priori on peptide and fragment level to ensure that the most intense fragments yielding the best signal with least interference are monitored. This makes transition selection a time-consuming process. Using synthetic peptides can accelerate assay development and allows high-throughput generation of transition lists (Picotti et al. 2010). Once a set of MRM transitions has been optimized, they can be reused and shared with the scientific community via dedicated databases (Institute for System Biology 2004; Picotti et al. 2009; Kusebauch et al. 2016).

In recent years a novel targeted method called parallel reaction monitoring (PRM) emerged (Bourmaud et al. 2016). It relies on instruments combining a quadrupole mass filter with a high-resolution mass analyzer, such as an orbitrap or a time-of-flight analyzer. Similar to MRM, the quadrupole acts as mass filter and lets only preselected precursor *m*/*z* pass. After fragmentation of the precursor, a full fragment ion spectrum is acquired in the high-resolution mass analyzer (Figs. 2b and 3b). MRM and PRM share many common features, but PRM has the advantage that suitable fragments can retrospectively be selected from the full product ion spectra, reducing the time required for selecting the best transitions. Additionally, peptide fragments are measured in a low-resolution, low mass accuracy quadrupole in MRM experiments which can lead to false-positive signals. Validation with synthetic peptides spiked into a sample might therefore be necessary. PRM offers an advantage here since more fragments are monitored and with higher resolution, which increases the information content of the spectra and reduces the risk for false positives.

A multitude of software is available to support the selection of peptide transitions and data analysis of MRM and PRM results. Among them are Skyline (MacLean et al. 2010) and SpectroDive [Biognosys (Ori et al. 2016)], as well as MS instrument vendor software. These software tools assist with automated data analysis and visualize transitions as peaks of co-eluting fragments. Peptides are

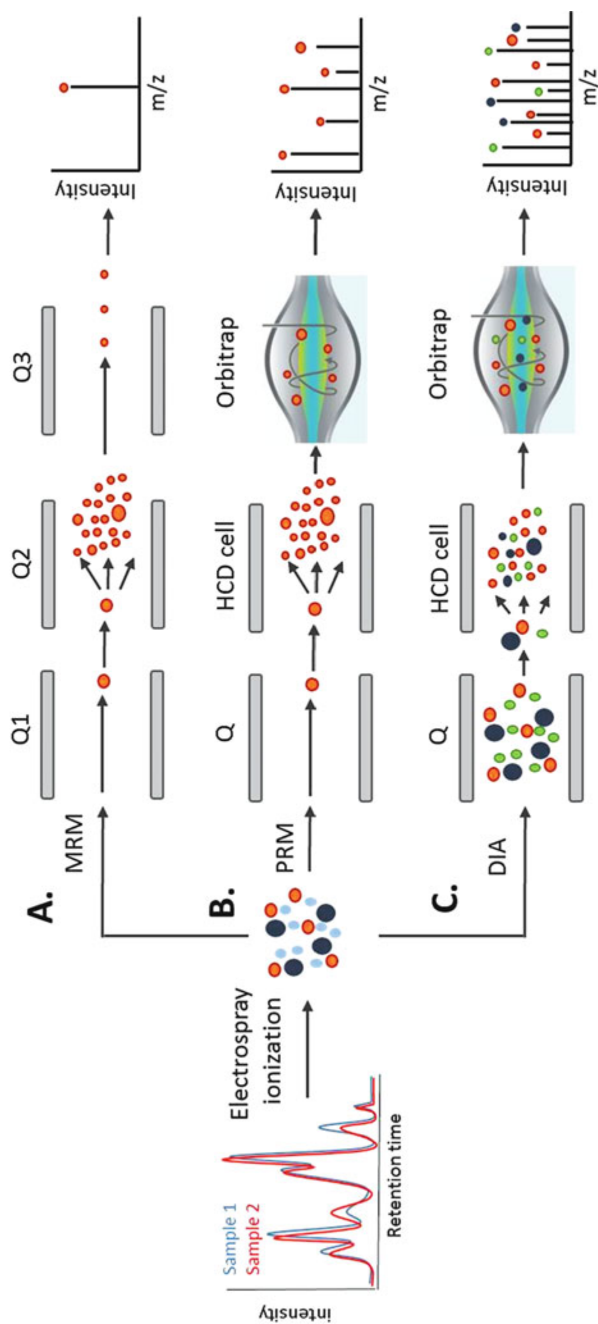

Fig. 3 Targeted and data-independent acquisition (DIA) mass spectrometry methods. (**a**) Multiple reaction monitoring (MRM), (**b**) parallel reaction monitoring (PRM), (**c**) data-independent acquisition

quantified by integrating the area under the most intense transition or the summed area of all transitions (Fig. 2b, lower panel).

The number of animal proteomics studies where targeted proteomics has been used is still limited, but both MRM and PRM have great potential. For example, PRM was recently used to discriminate and quantify cashmere fibers in commercial finished fabrics that have undergone heavy chemical treatment and as an alternative to light microscopy for assessing the authenticity of cashmere fibers and textiles (Li et al. 2016). Another interesting application is the detection of indirect markers of steroids treatment in bovine muscle using MRM (Stella et al. 2015). In this study, the authors developed a method to quantify 28 candidate protein markers whose expression is significantly altered after the administration of steroids.

In the two cited studies, only a few dozen peptides and proteins are monitored, but targeted techniques allow quantification of several hundred peptides in a single run. However, typically several peptides are monitored per protein of interest to ensure its unambiguous quantification. Thus, even when several hundred peptides are monitored, the number of proteins that can be quantified per run is limited. If a large number of proteins need to be quantified, DIA methods might be more suitable than targeted MS approaches.

3.2.3 Data-Independent Acquisition (DIA)

Within the last 5 years, DIA emerged as a new mass spectrometry approach which further expands the proteomic toolbox (Law and Lim 2013). DIA methods offer reliable peak detection and similar sensitivity and dynamic range as the targeted techniques. However, instead of measuring only a limited number of targets as in MRM and PRM, they strive to identify and quantify all peptides in a sample, potentially the whole proteome in one run. A further advantage of DIA over targeted MS methods is that it does not require prior knowledge of the sample content.

The term DIA summarizes a number of individual techniques such as hyper reaction monitoring (HRM), sequential windowed acquisition of all theoretical fragment ion mass spectra (SWATH), MSe and all-ion fragmentation, and many others (Gillet et al. 2012; Bilbao et al. 2015; Bruderer et al. 2015). While DDA and targeted MS methods isolate a single precursor for fragmentation (mass window, 1–2 Th), all DIA methods fragment larger mass windows (e.g., of 25 Th width) containing several peptide precursors. The fragments are acquired on a high-resolution mass analyzer (Fig. 2c), resulting in complex MS/MS spectra containing fragment ions from several precursors (Fig. 3c).

Usually, DIA experiments do include the acquisition of MS spectra but focus mainly on the acquisition of fragment ion spectra (MS/MS). As a consequence, peptide identification and quantification are typically based on MS/MS data. Data analysis is challenging since DIA spectra contain fragments from several precursors. Therefore, specialized software is used, such as Spectronaut (Biognosys), OpenSWATH, SWATH 2.0 (SCIEX), Skyline, or DIA-Umpire (Reiter et al.

2011; Bernhardt et al. 2012; Röst et al. 2014; Bilbao et al. 2015; Tsou et al. 2015; Hu et al. 2016). These programs group co-eluting fragments into MRM-like peaks, assign peptide identities to the data, and integrate peak areas for quantification, thus making data analysis easy for users.

Different methods exist for assigning peptide identities to the fragment ion peaks. The majority of workflows rely on a peptide-centric approach and spectral libraries for this task. A number of spectral libraries for different organisms and tissues are publicly available, and good results can be achieved using these resources (Institute for Systems Biology 2013; Muntel et al. 2015). Nonetheless, the best peptide identification results are attained with a spectral library generated from DDA runs acquired on the same samples as the DIA runs. Thus during the DDA runs for library generation, the sample type, sample preparation, chromatography setup, and instrument parameters should be as similar as possible to the conditions used in the final DIA experiment. Many of the abovementioned analysis software have built-in features to compile a spectral library from the resulting DDA fragment ion spectra and their database search results (e.g., from MaxQuant, Proteome Discoverer, Mascot, etc.). Other data analysis workflows do not require an existing spectral library: DIA-Umpire uses a data-centric approach to detect peaks and assemble fragments into pseudo-MS^2 spectra (Tsou et al. 2015). These spectra can be searched using conventional search engines that assign peptide identifications by matching with theoretical spectra. Moreover, DIA-Umpire can generate a spectral library from the pseudo-MS/MS spectra which can be used to extract quantitative data from the DIA runs.

DIA's advantages over the classic mass spectrometry approaches have led to DIA being termed "next-generation proteomics." DIA fragments and attempts to sequence all peaks eluting from the chromatography column. Thus, it allows a broader coverage of the proteome compared to targeted techniques which monitor only limited sets of targets, and to DDA where only the more intense precursors are fragmented. Additionally, DIA is not affected by inconsistent precursor selection, and missing peptide identification data points are rare in DIA compared to DDA experiments (Bruderer et al. 2015). The reliable peak picking makes DIA approaches suitable for monitoring the same peptides across many samples (e.g., replicates, conditions, timepoints).

Finally, DIA techniques offer excellent sensitivity (attomolar range) and a large dynamic range spanning four orders of magnitude (Gillet et al. 2012). These numbers are almost comparable to what can be achieved using targeted MS but without the need to decide a priori which peptides to monitor in an experiment.

A DIA approach (SWATH) has been recently applied to monitor the effects of high temperature on broiler production efficiency and meat yield (Tang et al. 2015). More than 2300 proteins were relatively quantified in broiler liver under different heat stress, leading to the identification of broiler self-regulation mechanisms. Another example where the SWATH method was applied is the identification of boar seminal plasma and the comparison of ejaculate portions, with the aim to identify proteins with potential roles as sperm function and fertility biomarkers (Perez-Patiño et al. 2016).

4 Absolute Quantification

All the strategies presented in the previous section allow straightforward relative quantification of proteins between several samples. Absolute quantification, however, requires that a molar or mass concentration (or even number of copy per cell) for each of the proteins of interest could be determined. To allow such a determination, targeted methods are usually used such as MRM and PRM in combination with an addition of a known amount of an internal standard during the sample preparation. This known amount of internal standard allows the calculation of the concentration of the peptides of interest. As the proteins are digested using specific protease (usually trypsin) before LC-MS/MS, the quantification is done at the peptide level, and the concentration of the protein is deduced from those of the proteotypic peptides chosen for this protein.

Absolute quantification is often applied to a limited number of proteins (validated biomarkers for clinical assay, proteins for food safety, etc.) although recently this number is increasing. In this context, Borchers and coworkers have quantified simultaneously 45 proteins in plasma (Kuzyk et al. 2009) and 136 proteins in urine (Percy et al. 2015).

The calculation of the peptide/protein concentration is done either by constructing a calibration curve or by single-point calibration. In proteomics, single-point calibration is most often used, mainly due to the unavailability of a blank matrix [a sample similar to the one studied but that does not contain the protein(s) of interest]. For single-point calibration, the amount of the internal standard is multiplied by the ratio of the area of the target analyte divided by the area of internal standard to calculate the amount of target peptide.

As mentioned in the section of experimental design, it is quite mandatory for absolute quantification to run quality control (QC) samples with each series of experiment. These QC samples will allow assessing whether, among others, recovery of the targets, digestion efficiency, and LC-MS/MS system have been performed equivalently between the series of analyses.

The internal standard should ideally have the same physicochemical properties as the analyte to quantify in order to account for any modification happening to the analyte during sample preparation (loss, adsorption, absorption, chemical modification, etc.) and for any variation encountered during mass spectrometric analysis (ionization efficiency, ion suppression due to matrix effect, etc.). For LC-MS/MS analyses, the golden standard is the isotopic dilution (ID) method where the internal standard is the same molecule as the target analyte but labeled with stable isotopes (^{13}C, ^{15}N, ^{18}O, ^{2}H, etc.). ID is usually combined with MRM analysis and now also PRM.

In isotopic dilution, the sooner the internal standard is added, the better it can account for loss, degradation, or any event happening from the sample preparation to the MS detection. Currently, proteins are quantified after protease digestion, and the constitutive peptides are measured in the mass spectrometer. The first isotopically labeled internal standard used in proteomics was thus a peptide. The isotopic

dilution method was first used for peptides in 1983 combined with fast-atom bombardment mass spectrometry (Desiderio and Kai 1983) and in 2003 in combination with MRM (Barnidge et al. 2003; Gerber et al. 2003). If the internal standard is produced in-house, purification, verification of identity, and quantification are the most important points to control to ensure its adequacy.

Historically, deuterium was the first isotope used for stable isotope labeling. However, the replacement of hydrogen by deuterium causes a change in the physicochemical properties of the molecule leading to, among others, a slight change in the retention time in reversed-phase chromatography (Ciccimaro and Blair 2010). As a consequence, the ^{13}C and ^{15}N stable isotopes are nowadays the most employed for isotopic labeling in the field of isotopic dilution coupled to LC-MS/MS.

The full panel of internal standards available in quantitative proteomics has been extensively reviewed recently (Kito and Ito 2008; Brun et al. 2009; Villanueva et al. 2014). In the next sections, only the most commonly used types of stable isotope-labeled standard are presented: peptides, synthetic concatemer proteins, and full-length proteins.

4.1 Stable Isotope-Labeled Peptides (AQUA)

As peptides are measured by LC-MS^2 for protein identification and quantification, the most evident internal standard is the isotopically labeled peptide corresponding to the target peptide. As peptides can be chemically synthetized, a stable isotope-labeled amino acid can be easily inserted into the peptide sequence. Deuterium was first used for this labeling, but now ^{13}C and ^{15}N are preferred because they do not induce a change in retention time in reversed-phase chromatography. Consequently, the peptide and its stable isotope-labeled standard perfectly co-elute. This is of great help in complex matrix to ensure that the proper target peptide is detected and quantified (Lange et al. 2008).

Peptides labeled with ^{13}C and ^{15}N were used by Gerber and coworkers to quantify both unmodified and phosphorylated peptides of phosphoproteins in whole cell lysate (Gerber et al. 2003). Those peptides are purified and accurately quantified using amino acid analysis (AAA). They named this strategy "AQUA" for absolute quantification, and this acronym, that later became a trade name, is now widely used by the proteomics community as a general reference to this kind of labeled peptides.

As trypsin is the most commonly used protease for digestion, the most used amino acids for isotope labeling are arginine (R) and lysine (K) as trypsin cleaves at the carboxyl side of both of these residues. This allows access to the complete series of *y* fragment ions (usually the most intense fragments in triple quadrupole instrument) that display a mass shift compared to the endogenous peptide fragments [mass shift of 8 Da for K ($^{13}C_6$,$^{15}N_2$) and 10 Da for R ($^{13}C_6$,$^{15}N_4$)].

The advantages of AQUA peptides are their ease of use and their commercial availability within a short delay of production (4–8 weeks) (Thermo Scientific 2017).

The drawbacks of AQUA peptides are:

- As they are peptides, they cannot be introduced in the beginning of the sample preparation and so cannot account neither for protein pre-fractionation/precipitation nor for digestion efficiency/completeness.
- Some peptides will present problems for quantitative resolubilization if provided dry (Brun et al. 2007; Mirzaei et al. 2008). To decrease this problem, ordering AQUA peptides in solution is recommended.
- Loss of peptides by adsorption of peptides onto the tube surfaces.
- Limitations due to chemical synthesis (length up to 15 amino acids, some patterns are not possible, etc.) (Brun et al. 2009).
- Costs are quite high (~$400 per peptide) if many proteins have to be quantified.

For larger study, a statistical tool for "sparse reference labeling" has been developed to reduce the number of necessary AQUA peptides (Chang et al. 2014). This tool is applicable when the targeted proteins have interactions between them and the research goal is well defined before designing the experiment.

4.2 Concatemers (QconCAT)

In 2005, Beynon and coworkers present a new strategy for absolute quantification of several proteins with only one internal standard: QconCAT, a synthetic protein made of a concatenation of tryptic peptides for quantification (Q peptides) (Beynon et al. 2005; Pratt et al. 2006). This synthetic protein is produced in *Escherichia coli* and could be easily labeled by supplying isotope-labeled amino acids in the growth media. The QconCAT is then purified and quantified using amino acids analysis.

In the original article, the QconCAT was first produced in media containing $^{15}NH_4Cl$ as sole nitrogen source (Beynon et al. 2005); this implies a mass shift different from peptide to peptide in function of their amino acid sequence. Now, as for AQUA peptides, stable isotope-labeled lysine and arginine are usually employed which enables the entire series of *y* fragment ions to be mass shifted from those of the target peptide.

QconCAT was applied in the domain of farm animals: a QconCAT was designed for chicken skeletal muscle proteins, i.e., triose-phosphate isomerase, glyceraldehyde 3-phosphate dehydrogenase, beta-enolase, alpha-actin, and actin polymerization inhibitor (Rivers et al. 2007). Another study using QconCAT in horses has been set up to detect acute phase proteins (Bislev et al. 2012; Bundgaard et al. 2014).

Similar strategies, such as PolySIS protein with labeled lysine residues (Anderson and Hunter 2006) and peptide-concatenated standard (PCS), which differs from QconCAT in the presence of the natural flanking sequences at both ends of each

peptide and in the production in a cell-free system (Kito et al. 2007), have been almost concomitantly developed.

The advantages of concatemers are:

- Some peptides that cannot be chemically synthetized are produced in QconCAT protein(s).
- Only one addition on internal standard for several peptides and proteins reduces pipetting inaccuracy impact. Moreover, all the peptides present in the QconCAT are released in a 1:1 ratio, therefore limiting the problem of solubility and/or adsorption on tubes of synthetic peptides. If different ranges of concentrations for different target proteins are expected, it would be possible to incorporate several copies of the same peptides to change this 1:1 ratio (Brun et al. 2009).
- Especially for the PCS type of concatemers, the digestion efficiency can be estimated between the different peptides if compared to unlabeled synthetic individual peptides.
- QconCATs are commercially available.

The drawbacks of the concatamers are as follows:

- It cannot be introduced at the beginning of process and so does not account for protein fractionation step (Zeiler et al. 2012).
- The concatamers have no folding to maximize their digestion efficiency, but doing so, they cannot reflect the same digestion efficiency as for the targeted protein(s) (Rivers et al. 2007; Brun et al. 2009; Brownridge and Beynon 2011; Lebert et al. 2015).
- QconCATs are more expensive than AQUA peptides if only a few proteins (around 8) have to be quantified (PolyQuant 2017).

4.3 Full-Length Labeled Protein (PSAQ)

Full-length labeled protein is the most appropriate internal standard for protein quantification as it can be introduced at the beginning of the process, just after protein extraction, and can therefore account for any loss, degradation, or modification happening during the whole process of the sample. Protein standard absolute quantification (PSAQ) was first introduced by Brun and coworkers in 2007 (Brun et al. 2007) for the quantification of two superantigenic toxins from *Staphylococcus aureus* (SEA and TSST-1). They produced these two proteins with all labeled-lysine and labeled-arginine amino acids in vitro and quantified these proteins after purification by AAA. In this study, they showed that the PSAQ strategy enables a better accuracy and inter-peptide consistency than both AQUA peptides, which showed diverging results between 2 peptides of the same protein, and QconCAT, which showed an underestimate of concentration certainly due to digestion efficiency difference between QconCAT and protein (Brun et al. 2007).

SILAC (stable isotope labeling by/with amino acids in cell culture) metabolomics labeling was adapted by Mann and coworkers to produce stable isotope-labeled proteins, a technique they named "Absolute SILAC" (Hanke et al. 2008). Full-length expressed stable isotope-labeled proteins for quantification (FlexiQUANT) is a variant of PSAQ method which includes a specific FLEX peptide tag in the produced protein for the accurate concentration determination of the produced protein after tryptic digestion (Singh et al. 2009). Recombinant isotope labeled and selenium quantified (RISQ) is another variant where the quantification of the produced labeled protein relies on ICP-MS through the replacement of all methionine residues by selenomethionine ones in the produced sequence (Zinn et al. 2010).

The major advantage that makes PSAQ strategy the golden standard for protein quantification relies in the fact that it can be introduced before protein fractionation to account for the recovery yield of the analyte and also that it should have the same digestion efficiency as the target protein. The latter point is not necessarily true because recombinant proteins do not always have the same folding as the natural proteins.

The drawbacks of PSAQ strategy are:

- It is not suited for PTM as the proteins are produced in cell-free media (Brun et al. 2009).
- This technique is generally restricted to soluble proteins (Zeiler et al. 2012).
- The folding of the labeled protein may not be the same as the natural protein. Therefore a difference in digestion efficiency could arise between the target and the standard (Villanueva et al. 2014).
- The price is quite high and a development/production is needed for each target protein (Villanueva et al. 2014).

5 The Case of PTMs in Quantitative Proteomics

Posttranslational modifications (PTMs) and their regulation are involved in the control of many cellular functions. Their study represents therefore an essential task in order to understand complex biological mechanisms. Besides prominent modifications such as phosphorylation, glycosylation, or acetylation, at least 200 PTMs were reported so far, making their study very complex, especially in the context of possible crosstalks between them.

From the analytical point of view, several challenges are present in quantitative PTM analyses. In contrast to proteins for which several peptides are generated, a given PTM can only be detected on one peptide, leading to both detection and identification/localization difficulties. This is especially true for highly regulated modifications such as phosphorylation that can be only transiently present on a given site.

This effect results in the frequent necessity to use enrichment strategies during sample preparation. This is however at the cost of high amounts of starting material. Depending on the PTMs studied, various approaches were proposed and rely on antibodies, other proteins (such as lectins used to enrich *N*-glycosylated proteins/peptides), or other materials like TiO_2 (for phosphopeptide enrichment). Enrichment is however not possible for all modifications.

Apart from this enrichment step, usual proteomic workflow can generally be applied in order to detect PTMs. Some precautions have however to be taken in order to generate peptides suitable for LC-MS/MS, and alternative proteolytic enzymes or chromatography phases can be requested. This can be, for instance, the case in phosphorylation and glycosylation studies in which unusually hydrophilic peptides can be generated.

In this section, we will review two very important PTMs, phosphorylation and glycosylation, and have an overview on several other modifications.

5.1 *Phosphoproteomics*

Phosphorylation is a key PTM involved in many biological regulation processes that can occur on serine, threonine, and tyrosine residues, representing around 1.85 million residues that could theoretically be phosphorylated in the human proteome (von Stechow et al. 2015). The main challenge in quantitative phosphoproteomics is therefore to detect changes in occupancy of particular sites (potentially several per proteins) as this modification exhibits in addition a highly dynamic nature. Moreover, low abundance of some phosphorylated peptides makes them difficult to detect and quantify in raw samples. This is due either to the low expression of some proteins or to the low phosphorylation yield of some more abundant proteins, for instance, if only a subset localized in a particular cell compartment is modified. As mentioned above, an enrichment step is therefore often mandatory in phosphorylation analyses. Two approaches were described so far, depending on modifications sites. Phospho-Ser and phospho-Thr residues containing peptides are enriched using metal ion (such as Fe^{3+} or Ti^{4+}) or metal oxide (TiO_2) affinity chromatography (Zhou et al. 2011). On the other hand, tyrosine phosphorylation (which is 3000 times less abundant than phospho-Ser and phospho-Thr) analysis is usually performed after antibody-based enrichment (Rush et al. 2005). Quantitative phosphorylation data are usually obtained by bottom-up approach potentially involving peptide labeling. In this case, labels are generally introduced by two ways. Several authors used previously described SILAC MS1-based methodology (well adapted to cell culture) (Oda et al. 1999), while others performed chemical peptide labeling using either MS1-based (dimethyl labeling) or MS2-based (ITRAQ or TMT) labels (Boersema et al. 2009).

These methods were successfully applied in the meat science field. For instance, dimethyl labeling was used in order to study postmortem phosphorylation status of porcine muscle proteins. Authors highlighted changes in phosphorylation of

184 sites belonging to 93 proteins within 24 h, potentially influencing the meat development (Huang et al. 2014).

Following enrichment and data collection, the interpretation of the data is the last but crucial step in the phosphoproteomic workflow. Numerous specialized tools are currently available allowing performing appropriate statistical analyses and thus highlighting activation of particular signaling pathways (Rudolph et al. 2016).

5.2 *Glycosylation*

Glycosylation is characterized by an intrinsic heterogeneity resulting from competition and kinetic effects during their biosynthesis in endoplasmic reticulum and Golgi apparatus. Their quantification takes therefore place at two levels: first, one can study the global glycosylation pattern of sample [composed of one or more protein(s)]. This glycomic approach is performed by isolating *N*- or *O*-glycans. Its drawback is the loss of localization information. On the other hand, the glycoproteomic approach aims at detecting and quantifying the heterogeneity at the glycosylation site level, but can potentially lead to very complex data treatments and interpretations. These two approaches are reviewed in this section.

5.2.1 Quantitative Glycomics

Glycans are released by hydrolysis of the glycosidic bond between its first monosaccharide and the lateral chain of its carrier amino acid residue. *N*-Glycans are usually enzymatically hydrolyzed using PNGase F, PNGase A, Endo H, etc., preserving the reducing end on the structure. This is an excellent reactive group that is usually used for fluorescent labeling (with compound such as 2-aminobenzamide, etc.). When coupled to liquid chromatography separation, this represents a major quantification strategy (Bigge et al. 1995), however, not allowing to properly identify compounds, which can be problematic in complex samples. On the other hand, *O*-glycan is generally chemically hydrolyzed using reductive β-elimination, irreversibly removing the reducing end of the glycan and therefore impeding any aldehyde-reactive labeling. Until recently, mass spectrometry-based quantitative glycomic experiments were performed using either label-free or mass shift isotopic labeling approaches, implying MS1-based quantification. In the last case, label is introduced in several ways using ^{13}C- or ^{2}H-labeled permethylation reagents (Alvarez-Manilla et al. 2006; Kang et al. 2007), PNGase digestion in ^{18}O water (Tao and Orlando 2014), or using the reducing end of *N*-glycans. This approach was also very recently applied on *O*-glycans relative quantification in an elegant protocol involving their nonreducing release (Wang et al. 2017). As described earlier for peptides, these two approaches suffer from several drawbacks. We can cite among others ionization efficiency variations (in label-free approach) and additional spectral spectra complexity (in mass shift approach).

In order to overcome these limitations, isobaric labeling-based quantification was recently implemented in the glycomic field. The aldehyde-reactive group of *N*-glycans was obviously used to add reagents such as aminoxyTMT (allowing 6-plex experiments) (Zhou et al. 2016), "QUANTITY" (Yang et al. 2015a), or iARTs (Yang et al. 2013). The principle of this methodology is similar to that which exists for peptides and allows a relative quantification at the MS2 level. However, no similar approaches allowing isobaric labeling of *O*-glycans have been reported so far. Apart from the untargeted experiments, efforts were made in parallel to develop targeted quantification of released glycans by MRM, in order to reach sensitivity and specificity of detection. This strategy was successively applied on both fluorescently labeled glycans (allowing detection limits around 100 attomoles to be reached) (Zhang et al. 2012a) and permethylated glycans (Zhou et al. 2015). This strategy is however still limited by the lack of standards which are only available for milk oligosaccharides and some *N*-glycans (Ruhaak and Lebrilla 2015), making absolute quantification difficult for many glycans.

5.2.2 Quantitative Glycoproteomics

As mentioned previously, glycosylation is characterized by an important heterogeneity. This means that a particular site may be occupied by different glycans but also that two sites even from the same protein may carry different glycan panels. To characterize these complex situations, the usual glycoproteomics workflow start as inproteomics with trypsin digestion. Then, due to their heterogeneity and their underrepresentation in peptide mixtures, glycopeptides are usually enriched using affinity (lectin, HILIC) chromatography even if strategies involving covalent binding (hydrazide chemistry, etc.) were also described. These methods were recently compared in a study showing that ion pairing ZIC-HILIC enrichment is particularly efficient for this purpose (Zhang et al. 2016). Two main parameters can then be assessed for every site: its occupancy and the glycosylation pattern in various biological conditions. Occupancy rate can be different in various conditions and for different sites. Even if it is generally high (so that the main variability comes from the glycans heterogeneity), it may be affected in various diseases explaining why several authors were interested in developing appropriate occupancy quantification strategies. These strategies are based on relative or absolute measurement of the deglycosylated *N*-glycopeptides (differing from unoccupied glycosites by the presence of E instead of N, after hydrolysis by PNGase). These methods are either relative [label-free or involving isotope labeling such as ITRAQ and ^{18}O labeling (Tian et al. 2011; Pan et al. 2012, 2014; Zhang et al. 2012b)] or absolute (Hülsmeier et al. 2007). Beside the occupancy, the second main challenge of quantitative glycoproteomics is the detection of sometimes small changes in the glycosylation pattern of one or more given site(s). Enriched glycopeptides are identified using CID fragmentation, and electron transfer dissociation (ETD) is often used for localization confirmation (Hoffmann et al. 2016). Quantitative information can be obtained using label-free comparison (with normalization performed site by site

and expressing data as percentage of each glycoform) (Plomp et al. 2014) or labeling methods. TMT labeling and quantification protocol was recently adapted by Ye et al. in order to increase labeling efficiency and reporter ions detection (Ye et al. 2013). Finally, a targeted approach involving MRM was also recently developed and uses oxonium ions and glycopeptide fragment Y1 ion to ensure specificity. This ion is generated by CID or HCD of glycopeptides and corresponds to peptide only retaining the first core *N*-acetylglucosamine residue. This ion is therefore sequence specific, while oxonium ions are glycan specific (Song et al. 2012; Sanda et al. 2013; Ruhaak and Lebrilla 2015).

5.3 Examples of MS-Based PTMs Quantitative Proteomics

Mass spectrometry has also proved its analytical power in the study of other PTMs. Even if it is not possible to do an exhaustive compilation of these applications here, some of them deserve additional illustrations. As model example, we can study modifications occurring on histones. Due to its importance in epigenetic regulation of gene expression, histones PTM through methylation, acetylation, ubiquitination, as well as phosphorylation and other PTMs have been investigated by many researchers and led to the concept of "histone code" (Strahl and Allis 2000; Kakutani et al. 2001), now known as "histone PTM crosstalk." Here again, mass spectrometry and proteomics approaches were used to decipher these complex modifications and their relationships. To date a bottom-up approach remains the most common approach in this task. However, ensuring reproducible digestion is a key point in order to obtain an accurate quantification of modified peptides. In histone particular case, some adaptions were done to the usual protocol to take into account the high proportion of basic residues in histones. In order to avoid generating too small or overlapping peptides, trypsin may be replaced by Glu-C. Alternatively, propionyl derivatization of free lysine residues followed by trypsin digestion was also described. This derivatization step also allows introducing labeled derivatives and therefore comparing histone PTMs in various conditions (Britton et al. 2011; Önder et al. 2015). An additional post-digestion derivatization using phenyl isocyanate was recently introduced to quantify low-abundant forms by ensuring a "sufficient" hydrophobicity and therefore improving retention and ionization of peptides before their LC-MS/MS analysis (Maile et al. 2015). SILAC is another widely used tool in histone biology and was, for instance, used by Bonenfant and coworkers to study various histones PTMs during cell cycle (Bonenfant et al. 2007). Apart from the histone case, substoichiometric non-labile PTMs such as acetylation (Choudhary et al. 2014) or ubiquitination (Rose et al. 2016) of proteins have also been studied using antibody-based enrichment of acetylated or diGly-containing peptides followed by their isotopic labeling. It is interesting to note that iTRAQ labeling was found to offer higher sensitivity than TMT in a wide acetylome study, highlighting possible influence of isobaric labels on peptide detection (Svinkina et al. 2015).

In heavily modified proteins like histones, the presence of PTM crosstalks represents an additional challenge that can be only partially addressed using bottom-up approach. Indeed, long-range crosstalks leading, for instance, to simultaneous presence of two modifications separated by a long sequence may not be detected using traditional protocols. Alternative methodologies of protein sequencing were developed, thanks to the capabilities of ETD, electron capture dissociation (ECD), or matrix-assisted laser desorption ionization in-source decay (MALDI-ISD) to fragment entire proteins (top-down) or large peptides (middle-down). For instance, ECD was successfully used to describe the more than 40 acetylated and methylated H4 histone forms and their occurrence throughout the cell cycle (Mackay et al. 2008; Pesavento et al. 2008). Knowing the ability of the alternative fragmentation techniques to preserve labile PTMs such as phosphorylation or glycosylation [reviewed in (Wiesner et al. 2008)], there is no doubt that these promising tools will be used more extensively in the future.

6 Promising Tools

6.1 MS Imaging and Guided Quantitative Proteomics

Proteomic analysis is often performed on samples in which the part bringing useful information may only constitute a few percents. This is in particular the case for heterogeneous samples (among which biopsies are a clear example) in which interesting signals are diluted by non-informative ones. Identifying regions of interest (ROIs) represent therefore an essential step in order to focus analyses on informative parts of the samples. These samples can be prepared in various ways. Formalin fixation combined with paraffin embedment (FFPE tissue processing) is a common method for long-term sample conservation in biobanks (Xu 2010; Wiśniewski et al. 2011; Magdeldin and Yamamoto 2012). FFPE tissue blocks indeed represent a gold mine for biomarkers discovery studies. Fresh frozen tissues can also be used and are actually easier to use in IMS. In both cases, a pre-purification at the sample level is a key step to improve limits of detection and quantification of low-abundant species. Until recently, there was an unmet need of downscaling the proteomic workflow, in particular the amount of laser microdissected tissues to cope with selective capture of specific tissue types. Several groups have optimized strategies to minimize the number of cells required for differential proteomics. A number of cells around 2500 are now commonly achieved yielding a correct number of protein identification and quantification (Longuespée et al. 2014, 2016) (Fig. 4).

The first step for micro-proteomics is the definition of the region to be laser microdissected. Different types of images can be used, among which the classical histochemical or immuno-histological staining ones (de Deus Moura et al. 2012). However, in the absence of known markers, a fully hypothesis-free method has to be used. MALDI imaging, coupled with the very high performances of new imaging mass spectrometer, represents such hypothesis-free visualization and can

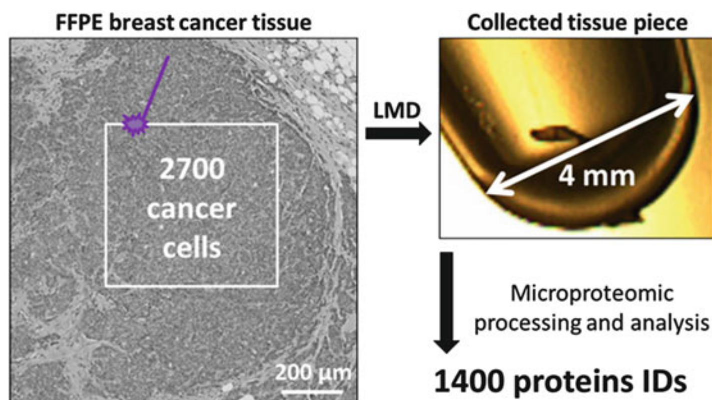

Fig. 4 Illustration of the collection of a tissue piece by LMD and the verification of its presence at the *bottom* of the tube through a binocular microscope. From a breast cancer tissue piece containing 2700 cells, it was possible to retrieve 1400 protein identifications (Longuespée et al. 2016)

therefore be viewed as the ultimate “molecular histology” method for wide applications in clinical proteomics and pharmacology.

MALDI imaging is now well described and 5–20 microns lateral resolution can be reached. After data acquisition, images can be reconstructed for every peak of the whole mass range, providing huge amount of molecular information. Experiments can be focused on low mass chemicals such as lipids and small molecules or on medium or higher mass, looking at peptides or proteins signals. For FFPE tissues, different protocols were proposed. It was shown that a first step of antigen retrieval improves the subsequent protein identification analysis (Shi et al. 2006).

Since 2016, faster MS imagers based on time-of-flight instruments can provide images in <1 h at lower lateral resolution (Prentice and Caprioli 2016). If needed, the images can be further acquired at higher lateral resolution to improve delimitation of ROIs, at the cost of higher acquisition times. In both case, the laser beam raster allows a better use of the sample material.

Several compounds may be emitted at the same nominal mass implying that many ionic species are present in the spectra. Therefore, the ultra-high mass resolution of Fourier transform instruments is a growing tool for in situ identification (Ly et al. 2016). Several additional strategies can improve both the identification and the limit of detection. The matrix choice is important as different biomolecules will be ionized using different matrices. In addition, additives may help ionization, and recently nanoparticles have also been used to increase the signal (Sugiura and Setou 2010).

Finally, ion mobility can help simplifying the spectra as it can separate conformers. Indeed, fully isobaric compounds may also be separated provided that the cross section resolution is sufficient. A new method called trapped ion mobility spectrometry (TIMS) can also selectively accumulate signals from low abundance analytes.

Beside MALDI imaging, the analyses of small molecules including lipids as well as peptides and intact small proteins can also be performed using desorption electrospray ionization (DESI) (Hsu et al. 2015). In its new version, lateral resolution has been improved (nanoDESI) and can be very useful on fresh tissue (Laskin et al. 2012).

6.2 *Top-Down Proteomics*

In this approach, the entire proteins are the biomolecules to quantify and identify. The difficulty to handle intact proteins and the relative low dynamic range for their quantification as intact analytes led to the "gel-free" proteomics. For a decade, progress in both separation methods and mass spectrometry has allowed the development of top-down proteomics in which intact proteins are identified and quantified. This approach is especially interesting in the previously discussed application of PTMs crosstalks.

From the mass spectrometry point of view, the production of highly charged ions and the introduction of high mass (more correctly high m/z) capacity allow detecting pseudo-molecular ions of medium to high mass intact proteins. This is of great value to characterize pure protein samples. Combined with activation methods such as electron capture dissociation (ECD) or electron transfer dissociation (ETD), electrospray ionization is a valuable tool to obtain structural characterization on intact proteins (Toby et al. 2016). Labile PTMs such as phosphorylation or glycosylation are kept intact allowing characterizing the heterogeneity of the protein batches. MS/MS spectra however are quite complex, presenting different charge states for various fragments. Without strong bioinformatics support, manual examination of spectra is very tedious and not compatible with high throughput. Turning to mixtures makes the situation more complicated. UPLC (or nanoUPLC) of proteins mixtures is still a challenge, and other separation methods such as capillary electrophoresis are quickly developing but do not yet constitute a mature standard method (Catherman et al. 2014). This presently limits top-down proteomics to low complexity mixtures or to targeted analysis intended at monitoring the status of PTMs (Gregorich et al. 2016). However, some clinical cases of protein profiling without the use of MS/MS have been described thanks to their isotopic signature (Xiao et al. 2017). Most of the examples described only utilize the highly accurate mass of specific intact proteoforms (Shortreed et al. 2016).

This is largely linked to the fact that MS technology has been driven for two decades toward efficient and automated peptides analysis, including post-acquisition data handling.

6.3 Middle-Down Proteomics, a Compromise or a Real Breakthrough?

Bottom-up and top-down proteomics present both advantages and limitations. The bottom-up approach is now mature, and a solid software basis exists. It requires however extensive sample preparation and has still difficulties with labile PTMs and with high extensive sequence coverage. Most mass spectrometers are designed and optimized for this approach. On the other hand, the top-down strategy requires less sample handling and is the method of choice to profile proteoforms in low complexity systems and for relatively low molecular weight proteins. Working with larger peptides arising from limited digestion with the so-called middle-down approach could take advantage from the two existing strategies (Cannon et al. 2010; Sidoli et al. 2015). Optimized conditions for enzymatic digestion have been described (Laskay et al. 2013). ECD is a candidate activation method to fragment the peptides keeping intact PTMs (Moradian et al. 2014; Sweredoski et al. 2015). With the improvement of bioinformatics tools, middle-down proteomics will find applications when the determination of PTMs is important.

6.4 Looking at Interactions

6.4.1 Hydrogen/Deuterium Exchange

Interactions between proteins can be analyzed using mass spectrometry. The first strategy aims at labeling regions which are accessible by deuterons present in bulk solvent, the interaction regions remaining unlabeled. The method is known as hydrogen deuterium exchange (HDX). This mapping allows probing both intra- and intermolecular interactions (Wales and Engen 2006; Tsutsui and Wintrode 2007). An alternative approach consists in exchanging all protons before interactions, monitoring the retro-exchange once the interaction is realized. The whole mass shift can be measured as a function of time to access to the kinetics of interactions. Regions of exchange can be observed at the peptide level after the enzymatic digestion and LC-MS. It requires quenching of the exchange reaction by adequate pH adjustment and working at low temperature to avoid scrambling. Localization of exchanged positions at the amino acid level can be achieved by MS/MS of peptides, but the scrambling in the H/D positions may occur and must be strictly controlled. HDX was namely used to probe the structure and the dynamics of interactions in complexes (Ramisetty and Washburn 2011; Li et al. 2014).

6.4.2 Cross-Linking

Another strategy to map interactions consists in "freezing" the partners in their interaction conformation using covalent linkers. The linkers are chemically designed to react with specific amino acids. Their lengths allow probing different distances between regions in interaction, being intra- or more often intermolecular. The digestion of the cross-linked (XL) species produces different peptide types that can be resolved by mass spectrometry. From the sequence of the peptides and the distances determined from the linkers lengths, a mapping of the interactions is possible. Quantitative XL-MS (qXL-MS) extends the capabilities of XL-MS. By introducing isotopically labeled linkers or SILAC-labeled samples, it allows for comparison across different experimental conditions and different biological states (Zhong et al. 2017).

References

Aebersold R, Mann M (2016) Mass-spectrometric exploration of proteome structure and function. Nature 537:347–355. https://doi.org/10.1038/nature19949

Al Shweiki MR, Mönchgesang S, Majovsky P, Thieme D, Trutschel D, Hoehenwarter W (2017) Assessment of label-free quantification in discovery proteomics and impact of technological factors and natural variability of protein abundance. J Proteome Res acs.jproteome.6b00645. https://doi.org/10.1021/acs.jproteome.6b00645

Almeida AM, Nanni P, Ferreira AM, Fortes C, Grossmann J, Bessa RJB, Costa P (2017) The longissimus thoracis muscle proteome in Alentejana bulls as affected by growth path. J Proteomics 152:206–215. https://doi.org/10.1016/j.jprot.2016.10.020

Alvarez-Manilla G, Warren NL, Atwood J III, Abney T, Azadi P, Pierce M, Orlando R (2006) Tools for glycomics: isotopic labeling of glycans with C-13 for relative quantitation. Glycobiology 16:677–687

Anderson L, Hunter CL (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol Cell Proteomics 5:573–588. https://doi.org/10.1074/mcp.M500331-MCP200

Bantscheff M, Schirle M, Sweetman G, Rick J, Kuster B (2007) Quantitative mass spectrometry in proteomics: a critical review. Anal Bioanal Chem 389:1017–1031. https://doi.org/10.1007/s00216-007-1486-6

Barnidge DR, Dratz EA, Martin T, Bonilla LE, Moran LB, Lindall A (2003) Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. Anal Chem 75:445–451. https://doi.org/10.1021/ac026154+

Bernhardt OM, Selevsek N, Gillet L, Rinner O et al (2012) Spectronaut: a fast and efficient algorithm for MRM-like processing of data-independent acquisition (SWATH) data. In: Proceedings of the 60th ASMS conference on mass spectrometry and allied topics, Vancouver

Beynon RJ, Doherty MK, Pratt JM, Gaskell SJ (2005) Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. Nat Methods 2:587–589. https://doi.org/10.1038/NMETH774

Bigge JC, Patel TP, Bruce JA, Goulding PN, Charles SM, Parekh RB (1995) Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid. Anal Biochem 230:229–238

Bilbao A, Varesio E, Luban J, Strambio-De-Castillia C, Hopfgartner G, Müller M, Lisacek F (2015) Processing strategies and software solutions for data-independent acquisition in mass spectrometry. Proteomics 15:964–980. https://doi.org/10.1002/pmic.201400323

Biniossek ML, Lechel A, Rudolph KL, Martens UM, Zimmermann S (2013) Quantitative proteomic profiling of tumor cell response to telomere dysfunction using isotope-coded protein labeling (ICPL) reveals interaction network of candidate senescence markers. J Proteomics 91:515–535. https://doi.org/10.1016/j.jprot.2013.08.007

Bislev SL, Kusebauch U, Codrea MC, Beynon RJ, Harman VM, Røntved CM, Aebersold R, Moritz RL, Bendixen E (2012) Quantotypic properties of QconCAT peptides targeting bovine host response to *Streptococcus uberis*. J Proteome Res 11:1832–1843. https://doi.org/10.1021/pr201064g

Boersema PJ, Aye TT, Van Veen TAB, Heck AJR, Mohammed S (2008) Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. Proteomics 8:4624–4632. https://doi.org/10.1002/pmic.200800297

Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJR (2009) Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nat Protoc 4:484–494. https://doi.org/10.1038/nprot.2009.21

Bonenfant D, Towbin H, Coulot M, Schindler P, Mueller DR, van Oostrum J (2007) Analysis of dynamic changes in post-translational modifications of human histones during cell cycle by mass spectrometry. Mol Cell Proteomics 6:1917–1932. https://doi.org/10.1074/mcp.M700070-MCP200

Bourmaud A, Gallien S, Domon B (2016) Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: principle and applications. Proteomics 16:2146–2159. https://doi.org/10.1002/pmic.201500543

Braun CR, Bird GH, Wühr M, Erickson BK, Rad R, Walensky LD, Gygi SP, Haas W (2015) Generation of multiple reporter ions from a single isobaric reagent increases multiplexing capacity for quantitative proteomics. Anal Chem 87:9855–9863. https://doi.org/10.1021/acs.analchem.5b02307

Britton L-MP, Gonzales-Cope M, Zee BM, Garcia BA (2011) Breaking the histone code with quantitative mass spectrometry. Expert Rev Proteomics 8:631–643. https://doi.org/10.1586/epr.11.47

Brownridge P, Beynon RJ (2011) The importance of the digest: proteolysis and absolute quantification in proteomics. Methods 54:351–360. https://doi.org/10.1016/j.ymeth.2011.05.005

Bruderer R, Bernhardt OM, Gandhi T, Miladinović SM, Cheng L-Y, Messner S, Ehrenberger T, Zanotelli V, Butscheid Y, Escher C, Vitek O, Rinner O, Reiter L (2015) Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. Mol Cell Proteomics 14:1400–1410. https://doi.org/10.1074/mcp.M114.044305

Brun V, Dupuis A, Adrait A, Marcellin M, Thomas D, Court M, Vandenesch F, Garin J (2007) Isotope-labeled protein standards: toward absolute quantitative proteomics. Mol Cell Proteomics 6:2139–2149. https://doi.org/10.1074/mcp.M700163-MCP200

Brun V, Masselon C, Garin J, Dupuis A (2009) Isotope dilution strategies for absolute quantitative proteomics. J Proteomics 72:740–749. https://doi.org/10.1016/j.jprot.2009.03.007

Bundgaard L, Jacobsen S, Dyrlund TF, Sørensen MA, Harman VM, Beynon RJ, Brownridge PJ, Petersen LJ, Bendixen E (2014) Development of a method for absolute quantification of equine acute phase proteins using concatenated peptide standards and selected reaction monitoring. J Proteome Res 13:5635–5647. https://doi.org/10.1021/pr500607s

Busch R, Neese RA, Awada M, Hayes GM, Hellerstein MK (2007) Measurement of cell proliferation by heavy water labeling. Nat Protoc 2:3045–3057. https://doi.org/10.1038/nprot.2007.420

Cannon J, Lohnes K, Wynne C, Wang Y, Edwards N, Fenselau C (2010) High-throughput middle-down analysis using an Orbitrap. J Proteome Res 9:3886–3890. https://doi.org/10.1021/pr1000994

Catherman AD, Skinner OS, Kelleher NL (2014) Top down proteomics: facts and perspectives. Biochem Biophys Res Commun 445:683–693. https://doi.org/10.1016/j.bbrc.2014.02.041

Chahrour O, Cobice D, Malone J (2015) Stable isotope labelling methods in mass spectrometry-based quantitative proteomics. J Pharm Biomed Anal 113:2–20. https://doi.org/10.1016/j.jpba.2015.04.013

Chang C, Sabidó E, Aebersold R, Vitek O (2014) Targeted protein quantification using sparse reference labeling. Nat Methods 11:301–304. https://doi.org/10.1038/nmeth.2806

Choi M, Filiz Eren-Dogu Z, Colangelo CM, Cottrell JS, Hoopmann MR, Kapp EA, Kim S, Lam H, Neubert TA, Palmblad M, Phinney BS, Weintraub ST, Maclean B, Vitek O (2016) ABRF proteome informatics research group (iPRG) 2015 study: detection of differentially abundant proteins in label-free quantitative LC-MS/MS experiments. J Proteome Res (just accepted manuscript). https://doi.org/10.1021/acs.jproteome.6b00881

Choudhary C, Weinert BT, Nishida Y, Verdin E, Mann M (2014) The growing landscape of lysine acetylation links metabolism and cell signalling. Nat Rev Mol Cell Biol 15:536–550. https://doi.org/10.1038/nrm3841

Christoforou AL, Lilley KS (2012) Isobaric tagging approaches in quantitative proteomics: the ups and downs. Anal Bioanal Chem 404:1029–1037. https://doi.org/10.1007/s00216-012-6012-9

Ciccimaro E, Blair IA (2010) Stable-isotope dilution LC–MS for quantitative biomarker analysis. Bioanalysis 2:311–341. https://doi.org/10.4155/bio.09.185

de Deus Moura R, Wludarski SCL, Carvalho FM, Bacchi CE (2012) Immunohistochemistry applied to the differential diagnosis between ductal and lobular carcinoma of the breast. Appl Immunohistochem Mol Morphol 1. https://doi.org/10.1097/PAI.0b013e318255bafa

Desiderio DM, Kai M (1983) Preparation of stable isotope-Incorporated peptide internal standards for field desorption mass spectrometry quantification of peptides in biologic tissue. Biol Mass Spectrom 10:471–479. https://doi.org/10.1002/bms.1200100806

Domon B, Aebersold R (2010) Options and considerations when selecting a quantitative proteomics strategy. Nat Biotechnol 28:710–721. https://doi.org/10.1038/nbt.1661

Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods 4:207–214. https://doi.org/10.1038/nmeth1019

Evans C, Noirel J, Ow SY, Salim M, Pereira-Medrano AG, Couto N, Pandhal J, Smith D, Pham TK, Karunakaran E, Zou X, Biggs CA, Wright PC (2012) An insight into iTRAQ: where do we stand now? Anal Bioanal Chem 404:1011–1027. https://doi.org/10.1007/s00216-012-5918-6

Fæste CK, Moen A, Schniedewind B, Haug Anonsen J, Klawitter J, Christians U (2016) Development of liquid chromatography-tandem mass spectrometry methods for the quantitation of Anisakis simplex proteins in fish. J Chromatogr A 1432:58–72. https://doi.org/10.1016/j.chroma.2016.01.002

Fisher RA (1926) The arrangement of field experiments. J Minist Agric Great Brit 33:503–513. https://doi.org/10.1007/978-1-4612-4380-9_8

Fisher SRA (1971) The design of experiments. Hafner Publishing Co., New York, pp 1–27

Gallego M, Mora L, Concepción Aristoy M, Toldrá F (2016a) The use of label-free mass spectrometry for relative quantification of sarcoplasmic proteins during the processing of dry-cured ham. Food Chem 196:437–444. https://doi.org/10.1016/j.foodchem.2015.09.062

Gallego M, Mora L, Toldrá F (2016b) Peptidomics as a tool for quality control in dry-cured ham processing. J Proteomics 147:98–107. https://doi.org/10.1016/j.jprot.2016.02.020

Gan CS, Chong PK, Pham TK, Wright PC (2007) Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). J Proteome Res 6:821–827. https://doi.org/10.1021/pr060474i

Geiger T, Wisniewski JR, Cox J, Zanivan S, Kruger M, Ishihama Y, Mann M (2011) Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. Nat Protoc 6:147–157. https://doi.org/10.1038/nprot.2010.192

Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci USA 100:6940–6945. https://doi.org/10.1073/pnas.0832254100

Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, Bonner R, Aebersold R (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics 11: O111.016717–O111.016717. https://doi.org/10.1074/mcp.O111.016717

Gregorich ZR, Peng Y, Cai W, Jin Y, Wei L, Chen AJ, McKiernan SH, Aiken JM, Moss RL, Diffee GM, Ge Y (2016) Top-down targeted proteomics reveals decrease in myosin regulatory light-chain phosphorylation that contributes to sarcopenic muscle dysfunction. J Proteome Res 15:2706–2716. https://doi.org/10.1021/acs.jproteome.6b00244

Gruhler A, Schulze WX, Matthiesen R, Mann M, Jensen ON (2005) Stable isotope labeling of arabidopsis thaliana cells and quantitative proteomics by mass spectrometry. Mol Cell Proteomics 4:1697–1709. https://doi.org/10.1074/mcp.M500190-MCP200

Guevel L, Lavoie JR, Perez-Iratxeta C, Rouger K, Dubreil L, Feron M, Talon S, Brand M, Megeney LA (2011) Quantitative proteomic analysis of dystrophic dog muscle. J Proteome Res 10:2465–2478. https://doi.org/10.1021/pr2001385

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17:994–999. https://doi.org/10.1038/13690

Hakimov HA, Walters S, Wright TC, Meidinger RG, Verschoor CP, Gadish M, Chiu DKY, Strömvik MV, Forsberg CW, Golovan SP (2009) Application of iTRAQ to catalogue the skeletal muscle proteome in pigs and assessment of effects of gender and diet dephytinization. Proteomics 9:4000–4016. https://doi.org/10.1002/pmic.200900049

Hanke S, Besir H, Oesterhelt D, Mann M (2008) Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. J Proteome Res 7:1118–1130. https://doi.org/10.1021/pr7007175

Hebert AS, Richards AL, Bailey DJ, Ulbrich A, Coughlin EE, Westphall MS, Coon JJ (2014) The one hour yeast proteome. Mol Cell Proteomics 13:339–347. https://doi.org/10.1074/mcp.M113.034769

Hernández-Castellano LE, Ferreira AM, Nanni P, Grossmann J, Argüello A, Capote J, Cai G, Lippolis J, Castro N, de Almeida AM (2016) The goat (*Capra hircus*) mammary gland secretory tissue proteome as influenced by weight loss: a study using label free proteomics. J Proteomics 145:60–69. https://doi.org/10.1016/j.jprot.2016.03.030

Hoffmann M, Marx K, Reichl U, Wuhrer M, Rapp E (2016) Site-specific O -glycosylation analysis of human blood plasma proteins. Mol Cell Proteomics 15:624–641. https://doi.org/10.1074/mcp.M115.053546

Hsu C-C, Chou P-T, Zare RN (2015) Imaging of proteins in tissue samples using nanospray desorption electrospray ionization mass spectrometry. Anal Chem 87:11171–11175. https://doi.org/10.1021/acs.analchem.5b03389

Hu A, Noble WS, Wolf-Yadlin A (2016) Technical advances in proteomics: new developments in data-independent acquisition. F1000Research 5:419. 10.12688/f1000research.7042.1

Huan P, Wang H, Liu B (2015) A label-free proteomic analysis on competent larvae and juveniles of the Pacific oyster *Crassostrea gigas*. PLoS One 10:e0135008. https://doi.org/10.1371/journal.pone.0135008

Huang H, Larsen MR, Palmisano G, Dai J, Lametsch R (2014) Quantitative phosphoproteomic analysis of porcine muscle within 24h postmortem. J Proteomics 106:125–139. https://doi.org/10.1016/j.jprot.2014.04.020

Hughes CS, Foehr S, Garfield DA, Furlong EE, Steinmetz LM, Krijgsveld J (2014) Ultrasensitive proteome analysis using paramagnetic bead technology. Mol Syst Biol 10:757–757. 10.15252/msb.20145625

Hülsmeier AJ, Paesold-Burda P, Hennet T (2007) N-glycosylation site occupancy in serum glycoproteins using multiple reaction monitoring liquid chromatography-mass spectrometry. Mol Cell Proteomics 6:2132–2138. https://doi.org/10.1074/mcp.M700361-MCP200

Institute for System Biology (2004) PeptideAtlas. http://www.peptideatlas.org/

Institute for Systems Biology (2013) SWATHAtlas. http://www.swathatlas.org/index.php

Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 4:1265–1272. https://doi.org/10.1074/mcp.M500061-MCP200

Kakutani T, Paszkowski J, Matzke AJ, Kooter JM, Benito MI, Walbot V, Haverkamp TH, Van Luenen HG, Stein P, Hayashi H, Schultz RM, Cogoni C, Faugeron G, Baskar R, Grossniklaus U, Springer PS, Holding DR, Groover A, Yordan C, Martienssen RA, Vinkenoog R, Adams S, Spielman M, Dickinson HG, Scott RJ, Byrne M, Timmermans M, Kidner C, Martienssen R, Cambareri EB, Aisner R, Carbon J, Sanmiguel P, Mauricio R, Cubas P, Vincent C, Coen E, Chen ZJ, Pikaard CS, Comai L, Lee HS, Richards E, Chandler V, Jacobsen S (2001) Translating the Histone Code. Science (80–) 293:1074–1081

Kang P, Mechref Y, Kyselova Z, Goetz JA, Novotny MV (2007) Comparative glycomic mapping through quantitative permethylation and stable-isotope labeling. Anal Chem 79:6064–6073. https://doi.org/10.1021/ac062098r

Karp NA, Lilley KS (2007) Design and analysis issues in quantitative proteomics studies. Proteomics. https://doi.org/10.1002/pmic.200700683

Kito K, Ito T (2008) Mass spectrometry-based approaches toward absolute quantitative proteomics. Curr Genomics 9:263–274. https://doi.org/10.2174/138920208784533647

Kito K, Ota K, Fujita T, Ito T (2007) A synthetic approach toward accurate mass spectrometric quantification of component stoichiometry of multiprotein complexes. J Proteome Res 6:792–800. https://doi.org/10.1021/pr060447s

Krüger M, Moser M, Ussar S, Thievessen I, Luber CA, Forner F, Schmidt S, Zanivan S, Fässler R, Mann M (2008) SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. Cell 134:353–364. https://doi.org/10.1016/j.cell.2008.05.033

Kusebauch U, Campbell DS, Deutsch EW, Chu CS, Spicer DA, Brusniak MY, Slagel J, Sun Z, Stevens J, Grimes B, Shteynberg D, Hoopmann MR, Blattmann P, Ratushny AV, Rinner O, Picotti P, Carapito C, Huang CY, Kapousouz M, Lam H, Tran T, Demir E, Aitchison JD, Sander C, Hood L, Aebersold R, Moritz RL (2016) Human SRMAtlas: a resource of targeted assays to quantify the complete human proteome. Cell 166:766–778. https://doi.org/10.1016/j.cell.2016.06.041

Kuzyk MA, Smith D, Yang J, Cross TJ, Jackson AM, Hardie DB, Anderson NL, Borchers CH (2009) Multiple reaction monitoring-based, multiplexed, absolute quantitation of 45 proteins in human plasma. Mol Cell Proteomics 8:1860–1877. https://doi.org/10.1074/mcp.M800540-MCP200

Lange V, Picotti P, Domon B, Aebersold R (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol. 4:222. https://doi.org/10.1038/msb.2008.61

Laskay UA, Lobas AA, Srzentic K, Gorshkov MV, Tsybin YO (2013) Proteome digestion specificity analysis for rational design of extended bottom-up and middle-down proteomics experiments. J Proteome Res 12(12):5558–5569. https://doi.org/10.1021/pr400522h

Laskin J, Heath BS, Roach PJ, Cazares L, Semmes OJ (2012) Tissue imaging using nanospray desorption electrospray ionization mass spectrometry. Anal Chem 84:141–148. https://doi.org/10.1021/ac2021322

Law KP, Lim YP (2013) Recent advances in mass spectrometry: data independent analysis and hyper reaction monitoring. Expert Rev Proteomics 10:551–566. https://doi.org/10.1586/14789450.2013.858022

Lebert D, Louwagie M, Goetze S, Picard G, Ossola R, Duquesne C, Basler K, Ferro M, Rinner O, Aebersold R, Garin JJ, Mouz N, Brunner E, Brun V (2015) DIGESTIF: a universal quality

standard for the control of bottom-up proteomics experiments. J Proteome Res 14:787–803. https://doi.org/10.1021/pr500834z

Leroy B, Rosier C, Erculisse V, Leys N, Mergeay M, Wattiez R (2010) Differential proteomic analysis using isotope-coded protein-labeling strategies: comparison, improvements and application to simulated microgravity effect on *Cupriavidus metallidurans* CH34. Proteomics 10:2281–2291. https://doi.org/10.1002/pmic.200900286

Li J, Steen H, Gygi SP (2003) Protein profiling with cleavable isotope-coded affinity tag (cICAT) reagents: the yeast salinity stress response. Mol Cell Proteomics 2:1198–1204. https://doi.org/10.1074/mcp.M300070-MCP200

Li J, Rodnin MV, Ladokhin AS, Gross ML (2014) Hydrogen–deuterium exchange and mass spectrometry reveal the pH-dependent conformational changes of Diphtheria toxin T domain. Biochemistry 53:6849–6856. https://doi.org/10.1021/bi500893y

Li S, Zhang Y, Wang J, Yang Y, Miao C, Guo Y, Zhang Z, Cao Q, Shui W (2016) Combining untargeted and targeted proteomic strategies for discrimination and quantification of cashmere fibers. PLoS One 11:e0147044. https://doi.org/10.1371/journal.pone.0147044

Liang H-C, Lahert E, Pike I, Ward M (2015) Quantitation of protein post-translational modifications using isobaric tandem mass tags. Bioanalysis 7:383–400. https://doi.org/10.4155/bio.14.296

Longuespée R, Fléron M, Pottier C, Quesada-Calvo F, Meuwis M-A, Baiwir D, Smargiasso N, Mazzucchelli G, De Pauw-Gillet M-C, Delvenne P, De Pauw E (2014) Tissue proteomics for the next decade? Towards a molecular dimension in histology. OMICS J Integr Biol 18:539–552. https://doi.org/10.1089/omi.2014.0033

Longuespée R, Alberts D, Pottier C, Smargiasso N, Mazzucchelli G, Baiwir D, Kriegsmann M, Herfs M, Kriegsmann J, Delvenne P, De Pauw E (2016) A laser microdissection-based workflow for FFPE tissue microproteomics: important considerations for small sample processing. Methods 104:154–162. https://doi.org/10.1016/j.ymeth.2015.12.008

Lottspeich F, Kellermann J (2011) ICPL labeling strategies for proteome research. In: Methods in molecular biology, pp 55–64

Lu P, Vogel C, Wang R, Yao X, Marcotte EM (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat Biotechnol 25:117–124. https://doi.org/10.1038/nbt1270

Lv Y, Zhang S, Wang J, Hu Y (2016) Quantitative proteomic analysis of wheat seeds during artificial ageing and priming using the isobaric tandem mass tag labeling. PLoS One. https://doi.org/10.1371/journal.pone.0162851

Ly A, Buck A, Balluff B, Sun N, Gorzolka K, Feuchtinger A, Janssen K-P, Kuppen PJK, van de Velde CJH, Weirich G, Erlmeier F, Langer R, Aubele M, Zitzelsberger H, McDonnell L, Aichler M, Walch A (2016) High-mass-resolution MALDI mass spectrometry imaging of metabolites from formalin-fixed paraffin-embedded tissue. Nat Protoc 11:1428–1443. https://doi.org/10.1038/nprot.2016.081

Maccarrone G, Lebar M, Martins-De-Souza D (2014) Brain quantitative proteomics combining gelc-ms and isotope-coded protein labeling (ICPL). In: Methods in molecular biology. pp 175–185

Mackay CL, Ramsahoye B, Burgess K, Cook K, Weidt S, Creanor J, Harrison D, Langridge-Smith P, Hupp T, Hayward L (2008) Sensitive, specific, and quantitative FTICR mass spectrometry of combinatorial post-translational modifications in intact histone H4. Anal Chem 80:4147–4153. https://doi.org/10.1021/ac702452d

MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26:966–968. https://doi.org/10.1093/bioinformatics/btq054

Maes E, Cho WC, Baggerman G (2015) Translating clinical proteomics: the importance of study design. Expert Rev Proteomics 12:217–219. https://doi.org/10.1586/14789450.2015.1041512

Maes E, Kelchtermans P, Bittremieux W, De Grave K, Degroeve S, Hooyberghs J, Mertens I, Baggerman G, Ramon J, Laukens K, Martens L, Valkenborg D (2016) Designing biomedical

proteomics experiments: state-of-the-art and future perspectives. Expert Rev Proteomics. https://doi.org/10.1586/14789450.2016.1172967

Magdeldin S, Yamamoto T (2012) Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues. Proteomics 12:1045–1058. https://doi.org/10.1002/pmic.201100550

Maile TM, Izrael-Tomasevic A, Cheung T, Guler GD, Tindell C, Masselot A, Liang J, Zhao F, Trojer P, Classon M, Arnott D (2015) Mass spectrometric quantification of histone post-translational modifications by a hybrid chemical labeling method. Mol Cell Proteomics 14:1148–1158. https://doi.org/10.1074/mcp.O114.046573

Mallick P, Kuster B (2010) Proteomics: a pragmatic perspective. Nat Biotechnol 28:695–709. https://doi.org/10.1038/nbt.1658

Mann M (2014) Fifteen years of stable isotope labeling by amino acids in cell culture (SILAC). Methods Mol Biol 1188:1–7. https://doi.org/10.1007/978-1-4939-1142-4_1

Mann M, Kulak NA, Nagaraj N, Cox J (2013) The coming age of complete, accurate, and ubiquitous proteomes. Mol Cell 49:583–590. https://doi.org/10.1016/j.molcel.2013.01.029

Martinez-Val A, Garcia F, Ximénez-Embún P, Ibarz N, Zarzuela E, Ruppen I, Mohammed S, Munoz J (2016) On the statistical significance of compressed ratios in isobaric labeling: a cross-platform comparison. J Proteome Res 15:3029–3038. https://doi.org/10.1021/acs.jproteome.6b00151

Mastroleo F, Van Houdt R, Leroy B, Benotmane MA, Janssen A, Mergeay M, Vanhavere F, Hendrickx L, Wattiez R, Leys N (2009) Experimental design and environmental parameters affect *Rhodospirillum rubrum* S1H response to space flight. ISME J 3:1402–1419. https://doi.org/10.1038/ismej.2009.74

May C, Brosseron F, Chartowski P, Meyer HE, Marcus K (2012) Differential proteome analysis using 2D-DIGE. In: Methods in molecular biology, pp 75–82

McAlister GC, Huttlin EL, Haas W, Ting L, Jedrychowski MP, Rogers JC, Kuhn K, Pike I, Grothe RA, Blethrow JD, Gygi SP (2012) Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. Anal Chem 84:7469–7478. https://doi.org/10.1021/ac301572t

Mirzaei H, McBee JK, Watts J, Aebersold R (2008) Comparative evaluation of current peptide production platforms used in absolute quantification in proteomics. Mol Cell Proteomics 7:813–823. https://doi.org/10.1074/mcp.M700495-MCP200

Moradian A, Kalli A, Sweredoski MJ, Hess S (2014) The top-down, middle-down, and bottom-up mass spectrometry approaches for characterization of histone variants and their post-translational modifications. Proteomics 14:489–497. https://doi.org/10.1002/pmic.201300256

Moritz B, Meyer HE (2003) Approaches for the quantification of protein concentration ratios. Proteomics 3:2208–2220. https://doi.org/10.1002/pmic.200300581

Mudaliar M, Tassi R, Thomas FC, McNeilly TN, Weidt SK, McLaughlin M, Wilson D, Burchmore R, Herzyk P, Eckersall PD, Zadoks RN (2016) Mastitomics, the integrated omics of bovine milk in an experimental model of *Streptococcus uberis* mastitis: 2. Label-free relative quantitative proteomics. Mol BioSyst 12:2748–2761. https://doi.org/10.1039/C6MB00290K

Muntel J, Xuan Y, Berger ST, Reiter L, Bachur R, Kentsis A, Steen H (2015) Advancing urinary protein biomarker discovery by data-independent acquisition on a quadrupole-orbitrap mass spectrometer. J Proteome Res 14:4752–4762. https://doi.org/10.1021/acs.jproteome.5b00826

Mustafa Ünlü, Morgan ME, Minden JS (1997) Difference gel electrophoresis. A single gel method for detecting changes in protein extracts. Electrophoresis 18:2071–2077. https://doi.org/10.1002/elps.1150181133

Neely BA, Soper JL, Gulland FMD, Bell PD, Kindy M, Arthur JM, Janech MG (2015) Proteomic analysis of cerebrospinal fluid in California sea lions (*Zalophus californianus*) with domoic acid toxicosis identifies proteins associated with neurodegeneration. Proteomics 15:4051–4063. https://doi.org/10.1002/pmic.201500167

Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, Lee A, Van Sluyter SC, Haynes PA (2011) Less label, more free: approaches in label-free quantitative mass spectrometry. Proteomics 11:535–553. https://doi.org/10.1002/pmic.201000553

Nesvizhskii AI, Vitek O, Aebersold R (2007) Analysis and validation of proteomic data generated by tandem mass spectrometry. Nat Methods 4:787–797. https://doi.org/10.1038/nmeth1088

Nikolov M, Schmidt C, Urlaub H (2012) Quantitative mass spectrometry-based proteomics: an overview. In: Methods in molecular biology, pp 85–100

NonLinear Dynamics Progenesis QI for Proteomics (2017) http://www.nonlinear.com/progenesis/qi-for-proteomics/

Oberg AL, Vitek O (2009) Statistical design of quantitative mass spectrometry-based proteomic experiments. J Proteome Res 8:2144–2156. https://doi.org/10.1021/pr8010099

Oda Y, Huang K, Cross FR, Cowburn D, Chait BT (1999) Accurate quantitation of protein expression and site-specific phosphorylation. Proc Natl Acad Sci USA 96:6591–6596. https://doi.org/10.1073/pnas.96.12.6591

Önder Ö, Sidoli S, Carroll M, Garcia BA (2015) Progress in epigenetic histone modification analysis by mass spectrometry for clinical investigations. Expert Rev Proteomics 12:499–517. https://doi.org/10.1586/14789450.2015.1084231

Ong S-E (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1:376–386. https://doi.org/10.1074/mcp.M200025-MCP200

Ong S-E, Mann M (2006) A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). Nat Protoc 1:2650–2660. https://doi.org/10.1038/nprot.2006.427

Ori A, Iskar M, Buczak K, Kastritis P, Parca L, Andrés-Pons A, Singer S, Bork P, Beck M (2016) Spatiotemporal variation of mammalian protein complex stoichiometries. Genome Biol 17:47. https://doi.org/10.1186/s13059-016-0912-5

Ow SY, Salim M, Noirel J, Evans C, Rehman I, Wright PC (2009) iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly.". J Proteome Res 8:5347–5355. https://doi.org/10.1021/pr900634c

Pan S, Tamura Y, Chen R, May D, McIntosh MW, Brentnall TA (2012) Large-scale quantitative glycoproteomics analysis of site-specific glycosylation occupancy. Mol BioSyst 8:2850–2856. https://doi.org/10.1039/c2mb25268f

Pan S, Chen R, Tamura Y, Crispin DA, Lai LA, May DH, Mcintosh MW, Goodlett DR, Brentnall TA (2014) Quantitative glycoproteomics analysis reveals changes in N-glycosylation level associated with pancreatic ductal adenocarcinoma. J Proteome Res:1293−1306

Paradela A, Marcilla M, Navajas R, Ferreira L, Ramos-Fernandez A, Fernández M, Mariscotti JF, Portillo FG, Albar JP (2010) Evaluation of isotope-coded protein labeling (ICPL) in the quantitative analysis of complex proteomes. Talanta 80:1496–1502. https://doi.org/10.1016/j.talanta.2009.06.083

Paulovich AG, Whiteaker JR, Hoofnagle AN, Wang P (2008) The interface between biomarker discovery and clinical validation: the tar pit of the protein biomarker pipeline. Proteomics Clin Appl 2:1386–1402. https://doi.org/10.1002/prca.200780174

Percy AJ, Yang J, Hardie DB, Chambers AG, Tamura-Wells J, Borchers CH (2015) Precise quantitation of 136 urinary proteins by LC/MRM-MS using stable isotope labeled peptides as internal standards for biomarker discovery and/or verification studies. Methods 81:24–33. https://doi.org/10.1016/j.ymeth.2015.04.001

Perez-Patiño C, Barranco I, Parrilla I, Valero ML, Martinez EA, Rodriguez-Martinez H, Roca J (2016) Characterization of the porcine seminal plasma proteome comparing ejaculate portions. J Proteomics 142:15–23. https://doi.org/10.1016/j.jprot.2016.04.026

Pesavento JJ, Bullock CR, LeDuc RD, Mizzen CA, Kelleher NL (2008) Combinatorial modification of human histone H4 quantitated by two-dimensional liquid chromatography coupled with top down mass spectrometry. J Biol Chem 283:14927–14937. https://doi.org/10.1074/jbc.M709796200

Picotti P, Aebersold R (2012) Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nat Methods 9:555–566. https://doi.org/10.1038/nmeth.2015

Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R (2009) Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. Cell 138:795–806. https://doi.org/10.1016/j.cell.2009.05.051

Picotti P, Rinner O, Stallmach R, Dautel F, Farrah T, Domon B, Wenschuh H, Aebersold R (2010) High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. Nat Methods 7:43–46. https://doi.org/10.1038/nmeth.1408

Piovesana S, Capriotti AL, Caruso G, Cavaliere C, La Barbera G, Zenezini Chiozzi R, Laganà A (2016) Labeling and label free shotgun proteomics approaches to characterize muscle tissue from farmed and wild gilthead sea bream (*Sparus aurata*). J Chromatogr A 1428:193–201. https://doi.org/10.1016/j.chroma.2015.07.049

Plomp R, Hensbergen PJ, Rombouts Y, Zauner G, Dragan I, Koeleman CAM, Deelder AM, Wuhrer M (2014) Site-specific N-glycosylation analysis of human immunoglobulin e. J Proteome Res 13:536–546. https://doi.org/10.1021/pr400714w

PolyQuant QconCATs or synthetic/AQUA peptides? (2017) http://www.polyquant.com/qconcat-technology/qconcats-or-syntheticaqua-peptides/

Pratt JM, Simpson DM, Doherty MK, Rivers J, Gaskell SJ, Beynon RJ (2006) Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. Nat Protoc 1:1029–1043. https://doi.org/10.1038/nprot.2006.129

Prentice BM, Caprioli RM (2016) The need for speed in matrix-assisted laser desorption/ionization imaging mass spectrometry. Postdoc J 4:3–13

Raijmakers R, Berkers CR, de Jong A, Ovaa H, Heck AJR, Mohammed S (2008) Automated online sequential isotope labeling for protein quantitation applied to proteasome tissue-specific diversity. Mol Cell Proteomics 7:1755–1762. https://doi.org/10.1074/mcp.M800093-MCP200

Rainczuk A, Condina M, Pelzing M, Dolman S, Rao J, Fairweather N, Jobling T, Stephens AN (2013) The utility of isotope-coded protein labeling for prioritization of proteins found in ovarian cancer patient urine. J Proteome Res 12:4074–4088. https://doi.org/10.1021/pr400618v

Ramisetty SR, Washburn MP (2011) Unraveling the dynamics of protein interactions with quantitative mass spectrometry. Crit Rev Biochem Mol Biol 46:216–228. https://doi.org/10.3109/10409238.2011.567244

Rauniyar N, McClatchy DB, Yates JR (2013) Stable isotope labeling of mammals (SILAM) for in vivo quantitative proteomic analysis. Methods 61:260–268. https://doi.org/10.1016/j.ymeth.2013.03.008

Reiter L, Rinner O, Picotti P, Hüttenhain R, Beck M, Brusniak M-Y, Hengartner MO, Aebersold R (2011) mProphet: automated data processing and statistical validation for large-scale SRM experiments. Nat Methods 8(5):430. https://doi.org/10.1038/nmeth.1584

Richards AL, Merrill AE, Coon JJ (2015) Proteome sequencing goes deep. Curr Opin Chem Biol 24:11–17. https://doi.org/10.1016/j.cbpa.2014.10.017

Rivers J, Simpson DM, Robertson DHL, Gaskell SJ, Beynon RJ (2007) Absolute multiplexed quantitative analysis of protein expression during muscle development using QconCAT. Mol Cell Proteomics 6:1416–1427. https://doi.org/10.1074/mcp.M600456-MCP200

Rose CM, Isasa M, Ordureau A, Prado MA, Beausoleil SA, Jedrychowski MP, Finley DJ, Harper JW, Gygi SP (2016) Highly multiplexed quantitative mass spectrometry analysis of ubiquitylomes. Cell Syst 3:395–403. https://doi.org/10.1016/j.cels.2016.08.009

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3:1154–1169. https://doi.org/10.1074/mcp.M400129-MCP200

Röst HL, Rosenberger G, Navarro P, Gillet L, Miladinović SM, Schubert OT, Wolski W, Collins BC, Malmström J, Malmström L, Aebersold R (2014) OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. Nat Biotechnol 32:219–223. https://doi.org/10.1038/nbt.2841

Ruderman D (2017) Designing successful proteomics experiments. Methods Mol Biol. https://doi.org/10.1007/978-1-4939-6747-6_19

Rudolph JD, de Graauw M, van de Water B, Geiger T, Sharan R (2016) Elucidation of signaling pathways from large-scale phosphoproteomic data using protein interaction networks. Cell Syst 3:585–593. https://doi.org/10.1016/j.cels.2016.11.005

Ruhaak LR, Lebrilla CB (2015) Applications of multiple reaction monitoring to clinical glycomics. Chromatographia 78:335–342. https://doi.org/10.1007/s10337-014-2783-9

Rush J, Moritz A, Lee KA, Guo A, Goss VL, Spek EJ, Zhang H, Zha XM, Polakiewicz RD, Comb MJ (2005) Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. Nat Biotechnol 23:94–101. https://doi.org/10.1038/nbt1046

Russell MR, Lilley KS (2012) Pipeline to assess the greatest source of technical variance in quantitative proteomics using metabolic labelling. J Proteomics. https://doi.org/10.1016/j.jprot.2012.09.020

Sanda M, Pompach P, Brnakova Z, Wu J, Makambi K, Goldman R (2013) Quantitative liquid chromatography-mass spectrometry-multiple reaction monitoring (LC-MS-MRM) analysis of site-specific glycoforms of haptoglobin in liver disease. Mol Cell Proteomics 12:1294–1305. https://doi.org/10.1074/mcp.M112.023325

Sap KA, Demmers JAA (2012) Labeling methods in mass spectrometry based quantitative proteomics. In: Integrative proteomics. InTech, pp 111–132

Scheerlinck E, Dhaenens M, Van Soom A, Peelman L, De Sutter P, Van Steendam K, Deforce D (2015) Minimizing technical variation during sample preparation prior to label-free quantitative mass spectrometry. Anal Biochem 490:14–19. https://doi.org/10.1016/j.ab.2015.08.018

Schmidt A, Kellermann J, Lottspeich F (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. Proteomics 5:4–15. https://doi.org/10.1002/pmic.200400873

Shao S, Guo T, Koh CC, Gillessen S, Joerger M, Jochum W, Aebersold R (2015) Minimal sample requirement for highly multiplexed protein quantification in cell lines and tissues by PCT-SWATH mass spectrometry. Proteomics 15:3711–3721. https://doi.org/10.1002/pmic.201500161

Shi S-R, Liu C, Balgley BM, Lee C, Taylor CR (2006) Protein extraction from formalin-fixed, paraffin-embedded tissue sections: quality evaluation by mass spectrometry. J Histochem Cytochem 54:739–743. https://doi.org/10.1369/jhc.5B6851.2006

Shi X, Li C, Cao M, Xu X, Zhou G, Xiong YL (2016) Comparative proteomic analysis of longissimus dorsi muscle in immuno- and surgically castrated male pigs. Food Chem 199:885–892. https://doi.org/10.1016/j.foodchem.2015.11.059

Shortreed MR, Frey BL, Scalf M, Knoener RA, Cesnik AJ, Smith LM (2016) Elucidating proteoform families from proteoform intact-mass and lysine-count measurements. J Proteome Res 15:1213–1221. https://doi.org/10.1021/acs.jproteome.5b01090

Sidoli S, Lin S, Karch KR, Garcia BA (2015) Bottom-up and middle-down proteomics have comparable accuracies in defining histone post-translational modification relative abundance and stoichiometry. Anal Chem 87:3129–3133. https://doi.org/10.1021/acs.analchem.5b00072

Singh S, Springer M, Steen J, Kirschner MW, Steen H (2009) FLEXIQuant: a novel tool for the absolute quantification of proteins, and the simultaneous identification and quantification of potentially modified peptides. J Proteome Res 8:2201–2210. https://doi.org/10.1021/pr800654s

Smith RD, Veenstra TD (2003) Proteome characterization and proteomics, volume 65. Academic Press

Soares R, Franco C, Pires E, Ventosa M, Palhinhas R, Koci K, Martinho de Almeida A, Varela Coelho A (2012) Mass spectrometry and animal science: protein identification strategies and particularities of farm animal species. J Proteomics 75:4190–4206. https://doi.org/10.1016/j.jprot.2012.04.009

Song E, Pyreddy S, Mechref Y (2012) Quantification of glycopeptides by multiple reaction monitoring liquid chromatography/tandem mass spectrometry. Rapid Commun Mass Spectrom 26:1941–1954. https://doi.org/10.1002/rcm.6290

Stella R, Biancotto G, Arrigoni G, Barrucci F, Angeletti R, James P (2015) Proteomics for the detection of indirect markers of steroids treatment in bovine muscle. Proteomics 15:2332–2341. https://doi.org/10.1002/pmic.201400468

Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403:41–45. https://doi.org/10.1038/47412

Sturm M, Bertsch A, Gröpl C, Hildebrandt A, Hussong R, Lange E, Pfeifer N, Schulz-Trieglaff O, Zerck A, Reinert K, Kohlbacher O (2008) OpenMS – an open-source software framework for mass spectrometry. BMC Bioinf 9:163. https://doi.org/10.1186/1471-2105-9-163

Sugiura Y, Setou M (2010) Matrix-assisted laser desorption/ionization and nanoparticle-based imaging mass spectrometry for small metabolites: a practical protocol. In: Rubakhin SS, Sweedler JV (eds) Mass spectrometry imaging: principles and protocols. Humana Press, Totowa, NJ, pp 173–195

Sury MD, Chen J-X, Selbach M (2010) The SILAC fly allows for accurate protein quantification in vivo. Mol Cell Proteomics 9:2173–2183. https://doi.org/10.1074/mcp.M110.000323

Svinkina T, Gu H, Silva JC, Mertins P, Qiao J, Fereshetian S, Jaffe JD, Kuhn E, Udeshi ND, Carr SA (2015) Deep, quantitative coverage of the lysine acetylome using novel anti-acetyl-lysine antibodies and an optimized proteomic workflow. Mol Cell Proteomics 14:2429–2440. https://doi.org/10.1074/mcp.O114.047555

Sweredoski MJ, Moradian A, Raedle M, Franco C, Hess S (2015) High resolution parallel reaction monitoring with electron transfer dissociation for middle-down proteomics. Anal Chem 87:8360–8366. https://doi.org/10.1021/acs.analchem.5b01542

Synowsky SA, van Wijk M, Raijmakers R, Heck AJR (2009) Comparative multiplexed mass spectrometric analyses of endogenously expressed yeast nuclear and cytoplasmic exosomes. J Mol Biol 385:1300–1313. https://doi.org/10.1016/j.jmb.2008.11.011

Tang X, Meng Q, Gao J, Zhang S, Zhang H, Zhang M (2015) Label-free quantitative analysis of changes in Broiler liver proteins under heat stress using SWATH-MS technology. Sci Rep 5:15119. https://doi.org/10.1038/srep15119

Tao S, Orlando R (2014) A novel method for relative quantitation of n-glycans by isotopic labeling using 18O-water. J Biomol Tech 25:111–117. https://doi.org/10.7171/jbt.14-2504-003

Thermo Scientific Peptides for Targeted Quantitation (2017) https://www.thermofisher.com/be/en/home/life-science/protein-biology/peptides-proteins/custom-peptide-synthesis-services/peptides-targeted-quantitation.html

Thompson A, Scha R, Kuhn K, Kienle S, Schwarz J, Schmidt N, Neumann T, Hamon C (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. Anal Chem 75:1895–1904. https://doi.org/10.1021/ac0262560

Tian Y, Bova GS, Zhang H (2011) Quantitative glycoproteomic analysis of optimal cutting temperature-embedded frozen tissues identifying glycoproteins associated with aggressive prostate cancer. Anal Chem 83:7013–7019. https://doi.org/10.1021/ac200815q

Ting L, Rad R, Gygi SP, Haas W (2011) MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. Nat Methods 8:937–940. https://doi.org/10.1038/nmeth.1714

Toby TK, Fornelli L, Kelleher NL (2016) Progress in top-down proteomics and the analysis of proteoforms. Annu Rev Anal Chem 9:499–519. https://doi.org/10.1146/annurev-anchem-071015-041550

Tsou C-C, Avtonomov D, Larsen B, Tucholska M, Choi H, Gingras A-C, Nesvizhskii AI (2015) DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. Nat Methods 12:258–264, 7 p following 264. https://doi.org/10.1038/nmeth.3255

Tsutsui Y, Wintrode PL (2007) Hydrogen/deuterium exchange-mass spectrometry: a powerful tool for probing protein structure, dynamics and interactions. Curr Med Chem 14:2344–2358

Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc 11:2301–2319. https://doi.org/10.1038/nprot.2016.136

Van Hoof D, Pinkse MW, Oostwaard DW, Mummery CL, Heck AJ, Krijgsveld J (2007) An experimental correction for arginine-to-proline conversion artifacts in SILAC-based quantitative proteomics. Nat Methods 4:677–678

Villanueva J, Carrascal M, Abian J (2014) Isotope dilution mass spectrometry for absolute quantification in proteomics: Concepts and strategies. J Proteomics 96:184–199. https://doi.org/10.1016/j.jprot.2013.11.004

von Stechow L, Francavilla C, Olsen JV (2015) Recent findings and technological advances in phosphoproteomics for cells and tissues. Expert Rev Proteomics 12:469–487. https://doi.org/10.1586/14789450.2015.1078730

Wales TE, Engen JR (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass Spectrom Rev 25:158–170. https://doi.org/10.1002/mas.20064

Wang G, WW W, Zeng W, Chou C-L, Shen R-F (2006) Label-free protein quantification using LC-coupled ion trap or FT mass spectrometry: reproducibility, linearity, and application with complex proteomes. J Proteome Res 5:1214–1223. https://doi.org/10.1021/pr050406g

Wang C, Zhang P, Jin W, Li L, Qiang S, Zhang Y, Huang L, Wang Z (2017) Quantitative O-glycomics based on improvement of the one-pot method for nonreductive O-glycan release and simultaneous stable isotope labeling with 1-(d0/d5)phenyl-3-methyl-5-pyrazolone followed by mass spectrometric analysis. J Proteomics 150:18–30. https://doi.org/10.1016/j.jprot.2016.08.012

Washburn MP, Wolters D, Yates JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19:242–247. https://doi.org/10.1038/85686

Wenger CD, Lee MV, Hebert AS, McAlister GC, Phanstiel DH, Westphall MS, Coon JJ (2011) Gas-phase purification enables accurate, multiplexed proteome quantification with isobaric tagging. Nat Methods 8:933–935. https://doi.org/10.1038/nmeth.1716

Werner T, Becher I, Sweetman G, Doce C, Savitski MM, Bantscheff M (2012) High-resolution enabled TMT 8-plexing. Anal Chem 84:7188–7194. https://doi.org/10.1021/ac301553x

Werner T, Sweetman G, Savitski MF, Mathieson T, Bantscheff M, Savitski MM (2014) Ion coalescence of neutron encoded TMT 10-plex reporter ions. Anal Chem 86:3594–3601. https://doi.org/10.1021/ac500140s

Westbrook JA, Noirel J, Brown JE, Wright PC, Evans CA (2015) Quantitation with chemical tagging reagents in biomarker studies. Proteomics Clin Appl 9:295–300. https://doi.org/10.1002/prca.201400120

Wiesner J, Premsler T, Sickmann A (2008) Application of electron transfer dissociation (ETD) for the analysis of posttranslational modifications. Proteomics 8:4466–4483. https://doi.org/10.1002/pmic.200800329

Wiśniewski JR, Ostasiewicz P, Mann M (2011) High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers. J Proteome Res 10:3040–3049. https://doi.org/10.1021/pr200019m

Wu Y, Wang F, Liu Z, Qin H, Song C, Huang J, Bian Y, Wei X, Dong J, Zou H (2014) Five-plex isotope dimethyl labeling for quantitative proteomics. Chem Commun 50:1708. https://doi.org/10.1039/c3cc47998f

Xiao K, Yu F, Tian Z (2017) Top-down protein identification using isotopic envelope fingerprinting. J Proteomics 152:41–47. https://doi.org/10.1016/j.jprot.2016.10.010

Xie F, Liu T, Qian W-J, Petyuk VA, Smith RD (2011) Liquid chromatography-mass spectrometry-based quantitative proteomics. J Biol Chem 286:25443–25449. https://doi.org/10.1074/jbc.R110.199703

Xu BJ (2010) Combining laser capture microdissection and proteomics: methodologies and clinical applications. Proteomics Clin Appl 4:116–123. https://doi.org/10.1002/prca.200900138

Yang S, Yuan W, Yang W, Zhou J, Harlan R, Edwards J, Li S, Zhang H (2013) Glycan analysis by isobaric aldehyde reactive tags and mass spectrometry. Anal Chem 85:8188–8195. https://doi.org/10.1021/ac401226d

Yang S, Wang M, Chen L, Yin B, Song G, Turko IV, Phinney KW, Betenbaugh MJ, Zhang H, Li S (2015a) QUANTITY: an isobaric tag for quantitative glycomics. Sci Rep 5:17585. https://doi.org/10.1038/srep17585

Yang Y, Zheng N, Zhao X, Zhang Y, Han R, Ma L, Zhao S, Li S, Guo T, Wang J (2015b) Proteomic characterization and comparison of mammalian milk fat globule proteomes by iTRAQ analysis. J Proteomics 116:34–43. https://doi.org/10.1016/j.jprot.2014.12.017

Yang M, Cong M, Peng X, Wu J, Wu R, Liu B, Ye W, Yue X (2016) Quantitative proteomic analysis of milk fat globule membrane (MFGM) proteins in human and bovine colostrum and mature milk samples through iTRAQ labeling. Food Funct 7:2438–2450. https://doi.org/10.1039/c6fo00083e

Ye X, Luke B, Andresson T, Blonder J (2009) 18O stable isotope labeling in MS-based proteomics. Brief Funct Genomic Proteomic 8:136–144. https://doi.org/10.1093/bfgp/eln055

Ye H, Boyne MT, Buhse LF, Hill J (2013) Direct approach for qualitative and quantitative characterization of glycoproteins using tandem mass tags and an LTQ orbitrap XL electron transfer dissociation hybrid mass spectrometer. Anal Chem 85:1531–1539. https://doi.org/10.1021/ac3026465

Zeiler M, Straube WL, Lundberg E, Uhlen M, Mann M (2012) A protein epitope signature tag (PrEST) library allows SILAC-based absolute quantification and multiplexed determination of protein copy numbers in cell lines. Mol Cell Proteomics 11:O111.009613–O111.009613. https://doi.org/10.1074/mcp.O111.009613

Zhang R, Sioma CS, Wang S, Regnier FE (2001) Fractionation of isotopically labeled peptides in quantitative proteomics. Anal Chem 73:5142–5149. https://doi.org/10.1021/ac010583a

Zhang B, VerBerkmoes NC, Langston MA, Uberbacher E, Hettich RL, Samatova NF (2006) Detecting differential and correlated protein expression in label-free shotgun proteomics. J Proteome Res 5:2909–2918. https://doi.org/10.1021/pr0600273

Zhang H, Wang Z, Stupak J, Ghribi O, Geiger JD, Liu QY, Li J (2012a) Targeted glycomics by selected reaction monitoring for highly sensitive glycan compositional analysis. Proteomics 12:2510–2522. https://doi.org/10.1002/pmic.201100567

Zhang Y, Yin H, Lu H (2012b) Recent progress in quantitative glycoproteomics. Glycoconj J 29:249–258. https://doi.org/10.1007/s10719-012-9398-x

Zhang C, Ye Z, Xue P, Shu Q, Zhou Y, Ji Y, Fu Y, Wang J, Yang F (2016) Evaluation of different N-glycopeptide enrichment methods for N-glycosylation sites mapping in mouse brain. J Proteome Res 15:2960–2968. https://doi.org/10.1021/acs.jproteome.6b00098

Zhong X, Navare AT, Chavez JD, Eng JK, Schweppe DK, Bruce JE (2017) Large-scale and targeted quantitative cross-linking MS using isotope-labeled protein interaction reporter (PIR) cross-linkers. J Proteome Res 16:720–727. https://doi.org/10.1021/acs.jproteome.6b00752

Zhou H, Low TY, Hennrich ML, van der Toorn H, Schwend T, Zou H, Mohammed S, Heck AJR (2011) Enhancing the identification of phosphopeptides from putative basophilic kinase substrates using Ti (IV) based IMAC enrichment. Mol Cell Proteomics 10:M110 006452. https://doi.org/10.1074/mcp.M110.006452

Zhou S, Hu Y, Desantos-Garcia JL, Mechref Y (2015) Quantitation of permethylated n-glycans through multiple-reaction monitoring (MRM) LC-MS/MS. J Am Soc Mass Spectrom 26:596–603. https://doi.org/10.1007/s13361-014-1054-1

Zhou S, Hu Y, Veillon L, Snovida SI, Rogers JC, Saba J, Mechref Y (2016) Quantitative LC-MS/MS glycomic analysis of biological samples using aminoxyTMT. Anal Chem 88:7515–7522. https://doi.org/10.1021/acs.analchem.6b00465

Zhu W, Smith JW, Huang C-MM (2010) Mass spectrometry-based label-free quantitative proteomics. J Biomed Biotechnol 2010:840518. https://doi.org/10.1155/2010/840518

Zinn N, Winter D, Lehmann WD (2010) Recombinant isotope labeled and selenium quantified proteins for absolute protein quantification. Anal Chem 82:2334–2340. https://doi.org/10.1021/ac9025412

Zougman A, Nagaraj N, Mann M, Wiśniewski JR (2009) Universal sample preparation method for proteome analysis. Nat Methods 6:359–362. https://doi.org/10.1038/nmeth.1322

Proteomic Research in Farm Animal Serum and Plasma

Laura Soler and Ingrid Miller

Abstract Proteomics is one of the modern tools for in-depth study of the protein and peptide composition of complex protein mixtures and also has major applications in the field of animal science. Blood-derived fluids such as serum and plasma are rather unique biological samples as their protein content is contributed by the summation of all cellular proteome sets in the organism. Thus, they are a valuable source of information, reflecting the physiopathological status of the individuals. Being easy to obtain, they were and still are often the samples of choice to study physiology, investigate or diagnose diseases, as well as monitor and compare the influence of potentially harmful substances in different species.

Keywords Serum/plasma • Proteomics • Farm animals • Veterinary medicine • Animal physiology

Abbreviations

2DE	Two-dimensional electrophoresis
Apo A-I	Apolipoprotein A-I
DIGE	Fluorescence difference gel electrophoresis
GPA	Growth-promoting agent
Hp	Haptoglobin
MS	Mass spectrometry
PTM	Posttranslational modification
SAA	Serum amyloid A
SELDI	Surface-enhanced laser desorption/ionization
SRM	Selected reaction monitoring

L. Soler
Toxalim (Research Center in Food Toxicology), Université de Toulouse, INRA, ENVT, INP-PURPAN, UPS, Toulouse Cedex 03, France

I. Miller (✉)
Institute of Medical Biochemistry, University of Veterinary Medicine Vienna, Vienna, Austria
e-mail: ingrid.miller@vetmeduni.ac.at

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_6

1 Methodological Aspects

1.1 Proteomic Methods

Nowadays, two main methodological approaches exist in proteomics: gel-based and gel-free. Historically, gel-based methods were the first developed, relying on two-dimensional electrophoresis (2DE) as a combination of isoelectric focusing (separation according to protein charge) and SDS-PAGE (separation according to protein size), both under denaturing and reducing conditions (Miller 2011). Serum and plasma were among the first samples for which identification maps were established, in the beginning based on specific or cross-reactive antibodies, later on identifying proteins by mass spectrometric (MS) methods. Among the species of interest were several farm animals—bovine (Wait et al. 2002; Talamo et al. 2003), porcine (Miller et al. 2009), ovine (Chiaradia et al. 2012), chicken (Huang et al. 2006), and lately also fish (Braceland et al. 2013). These maps of healthy animals were created to serve as reference patterns for comparison with serum/plasma of diseased animals and usually showed patterns of serum/plasma without any pretreatment or pre-fractionation. One application for the bovine serum map was a study on several acute phase proteins during development of clinical and subclinical mastitis. Data showed positive correlation of results of proteomic- and immunoassay-based quantification for the selected proteins (Alonso-Fauste et al. 2012).

The more recent proteomic approach is comparison of samples from two or more different pathophysiological states or breeds. Investigating specimens with similar analytical history but from different time points, healthy/diseased or exposed/nonexposed animals, allows flexibility in sample preparation and analysis/separation. In the electrophoretic field, this trend is nowadays best utilized by fluorescence difference gel electrophoresis (DIGE), where up to three previously fluorophore-labeled samples may be separated in one gel. This minimizes gel to gel variation and increases sensitivity and accuracy in quantitative comparisons (Timms and Cramer 2008; Miller 2012).

Due to the rapid development of devices and equipment, gel-free MS methods today show high sensitivity and a large variety for different applications. This, together with more refined sample pretreatment, was, for instance, able to extend the list of bovine serum proteins (Faulkner et al. 2012; Henning et al. 2014; Skrzypczak et al. 2011). Combination of MS with high-resolution multidimensional chromatography prevails, but inclusion of other methods is possible as well (Mansor et al. 2013). Quantification is based on isotopic or chemical labels or performed in a label-free mode (Gevaert and Vandekerckhove 2011; Bantscheff et al. 2012).

Identification of peptides, either from tryptically digested protein spots or from separated peptide mixtures, is achieved by searches in dedicated databases. Protein databases [UniProtKB (http://www.uniprot.org/uniprot/), NCBI database (http://www.ncbi.nlm.nih.gov/guide/proteins/)] encompass a wide variety of organisms, including among them the most common proteins from farm animal species. Still,

in some cases it may be necessary to identify proteins by de novo sequencing or as homologues of related species (Soares et al. 2012). Repositories have been set up lately to collect protein/peptide data and to use them as a source for further progress in the field—PRIDE (Vizcaino et al. 2009), the Global Proteome Machine Database (Craig et al. 2004), and PeptideAtlas (Deutsch et al. 2008). The PeptideAtlas contains a large set of farm animal protein data, for example, on cow (Bislev et al. 2012a) and pig (Hesselager et al. 2016) proteomes.

Whereas general proteomic investigations are holistic and inductive, formulating hypotheses only based on study results, the latest developments in the MS field also allow targeting of specific analyte sets by selected reaction monitoring (SRM). Based on their unique protein-specific peptides, they may become helpful tools for future clinical use and also in veterinary applications (Bislev et al. 2012b; Marco-Ramell et al. 2013).

Also in animal studies, the number of studies applying MS methods has meanwhile outperformed the number of investigations performed by 2DE. They benefit from high degree of automatization, high sensitivity, and growing databases. Nevertheless, previous studies, including also human serum, have shown that the "proteome" detected by various methods differs, that there is a considerable degree of proteins monitored only with one type of approach, and, overall, that MS- and gel-based methods are complementary (Anderson et al. 2004). Additionally, the search for disease biomarkers revealed that there are cases where not only the concentration of one or more specific proteins is important but also the type of modification or size of the molecule (protein species or proteoforms). This may be easier to detect (but not to identify) with gel-based methods, as this often changes the spot pattern (Rogowska-Wrzesinska et al. 2013; Oliveira et al. 2014).

Surface-enhanced laser desorption/ionization (SELDI) is an MS method where intact analytes are enriched on chips based on their affinity to specifically modified surface. It allows profiling of samples and search for disease-specific markers by comparing MS peak patterns under different conditions. With particular care to avoid interferences, it may be a helpful tool in the search for marker proteins, specifically smaller polypeptides (Poon 2007).

Serum or plasma may additionally serve as immunological "reagent," for screening for antibodies or autoantibodies in the respective individual. These immunoproteomic applications are helpful to detect previous or present contact with bacteria, viruses, parasites, or allergenic substances and may be combined with a large panel of methods.

1.2 Serum versus Plasma

Serum is obtained from blood after coagulation of fibrin along with the cellular components and their removal by centrifugation, while for plasma preparation anticoagulants (citrate, heparin, EDTA) are added to the collection tubes or used to pre-coat the sample tubes. Thus, the main differences in the proteome are

coagulation-related proteins, mainly the presence/absence of fibrinogen (Kim et al. 2007; Miller et al. 2010). Though both sample types may be used for evaluation of protein patterns, reference ranges or concentrations of particular proteins/protein families may be sample-type specific [for instance, lipoprotein fractions (Collins and Olivier 2010)]. Investigations of human serum vs. plasma samples have detected considerable differences in the low-molecular-weight proteome and peptidome, due to release of peptides derived from cellular components or the clot (Tammen et al. 2005). Similarly, pre-analytical factors such as patient fasting, hemolysis, sample handling (clotting time, time lag, centrifugation speed), and storage (temperature, freezing/thawing) markedly influence protein patterns and call for strict standardization (Hsieh et al. 2006; Hassis et al. 2015).

1.3 Species Specificity

Serum/plasma consists of a multitude of different proteins with a large variety of functions: carriers for metals, hormones, lipids, vitamins, drugs; enzymes, inhibitors, hormones, defense against harmful agents, buffers. Single components (may) vary depending on species and physiopathological conditions (Eckersall 2008). Though the main proteins are usually present throughout animal species, protein properties (isoelectric point, size, modifications) and concentrations may show species-dependent particularities (examples given below). This limits comparability of patterns between species and makes investigation of single species mandatory.

One of the most important acute phase markers in most species is haptoglobin (Hp), a glycosylated protein whose levels in serum are consistently increased during infections. It forms a tetramer of 2α- and 2β-chains, but for humans two different types of α-chains exist (two genotypes, resulting in three phenotypes of the protein, Fig. 1, top panel). Vertebrates show corresponding types but either with the α_2 (ruminants-) or the α_1-chain (other species) (Meier et al. 1980). Pattern in two-dimensional electrophoresis reflects the differing protein properties (Fig. 1, bottom panels). Additionally, Hp concentration in healthy bovines is very low (Wait et al. 2002), whereas the protein is well detectable in baseline status of other species [Fig. 1, (Miller 2011)].

The acute phase protein serum amyloid A (SAA) is another example for species specificity. In most species the SAA family consists of two main forms (SAA1 and SAA2) with neutral pI, circulating in serum and produced in the liver, and an additional alkaline family member, SAA3, mainly produced locally and not contributing to circulating levels. In the pig, however, the main circulating form is biochemically more similar to the SAA3 homologue of those other species, which may explain why most antibodies did not cross-react with pig SAA and also why the pig is resistant to AA-amyloidosis (Soler et al. 2013).

Various acute phase proteins are known, depending on the animal species, and may respond in a species-dependent way and time course (Eckersall and Bell 2010). For instance, α_1-acid glycoprotein has been widely reported as a positive acute

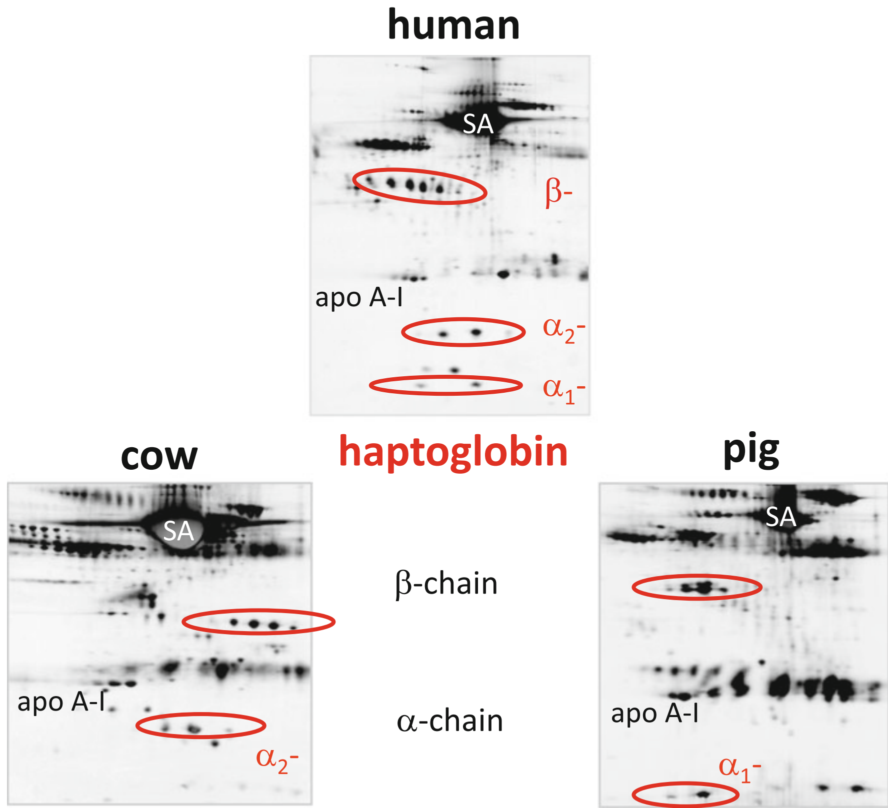

Fig. 1 Haptoglobin chains in serum 2DE patterns of different species. Serum samples of different species: (**a**) human (mixed haptoglobin phenotype), (**b**) bovine (cow with inflammation), and (**c**) pig. Detailed views from gels of IPG pH 4-10NL and 10–15% gradient gel for SDS-PAGE, silver stain, 0.5 μl serum. For details, see Miller et al. (2009). *SA* serum albumin, *apo A-I* apolipoprotein A-I, single haptoglobin chains are indicated (α_1-, α_2-, β-chain)

phase marker in many species, notably in the chicken, where it is one of the major markers of inflammation (O'Reilly and Eckersall 2014). However, this protein is a negative acute phase protein during a range of experimental infections and aseptic inflammation in the pig (Heegaard et al. 2013), whereas in the same species, increased serum levels have been described during obesity (Rødgaard et al. 2013) and treatment with antimicrobial growth promotors (Soler et al. 2016).

1.4 Proteoforms

Proteomics allows a more detailed characterization of protein heterogeneity, in other words the detection of different proteoforms, often due to posttranslational

modifications (PTM; in serum frequently due to glycosylation). Change of PTM may be a sign of disease or predisposition; thus, monitoring of proteoforms may help arrive at biomarker candidates with a better diagnostic performance than the overall protein concentration of the given protein.

Two examples of pig acute phase proteins show the high resolution and usefulness of modern proteomic methods: Hp was characterized regarding the glycosylation of its β-subunit (O- and N-glycosylated, contrary to human). In a viral infection by PCV2-SD, all proteoforms were upregulated to a similar extent, thus confirming that determination of overall Hp concentration was a good enough test (Marco-Ramell et al. 2014). At the same time, the negative acute phase marker apolipoprotein A-I (Apo A-I) showed differential regulation of the protein's three main spots in 2DE (corresponding to the immature form and the two mature forms, where no difference in PTM could be detected). Indications for a correlation between pattern changes and severity of infection or type of infectious agent were found (Marco-Ramell et al. 2015).

1.5 Reducing Serum/Plasma Complexity

Protein composition of serum/plasma shows high complexity, both in number and concentration of constituents. Concentrations span over more than ten orders of magnitude between highly abundant and trace proteins (Anderson and Anderson 2002). Low-abundance and trace proteins are regarded as typical leakage products from diseased, dying, or dead cells originating from different organs of the body. This makes them particularly interesting for investigations aiming at diagnosis and health monitoring (instead of more invasive sample sources, like biopsies). Proteomics, as a sensitive method able to discern also different often organ- or tissue-specific modifications of proteins, has proven to be an appropriate tool for this task.

The large dynamic range of serum/plasma proteins is a challenge for in-depth investigations, as this is much more than modern analytical techniques can cover (Altelaar and Heck 2012). The general recommendation is removal of major serum proteins or sample pre-fractionation, to enable detection of minor components. Still, analysis of medium- to high-abundance proteins has its benefits, for instance, evaluating level and pattern of acute phase proteins or modifications or isoforms characteristic for diseases/disorders (Marco-Ramell et al. 2014, 2015). Sample pretreatment bears the risk of unintentionally modifying ratios between proteins and proteoforms or of removing potentially interesting marker proteins (Gianazza et al. 2016). Studies on the "albuminome" have shown that even albumin-depletion methods may remove many more proteins than intended (Holewinski et al. 2013).

A wide range of methods has been suggested for sample treatment prior to proteomic experiments. This includes removal of high- to medium-abundance proteins, for instance, by immunoaffinity methods using up to 14 specific antibodies immobilized on beads (Liu et al. 2006). These devices have been developed for application on human samples, and their use in animal sera has to be critically

evaluated before use (Bandow 2010). An often suggested first step is depletion of albumin and immunoglobulins (Faulkner et al. 2011), but as they comprise about 60–70% of overall serum protein, this may not be sufficient. Other or additional ways to simplify the serum proteome are salting out, precipitation at low pH or with organic solvents (Fernández et al. 2011; Henning et al. 2015), and different electrophoretic or chromatographic steps (ion exchange, reversed phase chromatography) (Miller 2011; Whiteaker et al. 2007). The development of combinatorial peptide ligand libraries (CPLL) has added another method that allows species-independent enrichment of minor serum components (Righetti et al. 2006; Marco-Ramell and Bassols 2010; Di Girolamo et al. 2014). Often, a combination of different methods either performed in parallel or consecutively may help to mine deeper into the proteome (Soler et al. 2016; Jmeian and El Rassi 2009; Puangpila et al. 2015; Stalder et al. 2008). Examples for three of these methods are given in Fig. 2, using pig serum as source: dialysis at low pH to decrease the concentration

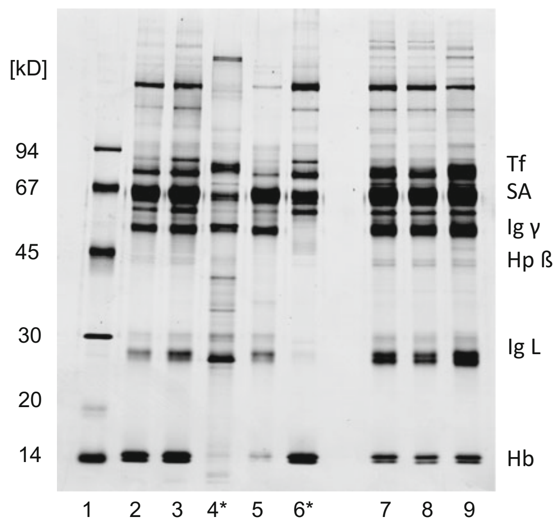

Fig. 2 Pretreatment of serum samples for depletion of major proteins or enrichment of minor components: (1) molecular weight marker (LMW, GE Healthcare Life Sciences, Munich, Germany). (2) Pig serum 1 (starting material for samples in lanes 3–6). ProteoMiner (Bio-Rad, Hercules, CA, USA): (3) fraction of major proteins that were removed (fall-through peak of cartridge) and (4) enriched minor proteins (bound to resin). For details, see Soler et al. (2016). ProteoPrep Albumin and IgG Depletion Sample Prep Kit (Protea, Morgantown, WV, USA): (5) fraction of major proteins that were removed (bound to resin) and (6) albumin/IgG-depleted fraction (fall-through peak of cartridge). For details, see Soler et al. (2016). (7) Pig serum 2 (starting material for samples in lanes 8–9). Dialysis of serum at pH 4.2 according to Henning et al. (2015): (8) supernatant and (9) redissolved precipitate. Samples marked with asterisks have lower concentrations of major serum components. SDS-PAGE in 10–15% gradient gels (140 × 140 × 1.5 mm), silver stain. The main serum proteins are indicated: *Hb* hemoglobin, *Hp ß* haptoglobin ß-chain, *Ig L* immunoglobulin light chain, *Ig γ* IgG H(γ)-chain, *SA* serum albumin, *Tf* transferrin

of a few highly abundant proteins (lanes 8, 9), removal of albumin and IgG (lanes 5, 6), and enrichment of low-abundance proteins (lanes 3, 4). As to be seen, in our hands low-pH dialysis did not markedly change the band pattern of pig serum (in contrast to Henning et al. 2015), whereas the two affinity methods were able to evidently reduce the abundance of albumin and transferrin. Combining data from both enrichment steps with analysis of unfractionated serum and thus investigating three serum proteome (sub)sets from the same samples allowed covering a wider range of analytes in a feeding study with oxytetracycline (Soler et al. 2016).

Alternatively, fractionation according to isoelectric point may be performed either before the proteomic separation (Miller 2011; Westbrook et al. 2001) or by applying the complex sample on so-called zoom gels, i.e., two-dimensional electrophoresis where the first dimensional IPG run is performed in narrow range pH gradients (Hoving et al. 2000; Marco-Ramell et al. 2014). Sample fractionation may also be achieved on the peptide level, after tryptic digestion of the proteins, utilizing different properties of the peptides.

2 Topics of Serum/Plasma Proteomics in Farm Animal Species

Advantages and limitations of serum/plasma proteomics in farm animals have been extensively reviewed (Almeida et al. 2015; Bendixen et al. 2011; de Almeida and Bendixen 2012; Di Girolamo et al. 2014; Eckersall et al. 2012), mainly in the areas discussed below.

2.1 Animal Health and Disease

Serum and plasma represent the most convenient sample types to collect in order to investigate the pathogenesis and/or the time course of a disease and so for the detection of biomarkers. Serum proteomics can help unraveling the alterations associated with different infections of relevance for veterinary sciences. This is the case of bovine paratuberculosis and avian coccidiosis, both diseases leading to important economic losses. Serum proteomics has helped identifying some proteins possibly implicated in the pathogenesis and response against these infections, as well as the stage of the infection. More specifically, transferrin, gelsolin isoforms α and β, some complement components, a copper amine oxidase, coagulation factors II and XIII B polypeptide, as well as fibrinogen γ-chain and its precursor were identified as showing differential abundances in paratuberculosis (You et al. 2012), whereas malate dehydrogenase 2, a subunit of NADH dehydrogenase 1 alpha subcomplex, and an ATP synthase were detected as potential early diagnostic markers for avian coccidiosis (Gilbert et al. 2011). In fish, two 2DE-based

proteomic studies have been investigating host responses in search of health biomarkers. The first study employed a model of trauma in the loach to identify possible markers of inflammation, such as parvalbumin, gastrin 71, and the signal recognition protein (Wu et al. 2004). The second study aimed at investigating the host response against the salmonid alphavirus, the etiological agent of pancreas disease in marine Atlantic salmon (Braceland et al. 2013). Results identified some proteins (creatine kinase, enolase, and malate dehydrogenase) that changed in abundance in correlation with tissue pathology as well as others that were correlated with the stage of the infection (hemopexin, transferrin, and apolipoprotein) (Braceland et al. 2013). 2DE coupled with MS has also helped describing the serum proteome changes related with the inflammatory, oxidative, and lipid metabolism alterations associated with the development of laminitis in cows (Dong et al. 2015); the exposure to swine fever virus causing modifications in pig proteins associated with inflammation, blood coagulation, and angiogenesis (Sun et al. 2011); or the acute phase proteins changing their abundance in the course of ovine rhino-tracheobronchitis (Chiaradia et al. 2012).

Early changes in the serum proteome associated with the predisposition to develop a particular pathology may serve as prognostic biomarkers. Serum proteomics has been employed in order to identify cows and sheep prone to develop postpartum uterine infection; finding suitable markers would represent an important tool for reproductive postpartum disease management (Cairoli et al. 2006; Chiaradia et al. 2012). In both species the serum proteome changes occurring at peripartum showed that levels of α_1-acid glycoprotein could be employed as a prognostic biomarker of endometritis predisposition (Cairoli et al. 2006; Chiaradia et al. 2012).

Adding gel-based proteomic data to results from other -omic methods has often revealed a broader, more detailed picture. The combination of proteomic and metabolomic data has proven superior to results from single methods, for example, for providing a prognosis for the outcome of bovine respiratory disease (Aich et al. 2009). Likewise, the same method combination was able to better describe the complex biological events associated with combined stress and viral infections in cattle (Aich et al. 2007). A good example for the complementarity even of different proteomic approaches is the comparison of the serum proteome of healthy cows with those from animals suffering from clinical and subclinical mastitis (Turk et al. 2012). The combination of peptidic profiling using MALDI-TOF, 2DE, and shotgun proteomics provided at first an overview of the quality of samples for proteomic evaluation, a factor that is often neglected but is in fact crucial for a correct interpretation of results. Combined results from proteomic analyses gave an overview of all inflammatory and oxidative changes associated with this condition but more interestingly identified the protein vitronectin as a very promising biomarker for the detection of subclinical mastitis, a condition that is hard to detect at the farm level (Turk et al. 2012).

The use of technical approaches other than 2DE proteomics is less common in animal clinical sciences, although they have been applied successfully and are becoming more popular. Shotgun proteomics of plasma has been used in a study

of foot rot in the dairy cow in search for possible factors associated with this disease, such as the peptidoglycan recognition protein L and keratin sulfate proteoglycan (Sun et al. 2013). Shotgun proteomics could also distinguish the host response toward two pathogens that are difficult to detect separately in the cow: *Mycobacterium tuberculosis* and *Mycobacterium paratuberculosis* (Seth et al. 2009). It also identified some proteins that could be considered possible biomarkers of foot-and-mouth disease virus-infected pigs, such as apolipoprotein A-IV precursor, haptoglobin, and probable chemoreceptor glutamine deamidase cheD (Liu et al. 2011). Existing shotgun proteomic and traditional biochemistry data may be employed in serum proteomics to design targeted assays based on the use of LC-MS/MS technology for the multiple quantification of already known biomarkers. Though up to now scarce in veterinary sciences, data from robust targeted proteomic assays for a given protein have been described to correlate very well with immunometric quantifications (Bislev et al. 2012b). Also a targeted SRM assay for quantification of different acute phase proteins in pig serum proved to be a good alternative to commercial kits for the multiple quantification of biomarkers (Marco-Ramell et al. 2013).

The combination of different proteomic biomarkers or “biomarker profiles” can give a combined measure associated with the studied condition. Biomarker profiles are potentially better at defining a complex pathophysiological condition than single biomarkers, and clinical proteomics is moving its focus to the identification of such panels as fingerprints that can define a diagnostic phenotype (Gruys et al. 2006; Koene et al. 2012). In animal sciences, there is a growing interest in identifying and implementing the use of disease biomarker profiles, for example, as established by SELDI. It has been employed in ovine serum from animals diagnosed with paratuberculosis, finding that the proteins transthyretin and α-hemoglobin were differentially abundant in healthy and diseased animals (Zhong et al. 2011). In pigs, serum SELDI profiles were useful for the early detection and identification of viral infections in an experimental model for porcine circovirus-associated disease (PCVAD), consisting of porcine circovirus type 2 (PCV2) infection in combination with either porcine parvovirus (PPV) or porcine reproductive and respiratory syndrome virus (PRRSV; Koene et al. 2012).

2.2 Animal Physiology, Stress, and Nutrition

Applications of serum/plasma proteomics in farm animals go beyond its use in clinical sciences and also include studies on physiology, welfare, and nutrition in farm animals (Bendixen et al. 2011; Eckersall et al. 2012). Serum proteomic analysis is in this sense a powerful tool to understand connections between genotype, phenotype, and environmental and managing factors. These remain poorly understood and need to be known to balance production, health, and well-being. Identification of biomarkers through proteomics for feed efficiency in livestock as well as the identification of pathways with distinct expression patterns between

animals that differ in feed efficiency could therefore be useful for animal production. In this sense, serum proteomics has been employed to identify biomarkers such as gelsolin, vitronectin, and serpin A3 linked with increased swine feed efficiency (Grubbs et al. 2016). Likewise, the proteins plasminogen and vitamin D-binding protein precursor were found in hens as highly correlated with early sexual maturity and increased egg production (Huang et al. 2006).

Ethical production of food is nowadays a main concern of consumers. In parallel, it has been confirmed that handling procedures creating stress in farm animals have a negative impact on the quality of the final product and on the cost of production. The objective evaluation of stress is required in order to improve animal care and management in farm animals. This aspect of animal production is discussed in detail in this book in the chapter "Proteomics and Animal Stress." Serum proteomics has proven to be a good tool for the discovery and use of biochemical biomarkers to evaluate the stress response in farm animals. Some examples in the literature include the evaluation of the stress response toward individual housing of sows (Marco-Ramell et al. 2016), housing at high stocking density in pigs (Marco-Ramell et al. 2011), the response to heat stress in lactating dairy cows (Fuquay 1981), fish handling stress (Liu et al. 2008), and the adaptation of cows to harsh breeding conditions (Marco-Ramell et al. 2012).

2.3 Food Safety

The use of growth-promoting agents (GPAs) in meat production to increase feed conversion and lean meat production has been a common practice in the past. The use of such substances, among them anabolic and sexual steroids (mainly androgens and estrogens), β_2-agonists, and antimicrobials, pose a serious risk for public health, and many countries have banned their use (Brameld and Parr 2016). This has created, on one hand, the need of finding natural, harmless alternatives to GPAs and, on the other hand, the need of accurately identifying the illicit use of these treatments. Serum/plasma proteomic analysis can help in both fields. When looking for natural alternatives that help boosting the performance of farm animals, the first step is understanding the mechanism of action by which antimicrobial substances employed at subclinical dose worked as growth promoters (Brown et al. 2017). The serum proteome of recently weaned pigs fed with oxytetracycline showed some signs that helped explain the switch to an energy-saving phenotype thanks to an anti-inflammatory effect as well as a protective effect against lipid oxidation (Soler et al. 2016). As a natural alternative to antimicrobials, probiotics have been proposed. The effect of exposing the rainbow trout to two different probiotics (*Aeromonas sobria* GC2 and *Bacillus* sp. JB-1) has been explored in the serum proteome. Opposite to expected, results showed that the use of these probiotics was associated with the development of an acute phase response (Brunt et al. 2008).

The second consequence of the banning of GPAs is the present need to find sensitive ways of detecting their use. The analytical detection of these substances is

often very complicated, since they are present in low doses in the analytical samples and have a structure that is similar to some natural hormones. In turn, the evaluation of the molecular changes associated with their use could be more suitable as a means to detect their use. Serum proteomics has been employed for the latter in an extensive way, as it offers an overview of the changes linked with the use of GPAs. The general conclusion is that a combination of markers of different biological functions should be employed, as the use of GPAs affects different physiological processes including an altered hepatic metabolism, an acute phase reaction activity, the response to sexual hormones, and the action on the musculoskeletal system (Kinkead et al. 2015; McGrath et al. 2013). A relationship between the treatment with different substances and an increase of proteins with phospholipid binding function, among them β_2-glycoprotein 1 (Della Donna et al. 2009), a specific truncated form of Apo A-I (Draisci et al. 2007), and serum paraoxonase 1 (Guglielmetti et al. 2014), has been found associated with the use of different illicit substances.

2.4 Animal Models of Human Medicine

Large animal models of disease for human medicine are increasingly employed to overcome some disadvantages of rodent models. Differences in size, anatomy, physiology, etc. between rodents and men have forced researchers to look for alternative models, of which the pig seems to be one of the most suitable (de Almeida and Bendixen 2012). Some studies have successfully applied serum proteomics in pig models of disease, such as models of acute sepsis (Thongboonkerd et al. 2009) or nonalcoholic steatohepatitis (Bell et al. 2010) to decipher the molecular response to a given condition with the objective of translating results to human medicine.

3 Conclusions and Future Perspectives

Achievements of the last decades clearly show the added value of proteomics in biological research and now also in the field of animal proteomics. However, over the years proteomic methods have been and are still being refined, and newer research reveals the advantages of determining proteoforms both in research discovery phase and in diagnosis. Additionally, new technical developments in the field increase sensitivity and allow mining deeper into the very low-abundance protein population of serum/plasma. This, together with the growing interest in animal-related studies, will help to further enhance in-depth knowledge in the future, even in comparatively well-studied samples like serum/plasma.

References

Aich P, Jalal S, Czuba C et al (2007) Comparative approaches to the investigation of responses to stress and viral infection in cattle. OMICS 11:413–434

Aich P, Babiuk LA, Potter AA et al (2009) Biomarkers for prediction of bovine respiratory disease outcome. OMICS 13:199–209

Almeida AM, Bassols A, Bendixen E et al (2015) Animal board invited review: advances in proteomics for animal and food sciences. Animal 9:1–17

Alonso-Fauste I, Andrés M, Iturralde M et al (2012) Proteomic characterization by 2-DE in bovine serum and whey from healthy and mastitis affected farm animals. J Proteomics 75:3015–3030

Altelaar AF, Heck AJ (2012) Trends in ultrasensitive proteomics. Curr Opin Chem Biol 16:206–213

Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 1:845–867

Anderson NL, Polanski M, Pieper R et al (2004) The human plasma proteome. A nonredundant list developed by combination of four separate sources. Mol Cell Proteomics 3:311–326

Bandow JE (2010) Comparison of protein enrichment strategies for proteome analysis of plasma. Proteomics 10:1416–1425

Bantscheff M, Lemeer S, Savitski MM et al (2012) Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. Anal Bioanal Chem 404:939–965

Bell LN, Lee L, Saxena R et al (2010) Serum proteomic analysis of diet-induced steatohepatitis and metabolic syndrome in the Ossabaw miniature swine. Am J Physiol Gastrointest Liver Physiol 298:G746–G754

Bendixen E, Danielsen M, Hollung K, Gianazza E, Miller I (2011) Farm animal proteomics—a review. J Proteomics 74:282–293

Bislev SL, Deutsch EW, Sun Z et al (2012a) A Bovine PeptideAtlas of milk and mammary gland proteomes. Proteomics 12:2895–2899

Bislev SL, Kusebauch U, Codrea MC et al (2012b) Quantotypic properties of QconCAT peptides targeting bovine host response to *Streptococcus uberis*. J Proteome Res 11:1832–1843

Braceland M, Bickerdike R, Tinsley J et al (2013) The serum proteome of Atlantic salmon, Salmo salar, during pancreas disease (PD) following infection with salmonid alphavirus subtype 3 (SAV3). J Proteomics 94:423–436

Brameld JM, Parr T (2016) Improving efficiency in meat production. Proc Nutr Soc 75:242–246

Brown K, Uwiera RRE, Kalmokoff ML et al (2017) Antimicrobial growth promoter use in livestock: a requirement to understand their modes of action to develop effective alternatives. Int J Antimicrob Agents 49:12–24

Brunt J, Hansen R, Jamieson DJ et al (2008) Proteomic analysis of rainbow trout (Oncorhynchus mykiss, Walbaum) serum after administration of probiotics in diets. Vet Immunol Immunopathol 121:199–205

Cairoli F, Battocchio M, Veronesi MC et al (2006) Serum protein pattern during cow pregnancy: acute-phase proteins increase in the peripartum period. Electrophoresis 27:1617–1625

Chiaradia E, Avellini L, Tartaglia M et al (2012) Proteomic evaluation of sheep serum proteins. BMC Vet Res 8:66

Collins LA, Olivier M (2010) Quantitative comparison of lipoprotein fractions derived from human plasma and serum by liquid chromatography-tandem mass spectrometry. Proteome Sci 8:42

Craig R, Cortens JP, Beavis RC (2004) Open source system for analyzing, validating, and storing protein identification data. J Proteome Res 3:1234–1242

de Almeida AM, Bendixen E (2012) Pig proteomics: a review of a species in the crossroad between biomedical and food sciences. J Proteomics 75:4296–4314

Della Donna L, Ronci M, Sacchetta P et al (2009) A food safety control low mass-range proteomics platform for the detection of illicit treatments in veal calves by MALDI-TOF-MS serum profiling. Biotechnol J 4:1596–1609

Deutsch EW, Lam H, Aebersold R (2008) PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows. EMBO Rep 9:429–434

Di Girolamo F, D'Amato A, Lante I et al (2014) Farm animal serum proteomics and impact on human health. Int J Mol Sci 15:15396–15411

Dong SW, Zhang SD, Wang DS et al (2015) Comparative proteomics analysis provide novel insight into laminitis in Chinese Holstein cows. BMC Vet Res 11:161

Draisci R, Montesissa C, Santamaria B et al (2007) Integrated analytical approach in veal calves administered the anabolic androgenic steroids boldenone and boldione: urine and plasma kinetic profile and changes in plasma protein expression. Proteomics 7:3184–3193

Eckersall PD (2008) Chapter 5: Proteins, proteomics, and the dysproteinemias. In: Kaneko JJ, Harvey JW, Bruss ML (eds) Clinical biochemistry of domestic animals, 6th edn. Elsevier Academic Press, Amsterdam, pp 117–156

Eckersall PD, Bell R (2010) Acute phase proteins: biomarkers of infection and inflammation in veterinary medicine. Vet J 185:23–27

Eckersall PD, de Almeida AM, Miller I (2012) Proteomics, a new tool for farm animal science. J Proteomics 75:4187–4189

Faulkner S, Elia G, Hillard M et al (2011) Immunodepletion of albumin and immunoglobulin G from bovine plasma. Proteomics 11:2329–2335

Faulkner S, Elia G, Mullen MP et al (2012) A comparison of the bovine uterine and plasma proteome using iTRAQ proteomics. Proteomics 12:2014–2023

Fernández C, Santos HM, Ruíz-Romero C (2011) A comparison of depletion versus equalization for reducing high-abundance proteins in human serum. Electrophoresis 2011(32):2966–2974

Fuquay JW (1981) Heat stress as it affects animal production. J Anim Sci 52:64–74

Gevaert K, Vandekerckhove J (eds) (2011) Gel-free proteomics—Methods and protocols, Methods in Molecular Biology, vol 753. Springer, Heidelberg

Gianazza E, Miller I, Palazzolo L et al (2016) With or without you — proteomics with or without major plasma/serum proteins. J Proteomics 140:62–80

Gilbert ER, Cox CM, Williams PM et al (2011) Eimeria species and genetic background influence the serum protein profile of broilers with coccidiosis. PLoS ONE 6:e14636

Grubbs JK, Dekkers JCM, Huff-Lonergan E et al (2016) Identification of potential serum biomarkers to predict feed efficiency in young pigs. J Anim Sci 94:1482

Gruys E, Toussaint MJM, Niewold T et al (2006) Monitoring health by values of acute phase proteins. Acta Histochem 108:229–232

Guglielmetti C, Mazza M, Pagano M et al (2014) Identification by a proteomic approach of a plasma protein as a possible biomarker of illicit dexamethasone treatment in veal calves. Food Addit Contam Part A Chem Anal Control Expo Risk Assess 31:833–838

Hassis ME, Niles RK, Braten MN et al (2015) Evaluating the effects of preanalytical variables on the stability of the human plasma proteome. Anal Biochem 478:14–22

Heegaard PMH, Miller I, Sorensen NS et al (2013) Pig α_1-acid glycoprotein: characterization and first description in any species as a negative acute phase protein. PLoS One 8:e68110

Henning AK, Groschup MH, Mettenleiter TC (2014) Analysis of the bovine plasma proteome by matrix-assisted laser desorption/ionisation time-of-flight tandem mass spectrometry. Vet J 199:175–180

Henning A, Albrecht D, Riedel K et al (2015) An alternative method for serum protein depletion/enrichment by precipitation at mildly acidic pH values and low ionic strength. Proteomics 15:1935–1940

Hesselager MO, Codrea MC, Sun Z et al (2016) The Pig PeptideAtlas: a resource for systems biology in animal production and biomedicine. Proteomics 16:634–644

Holewinski JR, Jin Z, Powell MJ et al (2013) A fast and reproducible method for albumin isolation and depletion from serum and cerebrospinal fluid. Proteomics 13:743–750

Hoving S, Voshol H, van Oostrum J (2000) Towards high performance two-dimensional gel electrophoresis using ultrazoom gels. Electrophoresis 21:2617–2621

Hsieh SY, Chen RK, Pan YH et al (2006) Systematical evaluation of the effects of sample collection procedures on low molecular- weight serum/plasma proteome profiling. Proteomics 6:3189–3198

Huang SY, Lin JH, Chen YH et al (2006) Analysis of chicken serum proteome and differential protein expression during development in single-comb White Leghorn hens. Proteomic 6:2217–2224

Jmeian Y, El Rassi Z (2009) Liquid-phase-based separation systems for depletion, prefractionation and enrichment of proteins in biological fluids for in-depth proteomics analysis. Electrophoresis 30:249–261

Kim H, Kim M, So E et al (2007) Comparison of proteomes in various human plasma preparations by two-dimensional gel electrophoresis. J Biochem Biophys Methods 70:619–625

Kinkead RA, Elliott CT, Cannizzo FT et al (2015) Proteomic identification of plasma proteins as markers of growth promoter abuse in cattle. Anal Bioanal Chem 407:4495–4507

Koene MG, Mulder HA, Stockhofe-Zurwieden N et al (2012) Serum protein profiles as potential biomarkers for infectious disease status in pigs. BMC Vet Res 8:32

Liu T, Qian W, Mottaz HM et al (2006) Evaluation of multiprotein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. Mol Cell Proteomics 5:2167–2174

Liu X, Afonso L, Altman E et al (2008) O-acetylation of sialic acids in N-glycans of Atlantic salmon (Salmo salar) serum is altered by handling stress. Proteomics 8:2849–2857

Liu Y, Zhang K, Zheng H et al (2011) Proteomics analysis of porcine serum proteins by LC-MS/MS after foot-and-mouth disease virus (FMDV) infection. J Vet Med Sci 73:1569–1572

Mansor R, Mullen W, Albalat A et al (2013) A peptidomic approach to biomarker discovery for bovine mastitis. J Proteomics 85:89–98

Marco-Ramell A, Bassols A (2010) Enrichment of low-abundance proteins from bovine and porcine serum samples for proteomic studies. Res Vet Sci 89:340–343

Marco-Ramell A, Pato R, Peña R et al (2011) Identification of serum stress biomarkers in pigs housed at different stocking densities. Vet J 190:e66–e71

Marco-Ramell A, Arroyo L, Saco Y et al (2012) Proteomic analysis reveals oxidative stress response as the main adaptative physiological mechanism in cows under different production systems. J Proteomics 75:4399–4411

Marco-Ramell A, Dyrlund TF, Bislev S et al (2013) Targeted proteomics as a tool for porcine acute phase proteins measurements. In: de Almeida A, Eckersall D, Bencurova E, Dolinska S, Mlynarcik P, Vincova M, Bhide M (eds) Farm animal proteomics. Wageningen Academic Publishers, Wageningen, pp 217–220

Marco-Ramell A, Miller I, Nöbauer K et al (2014) Proteomics on porcine haptoglobin and IgG/IgA show protein species distribution and glycosylation pattern to remain similar in PCV2-SD infection. J Proteomics 101:205–216

Marco-Ramell A, Hummel K, Razzazi-Fazeli E et al (2015) Concentration and pattern changes of porcine serum apolipoprotein A-I in four different infectious diseases. Electrophoresis 36:543–551

Marco-Ramell A, Arroyo L, Peña R et al (2016) Biochemical and proteomic analyses of the physiological response induced by individual housing in gilts provide new potential stress markers. BMC Vet Res 12:265

McGrath TF, van Meeuwen JA, Massart AC et al (2013) Effect-based proteomic detection of growth promoter abuse. Anal Bioanal Chem 405:1171–1179

Meier F, Seidel B, Geserick G et al (1980) Haptoglobintypisierung von Serumproben ausgewählter Säugetiere mittels Stärkegelelektrophorese. Mh Vet-Med 35:617–620

Miller I (2011) Protein Separation Strategies. In: Eckersall PD, Whitfield PD (eds) Methods in animal proteomics. Wiley-Blackwell, Oxford, pp 41–76

Miller I (2012) Application of 2D-DIGE in animal proteomics. Methods Mol Biol 854:373–396

Miller I, Wait R, Sipos W et al (2009) A proteomic reference map for pig serum proteins as a prerequisite for diagnostic applications. Res Vet Sci 86:362–367

Miller I, Gianazza E, Gemeiner M (2010) Any use in proteomics for low-tech approaches? Detecting fibrinogen chains of different animal species in two-dimensional electrophoresis patterns. J Chromatogr B 878:2314–2318

O'Reilly EL, Eckersall PD (2014) Acute phase proteins: a review of their function, behaviour and measurement in chickens. Worlds Poult Sci J 70:27–44

Oliveira BM, Coorssen JR, Martins-de-Souza D (2014) 2DE: the phoenix of proteomics. J Proteomics 104:140–150

Poon TCW (2007) Opportunities and limitations of SELDI-TOF-MS in biomedical research: practical advices. Exp Rev Proteomics 4:51–65

Puangpila C, Mayadunne E, El Rassi Z (2015) Liquid phase based separation systems for depletion, prefractionation, and enrichment of proteins in biological fluids and matrices for in-depth proteomics analysis—an update covering the period 2011–2014. Electrophoresis 36:238–252

Righetti PG, Boschetti E, Lomas L et al (2006) Protein equalizer technology: the quest for a "democratic proteome". Proteomics 6:3980–3992

Rødgaard T, Stagsted J, Christoffersen BØ et al (2013) Orosomucoid expression profiles in liver, adipose tissues and serum of lean and obese domestic pigs, Göttingen minipigs and Ossabaw minipigs. Vet Immunol Immunopathol 151:325–330

Rogowska-Wrzesinska A, Le Bihan M, Thaysen-Andersen M et al (2013) 2D gels still have a niche in proteomics. J Proteomics 88:4–13

Seth M, Lamont EA, Janagama HK et al (2009) Biomarker discovery in subclinical mycobacterial infections of cattle. PloS One 4:e5478

Skrzypczak WF, Ozgo M, Lepczynski A et al (2011) Defining the blood plasma protein repertoire of seven day old dairy calves - a preliminary study. J Physiol Pharmacol 62:313–319

Soares R, Franco C, Pires E et al (2012) Mass spectrometry and animal science: protein identification strategies and particularities of farm animal species. J Proteomics 75:4190–4206

Soler L, Molenaar A, Merola N et al (2013) Why working with porcine circulating serum amyloid A is a pig of a job. J Theor Biol 317:119–125

Soler L, Miller I, Hummel K et al (2016) Growth promotion in pigs by oxytetracycline coincides with down regulation of serum inflammatory parameters and of hibernation-associated protein HP-27. Electrophoresis 37:1277–1286

Stalder D, Haeberli A, Heller M (2008) Evaluation of reproducibility of protein identification results after multidimensional human serum protein separation. Proteomics 8:414–424

Sun J, Shi Z, Guo H et al (2011) Proteomic analysis of swine serum following highly virulent classical swine fever virus infection. Virol J 8:107

Sun D, Zhang H, Guo D et al (2013) Shotgun proteomic analysis of plasma from dairy cattle suffering from footrot: characterization of potential disease-associated factors. PLoS ONE 8: e55973

Talamo F, D'Ambrosio C, Arena S et al (2003) Proteins from bovine tissues and biological fluids: defining a reference electrophoresis map for liver, kidney, muscle, plasma and red blood cells. Proteomics 3:440–460

Tammen H, Schulte I, Hess R et al (2005) Peptidomic analysis of human blood specimens: comparison between plasma specimens and serum by differential peptide display. Proteomics 5:3414–3422

Thongboonkerd V, Chiangjong W, Mares J et al (2009) Altered plasma proteome during an early phase of peritonitis-induced sepsis. Clin Sci (Lond) 116:721–730

Timms JF, Cramer R (2008) Difference gel electrophoresis. Proteomics 8:4886–4897

Turk R, Piras C, Kovačić M et al (2012) Proteomics of inflammatory and oxidative stress response in cows with subclinical and clinical mastitis. J Proteomics 75:4412–4428

Vizcaino JA, Cote R, Reisinger F et al (2009) A guide to the proteomics identifications database proteomics data repository. Proteomics 9:4276–4283

Wait R, Miller I, Eberini I et al (2002) Strategies for proteomics with incompletely characterized genomes: the proteome of Bos taurus serum. Electrophoresis 23:3418–3427

Westbrook JA, Yan JX, Wait R et al (2001) Zooming - in on the proteome: very narrow - range immobilized pH gradients reveal more protein species and isoforms. Electrophoresis 22:2865–2871
Whiteaker JR, Zhang H, Eng JK et al (2007) Head-to-head comparison of serum fractionation techniques. J Proteome Res 6:828–836
Wu Y, Wang S, Peng X (2004) Serum acute phase response (APR)-related proteome of loach to trauma. Fish Shellfish Immunol 16:381–389
You Q, Verschoor CP, Pant SD et al (2012) Proteomic analysis of plasma from Holstein cows testing positive for mycobacterium avium subsp. Paratuberculosis (MAP). Vet Immunol Immunopathol 148:243–251
Zhong L, Taylor D, Begg DJ et al (2011) Biomarker discovery for ovine paratuberculosis (Johne's disease) by proteomic serum profiling. Comp Immunol Microbiol Infect Dis 34:315–326

Proteomic Research in Urine and Other Fluids

Gloria Isani, Enea Ferlizza, and Jarlath E. Nally

Abstract Urine is an ideal biological sample. It is an important source of clinical biomarkers of systemic and renal diseases and can be obtained repeatedly without causing stress. Different analytical protocols for the study of the urine proteome have been used, from the simplest electrophoresis on agarose gels to the most sensitive and complex 2DE. As with urine, a range of bodily fluids (saliva, cerebrospinal, bronchoalveolar, amniotic fluids) are amenable to the tools of proteomics. The identification and characterization of the protein content of such fluids are being used to provide novel insights into the evolutionary adaptations of farm and domestic animal species and to characterize their normal physiological state. A comprehensive review on the application of proteomics to these biological samples is presented.

Keywords Saliva • Cerebrospinal fluid • Bronchoalveolar fluid • Amniotic fluid • Electrophoresis • Veterinary medicine • Farm animals

Abbreviations

2DE	Two-dimensional electrophoresis
Ig	Immunoglobulins
MS	Mass spectrometry
AKI	Acute kidney injury
DIGE	Difference gel electrophoresis
BSE	Bovine spongiform encephalopathy

G. Isani (✉) • E. Ferlizza
Department of Veterinary Medical Sciences (DIMEVET), Alma Mater Studiorum-University of Bologna, Bologna, Italy
e-mail: gloria.isani@unibo.it

J.E. Nally
Bacterial Diseases of Livestock Research Unit, National Animal Disease Center, Agricultural Research Service (ARS), United States Department of Agriculture (USDA), Ames, IA, USA
e-mail: Jarlath.Nally@ars.usda.gov

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_7

AGE	Agarose gel electrophoresis
CAM	Cellulose acetate membrane
HRE	High-resolution electrophoresis
PrPC	Physiological prion protein
GO	Gene ontology
CSF	Cerebrospinal fluid
BALF	Bronchoalveolar fluid

1 Urine

Urine, produced by healthy kidneys under normal physiological conditions, contains a relatively small amount of protein. In general, molecules with an effective radius exceeding ~4 nm are not able to pass via the glomerular filter, and most proteins and protein fragments are reabsorbed along the tubule (Geckle 2007). A few proteins are secreted by the kidney, e.g., uromodulin. The abnormal presence of protein in urine is defined as *proteinuria*. Proteinuria provides insights into the systemic (presence of proteins derived from plasma) and local (proteins of kidney origin) pathophysiology from a clinical perspective. As a consequence, urine is considered an ideal source for biomarkers of clinical significance. Moreover, urine samples can be obtained repeatedly, in large quantities and noninvasively.

As the study of the urinary proteome originated and developed in human medicine, a short introduction summarizing the fundamental stages of urinary proteome characterization in humans is presented. Subsequently, different analytical protocols for the study of the urinary proteome will be discussed, from the simplest assays used during clinical assessments to the most sophisticated used only for research, with some reference to their use in veterinary medicine, and a possible standard protocol for the analysis of urine applicable to normal and proteinuric samples.

1.1 Urinary Proteomics: A Long-Standing Tradition

The first studies on the protein composition of normal urine were performed in 1979, applying two-dimensional electrophoresis (2DE) at a time when this proteomic technique was still in its infancy (Anderson et al. 1979). In 2004 Oh et al. (2004) defined a preliminary urine proteome map and annotated 113 different proteins on a 2D gel by peptide mass fingerprinting. Many other studies followed in subsequent years, and recently, a number of large-scale proteomic studies have been carried out to improve our knowledge on the urinary proteome from healthy individuals. By combining data from different studies published to date, at least 3000 different proteins and peptides have been identified in urine (Molina et al. 2011; Di Meo et al. 2017). Marimuthu et al. (2011) compared the proteins identified in their study with two other large-scale studies previously performed by Adachi

et al. (2006) and by Li et al. (2010). Six hundred and fifty-eight proteins were commonly identified in all three studies and thus represent a universal panel of proteins physiologically present in human urine. Out of these, albumin, uromodulin, heavy and light chains of immunoglobulins (Ig), and transferrin are the most represented. These proteins are generally classified following nephronal localization and functional criteria. Kidney secretory and structural proteins, serum and transport proteins, coagulation and complement factors, immunoglobulins, and other immune proteins, enzymes, metal-binding proteins, and lipoproteins are some of the possible functional categories reported by Candiano et al. (2010). The most abundant proteins in human healthy urine are reported in Table 1.

Wide variability of the urine proteome from healthy individuals has been related to gender, age, diet, physical exercise, drug consumption, and environmental conditions. The extent and sources of this intra- and interindividual variability have been investigated by several studies (Oh et al. 2004; Thongboonkerd et al. 2006; Nagaraj and Mann 2011; Molina et al. 2011; Liu et al. 2012; He et al. 2012; Guo et al. 2015; Pastushkova et al. 2016).

In addition, the urinary proteome is characterized by a highly dynamic range of almost five orders of magnitude (Nagaraj and Mann 2011). Therefore, the most challenging proteins to identify are low abundance proteins, and many of them are yet to be characterized. One innovative technique that has been developed to overcome this limitation is the use of beads coated with hexameric peptide ligand libraries; this allows for the enrichment of low abundant proteins while removing the most abundant ones (Decramer et al. 2008; Candiano et al. 2010; Filip et al. 2015).

Many of the proteins identified in large-scale proteomic analysis on normal human urine are now under investigation in clinical trials to validate their use as sensitive and specific biomarkers of disease. Chronic kidney disease, acute kidney injury (AKI), and diabetic nephropathy are only three of the diseases studied extensively in human medicine using proteomic techniques on urine (Andersen et al. 2010; Lhotta 2010; Devarajan 2011). Such diseases are also highly prevalent in animal medicine, which could benefit from the discovery of novel biomarkers of health.

1.2 Urinary Proteomics in Veterinary Medicine: An Unexplored but Promising Field

Despite the importance of proteomics techniques and their potential for biomarker discovery and diagnostics, high costs and limited interest by animal researchers and clinicians have resulted in few applications of proteomics to urine samples in veterinary medicine, and, as a consequence, fragmentary data are present in the literature. Pyo et al. (2003) studied the urinary proteome in cows by 2DE to search for potential biomarkers of early pregnancy. Simon et al. (2008) applied difference

Table 1 Most abundant proteins found in urine from healthy humans (Molina et al. 2011; He et al. 2012)

He et al. (2012)		Molina et al. (2011)	
A1BG	Alpha-1B-glycoprotein	A1BG	Alpha-1B-glycoprotein
AHSG	Alpha-2-HS-glycoprotein	ACP2	Lysosomal acid phosphatase
AMBP	α1-Microglobulin/bikunin precursor	ALB	Serum albumin
AZGP1	Zinc alpha-2-glycoprotein 1	AMBP	α1-Microglobulin/bikunin precursor
CD14	Monocyte differentiation antigen CD14	AZGP1	Zinc-alpha-2-glycoprotein
COL6A1	Collagen alpha-1(VI) chain	CADM4	Cell adhesion molecule 4
CTSD	Cathepsin D	CD14	Monocyte differentiation antigen CD14
DNASE1	Deoxyribonuclease-1	CDH1	Epithelial cadherin
GAA	Lysosomal alpha-glucosidase	GAA	Lysosomal alpha-glucosidase
GC	Vitamin D-binding protein	GSN	Gelsolin
HPX	Hemopexin	HSPG2	Basement membrane-heparan sulfate proteoglycan core protein
KNG1	Kininogen 1	KLK3	Prostate-specific antigen
ORM1	Orosomucoid 1	KNG1	Kininogen-1
PIGR	Polymeric immunoglobulin receptor	MASP2	Mannan-binding lectin serine protease 2
PTGDS	Prostaglandin-H2 D-isomerase	PEBP4	Phosphatidylethanolamine-binding protein
RBP4	Plasma retinol-binding protein	PRSS3	Trypsinogen 4
SERPING1	Plasma protease C1 inhibitor	PTGDS	Prostaglandin -H2 D-isomerase
UMOD	Uromodulin	QPCT	Glutaminyl-peptide cyclotransferase
		SDR9C7	Orphan short-chain dehydrogenase/reductase
		SECTM1	Secreted and transmembrane protein 1
		SERPINA1	Alpha-1-antitrypsin
		TF	Serotransferrin
		UMOD	Uromodulin

gel electrophoresis (DIGE) to discover biomarkers of bovine spongiform encephalopathy (BSE) in bovine urine. Nabity et al. (2011) applied DIGE and SELDI-TOF to urine samples from dogs affected by hereditary nephropathy.

More recent studies have applied proteomic techniques to explore the urinary proteome in healthy dogs, cows, and cats (Brandt et al. 2014; Bathla et al. 2015; Ferlizza et al. 2015), while the urinary proteome of pigs and other farm animals remains mostly unexplored. These preliminary studies are a starting point for future

research that can focus on the proteome of urine under physiological conditions compared to that during specific clinical pathologies.

1.3 Methods and Protocols for Sample Preparation

The prerequisite for each analytical technique to obtain reliable results is the appropriate treatment of samples in order to preserve their integrity and, at the same time, avoid/remove any interference. Therefore, one of the priorities in the field of urinary proteomics is to optimize the protocol for sample preparation, due to the presence of interfering molecules and ions, of a wide dynamic range and of the differences in protein concentration between healthy/non-proteinuric and diseased/proteinuric animals.

1.3.1 Sample Collection and Stability

The discussion on sample collection procedures (free catch, cystocentesis, catheterization) is out of the scope of the present chapter, but it is important to bear in mind that they can modify the urine protein composition, particularly in animals, because it is not possible to collect 24-h urine as routinely done in humans. Although in humans it has been reported that the variation of intraday samples is limited (Khan and Packer 2006; Thongboonkerd et al. 2006), where possible, we suggest collecting a morning urine sample by free catch in order to limit variables and stress for animals.

The literature reports that urine is stable for hours at room temperature, a few days in the fridge (+4 °C), and for years in the freezer (−20 °C) (Candiano et al. 2010). Molina et al. (2011) found differences in the number of protein spots in urine samples frozen at −20 °C (without protease inhibitors) compared to fresh urine samples and to samples stored at −80 °C; therefore, we recommend storing samples at −80 °C. The use of protease inhibitors is not recommended because they may change the urinary proteome, in particular in healthy/non-proteinuric samples which contain few proteases.

1.3.2 Sample Treatment and Protein Quantification

Different protocols have been proposed to concentrate and desalt urine from healthy individuals before 2DE or SDS-PAGE, including ultrafiltration, precipitation, lyophilization, and ultracentrifugation (Thongboonkerd et al. 2006; Martin-Lorenzo et al. 2014). We recommend precipitation with trichloroacetic acid and washing the pellet with ethanol or 80% acetone as the method of choice for 2D urine proteome analysis, because this protocol provides high protein recovery yield and a large number of protein spots that can be visualized in 2D gels. An interesting alternative can be offered by centrifugal filtration with 3 kDa cutoff filters; it is less time-consuming and

allows the elimination of interfering ions and metabolites and is easy to perform without chemical treatment. Different methods are available for protein quantification, from automated analysis used in the clinical laboratory to protocols applied mainly in research. In the clinical routine, urine proteins are determined by a spectrophotometric automated method based on pyrogallol red, while other methods are the preferred choice in the research field (e.g., Bradford or Lowry).

1.4 Gel-Based Proteomics: From Zone Electrophoresis to DIGE

Different electrophoretic approaches are available, characterized by increasing analytical sensitivity and protein separation performances from semiautomated agarose gel electrophoresis to 2D-DIGE. All these techniques have advantages and disadvantages when applied to urine samples obtained from farm animals.

1.4.1 Agarose Gel Electrophoresis

Agarose gel electrophoresis (AGE) and, less frequently, cellulose acetate membrane (CAM) electrophoresis are used to separate serum proteins in routine clinical analysis, but they are rarely applied to urine. Zone electrophoresis on CAM or AGE allows the separation of proteins in native conditions on the basis of the mass/charge ratio yielding the identification of four to six zones (albumin, alpha 1, alpha 2, beta 1, beta 2, and gamma globulins) (Eckersall 2008). CAM is currently seldom applied to urine samples because separated zones are relatively wide, while the resolution and limit of detection are low; as a consequence, in many laboratories, it has been replaced by AGE. AGE, generally performed on semiautomated systems, is characterized by good reproducibility for both qualitative and quantitative analysis and can be considered to be the easiest, cheapest, and fastest approach to separate, visualize, and quantify urine albumin and globulins. Ferlizza et al. (2017) applied AGE on high-resolution gels (HRE) to urine samples collected from healthy and diseased cats. The same analytical approach was previously used by Giori et al. (2011) for canine urine. These results suggest the use of HRE-AGE in clinical practice to obtain additional information useful for the evaluation of kidney function. Nevertheless, this technique presents a few disadvantages; the first relates to the low sensitivity of staining based on acid violet that limits the analysis of urine samples with protein concentrations lower than 40 mg/dL (Fig. 1). The second limitation relates to the difficulty in defining the limit of detection and thus quantification. Despite these disadvantages, AGE can be considered to be a useful diagnostic tool offering additional clinical information to urinalysis.

A method complementary to AGE, which can provide additional information to improve characterization of the urinary proteome, is SDS-AGE, also performed by semiautomated systems. This method separates proteins according to their

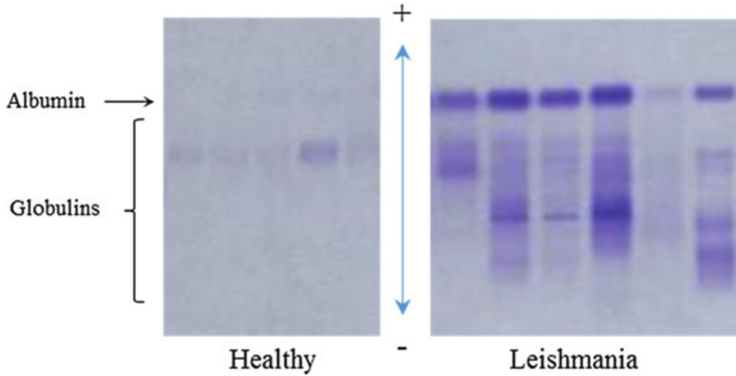

Fig. 1 Examples of urine samples analyzed by HRE-AGE. Samples were collected from healthy dogs and from dogs affected by leishmaniasis

molecular mass and allows the visualization of different protein profiles related to their nephronal origin: glomerular, tubular, or both. The glomerular pattern is characterized by high molecular mass bands (>70 kDa) that contain proteins derived from plasma due to damage of the glomerular filtration barrier. On the other hand, the tubular pattern is characterized by low molecular mass bands (<70 kDa) that contain proteins derived from impaired tubular reabsorption; a mixed pattern is characterized by the presence of both high and low molecular mass proteins. SDS-AGE has been applied by different authors to evaluate kidney function in dogs and cats affected by several pathologic conditions (Bonfanti et al. 2004; Abate et al. 2005; Schellenberg et al. 2008; Gerber et al. 2009; Giori et al. 2011; Paltrinieri et al. 2015; Lavoué et al. 2015). The predominant limitations of analysis are time-consuming protocols, the inability to cut bands for the subsequent identification by mass spectrometry (MS), and the relatively high cost of the analysis.

1.4.2 1D SDS-PAGE

The classical approach for protein separation is based on 1D SDS-PAGE; in this case, proteins are separated on the basis of their molecular mass under denaturing conditions. This method provides for veterinary medicine the best quality/price ratio, with a good resolution for protein separation, high reproducibility and high or very high sensitivity, depending on the staining used (Coomassie, silver nitrate or fluorescent), and relatively low cost (Fig. 2, Isani et al. unpublished results). This technique was used to investigate urine protein composition particularly in companion animals (Raila et al. 2007; Miyazaki et al. 2007; Schaefer et al. 2011; Lemberger et al. 2011; Brandt et al. 2014; Ferlizza et al. 2015) and to a lesser extent in farm animals and nonconventional species (Halbmayr and Schusser 2002; McLean et al. 2007; Alhaider et al. 2012). The higher sensitivity of the staining

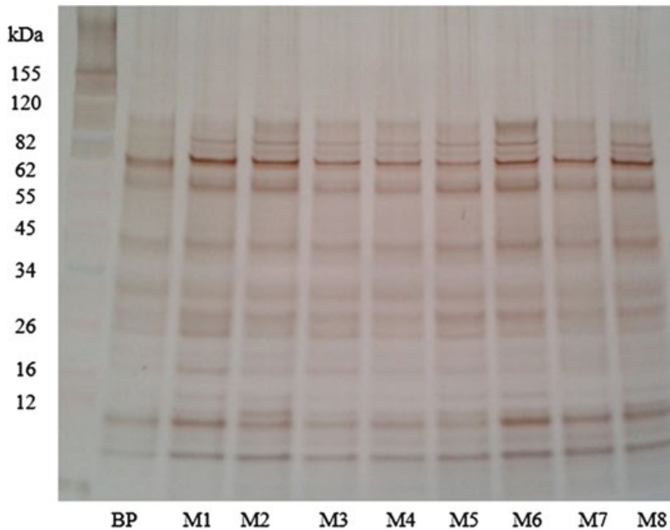

Fig. 2 Example of urine samples collected from cows analyzed by 1D SDS-PAGE. Urine samples collected from one cow before (BP) and during pregnancy (months [M] 1-2-3-4-5-6-7-8). Two μg of proteins were loaded, separated on 12% gel in MOPS buffer, and stained with silver staining

protocols allows relatively easy characterization of a healthy urine profile. SDS-PAGE allows for the visualization of different patterns (tubular, glomerular, or mixed) to identify the nephronal origin of proteins in urine. Moreover, since this technique can be performed also in nonreducing conditions, it is possible to separate proteins according to their original size allowing a better visualization of the native proteome, and nonreducing SDS-PAGE can be performed to better discriminate the different kinds of gammopathies. Furthermore, SDS-PAGE can be performed as a preparative step for subsequent protein identification by MS allowing an accurate characterization of the protein profile and the respective identification of putative biomarkers of nephropathy.

The main disadvantage of SDS-PAGE is that the electrophoretic profile of proteinuric samples can contain many protein bands, indicating a proteome too complex to allow complete separation and characterization.

1.4.3 2DE in Combination with MS, Including DIGE

The best method to separate and visualize the complexity of the urine proteome by means of gel-based proteomics is 2DE. However, in veterinary medicine, its application to urine samples is still limited and restricted to a few species (Pyo et al. 2003; Miller et al. 2004, 2014; Bathla et al. 2015; Ferlizza et al. 2015). Low reproducibility is one of the major disadvantages affecting 2DE, but this has been

significantly improved in recent years. 2D-DIGE overcomes this limitation but has rarely been applied in veterinary medicine research. 2D-DIGE has been used to identify markers of BSE in bovine urine samples (Simon et al. 2008; Plews et al. 2011). In 2011, Nabity et al. applied 2D-DIGE on urine samples of dogs affected by X-linked hereditary nephropathy identifying putative markers of nephropathy.

Figure 3 presents a simplified scheme of a standard protocol used in our laboratory (Isani and Ferlizza) for urine analysis from sample collection to protein identification by mass spectrometry.

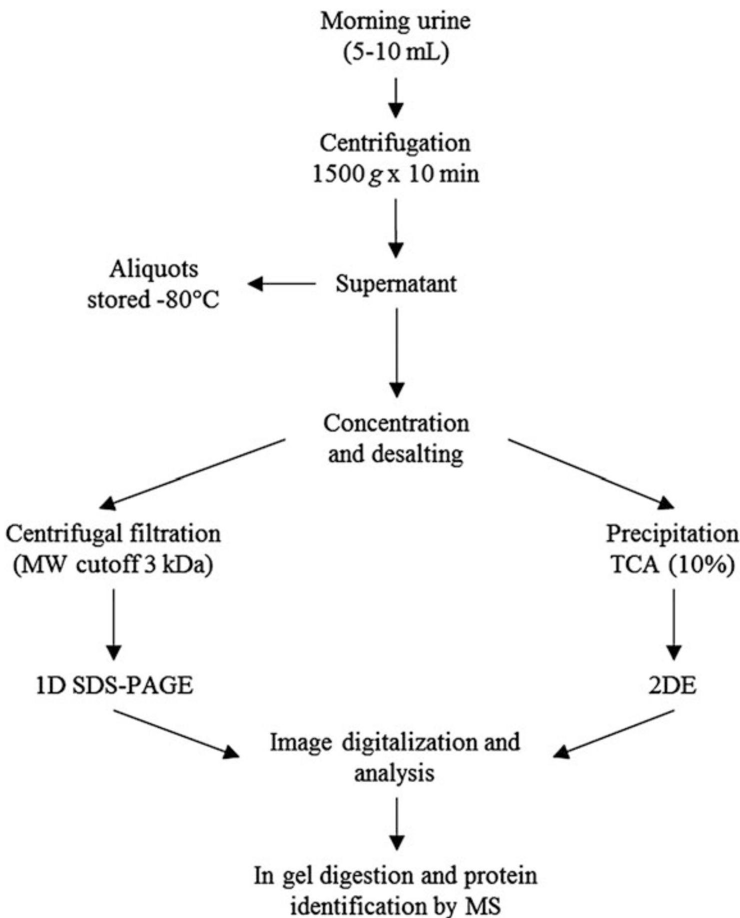

Fig. 3 Standard protocol for urine proteome separation. 1D SDS-PAGE: optimal protein amount of 5–10 μg for minigel (8 × 8 cm) and Coomassie staining. Depending on urine protein concentration, centrifugal filtration (3 kDa cutoff) in case of healthy urine, or dilution with ultrapure water for proteinuric samples. 2DE: optimal protein amount of 250–300 μg for 17 cm strip and Coomassie staining. After 10%TCA precipitation, wash pellet with ethanol or 80% acetone, air-dry, and store at −20 °C

1.5 Urine Proteome and Clinical Application in Farm Animals

Despite the economic importance and the worldwide distribution of farm animals, limited data is available on the characterization of the urinary proteome of healthy animals. The authors are not aware of a complete map of the urinary proteome for any species of veterinary interest, including farm animals (cow, sheep, goat, horse). On the other hand, proteomic techniques have been applied to the characterization of other biological samples from farm animals, e.g., milk and meat, as recently reviewed (Almeida et al. 2015).

1.5.1 The Urinary Proteome of Bovine and Other Ruminants

Under normal physiological conditions, ruminant urine contains a low amount of proteins and has a basic pH that can interfere not only with the semiquantitative dipstick test used in a clinical setting but also with other analytical methods. As a consequence, attention should be paid when determining total proteins in urine of ruminants, and the use of pyrogallol red is suggested. In healthy dairy cows and heifers, mean values of 13.4 mg/dL and 19.4 mg/dL total urine proteins were, respectively, determined by the authors.

Typical 1DE profiles of urines from Friesian heifers and cows (*Bos taurus*) are reported in Fig. 4 (Isani et al., unpublished results). There are common bands with apparent molecular mass of 97, 86, 78, 70, 59, 38, 27, and $<$12 kDa, the most abundant of which probably corresponds to albumin (70 kDa).

The first characterization of bovine urine was reported by Pyo et al. (2003) applying 2DE on urine samples collected from 30 pregnant and 20 nonpregnant cows. However, of the over 200 spots detected and selected for MS identification, less than 5% could be identified due to the incomplete database annotation for the bovine species, whose genome was not sequenced at that time. Nevertheless, one of the protein spots was identified as bovine pregnancy-associated protein which showed a high correlation to pregnancy and was considered by the authors as a useful diagnostic biomarker of pregnancy in cows.

In 2015, the urine from Karan Fries cows was analyzed by 2DE leading to the identification of more than 1500 proteins (Bathla et al. 2015). These proteins were categorized by gene ontology (GO) classification. On the basis of cellular localization, identified proteins were mainly cytoplasmic (29%), extracellular (20%), or belonged to an organelle fraction (20%). Classification by molecular function indicated that a large majority of these proteins were involved in catalytic activity (32%) and binding (30%); classification by biological process categorization indicated that they were mainly involved in metabolism (25%) and cellular processes (20%). The study of Bathla et al. (2015) was the first high-throughput approach to characterize the bovine urinary proteome which provided a database and considered to be the starting point for future studies focused on biomarker discovery in cows.

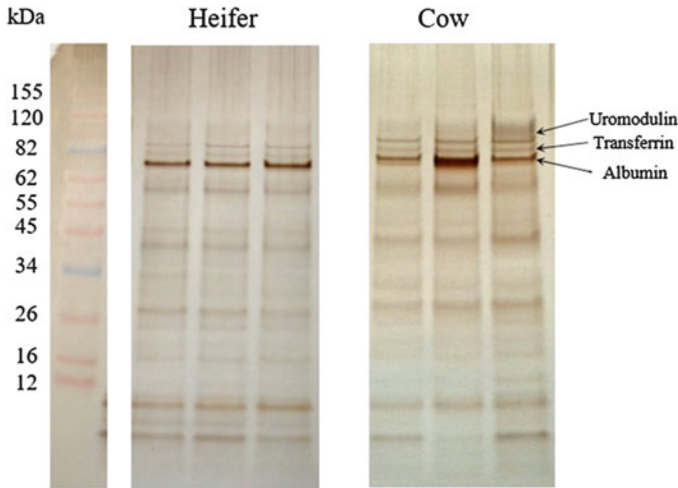

Fig. 4 Examples of urine samples collected from healthy Friesian heifers and cows analyzed by 1D SDS-PAGE. Two μg of proteins were loaded, separated on 4–12% gel in MES buffer, and stained with silver staining. Putative uromodulin, transferrin, and albumin bands, identified on the basis of apparent molecular mass, are indicated

Other studies have exploited and applied the potential of proteomics on urine samples aimed at discovering alternative biomarkers of BSE applicable to the development of diagnostic tests. In 2008, 2D-DIGE lead to the identification of a cluster of 16 spots containing 5 proteins, namely, clusterin, Ig gamma-2 chain C region, similar to uroguanylin, cystatin E/M, and cathelicidin 1, with differential abundance in urine of control versus BSE-infected cows (Simon et al. 2008). These proteins were considered specific for monitoring disease progression, and the authors suggested that they may also provide new information on the biochemical basis of BSE infection. Subsequently, the influence of sex, breed, and age was investigated. Results demonstrated that, at the clinical stage of the disease, these factors significantly affected the urinary proteome, through a specific panel of proteins which showed high accuracy to discriminate between control and infected cattle (Plews et al. 2011). In addition to BSE, ketosis, an important metabolic disorder affecting dairy cows during the parturition/first days of lactation periods, is a big issue in farm animal medicine that still needs early predictors. The study of Xu et al. (2015) was the first to report SELDI-TOF-MS application to the proteome of urine in large animals, leading to the identification of 11 proteins underrepresented in urine of cows affected by ketosis. Fibrinogen, C1 inhibitor, and osteopontin were associated with the inflammatory response, VGF nerve growth factor inducible and amyloid β were associated with the neuronal function, while serum amyloid A and apolipoprotein CIII were associated with lipid metabolism; the remaining four were cystatin C, transthyretin, hepcidin, and human neutrophil peptides.

Few studies have applied proteomics on urine samples from goats (*Capra hircus*) and sheep (*Ovis aries*). Ozgo et al. (2009) studied the urine proteins in young goats by SDS-PAGE to evaluate the maturation of the glomerular filtration barrier. Neonatal proteinuria and its changes during the first month were related to the presence of low molecular mass proteins (<67 kDa). The authors did not identify these proteins, but hypothesized that proteinuria was due to an adaptive mechanism for the removal of excess protein taken up with colostrum as also reported in foals (Jeffcott and Jeffcott 1974).

In 2012, Palviainen et al. (2012a) applied 2DE to urine from 12 sheep to search for novel biomarkers of acute kidney injury (AKI). AKI was induced in sheep by the administration of ketoprofen via catheterization, and urine samples were collected and analyzed by 2DE. Calbindin-D28k, retinol-binding protein, and antigen-presenting glycoprotein CD1d were detected only in the urine of the AKI group. The presence of these proteins in urine confirmed the tubular involvement during AKI in sheep, suggesting that calbindin-D28k and antigen-presenting glycoprotein CD1d proteins may represent sensitive and specific biomarkers of kidney injury. Sheep, similar to cows and humans, is a target of transmissible spongiform encephalopathies, which in this species is known as scrapie. Andrievskaia et al. (2008) applied SDS-PAGE and Western blotting on urine samples from healthy and naturally scrapie-infected sheep to detect and quantify the physiological prion protein (PrP^C). PrP^C was identified in very low levels in urine of healthy control animals (0.04 ng/mL), while 21% of scrapie-infected sheep presented elevated concentrations (0.6–4.7 ng/mL).

1.5.2 The Urinary Proteome of Other Farm Animals

Though the pig (*Sus scrofa*) is an economically important species and a consolidated model for translational medicine, data on the characterization of its urinary proteome is lacking. Our preliminary data on pig urine proteome is reported in Fig. 5 (Forni and Isani, unpublished results). Noteworthy is the low abundance of high molecular mass proteins and the presence of putative albumin at 67 kDa. The putative uromodulin at 90–100 kDa is also present in urine from healthy specimens along with many different low molecular mass proteins.

Studies on the characterization of the urinary proteome from horses (*Equus caballus*) are scarce. SDS-PAGE was applied for qualitative and quantitative evaluation of urine proteins in healthy and diseased horses (Halbmayr and Schusser 2002) and before and after exercise (Scarpa et al. 2007). Data were quite similar to those obtained by the authors, who analyzed urine from healthy and diseased horses (Isani et al., unpublished results) (Fig. 6). Noticeable is the presence of an abundant band at 90–100 kDa attributable to uromodulin on the basis of molecular mass, followed by putative transferrin and albumin. The presence of different thin bands between 73 and 61 kDa is intriguing; they might be albumin isoforms due to posttranslational modifications as reported in human urine (Barratt and Topham 2007). The bands at 58 and 27 kDa might be assigned to the heavy and light chains

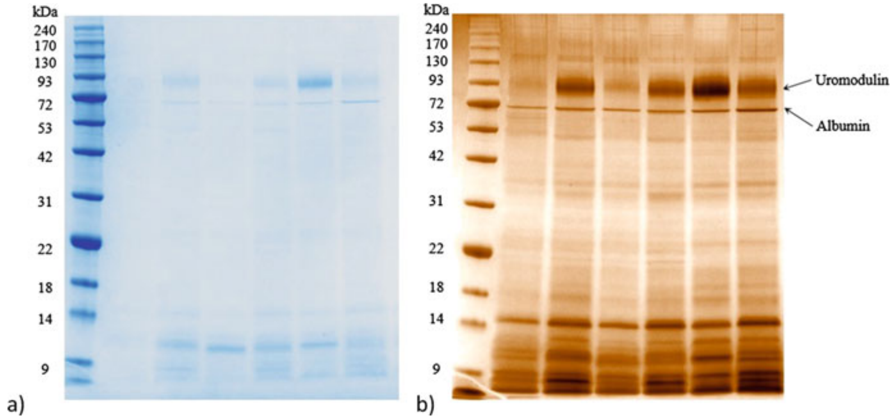

Fig. 5 SDS-PAGE of urine samples collected from piglets (3 months old). Three μg of proteins were loaded on a 12% gel in MOPS buffer and stained with (a) Coomassie brilliant blue staining or (b) silver staining. Putative uromodulin and albumin bands, identified on the basis of apparent molecular mass, are indicated

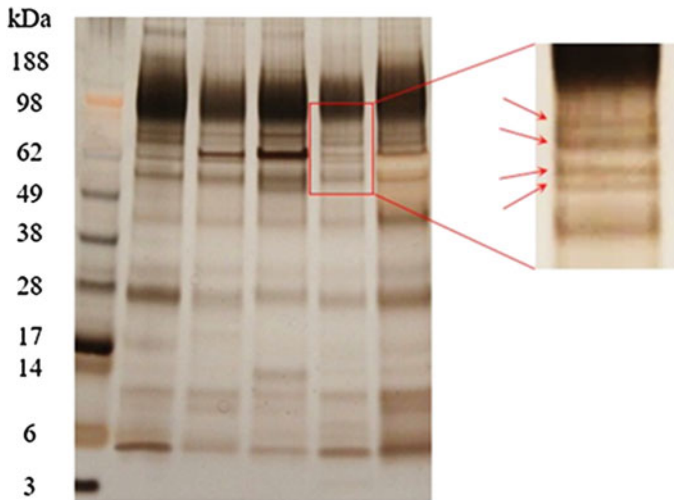

Fig. 6 SDS-PAGE of urine samples collected from healthy horses. Two μg of proteins were loaded on 4–12% gel in MES buffer and stained with silver staining. The magnification highlights the thin bands between 73 and 61 kDa

of IgG, respectively. In specimens affected by renal, gastrointestinal, and orthopedic diseases, the concentration of putative uromodulin declines, similar to observations reported also by Halbmayr and Schusser (2002). Also in horses, the decrease of uromodulin in urine seems to be a common sign of renal impairment, as reported in humans and dogs (Lhotta 2010; Raila et al. 2014). Other studies applied acetate and agarose gel electrophoresis to equine urine as a diagnostic tool

for the evaluation of peripartum stress and multiple myeloma. Jeffcott and Jeffcott (1974) analyzed proteinuria in foals during the first 24 h of life reporting the presence of low molecular mass proteins and the absence of gamma globulins. Geelen et al. (1997) evidenced an increase of IgG, whereas Pusterla et al. (2004) reported that the electrophoretic pattern presented peaks in the albumin and gamma globulin zones and that the gamma globulin fraction consisted mainly of IgA.

1.5.3 The Urinary Proteome of Companion Animals

Companion animals, though less investigated than humans, can be considered good models for future application of proteomic techniques to urine of farm animals. In recent years, research has shed light on the urinary proteome of a man's best friend. Urine of healthy dogs (*Canis familiaris*) is characterized by the presence of a few abundant proteins, namely, albumin, uromodulin, and immunoglobulin light chains, and in entire male adult dogs, of the prostate-specific proteins (Miller et al. 2014). Differently from other animal species, sex-specific proteins, e.g., arginine esterase similar to human prostate-specific antigen, represent most of total urine proteins in adult entire males, in contrast to females and castrated males. Similarly, in human urine, a total of 20 protein spots with more than a twofold change were identified and related to the prostate origin in males (Guo et al. 2015). The most complete characterization of the normal urinary proteome in dogs was recently performed by Brandt et al. (2014), who investigated exosomal and soluble proteins from dog urine by 1D SDS-PAGE combined with LC-MS/MS. Authors identified 391 and 214 proteins in exosomal pellets and urine supernatants, respectively. Interestingly, 52% of the exosomal fraction and 62% of soluble proteins identified in this study have been already identified in the human urinary proteome, and some of these proteins, e.g., uromodulin and retinol-binding protein, are already being used as urinary biomarkers for renal and extrarenal diseases in humans, pointing out the importance of the dog as an animal model for translational medicine.

Different studies have also applied proteomic techniques to dog urine to identify novel biomarkers of disease. In particular, the urinary proteome was investigated in dogs affected by nephropathy (Miller et al. 2004, 2014; Forterre et al. 2004; Nabity et al. 2011), different types of urinary stones (Forterre et al. 2006), transitional cell carcinoma (Bracha et al. 2014), envenomation by the bite of *Vipera berus berus* (Palviainen et al. 2012b), or infected by *Dirofilaria immitis* (Hormaeche et al. 2014) and by *Leishmania infantum* (Zaragoza et al. 2003; Ferlizza et al. 2013).

On the other hand, urine proteome in the domestic cat is still to be completely characterized. Only in recent years, two studies have analyzed the urinary proteome of healthy cats by 1D SDS-PAGE and 2DE (McLean et al. 2007; Ferlizza et al. 2015), yielding a preliminary proteomic map. The most abundant proteins are albumin, transferrin, uromodulin, and, in entire males, cauxin. Moreover, other low molecular mass proteins (<70 kDa) were identified as haptoglobin and Ig light chains and high molecular mass proteins (>70 kDa) as alpha-2-macroglobulin. Finally, few studies applied proteomic techniques to search for novel biomarkers in

urine of cats affected by chronic kidney disease (Ferlizza et al. 2015), tubular nephropathy (Miyazaki et al. 2007), urinary tract infection, idiopathic cystitis and urolithiasis (Lemberger et al. 2011), and azotemia (Jepson et al. 2013).

2 Proteomics of Saliva and Other Bodily Fluids in Farm Animals

As with urine, a range of bodily fluids are amenable to proteomics tools. The identification and characterization of the protein content of such fluids are being used to provide novel insights into the evolutionary adaptations of farm and domestic animal species and to characterize their normal physiological state. Results can be used to understand animal physiology, improve livestock productivity and welfare, and provide biomarkers of stress and animal disease.

2.1 Proteomics of Saliva

The identification and characterization of the proteome of saliva has been pioneered in human samples, to define pathophysiological-related changes relevant to disease pathogenesis and diagnostics. Such studies provide insights into the proteome and diagnostic potential of saliva and identify it as an amenable sample readily available from farm and domestic animals for detailed proteomic analysis (Mavromati et al. 2014). However, it is clear that there is significant variation in the protein content of saliva from different mammalian species, which likely reflects distinct evolutionary and environmental lineages (de Sousa-Pereira et al. 2013).

Gland-specific saliva can be collected directly from individual salivary glands which include the parotid, submandibular, sublingual, and minor salivary glands. Their respective characterizations allow for insights into gland-specific pathology. Whole (mixed) saliva comprises secretions from each of these salivary glands in addition to constituents of non-salivary origin readily detected in the oral cavity; these include gingival crevicular fluid, expectorated bronchial and nasal secretions, serum and blood derivatives from oral wounds, bacteria and their respective products, viruses, fungi, cellular components, and food debris. Collectively, the components of whole saliva comprise a bodily fluid essential for the health of the oral cavity and which functions to facilitate maintenance of homeostasis, food digestion, and protection from pathogens. Saliva comprises a complex hypotonic aqueous solution composed of more than 99% water that contains a range of proteins, peptides, enzymes, hormones, sugars, lipids, electrolytes, and a plethora of other compounds. Saliva is a relatively easy sample to collect; collection is noninvasive and stress-free. It can be performed multiple times on the same animal and in a relatively short time. Collectively, these factors make it an ideal candidate to screen for the identification of protein/peptide biomarkers of health status.

A cross-species comparison of the proteome of saliva was performed on seven mammalian species classified into four major evolutionary groups: ungulates (which include the order Artiodactyla, sheep and cattle, and the order Perissodactyla, horse), glires (which include rodents and lagomorphs: rat and rabbit), carnivores (dog), and primates (human) (de Sousa-Pereira et al. 2015). Distinct protein profiles were observed for each group, and only seven protein families were common to all groups; common protein families included carbonic anhydrase, albumin, polymeric immunoglobulin receptor, prolactin-inducible protein, lactoperoxidase, glutathione-S-transferase P, and keratin. Collectively, these protein families have antibacterial activity, buffer capacity, and lubrication activity. This comparative approach is defining a proteomic signature characteristic of each species and their respective environmental adaptation. For example, amylase, which plays a key role during digestion since it catalyzes the hydrolysis of glycosidic bonds from diet polysaccharides such as starch, was identified in the saliva of rodents, lagomorphs, carnivores, and ungulates. In contrast, latherin was identified only in horse saliva; this is a protein family that helps mastication of large quantities of dry food and reflects an adaptive mechanism to environmental constraints for horses. Goats have a higher tolerance than sheep to plant allelochemicals in their diet which has been hypothesized to correlate with the existence of tannin-binding proteins identified in goat, but not sheep, saliva (Austin et al. 1989; Narjisse et al. 1995).

Although saliva is amenable to the tools of proteomics, several studies have highlighted the importance of standardizing the collection and handling of saliva prior to analysis. Inherent variability and experimental artifacts during sample preparation and analysis have led to the proposal and development of standardized methodologies to collect, process, and extract samples, e.g., centrifugation for removal of food debris (Vitorino et al. 2012a, b; de Sousa-Pereira et al. 2015). Consideration must also be given to the fact that the secretion of saliva into the oral cavity results in the exposure of saliva to a large number of proteolytic activities that can be of endogenous and exogenous origin and that can rapidly process and modify salivary proteins. Protease inhibitors can be advantageous, but their use can also increase sample complexity and interfere with mass spectrometry. These variables are further compounded in animal studies compared to human studies when one must also consider the variation of animal diets, the mode of digestion, and the logistics of sample collection. Such considerations include inter- and intraindividual variability (Jehmlich et al. 2013), response to age and different environmental conditions, variability due to genetic polymorphisms, and the collection of stimulated versus unstimulated samples. Saliva composition is mainly under autonomic control and changes rapidly in response to a wide range of different chemical and mechanical stimuli (D'Amato et al. 1988; Edwards and Titchen 2002, 2003; Scocco et al. 2011). All samples of saliva should be processed immediately upon collection. Analysis of porcine saliva showed that its protein composition changes during storage at −20 °C, so preservation of samples at −80 °C is recommended (Messana et al. 2008; Gutiérrez et al. 2011a).

Porcine saliva has long been used as an analytical sample for specific pathogen disease diagnosis, either by direct detection of pathogen or indirectly via the

detection of specific antibodies (Gutiérrez et al. 2014). Multiple analytes have been detected and measured in porcine saliva and include haptoglobin (Hiss et al. 2003), cortisol (Ruis et al. 1997), estrone sulfate (Ohtaki et al. 1997), progesterone (Moriyoshi et al. 1996), and immunoglobulin (Van der Stede et al. 2001; Devillers et al. 2004). With the advent of increasingly sensitive instrumentation, C-reactive protein (CRP) was quantified in samples from healthy compared to diseased pigs as an indicator of pig health and welfare monitoring. The limit of detection for salivary CRP was 0.47 ng/mL, and salivary concentrations were ~322-fold lower than that of serum. Saliva samples could be stored at 20 °C for several days without compromising test results, and levels of CRP could differentiate healthy pigs from those with a variety of inflammatory conditions (Gutiérrez et al. 2009). Subsequent studies determined that while amounts of CRP were not influenced by circadian pattern, the amounts of other acute phase proteins such as haptoglobin differed according to early morning versus late afternoon sample collection; this highlights another variable (circadian rhythms) to consider for sample analysis as well as the need to report components of saliva as ranges rather than absolute values (Gutiérrez et al. 2013a). Initial studies using one- and two-dimensional electrophoresis and immunoblotting confirmed the complex dynamic nature of the porcine salivary proteome and provided 2D gel proteomic maps of salivary proteins including immunoglobulins (Gutiérrez et al. 2011b). Multiple protein isoforms are attributed to fragmentation, bacterial degradation, endogenous truncation, glycosylation, and phosphorylation. Thirteen proteins were identified by mass spectrometry (Gutiérrez et al. 2011b). Improved databases and comparative analyses increased this to 20 unique proteins and provided normalization strategies to identify a panel of biomarkers that could discriminate clinically infected pigs compared to their healthy counterparts (Gutiérrez et al. 2013b). Porcine saliva and serum share several proteins highlighting the potential for these to act as markers of general health. The most abundant porcine salivary proteins include lipocalins, in contrast to amylase which comprises up to 25% of the human salivary proteome. Porcine saliva also contains proline-rich proteins, but note that their detection on gels is stain dependent (Miller et al. 2006). Enrichment for glycoproteins using boronic acid allowed for the detection of differential expression of carbonic anhydrase VI, α-1-antichymotrypsin, and haptoglobin by 2D gel electrophoresis of saliva derived from pigs suffering rectal prolapse; interestingly, glycopatterns of haptoglobin in saliva differed to that of haptoglobin derived from serum (Gutiérrez et al. 2016).

Cows secrete on the order of 150 liters of saliva per day (Bailey 1961). A comprehensive proteome analysis of bovine saliva identified 402 salivary proteins and 45 N-linked glycoproteins using three different strategies: nontargeted, targeted, and glycocapture (Ang et al. 2011). In the nontargeted approach, four prefractionation methodologies were performed based on different physiochemical properties and included SDS-PAGE (electrophoretic mobility), off-gel fractionation (protein isoelectric point), reverse-phase HPLC (protein hydrophobicity), and strong cation exchange HPLC (peptide charge). This nontargeted approach facilitated the identification of 396 proteins by MS, only 10% (42) of which were identified in all approaches, highlighting the need for multiple complementary strategies for comprehensive

analyses. The targeted approach used publicly available information to create a list of peptides for identification by mass accuracy and resolution on the mass spectrometer; it is limited by the magnitude of selected databases used to generate a user list. The glycocapture approach utilized hydrazide coupling methodology to enrich for glycosylated peptides prior to MS and provides the complete N-glycoprotein profile of bovine saliva. The most abundant proteins in bovine saliva are similar to that of human with the notable absence of amylase and the presence of short palate, lung and nasal epithelium carcinoma-associated protein 2A (BSP30), and odorant-binding protein (Ang et al. 2011). In buffalo, protein content of whole saliva has been monitored through all stages of the estrus cycle (Muthukumar et al. 2014). Collectively, 179 proteins were identified from saliva taken at different stages of the estrus cycle, but 37 of these were expressed exclusively during estrus; the expression of β-enolase and TLR 4 in saliva was validated by immunoblot as an indicator of estrus in buffalo (Muthukumar et al. 2014).

Sheep secrete on the order of 10 liters of saliva per day (Kay 1960). Separation of the ovine salivary proteome by SDS-PAGE followed by MS identified 319 protein families and a predominance of proteins associated with metabolism and inflammation (de Sousa-Pereira et al. 2015). Three hundred and twenty-four unique proteins were identified compared to only 74 unique proteins that were identified in bovine saliva when processed in the same way; for example, CATHL1 and CATHL2, members of the cathelicidin family, were identified in sheep saliva, whereas only CATHL4 was identified in bovine saliva, suggesting differential evolutionary pressures. Analysis of ovine saliva collected directly from the parotid gland comprised 260 protein spots evident by 2D gel electrophoresis, 106 of which were identified using MALDI-TOF MS and 11 by LC-MS/MS totaling 26 different proteins, several of which show evidence of phosphorylation and/or glycosylation (Lamy et al. 2009). Identified proteins function primarily as transporters (i.e., annexin, apolipoprotein, haptoglobin, serum albumin, serotransferrin, transthyretin, vitamin D-binding protein, hemoglobin, lactoferrin, lactoglobulin, casein) followed by immune modulation (e.g., immunoglobulin and antimicrobials). As with all complex protein samples, multiple fractionation and identification strategies are advised for a more comprehensive analysis of protein content. Comparison of sheep parotid gland secretions with that of goats was very similar by 2DE; only three proteins were identified as unique to goats, apolipoprotein A-IV, hemoglobin, and cathelicidin-3 precursor, compared to two proteins which were unique to sheep, clusterin and transthyretin precursor (Lamy et al. 2009). Sheep and goats that were conditioned on Quebracho tannin extract showed increased total protein content and proteomic changes in parotid saliva (Lamy et al. 2011); however, no constitutive changes in tannin-binding salivary proteins were detected.

LC-MS/MS has also been applied to saliva derived from horses (de Sousa-Pereira et al. 2015). Of the 195 unique proteins identified, 57 were found only in saliva from horses with systemic inflammation which included acute phase proteins such as serum amyloid A, fibrinogen, haptoglobin, and alpha1-acid glycoprotein (Jacobsen et al. 2014).

2.2 Proteomics of Other Bodily Fluids

Proteomic analysis has been pioneered on an assortment of alternative bodily fluids in a subset of domestic animal species including cerebrospinal fluid, bronchoalveolar lavage fluid, and amniotic fluids.

Cerebrospinal fluid (CSF) is a potential source of biomarkers of aging and neurodegenerative disorders which can be detected and characterized using the tools of proteomics. Since CSF surrounds the brain and spinal cord and acts as an intermediate between blood and nervous tissue, its proteomic content reflects the metabolic state of normal versus diseased brain tissue. Analysis of bovine CSF from 75 suspect cases of bovine spongiform encephalopathy (BSE) compared with 38 normal cattle by 2D gel electrophoresis showed consistent differences including the presence of apolipoprotein E and two unidentified proteins of 35 and 36 kDa with a pI of 5.5 (Jones et al. 1996) and as similarly observed by Hochstrasser et al. (1997). A comprehensive reference map of bovine CSF proteins has been generated and includes the identification of 66 different proteins, 58 of which had not been previously identified in bovine CSF (Brenn et al. 2009). Charge isoforms were identified for nearly all proteins. The high salt content and low amounts of protein in CSF require optimal sample preparation. The effect of four different sample treatments to ovine CSF was compared; protein precipitation with acetone or using a 2D cleanup kit showed best sample recovery in terms of protein gel spots of 2DE gels compared to the use of two different spin columns. However, the 53 kDa transthyretin tetramer was not identified in samples treated with the 2D cleanup kit but was retained in samples treated with acetone or spin columns (Chen et al. 2006). Ovine CSF has been used to study age-related changes in composition; results indicate that there is an age-related reduction in CSF turnover which has a concentrating effect, and therefore, CFS protein concentrations should be normalized according to their age-specific turnovers before comparison with samples from different age groups (Chen et al. 2010). Among 103 proteins identified in ovine CSF, 41 were differentially regulated according to photoperiod; 18 were more abundant during long days, and 23 were more abundant in short days (Teixeira-Gomes et al. 2015). Equine CSF had been analyzed for the presence of apolipoproteins associated with high-density lipoproteins; as is the case in plasma, apo A-II circulates as a homodimer, whereas there appears to be a higher percentage of acylated apo A-I in CSF compared to plasma (Puppione et al. 2012).

The respiratory health of domestic farm animals can be determined via the proteome content of bronchoalveolar lavage fluid (BALF), which contains secreted peptides and proteins derived from airway mucosa and alveolar surfaces that contribute to host defenses. The antimicrobial peptides, prophenin-2 and PR-39, and the calcium-binding protein calgranulin C were reproducibly upregulated in BALF of pigs chronically infected with *Actinobacillus pleuropneumoniae* (Hennig-Pauka et al. 2006). Concentrations of PR-39 were significantly ($P < 0.05$) increased in BALF (median of 4.8 nM) but not in serum (median of 2.5 nM) on day 21 after infection. With the advent of the *Sus scrofa* genome database, a

comprehensive analysis of porcine BALF and methacholine-induced tracheal secretions identified 3858 porcine-specific proteins that encompass a diverse array of functions that include host defense, molecular transport, cell communication, and cytoskeletal and metabolic functions (Bartlett et al. 2013). Cattle that are subjected to stressors such as transportation, weaning, and comingling demonstrate an increased susceptibility to bacterial pneumonia which in turn is associated with elevated levels of endogenous glucocorticoids. Consistent with the initiation of an acute phase response, dexamethasone-treated cattle have increased levels of alpha-1-acid glycoprotein and alpha-1-antitrypsin in BALF which was detected by 2DE, while levels of alpha-HS-glycoprotein were decreased (Mitchell et al. 2007). Treatment with dexamethasone also induced adipocyte fatty acid-binding protein and odorant-binding protein as well as alpha-enolase, cofilin-1, and immunoglobulin J chain. Similar results were obtained in BALF samples collected from weaned and transported calves (Mitchell et al. 2008). LC-MS/MS of bovine BALF identified 88 unique proteins, of which 20 were only detected in samples collected from steers with experimentally induced clinical pneumonia with *Mannheimia haemolytica* (Boehmer et al. 2011). Differences in protein pattern/concentration were detected between sham- and *M. haemolytica*-infected steers for haptoglobin, as well as the antimicrobial peptides cathelicidin-1 and cathelicidin-4 and inter-α-trypsin inhibitor heavy chain-4. A genetic linkage study linked eight candidate genes with proteins in BALF collected from healthy and recurrent airway obstruction-affected horses (Racine et al. 2011). The equine genome allowed for the identification of 582 proteins in normal cell-free equine BALF confirming genome annotation and providing functional annotation and as well as a framework for continued analysis of the biological significance of BALF protein content (Bright et al. 2011).

Characterization of the equine amniotic fluid proteome is a prerequisite to study changes during disease associated with pregnancy and thus to identify biomarkers of health status or embryonic abnormalities (Isani et al. 2016). Thirty-four proteins were identified of which 12 were associated with the extracellular matrix highlighting their important role during the development of fetal tissues. The three most abundant proteins were albumin, major allergen Eqc1, and fibronectin. No immunoglobulin was detected in equine amniotic fluid. A reference map for the equine amniotic membrane has been established (Galera et al. 2015). From gel-based proteomics, 49 spots were excised and 43 proteins were identified by LC-MS/MS. Shotgun proteomics allowed the identification of 116 proteins; only 10 protein identifications were common to both analyses. A comparison of the proteomic profile of the chorioamnion and corresponding caruncle for buffalo embryos associated with normal or retarded development by 2D-DIGE identified decreased concentration of proteins involved with protein folding (HSP 90-alpha, calreticulin), calcium binding (annexin A1, annexin A2), and coagulation (fibrinogen alpha-chain) ($P < 0.05$), whereas proteins involved in protease inhibition (alpha-1-antiproteinase, serpin H1, serpin A3-8), DNA and RNA binding (heterogeneous nuclear ribonucleoproteins A2/B1 and K), chromosome segregation (serine/threonine-protein phosphatase 2A), cytoskeletal organization (ezrin), cell

redox homeostasis (amine oxidase-A), and hemoglobin binding (haptoglobin) were upregulated ($P < 0.05$) (Balestrieri et al. 2013). The identity of 139 individual protein species was confirmed in bovine amniotic and allantoic fluids collected at day 45 of gestation using two strategies: first, 2DE combined with MALDI-TOF-MS/MS and LC-ESI-MS/MS analysis of individual protein spots and, second, a global protein snapshot of the enriched 5–50 kDa protein fraction by LC-ESI-MS/MS and LC-MALDI-TOF-MS/MS (Riding et al. 2008). Immunoglobulin was identified in bovine allantoic fluid at day 45 of gestation. Amniotic and allantoic fluids collected from pregnant ewes experimentally infected with *Chlamydophila abortus* or *Toxoplasma gondii* also contain immunoglobulin specific for antigens derived from their respective infectious agent (Marques et al. 2011, 2012). In both cases, results suggest both a maternal and fetal source of immunoglobulin.

3 Conclusions

Proteomics provides for the characterization of complex biological samples. Bodily fluids, including urine, saliva, CSF, and BALF, and amniotic fluids are readily amenable to the tools of proteomics. Though farm animals are mammals that are characterized by high homology at the genomic, proteomic, and metabolic level, it is clear that they are readily discriminated by differences in their respective proteomes. This is exemplified by the characterization of urine and saliva which provides unique insight into an animal's normal physiological state. The continued characterization of such fluids can be used to further provide insights into how animals respond to environmental pressures, as well as defining biomarkers of pathophysiological-related changes to health and fitness.

References

Abate O, Vittone V, Zanatta R et al (2005) Valutazione qualitativa della proteinuria mediante SDS-AGE ai fini della localizzazione del danno renale nel cane e nel gatto. Veterinaria 19:9–14

Adachi J, Kumar C, Zhang Y et al (2006) The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. Genome Biol 7:R80. https://doi.org/10.1186/gb-2006-7-9-r80

Alhaider A, Bayoumy N, Argo E et al (2012) Survey of the camel urinary proteome by shotgun proteomics using a multiple database search strategy. Proteomics 12:3403–3406. https://doi.org/10.1002/pmic.201100631

Almeida AM, Bassols A, Bendixen E et al (2015) Animal board invited review: advances in proteomics for animal and food sciences. Animal 9:1–17. https://doi.org/10.1017/S1751731114002602

Andersen S, Mischak H, Zürbig P, Parving HH, Rossing P (2010) Urinary proteome analysis enables assessment of renoprotective treatment in type 2 diabetic patients with microalbuminuria. BMC Nephrol 11:29. https://doi.org/10.1186/1471-2369-11-29.

Anderson NG, Anderson NL, Tollaksen SL et al (1979) Analytical techniques for cell fractions: XXV. Concentration and two-dimensional electrophoretic analysis of human urinary proteins. Anal Biochem 95:48–61. https://doi.org/10.1016/0003-2697(79)90184-2

Andrievskaia O, Algire J, Balachandran A, Nielsen K (2008) Prion protein in sheep urine. J Vet Diagnostic Investig 20:141–146. https://doi.org/10.1177/104063870802000201

Ang CS, Binos S, Knight MI et al (2011) Global survey of the bovine salivary proteome: integrating multidimensional prefractionation, targeted, and glycocapture strategies. J Proteome Res 10:5059–5069. https://doi.org/10.1021/pr200516d

Austin PJ, Suchar LA, Robbins CT, Hagerman AE (1989) Tannin-binding proteins in saliva of deer and their absence in saliva of sheep and cattle. J Chem Ecol 15:1335–1347. https://doi.org/10.1007/BF01014834

Bailey CB (1961) Saliva secretion and its relation to feeding in cattle. 3. The rate of secretion of mixed saliva in the cow during eating, with an estimate of the magnitude of the total daily secretion of mixed saliva. Br J Nutr 15:443–451. https://doi.org/10.1079/BJN19610062

Balestrieri ML, Gasparrini B, Neglia G et al (2013) Proteomic profiles of the embryonic chorioamnion and uterine caruncles in buffaloes (*Bubalus bubalis*) with normal and retarded embryonic development. Biol Reprod 88:1–14. https://doi.org/10.1095/biolreprod.113.108696

Barratt J, Topham P (2007) Urine proteomics: the present and future of measuring urinary protein components in disease. Can Med Assoc J 177:361–368. https://doi.org/10.1503/cmaj.061590

Bartlett JA, Albertolle ME, Wohlford-Lenane C et al (2013) Protein composition of bronchoalveolar lavage fluid and airway surface liquid from newborn pigs. Am J Physiol Lung Cell Mol Physiol 305:L256–L266. https://doi.org/10.1152/ajplung.00056.2013

Bathla S, Rawat P, Baithalu R et al (2015) Profiling of urinary proteins in Karan Fries cows reveals more than 1550 proteins. J Proteomics 127:193–201. https://doi.org/10.1016/j.jprot.2015.05.026

Boehmer JL, Degrasse JA, Lancaster VA et al (2011) Evaluation of protein expression in bovine bronchoalveolar fluid following challenge with *Mannheimia haemolytica*. Proteomics 11:3685–3697. https://doi.org/10.1002/pmic.201000710

Bonfanti U, Zini E, Minetti E, Zatelli A (2004) Free light-chain proteinuria and normal renal histopathology and function in 11 dogs exposed to *Leishmania infantum*, *Ehrlichia canis*, and *Babesia canis*. J Vet Intern Med 18:618–624. https://doi.org/10.1111/j.1939-1676.2004.tb02596.x

Bracha S, McNamara M, Hilgart I et al (2014) A multiplex biomarker approach for the diagnosis of transitional cell carcinoma from canine urine. Anal Biochem 455:41–47. https://doi.org/10.1016/j.ab.2014.03.017

Brandt LE, Ehrhart EJ, Scherman H et al (2014) Characterization of the canine urinary proteome. Vet Clin Pathol 43:193–205. https://doi.org/10.1111/vcp.12147

Brenn A, Karger A, Skiba M et al (2009) A comprehensive proteome map of bovine cerebrospinal fluid. Proteomics 9:5199–5205. https://doi.org/10.1002/pmic.200800906

Bright LA, Mujahid N, Nanduri B et al (2011) Functional modelling of an equine bronchoalveolar lavage fluid proteome provides experimental confirmation and functional annotation of equine genome sequences. Anim Genet 42:395–405. https://doi.org/10.1111/j.1365-2052.2010.02158.x

Candiano G, Santucci L, Petretto A et al (2010) 2D-electrophoresis and the urine proteome map: where do we stand? J Proteomics 73:829–844. https://doi.org/10.1016/j.jprot.2009.12.003

Chen RL, Sage EA, Dunn MJ et al (2006) Optimising ovine cerebrospinal fluid preparation for two-dimensional gel electrophoresis. Proteomics 6:3170–3175. https://doi.org/10.1002/pmic.200500623

Chen CPC, Chen RL, Preston JE (2010) The influence of cerebrospinal fluid turnover on age-related changes in cerebrospinal fluid protein concentrations. Neurosci Lett 476:138–141. https://doi.org/10.1016/j.neulet.2010.04.015

D'Amato M, De Beurme FA, Lefebvre RA (1988) Comparison of the effect of vasoactive intestinal polypeptide and non-adrenergic non-cholinergic neurone stimulation in the cat gastric fundus. Eur J Pharmacol 152:71–82

de Sousa-Pereira P, Abrantes J, Colaço B et al (2013) Characterization of thymosin β4 in mammals' saliva. Peptides 40:1–7. https://doi.org/10.1016/j.peptides.2012.12.007
de Sousa-Pereira P, Cova M, Abrantes J et al (2015) Cross-species comparison of mammalian saliva using an LC-MALDI based proteomic approach. Proteomics 15:1598–1607. https://doi.org/10.1002/pmic.201400083
Decramer S, Gonzalez de Peredo A, Breuil B et al (2008) Urine in clinical proteomics. Mol Cell Proteomics 7:1850–1862. https://doi.org/10.1074/mcp.R800001-MCP200
Devarajan P (2011) Biomarkers for the early detection of acute kidney injury. Curr Opin Pediatr 23:194–200. https://doi.org/10.1097/MOP.0b013e328343f4dd.
Devillers N, Farmer C, Mounier AM et al (2004) Hormones, IgG and lactose changes around parturition in plasma, and colostrum or saliva of multiparous sows. Reprod Nutr Dev 44:381–396. https://doi.org/10.1051/rnd:2004043
Di Meo A, Batruch I, Yousef AG et al (2017) An integrated proteomic and peptidomic assessment of the normal human urinome. Clin Chem Lab Med 55:237–247. https://doi.org/10.1515/cclm-2016-0390
Eckersall PD (2008) Proteins, proteomics, and the dysproteinemias. In: Kaneko JJ, Harvey JW, Bruss ML (eds) Clinical biochemistry of domestic animals, 6th edn. Elsevier Academic Press, Amsterdam, pp 117–155
Edwards A, Titchen D (2002) The effect of parasympathetic postganglionic denervation on parotid salivary protein secretion in anaesthetized sheep. Auton Neurosci 100:50–57. https://doi.org/10.1016/S1566-0702(02)00148-0
Edwards A, Titchen D (2003) Autonomic control of protein production by the parotid gland of the sheep. Auton Neurosci 103:38–49. https://doi.org/10.1016/S1566-0702(02)00147-9
Ferlizza E, Buono A, Fracassi F et al (2013) Urine proteome in dogs affected by leishmaniasis and Cushing's syndrome. Biochim Clin 37, Special Supplement. S219
Ferlizza E, Campos A, Neagu A et al (2015) The effect of chronic kidney disease on the urine proteome in the domestic cat (*Felis catus*). Vet J 204:73–81. https://doi.org/10.1016/j.tvjl.2015.01.023
Ferlizza E, Dondi F, Andreani G et al (2017) Validation of an electrophoretic method to detect albuminuria in cat. J Feline Med Surg 19:860–868. https://doi.org/10.1177/1098612X16664112
Filip S, Vougas K, Zoidakis J et al (2015) Comparison of depletion strategies for the enrichment of low-abundance proteins in urine. PLoS One 10:e0133773. https://doi.org/10.1371/journal.pone.0133773
Forterre S, Raila J, Schweigert FJ (2004) Protein profiling of urine from dogs with renal disease using ProteinChip analysis. J Vet Diagn Invest 16:271–277
Forterre S, Raila J, Kohn B et al (2006) Protein profiling of organic stone matrix and urine from dogs with urolithiasis. J Anim Physiol Anim Nutr (Berl) 90:192–199
Galera PD, Ribeiro CR, Sapp HL et al (2015) Proteomic analysis of equine amniotic membrane: characterization of proteins. Vet Ophthalmol 18:198–209. https://doi.org/10.1111/vop.12190
Geckle M (2007) Renal albumin handling: a look at the dark side of the filter. Kidney Int 71:479–481. https://doi.org/10.1038/sj.ki.5002123
Geelen SNJ, Bernadina WE, Grinwis GCM, Kalsbeek HC (1997) Monoclonal gammopathy in a Dutch warmblood mare. Vet Q 19:29–32. https://doi.org/10.1080/01652176.1997.9694734
Gerber B, Eichenberger S, Haug K et al (2009) Association of urine protein excretion and infection with *Borrelia burgdorferi* sensu lato in Bernese Mountain dogs. Vet J 182:487–488. https://doi.org/10.1016/j.tvjl.2008.08.022
Giori L, Tricomi FM, Zatelli A et al (2011) High-resolution gel electrophoresis and sodium dodecyl sulphate-agarose gel electrophoresis on urine samples for qualitative analysis of proteinuria in dogs. J Vet Diagn Invest 23:682–690. https://doi.org/10.1177/1040638711407900

Guo Z, Zhang Y, Zou L et al (2015) A proteomic analysis of individual and gender variations in normal human urine and cerebrospinal fluid using iTRAQ quantification. PLoS One 10: e0133270. https://doi.org/10.1371/journal.pone.0133270

Gutiérrez AM, Martínez-Subiela S, Eckersall PD, Cerón JJ (2009) C-reactive protein quantification in porcine saliva: a minimally invasive test for pig health monitoring. Vet J 181:261–265. https://doi.org/10.1016/j.tvjl.2008.03.021

Gutiérrez AM, Martínez-Subiela S, Cerón JJ (2011a) Evaluation of changes in haptoglobin and C-reactive protein concentrations caused by freezing of saliva and meat juice samples collected from healthy and diseased pigs. Am J Vet Res 72:11–17. https://doi.org/10.2460/ajvr.72.1.11

Gutiérrez AM, Miller I, Hummel K et al (2011b) Proteomic analysis of porcine saliva. Vet J 187:356–362. https://doi.org/10.1016/j.tvjl.2009.12.020

Gutiérrez AM, Escribano D, Fuentes M, Cerón JJ (2013a) Circadian pattern of acute phase proteins in the saliva of growing pigs. Vet J 196:167–170. https://doi.org/10.1016/j.tvjl.2012.10.003

Gutiérrez AM, Nöbauer K, Soler L et al (2013b) Detection of potential markers for systemic disease in saliva of pigs by proteomics: a pilot study. Vet Immunol Immunopathol 151:73–82. https://doi.org/10.1016/j.vetimm.2012.10.007

Gutiérrez AM, Cerón JJ, Fuentes-Rubio M et al (2014) A proteomic approach to porcine saliva. Curr Protein Pept Sci 15:56–63. https://doi.org/10.2174/1389203715666140221115704

Gutiérrez AM, Miller I, Kolarich D et al (2016) Detection and first characterization of an uncommon haptoglobin in porcine saliva of pigs with rectal prolapse by using boronic acid sample enrichment. Animal 1860:1–9. https://doi.org/10.1017/S1751731116002159

Halbmayr E, Schusser GF (2002) Quantitative and qualitative urinary protein excretion in healthy and sick horses. Tierarztl Prax Ausgabe G Grosstiere - Nutztiere 30:115–123

He W, Huang C, Luo G et al (2012) A stable panel comprising 18 urinary proteins in the human healthy population. Proteomics 12:1059–1072. https://doi.org/10.1002/pmic.201100400

Hennig-Pauka I, Jacobsen I, Blecha F et al (2006) Differential proteomic analysis reveals increased cathelicidin expression in porcine bronchoalveolar lavage fluid after an *Actinobacillus pleuropneumoniae* infection. Vet Res 37:75–87. https://doi.org/10.1051/vetres:2005043

Hiss S, Knura-Deszczka S, Regula G et al (2003) Development of an enzyme immuno assay for the determination of porcine haptoglobin in various body fluids: testing the significance of meat juice measurements for quality monitoring programs. Vet Immunol Immunopathol 96:73–82. https://doi.org/10.1016/S0165-2427(03)00141-7

Hochstrasser DF, Frutiger S, Wilkins MR et al (1997) Elevation of apolipoprotein E in the CSF of cattle affected by BSE. FEBS Lett 416:161–163. https://doi.org/10.1016/S0014-5793(97)01191-5

Hormaeche M, Carretón E, González-Miguel J et al (2014) Proteomic analysis of the urine of Dirofilaria immitis infected dogs. Vet Parasitol 203:241–246. https://doi.org/10.1016/j.vetpar.2014.01.025

Isani G, Ferlizza E, Cuoghi A et al (2016) Identification of the most abundant proteins in equine amniotic fluid by a proteomic approach. Anim Reprod Sci 174:150–160. https://doi.org/10.1016/j.anireprosci.2016.10.003

Jacobsen S, Top Adler DM, Bundgaard L et al (2014) The use of liquid chromatography tandem mass spectrometry to detect proteins in saliva from horses with and without systemic inflammation. Vet J 202:483–488. https://doi.org/10.1016/j.tvjl.2014.08.032

Jeffcott LB, Jeffcott TJ (1974) Studies on passive immunity in the foal: III. The characterization and significance of neonatal proteinuria. J Comp Pathol 84:455–465

Jehmlich N, Dinh KHD, Gesell-Salazar M et al (2013) Quantitative analysis of the intra- and inter-subject variability of the whole salivary proteome. J Periodontal Res 48:392–403. https://doi.org/10.1111/jre.12025

Jepson R, Coulton G, Cowan M et al (2013) Evaluation of mass spectrometry of urinary proteins and peptides as biomarkers for cats at risk of developing azotemia. Am J Vet Res 74:333–342

Jones V, Martin TC, Keyes P, Dawson M (1996) Protein markers in cerebrospinal fluid from bse-affected cattle. Vet Rec 139:360–363

Kay RNB (1960) The rate of flow and composition of various salivary secretions in sheep and calves. J Physiol 150:515–537. https://doi.org/10.1113/jphysiol.1960.sp006402

Khan A, Packer NH (2006) Simple urinary sample preparation for proteomic analysis. J Proteome Res 5:2824–2838. https://doi.org/10.1021/pr060305y

Lamy E, da Costa G, Santos R et al (2009) Sheep and goat saliva proteome analysis: a useful tool for ingestive behavior research? Physiol Behav 98:393–401. https://doi.org/10.1016/j.physbeh.2009.07.002

Lamy E, Da Costa G, Santos R et al (2011) Effect of condensed tannin ingestion in sheep and goat parotid saliva proteome. J Anim Physiol Anim Nutr 95:304–312. https://doi.org/10.1111/j.1439-0396.2010.01055.x

Lavoué R, Trumel C, Smets PMY et al (2015) Characterization of proteinuria in dogue de bordeaux dogs, a breed predisposed to a familial glomerulonephropathy: a retrospective study. PLoS One 10:e0133311. https://doi.org/10.1371/journal.pone.0133311

Lemberger SIK, Deeg CA, Hauck SM et al (2011) Comparison of urine protein profiles in cats without urinary tract disease and cats with idiopathic cystitis, bacterial urinary tract infection, or urolithiasis. Am J Vet Res 72:1407–1415. https://doi.org/10.2460/ajvr.72.10.1407

Lhotta K (2010) Uromodulin and chronic kidney disease. Kidney Blood Press Res 33:393–398. https://doi.org/10.1159/000320681

Li QR, Fan KX, Li RX et al (2010) A comprehensive and non-prefractionation on the protein level approach for the human urinary proteome: touching phosphorylation in urine. Rapid Commun Mass Spectrom 24:823–832. https://doi.org/10.1002/rcm.4441

Liu X, Shao C, Wei L et al (2012) An individual urinary proteome analysis in normal human beings to define the minimal sample number to represent the normal urinary proteome. Proteome Sci 10:70. https://doi.org/10.1186/1477-5956-10-70

Marimuthu A, O'Meally RN, Chaerkady R et al (2011) A comprehensive map of the human urinary proteome. J Proteome Res 10:2734–2743. https://doi.org/10.1021/pr2003038

Marques PX, O'Donovan J, Souda P et al (2011) Amniotic and allantoic fluids from experimentally infected sheep contain immunoglobulin specific for *Chlamydophila abortus*. Vet Immunol Immunopathol 140:1–9. https://doi.org/10.1016/j.vetimm.2010.11.002

Marques PX, O'Donovan J, Williams EJ et al (2012) Detection of *Toxoplasma gondii* antigens reactive with antibodies from serum, amniotic, and allantoic fluids from experimentally infected pregnant ewes. Vet Parasitol 185:91–100. https://doi.org/10.1016/j.vetpar.2011.10.028

Martin-Lorenzo M, Gonzalez-Calero L, Zubiri I et al (2014) Urine 2DE proteome analysis in healthy condition and kidney disease. Electrophoresis 35:2634–2641. https://doi.org/10.1002/elps.201300601

Mavromati J, Cash P, Restelli L, Soler L (2014) Proteomics and protein analyses of ovine and caprine body fluids: current studies and future promises. Curr Protein Pept Sci 15:45–55. https://doi.org/10.2174/1389203715666140221113544

McLean L, Hurst JL, Gaskell CJ et al (2007) Characterization of cauxin in the urine of domestic and big cats. J Chem Ecol 33:1997–2009. https://doi.org/10.1007/s10886-007-9354-6

Messana I, Inzitari R, Fanali C et al (2008) Facts and artifacts in proteomics of body fluids. What proteomics of saliva is telling us? J Sep Sci 31:1948–1963. https://doi.org/10.1002/jssc.200800100

Miller I, Teinfalt M, Leschnik M et al (2004) Nonreducing two-dimensional gel electrophoresis for the detection of Bence Jones proteins in serum and urine. Proteomics 4:257–260. https://doi.org/10.1002/pmic.200300546

Miller I, Crawford J, Gianazza E (2006) Protein stains for proteomic applications: which, when, why? Proteomics 6:5385–5408

Miller I, Preßlmayer-Hartler A, Wait R et al (2014) In between - Proteomics of dog biological fluids. J Proteomics 106:30–45. https://doi.org/10.1016/j.jprot.2014.04.016

Mitchell GB, Clark ME, Caswell JL (2007) Alterations in the bovine bronchoalveolar lavage proteome induced by dexamethasone. Vet Immunol Immunopathol 118:283–293. https://doi.org/10.1016/j.vetimm.2007.05.017

Mitchell GB, Clark ME, Siwicky M, Caswell JL (2008) Stress alters the cellular and proteomic compartments of bovine bronchoalveolar lavage fluid. Vet Immunol Immunopathol 125:111–125. https://doi.org/10.1016/j.vetimm.2008.05.005

Miyazaki M, Soeta S, Yamagishi N et al (2007) Tubulointerstitial nephritis causes decreased renal expression and urinary excretion of cauxin, a major urinary protein of the domestic cat. Res Vet Sci 82:76–79. https://doi.org/10.1016/j.rvsc.2006.06.009

Molina L, Salvetat N, Ben AR et al (2011) Analysis of the variability of human normal urine by 2D-GE reveals a "public" and a "private" proteome. J Proteomics 75:70–80. https://doi.org/10.1016/j.jprot.2011.06.031

Moriyoshi M, Tamaki M, Nakao T, Kawata K (1996) Early pregnancy diagnosis in the sow by saliva progesterone measurement using a bovine milk progesterone qualitative test EIA kit. J Vet Med Sci 58:737–741

Muthukumar S, Rajkumar R, Rajesh D et al (2014) Exploration of salivary proteins in buffalo: an approach to find marker proteins for estrus. FASEB J 28:4700–4709. https://doi.org/10.1096/fj.14-252288

Nabity MB, Lees GE, Dangott LJ et al (2011) Proteomic analysis of urine from male dogs during early stages of tubulointerstitial injury in a canine model of progressive glomerular disease. Vet Clin Pathol 40:222–236. https://doi.org/10.1111/j.1939-165X.2011.00307.x

Nagaraj N, Mann M (2011) Quantitative analysis of the intra-and inter-individual variability of the normal urinary proteome. J Proteome Res 10:637–645

Narjisse H, Elhonsali MA, Olsen JD (1995) Effects of oak (*Quercus ilex*) tannins on digestion and nitrogen balance in sheep and goats. Small Rumin Res 18:201–206. https://doi.org/10.1016/0921-4488(95)00700-0

Oh J, Pyo JH, Jo EH et al (2004) Establishment of a near-standard two-dimensional human urine proteomic map. Proteomics 4:3485–3497. https://doi.org/10.1002/pmic.200401018

Ohtaki T, Moriyoshi M, Nakada K et al (1997) Radioimmunoassay of saliva estrone sulfate in pregnant sows. J Vet Med Sci 59:759–763. https://doi.org/10.1292/jvms.59.759

Ozgo M, Skrzypczak W, Drzezdzon D et al (2009) Urinary excretion of low molecular weight proteins in goats. J Physiol Pharmacol 60:119–125

Paltrinieri S, Sironi G, Giori L et al (2015) Changes in serum and urine SAA concentrations and qualitative and quantitative proteinuria in abyssinian cats with familial amyloidosis: a five-year longitudinal study (2009-2014). J Vet Intern Med 29:505–512. https://doi.org/10.1111/jvim.12561

Palviainen M, Raekallio M, Rajamäki MM et al (2012a) Kidney-derived proteins in urine as biomarkers of induced acute kidney injury in sheep. Vet J 193:287–289. https://doi.org/10.1016/j.tvjl.2011.10.004

Palviainen M, Raekallio M, Vainionpää M et al (2012b) Proteomic profiling of dog urine after European adder (*Vipera berus berus*) envenomation by two-dimensional difference gel electrophoresis. Toxicon 60:1228–1234. https://doi.org/10.1016/j.toxicon.2012.08.010

Pastushkova LK, Kononikhin AS, Tiys ES et al (2016) Characteristics of age-dependent changes in urine proteome in healthy men. Adv Gerontol 6:123–128. https://doi.org/10.1134/S2079057016020107

Plews M, Lamoureux L, Simon SL et al (2011) Factors affecting the accuracy of urine-based biomarkers of BSE. Proteome Sci 9:6. https://doi.org/10.1186/1477-5956-9-6

Puppione DL, Della Donna L, Bassilian S et al (2012) Mass spectral measurements of the apoHDL in horse (*Equus caballus*) cerebrospinal fluid. Comp Biochem Physiol Part D Genomics Proteomics 7:172–174. https://doi.org/10.1016/j.cbd.2012.02.002

Pusterla N, Stacy BA, Vernau W et al (2004) Immunoglobulin A monoclonal gammopathy in two horses with multiple myeloma. Vet Rec 155:19–23. https://doi.org/10.1136/vr.155.1.19

Pyo J, Hwang SI, Oh J et al (2003) Characterization of a bovine pregnancy-associated protein using two-dimensional gel electrophoresis, N-terminal sequencing and mass spectrometry. Proteomics 3:2420–2427. https://doi.org/10.1002/pmic.200300618

Racine J, Gerber V, Feutz MM et al (2011) Comparison of genomic and proteomic data in recurrent airway obstruction affected horses using ingenuity pathway analysis®. BMC Vet Res 7:48. https://doi.org/10.2746/042516401776767412

Raila J, Aupperle H, Raila G (2007) Renal pathology and urinary protein excretion in a 14-Month-old Bernese mountain dog with chronic renal failure. J Vet Med A 54:131–135

Raila J, Schweigert FJ, Kohn B (2014) Relationship between urinary Tamm-Horsfall protein excretion and renal function in dogs with naturally occurring renal disease. Vet Clin Pathol 43:261–265. https://doi.org/10.1111/vcp.12143

Riding GA, Jones A, Holland MK et al (2008) Proteomic analysis of bovine conceptus fluids during early pregnancy. Proteomics 8:160–177. https://doi.org/10.1002/pmic.200700465

Ruis MAW, Te Brake JHA, Engel B et al (1997) The circadian rhythm of salivary cortisol in growing pigs: effects of age, gender, and stress. Physiol Behav 62:623–630. https://doi.org/10.1016/S0031-9384(97)00177-7

Scarpa P, Di Fabio V, Ramirez C et al (2007) Proteinuria, GGT index and fractional clearance of electrolytes in exercising athletic horses. Vet Res Commun 31:339–342

Schaefer H, Kohn B, Schweigert FJ, Raila J (2011) Quantitative and qualitative urine protein excretion in dogs with severe inflammatory response syndrome. J Vet Intern Med 25:1292–1297. https://doi.org/10.1111/j.1939-1676.2011.00829.x

Schellenberg S, Mettler M, Gentilini F et al (2008) The effects of hydrocortisone on systemic arterial blood pressure and urinary protein excretion in dogs. J Vet Intern Med 22:273–281

Scocco P, Aralla M, Catorci A et al (2011) Immunodetection of aquaporin 5 in sheep salivary glands related to pasture vegetative cycle. Folia Histochem Cytobiol 49:458–464. https://doi.org/10.5603/FHC.2011.0065

Simon SL, Lamoureux L, Plews M et al (2008) The identification of disease-induced biomarkers in the urine of BSE infected cattle. Proteome Sci 6:23. https://doi.org/10.1186/1477-5956-6-23

Teixeira-Gomes AP, Harichaux G, Gennetay D et al (2015) Photoperiod affects the cerebrospinal fluid proteome: a comparison between short day– and long day–treated ewes. Domest Anim Endocrinol 53:1–8. https://doi.org/10.1016/j.domaniend.2015.04.003

Thongboonkerd V, Chutipongtanate S, Kanlaya R (2006) Systematic evaluation of sample preparation methods for gel-based human urinary proteomics: quantity, quality, and variability. J Proteome Res 5:183–191. https://doi.org/10.1021/pr0502525

Van der Stede Y, Cox E, Van den broeck W, Goddeeris BM (2001) Enhanced induction of the IgA response in pigs by calcitriol after intramuscular immunization. Vaccine 19:1870–1878. https://doi.org/10.1016/S0264-410X(00)00440-0

Vitorino R, Barros AS, Caseiro A et al (2012a) Evaluation of different extraction procedures for salivary peptide analysis. Talanta 94:209–215. https://doi.org/10.1016/j.talanta.2012.03.023

Vitorino R, Guedes S, Manadas B et al (2012b) Toward a standardized saliva proteome analysis methodology. J Proteomics 75:5140–5165. https://doi.org/10.1016/j.jprot.2012.05.045

Xu C, Shu S, Xia C et al (2015) Mass spectral analysis of urine proteomic profiles of dairy cows suffering from clinical ketosis. Vet Q 35:133–141. https://doi.org/10.1080/01652176.2015.1055352

Zaragoza C, Barrera R, Centeno F (2003) SDS-PAGE and Western blot of urinary proteins in dogs with leishmaniasis. Vet Res 34:137–151. https://doi.org/10.1051/vetres

Colostrum Proteomics Research: A Complex Fluid with Multiple Physiological Functions

Lorenzo E. Hernández-Castellano, Craig R. Baumrucker, Josef Gross, Olga Wellnitz, and Rupert M. Bruckmaier

Abstract This chapter addresses several aspects related to colostrum, from its synthesis in the mammary gland to the absorption of the colostrum components by the newborn and its effects on the maturation of the immune and digestive system. This chapter also describes the different colostrum fractions, considering high and low abundant proteins. Additionally, the function of those proteins and their role on the passive immune transfer and nutrition in the newborn animal are also included. Finally, it will be described how different proteomics tools have been utilized in order to answer questions about colostrum associated immunology and nutrition in newborn ungulates.

Keywords Colostrum • Farm animals • Proteomics • Immune system • Mammary gland

Abbreviations

Igs	Immunoglobulins
MS	Mass spectrometry
MALDI	Matrix-assisted laser desorption/ionization
TOF	Time of flight
iTRAQ	Isobaric tag for relative and absolute quantitation
MFGM	Milk fat globule membrane

L.E. Hernández-Castellano (✉)
Veterinary Physiology, Vetsuisse Faculty, University of Bern, Bern, Switzerland

Faculty of Science and Technology, Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark
e-mail: lorenzo.hernandez@vetsuisse.unibe.ch; lhc@mbg.au.dk

C.R. Baumrucker
Department of Animal Science, Pennsylvania State University, PA, University Park, USA

J. Gross • O. Wellnitz • R.M. Bruckmaier
Veterinary Physiology, Vetsuisse Faculty, University of Bern, Bern, Switzerland

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_8

PIT	Passive immune transfer
TJ	Tight junctions
FcRn	Fc receptor of the neonate
IgG	Immunoglobulin G
SCNT	Somatic cell nuclear transfer
UPLC	Ultra performance liquid chromatography
LTQ	Linear trap quadrupole
SLS	Small litter size
LLS	Large litter size
BTF3	Basic transcription factor 3
MCM3	Minichromosome maintenance complex component 3
FABPs	Fatty acid-binding proteins
HSPs	Heat stress proteins
DIGE	Difference gel electrophoresis
APP	Acute-phase proteins
PGF	Placental growth factor
TNF-α	Tumor necrosis factor alpha
TNF-β	Tumor necrosis factor beta
VEGF	Vascular endothelial growth factor
HAP	High abundant proteins
LAP	Low abundant proteins
IGF-1	Insulin-like growth factor-1
EGF	Epidermal growth factor-like peptide
2DE	Two-dimensional electrophoresis
XDH/XO	Xanthine dehydrogenase/oxidase
C3	Complement component 3
LBP	Lipopolysaccharide-binding protein
FN1	Fibrinectin 1
SAA	Serum amyloid A
APO	Apolipoproteins

1 Introduction

Colostrum, the first postpartum secretion of the mammary gland, is the first source of nutritional components such as fat, lactose, vitamins, and minerals for newborn animals (Hernández-Castellano et al. 2014a). Colostrum is different from mature milk as it has higher protein and fat content. The higher protein content is principally associated with increased immunoglobulin (Ig) content. The early neonatal provision of Igs provides decreased risk of disease until its own immune system develops (Hernández-Castellano et al. 2014a; Ontsouka et al. 2003). While a number of mammalian species transfer dam Igs to the fetus by placental transfer, this is not the case for ungulate species. Therefore, colostrum is the only source of Igs to the ungulate. The relation between colostrum and newborn survival has been exhaustively characterized in many species such as cattle, sheep, goats, horses, and

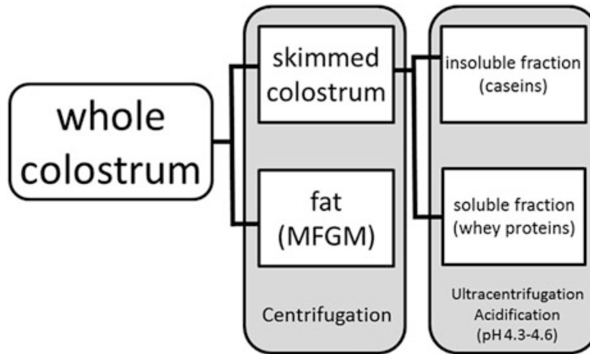

Fig. 1 Schematic procedure for the separation of the three main colostrum protein fractions. One fraction can be obtained upon centrifugation: fat and the surrounding milk fat globule membrane (MFGM). The skimmed colostrum can be treated with ultracentrifugation or acidification (or both) to obtain the last two fractions: caseins and whey proteins. Adapted from Hernández-Castellano et al. (2014a)

pigs (Argüello et al. 2004; Tyler et al. 1999; Hernández-Castellano et al. 2015a; Braun et al. 1983; Scoot et al. 1972). Additionally, colostrum contains a complex mixture of proteins that actively participate in the protection of the neonate against pathogens and other postpartum environmental challenges (Bendixen et al. 2011).

Because of the wide range of protein concentrations present in colostrum, the identification of the whole colostrum proteome is complex. In order to reduce complexity and in consequence increase the coverage of the colostrum proteome, it is necessary to pretreat colostrum and then study the different fractions separately [Fig. 1; (Ontsouka et al. 2003; Nissen et al. 2012)]. After that, two major proteomics strategies can be performed, either at the protein level [top-down approach; (Westermeier and Naven 2002)] or after the digestion of the protein mixture to peptides [bottom-up approach; (Gevaert and Vandekerckhove 2011)]. In the top-down approach, proteins are first separated according to their charge (isoelectric point) and then separated by molecular weight in a polyacrylamide gel and, finally, analyzed by mass spectrometry (MS) (Almeida et al. 2015). One of those analyses is based on the matrix-assisted laser desorption/ionization (MALDI)-time of flight (TOF) MS, which is a robust, sensitive, and relatively inexpensive method able to identify thousands of proteins in a specific tissue or fluid. In the bottom-up approach, intact proteins are first labeled, then enzymatically digested to peptides, separated by chromatography, and, finally, analyzed by MS. One of the most common approaches based on this strategy is the isobaric tag for relative and absolute quantitation (iTRAQ).

In the last years, proteomics has been used to characterize protein changes in the transition from colostrum to milk in several species (Nissen et al. 2012; Reinhardt and Lippolis 2008; Stelwagen et al. 2009; Hernández-Castellano et al. 2016a; Yang et al. 2016; Lu et al. 2016), to define the mammary gland proteome during different lactation stages (Cugno et al. 2016; Beddek et al. 2008; Hesselager et al. 2016; Hernández-Castellano et al. 2016b), to determine proteolytic processes in the

stomach and gut of piglets upon colostrum intake (Danielsen et al. 2011), or to describe changes in the blood proteome profile of lambs after colostrum intake (Hernández-Castellano et al. 2014b, 2015b).

2 Placentation, Mammary Gland, and Colostrogenesis in Eutherian Species

As can be observed in Table 1, the complexity and number of layers of the epitheliochorial [cows, horses, and pigs (Padua et al. 2010; Horsley et al. 1998; Porter 1973)] and synepitheliochorial [sheep and goats (Wooding and Burton 2008; Wooding et al. 1986)] placentas do not allow an adequate transfer of Igs from the dam to the fetus compared to the hemochorial placenta of other species such as humans or mice (Kirby and Bradbury 1965; Prystowsky 1958). Therefore, most of the newborn farm animals are considered agammaglobulinemic [calves, foals, camels, and piglets (Wooding and Burton 2008; Al-Busadah 2007; Erhard et al. 2001; Brignole and Stott 1979)] or hypogammaglobulinemic [lambs and kids (Hernández-Castellano et al. 2015a)] at birth. In those species, colostrum intake and the absorption of colostrum proteins play an essential role in the passive immune transfer (PIT) and ultimately on newborn survival rates (Stelwagen et al. 2009; Hernández-Castellano et al. 2015c; Hadorn et al. 1997).

During a short period before parturition, the mammary gland secretes diverse components either by local synthesis at the mammary gland or through the transfer of components from the maternal bloodstream into mammary secretions

Table 1 Placental layers and transfer of immunoglobulin G (IgG) in different species

	Cow	Horse	Pig	Goat/sheep	Dog	Human	Mice	Rabbit
Fetal layers								
Chorionic endothelium	X	X	X	X	X	X	X	X
Connective tissue	X	X	X	X	X	X		
Chorionic epithelium	X	X	X	X	X	X		
Maternal layers								
Uterine epithelium	X[a]	X	X					
Connective tissue	X	X	X	X				
Uterine endothelium	X	X	X	X	X			
Total number of layers[b]	6	6	6	5	4	3	1	1
IgG transfer (mainly)								
Uterus					X	X	X	X
Colostrum	X	X	X	X	X			

[a]The number of layers in cow placenta seems to be inconsistent as some authors have classified it as epitheliochorial [six layers (Padua et al. 2010)] and others as synepitheliochorial [five layers (Wooding and Burton 2008)]

[b]The transfer of IgG (and other immune components) through the placenta decreases as the number of placental layers increases

(Barrington and Parish 2001; Baumrucker and Bruckmaier 2014). This process is known as colostrogenesis and can be affected by the mammary gland development during pregnancy (Castro et al. 2011), which is regulated by several factors such as hormones (predominantly estradiol, prolactin, and progesterone), breed, nutrition, litter size, length of dry period, and health status of the mother (Castro et al. 2011; Lerias et al. 2014; Farmer 2013).

Colostrum components are secreted by different mechanisms (Baumrucker and Bruckmaier 2014); therefore, proteins present in colostrum are either directly synthesized in the mammary gland or are transferred from the bloodstream (or both). The appearance of factors in mammary colostrum secretions may occur by one or more of the nine different pathways shown in Fig. 2.

Pathway 1 shows a fluid-phase endocytosis process that occurs during colostrogenesis and is the major means of IgG1 entrance into cytoplasmic endosomes. As these endosomes become acidified (pathway 2), the IgG1 specifically binds to the Fc receptor of the neonate (FcRn) and can be recycled (pathway 3) to the extracellular fluid or transcytosed (pathway 4) to the colostrum secretions.

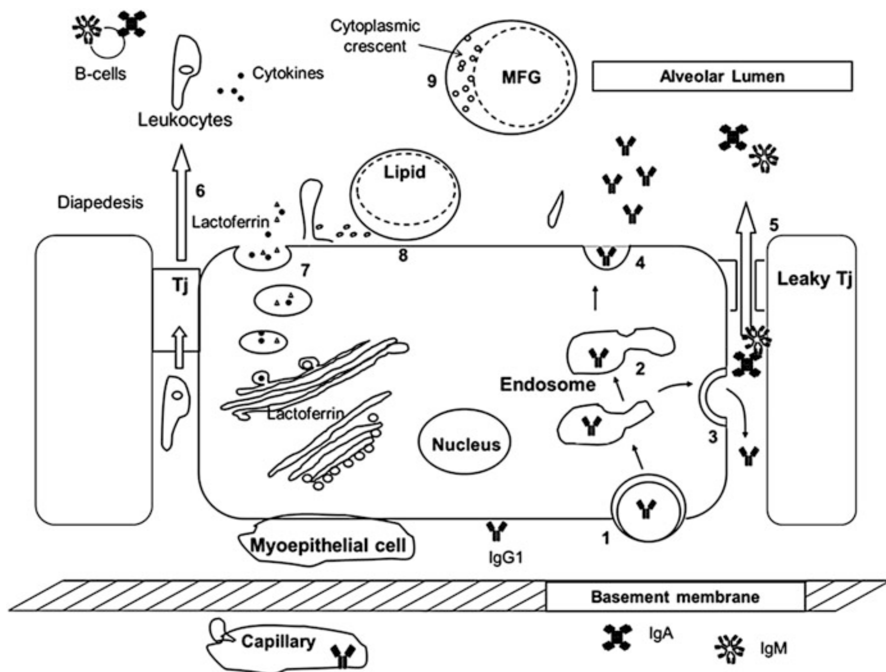

Fig. 2 Pathways of mammary epithelial cell transport during colostrogenesis. Pathway 1, fluid-phase endocytosis; pathway 2, endosome FcRn/IgG1 binding; pathway 3, FcRn/IgG1 recycling; pathway 4, FcRn/IgG1 transcytosis; pathway 5, passage through leaky tight junctions (TJ); pathway 6, diapedesis of blood cells; pathway 7, apocrine secretion of mammary synthesized components; pathway 8, merocrine secretion of fat; pathway 9, cytoplasmic crescents secreted within the milk fat globule (MFG)

The regulation of these two pathways occurs via specific small G proteins. Blood components may also enter the colostrum through leaking tight junctions (pathway 5) where serum albumin and perhaps blood cytokines and lactoferrin may enter. Pathway 6 shows that blood leukocytes and small amounts of B cells may enter colostrum by the process of diapedesis. These cells synthesize and secrete cytokines and contribute to colostrum regulatory factors. Mammary synthesized components' (lactoferrin/others) secretion occurs by an apocrine mechanism (pathway 7) and merocrine mechanism for fat secretion (pathway 8). Endocrine components that are fat soluble, such as steroid hormones, are known to pass biological membranes (enter cells) and may have cellular binding proteins that provide retention capacity in aqueous solutions or are secreted in the milk fat. Also secreted within the MFGM is a mixture of mammary cell cytoplasmic components (cytoplasmic crescent, pathway 9).

Regarding the use of proteomics, Talbot et al. (2010) used a 2D MALDI-TOF MS and an ion trap MS approach to describe the importance of annexins I, II, V, and VII during placentation in dairy cows. In the study of Talbot et al. (2010), trophectoderm cell lines from somatic cell nuclear transfer (SCNT) cows, known for having increased pregnancy failure (abortion and developmental abnormalities of the offspring), showed lower expression of those proteins compared to the control cows. Those proteins are involved in crucial processes during pregnancy: inhibition of placental blood coagulation (annexin V), tissue immunomodulation and apoptosis (annexin V), intracellular membrane fusion (annexins I, II, and VII), signal transduction for cell differentiation (annexins I and II), extracellular matrix remodeling via plasminogen activation (annexin II), and, indirectly, the matrix metalloproteinases and prostaglandin synthesis regulation. In pigs, Lee et al. (2015) using a nano acquity ultra performance liquid chromatography (UPLC) together with an Orbitrap linear trap quadrupole (LTQ) Velos MS studied placental proteomics differences between small litter size (SLS) and large litter size (LLS) sows. Those authors found 6 and 13 upregulated proteins in SLS and LLS, respectively. In the SLS group, some of the increased proteins were associated to proteolysis processes (calpain, cytosolic nonspecific dipeptidase, prenylcysteine oxidase, and ubiquitin). Increased proteolytic degradation has been associated to proteins undergoing oxidative damage (Lappas et al. 2011), affecting the placenta structure and functionality which could explain the reduced litter size of those piglets. From the 13 proteins identified in the LLS group, puromycin and RBP4 were the most relevant proteins as both of them have been related to the litter size in mice and humans, respectively (Osada et al. 2001; Chan et al. 2011).

In ruminants and pigs, developmental changes in the mammary gland during late pregnancy colostrogenesis have not been deeply studied using proteomics. Some studies performed in water buffalo (Jena et al. 2015) and mice (Davies et al. 2006) can provide some insight of the ongoing processes during colostrogenesis in domestic ungulates. Using a 2D MALDI-TOF MS approach, Davies et al. (2006) described how the basic transcription factor 3 (BTF3), increased during pregnancy, seems to affect the transcription of the estrogen receptor in the mammary gland during pregnancy, regulating mammary alveolar growth in mice. In this study, the minichromosome maintenance complex component 3 (MCM3), which is required

for DNA replication and cell proliferation (Ford and Chevalier 1995), was also increased during pregnancy. It seems that both proteins (BTF3 and MCM3) regulate mammary gland development in mice and therefore they may affect colostrogenesis.

Proteomics studies performed in heifers and early lactation water buffalo by Jena et al. (2015) using a two-dimensional difference gel electrophoresis (2-DIGE) MALDI-TOF/TOF MS approach showed the expression of diverse proteins involved in mammary gland development. Specifically, fatty acid-binding proteins (FABPs), ACSL1, and SLC2746, all involved in the regulation of fatty acid uptake and intracellular transport (Chmurzynska 2006), were increased in water buffaloes in peak lactation compared to heifers. As described above in placenta tissue, annexin II seems to also increase in mammary gland tissue at late pregnancy and progressively decrease during the early lactation period (Jena et al. 2015). In the mammary gland, this protein is involved in cell proliferation and mammary epithelial cell differentiation processes (Desrivieres et al. 2003). Early lactation water buffalo also showed upregulation of EEF1D, an essential protein during protein synthesis, compared to heifers. Heat stress proteins (HSPs) are chaperones that play an important role in proper folding and translocation of polypeptides across membranes (De Maio 1999). The increased regulation of HSPA5 and HSP27 in early lactation water buffalo compared to heifers suggests their active participation in the secretory pathway during colostrogenesis and lactation (Jena et al. 2015).

3 Colostrum Constituents

Colostrum is a complex mixture of water, proteins, carbohydrates, lipids, vitamins, and minerals (Koldovsky 1989; Blum and Hammon 2000). Immunoglobulins are certainly one of the most important and abundant proteins in colostrum, being IgG the most abundant immunoglobulin in ruminant and pig colostrum (Sasaki et al. 1976; Vaerman et al. 1970). As shown in Fig. 3, colostrum also contains other substances such as antimicrobial and anti-inflammatory agents (Stelwagen et al. 2009; Hernández-Castellano et al. 2016a; Ebrahim et al. 2014), growth factors that control early gastrointestinal development (Purup et al. 2007), cytokines, enzymes, and several different peptides (Blum and Hammon 2000; Koldovsky 1980). Interestingly, most of those bioactive compounds are present in colostrum in higher concentration than in milk, notably Igs, lactoferrin, and growth factors, indicating the importance of colostrum on the health status of the newborn animal (Scammel 2001; Pakkanen and Aalto 1997).

Several proteomics studies were performed in order to describe the colostrum proteome. Because of the wide range of protein concentrations and subcellular locations in colostrum, a few high abundant proteins (HAP), usually studied by traditional biochemical methods, mask the presence of the low abundant proteins (LAP). Therefore, most of the proteomics studies performed in colostrum have been focused in any of the three main physicochemical fractions (whey, milk fat globule membrane, and casein).

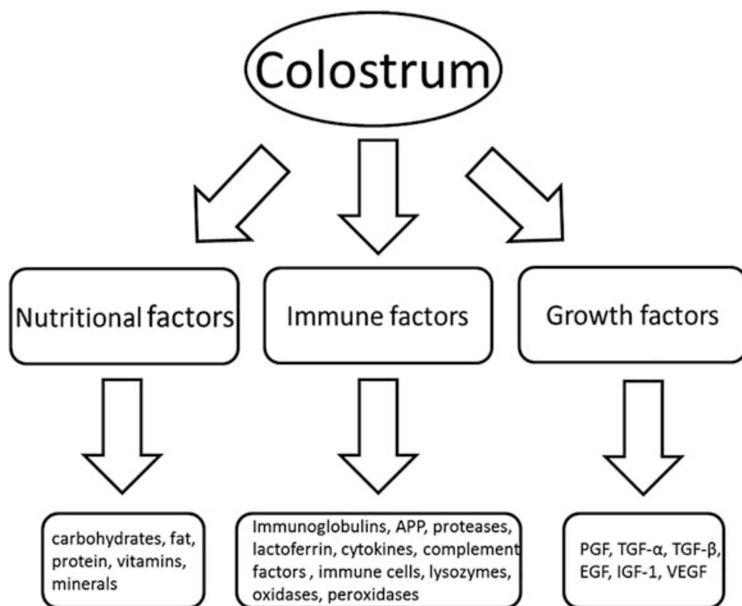

Fig. 3 Colostrum composition. *APP* acute-phase proteins, *PGF* placental growth factor, *TNF-α* tumor necrosis factor alpha, *TNF-β* tumor necrosis factor beta, *EGF* epidermal growth factor-like peptide, *IGF-1* insulin-like growth factor-1, *VEGF* vascular endothelial growth factor

3.1 Proteins in Whey Colostrum

Soluble proteins compose the whey fraction in colostrum. In cows and sows, this fraction represents approximately 90% of the total proteins present in colostrum (Sgarbieri 2004; Klobasa et al. 1987). Colostrum whey contains hundreds of proteins, being β-lactoglobulin, α-lactalbumin, serum albumin, Igs, lactoferrin, and lactoperoxidase the major constituents in colostrum whey (Ontsouka et al. 2003; Sgarbieri 2004; Theil and Hurley 2016). While Igs play one of the most important roles in the PIT, there are other proteins in colostrum whey that actively contribute to the immune response in the newborn animal. Lactoferrin, for instance, is able to increase the activity of several leucocytes such as lymphocytes, neutrophils, and monocytes (Actor et al. 2009) and has been shown to actively contribute to the reaction to infections caused by *E. coli* in humans (Bullen et al. 1972). Furthermore, Hernandez-Ledesma et al. (2008) described how bioactive peptides derived from either alcalase-, pepsin-, or trypsin-digested β-lactoglobulin act as a bacteriostatic against *E. coli*, *Bacillus subtilis*, and *Staphylococcus aureus* (El-Zahar et al. 2004; Pihlanto-Leppala et al. 1999). Colostrum whey also contains growth factors which stimulate the proliferation and differentiation of the intestinal epithelium (Penchev-Georgiev 2008), inhibit apoptosis (Blum and Baumrucker 2002), and promote skeletal muscle development (Marnila and Korhonen 2011). Some of the most abundant growth factors present in colostrum are insulin-like growth factor-1 (IGF-1), epidermal growth factor-like peptide (EGF), and TGF-β2 (Theil and

Hurley 2016; Elfstrand et al. 2002; Nguyen et al. 2007; Kolb and Kaskous 2003; Blum and Baumrucker 2008).

Most of the proteomics studies have focused on the colostrum whey fraction from cows. Smolenski et al. (2007) using two-dimensional electrophoresis (2DE) MALDI-TOF/TOF MS approach found higher concentrations of nucleobindin in cow colostrum whey compared to milk whey. This protein contributes to the already described immune properties of the colostrum, promoting the activation of lymphocytes in humans (Kubota et al. 1998) and mice (Kanai et al. 1993). In another study using a similar proteomics approach, Senda et al. (2011) found proteins such as the bovine vitamin D-binding protein or different IgG chains present only in cow colostrum whey and not in milk whey. Those findings reflect the importance of colostrum not only in the immune protection of the newborn animal but also on the normal bone formation.

Further advancement in new proteomics approaches with higher sensibilities has allowed the identification of more components in the colostrum whey. Using an Orbitrap LTQ MS coupled to a HPLC system, Le et al. (2011) found 36 unique proteins in cow colostrum whey that were not present in milk whey. Additionally, those authors also described the presence of 217 proteins in both colostrum and milk whey. Interestingly, some of those proteins participate in complement activation (CFH, C3, MBL2, C9, CFD, and CFB), immune response (AZGP1, B2M, PGLYRP1, CD14, C6, C4BPA, MIF, C1S, and CD46), acute-phase response (F2, AGP, ITIH4, AHSG, SERPINF2, and SAA), and defensive response against bacterial infections (LTF, CSN1S2, CAMP, CATHL2, CATHL5, and CATHL6). Those results were confirmed using a similar proteomics approach by Zhang et al. (2015) who identified 208 proteins in colostrum and milk whey from cows during the first 9 days postpartum, with 64 proteins showing increased regulation in colostrum whey compared to milk whey. Most of those proteins (47) participate in different biological processes such as the complement system (C1R, C1S, C3, C6, C7, C8B, C9, CFB, CFD, CFH, CFI), immune response against bacterial infections (CATHL1, CATHL2, CATHL3, CATHL4, CATHL5, CATHL6, CATHL7, LPO, PGLYRP1), immunoglobulin-like proteins (A1BG, AHSG, B2M, BoLA, PIGR, IGJ, IGK, IGLL1), and acute-phase proteins (ORM1, F2, C3, FN1, SERPINF2, ITIH4, SAA1, SAA3). In addition, other immune-related proteins were identified in this study (AZGP1, B4GALT1, BOLA-NC1, MUC15, CHI3L1, CLU, CRISP3, GLYCAM1, GP2, HP, RNASE4).

Those results reflect the great and wide range of LAP present in colostrum whey that can participate in several immune processes and actively contribute to the defense of the newborn animal.

3.2 Proteins in the Milk Fat Globule Membrane

As can be observed in the Fig. 4, the MFGM is a three-layer membrane composed of proteins and phospholipids that covers the milk fat globule (Keenan and

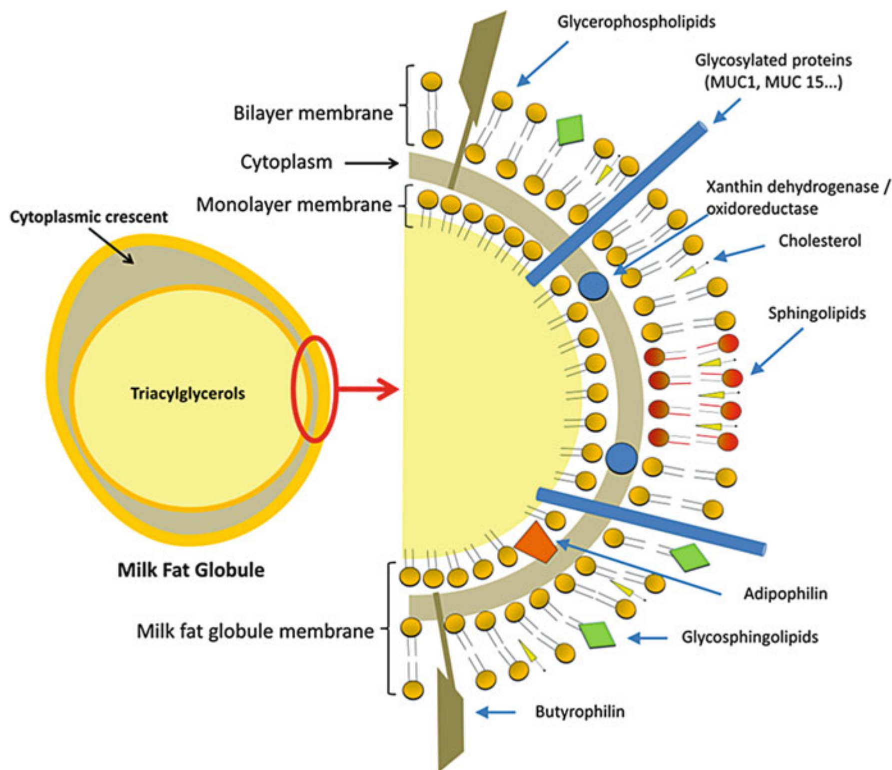

Fig. 4 Structure of the milk fat globule and schematic representation of the milk fat globule membrane. Adapted from (Lopez 2011)

Dylewski 1995). The MFGM mostly originated from the endoplasmic reticulum membrane and the plasma membrane (Cebo 2012). As described by Mather (2000), there are eight HAP in the MFGM that can be measured by traditional biochemical methods. From those HAP, xanthine dehydrogenase/oxidase (XDH/XO) and mucin-1 have been shown to participate in the protection against pathogens (both proteins are shown in Fig. 4). Particularly, XDH/XO is a protein essentially synthesized in the mammary gland (Parks and Granger 1986), reaching maximum concentrations during late pregnancy and early lactation in mice (Kurosaki et al. 1996) and cows (Zhang et al. 2015). In addition to the secondary antibacterial function of this protein as a source of H_2O_2 for lactoperoxidase, XDH/XO induces the expression of genes encoding, for instance, adhesive proteins, cell receptors, and other components of the immune system (Mather 2000). MUC-1 is a glycoprotein with the capacity of binding and sequestering pathogenic microorganisms in the gastrointestinal tract (Peterson et al. 1998; Schroten et al. 1992); therefore, the presence of this protein in colostrum can protect the suckling neonate against external microorganisms.

Besides HAP, LAP present in colostrum MFGM seems to play crucial roles in the immune protection, in the nutrient absorption, and in the bone and musculoskeletal development. In this way, Lu et al. (2016) using an EASY-nLC-Orbitrap LTQ Velos approach observed several acute-phase proteins that were increased in goat colostrum MFGM compared to milk MFGM. Increased proteins with immune-related properties were identified as complement component 3 (C3), lipopolysaccharide-binding protein (LPB), fibrinectin 1 (FN1), and serum amyloid A (SAA). Similar studies performed by Reinhardt and Lippolis (2008) using a shotgun proteomics approach coupled with iTRAQ labeling of peptides quantified protein changes in colostrum and milk MFGM from cows. Those authors identified 138 proteins in the colostrum MFGM from which 19 were upregulated compared to milk MFGM on day 7 postpartum. Some of those increased proteins were identified as apolipoproteins (APO), specifically APO A-I, C-III, E, and A-IV. Apolipoproteins participate in the lipoprotein metabolism through the absorption, transport, and redistribution of lipids among various cells and tissues (Mahley et al. 1984). The presence of these proteins in colostrum may increase and facilitate the absorption efficiency of nutrient which is crucial in newborn animals. Additionally, other proteins with immune functions that are usually found in the whey fraction were also found in the MFGM (the Ig mu chain region, lactoferrin, or the FC fragment of IgG). In another study, Yang et al. (2016) using iTRAQ labeling and a LC-MS/MS approach found 21 proteins increased in cow colostrum MFGM compared to milk MFGM. Besides APO A-I and E, in which functions have been described above, serine protease HTRA1 and ribonuclease-4 were also upregulated.

Serine protease HTRA1 has been described to participate in the osteogenesis through modification of proteins within the extracellular matrix, and it seems to be also implicated in musculoskeletal development in humans (Tiaden et al. 2012). On the other hand, ribonuclease-4 seems to contribute to the organism's response against infections caused by diverse pathogens. (Harris et al. 2010) described how ribonuclease-4 extracted from cow milk was able to reduce the viability of *Candida albicans* in vitro. Additionally, (Murata et al. 2013) observed that the antimicrobial effects of lactoferrin and lactoferricin against *E. coli* were enhanced by co-incubation with a mixture of ribonuclease-4 and ribonuclease-5.

3.3 Caseins in Colostrum

Caseins are phosphoproteins found in colostrum and milk in large colloidal particles (20–300 nm diameter), known as casein micelles. Caseins may be classified into four main types (αs1, αs2, β, and κ). Those proteins participate in several important biological functions such as ion carriers (calcium, phosphate, iron, zinc, and copper), bioactive peptide precursors, and immunomodulators (Korhonen and Pihlanto 2007). According to Clare and Swaisgood (2000), casein-derived peptides such as casecidin, casocidin I, or isracidin have several antimicrobial activities against infections caused by Gram-positive bacteria (i.e., *Staphylococcus aureus*),

Gram-negative bacteria (i.e., *E. coli*), and yeast (i.e., *Candida albicans*). Consequently, peptides derived from caseins are receiving attention as possible sources of natural bioactivity with health benefits (Silva and Malcata 2005), probably because they stimulate the innate immune system within the mammary gland and prevent udder infections during the dry phase (Silanikove et al. 2005).

4 Colostrum Absorption and Maturation of the Immune and Digestive System of the Newborn

In newborn ungulates, colostrum protein absorption mainly consists of Igs (IgG, IgM, and IgA, mainly) which are the main absorbed components from colostrum; however, the absorption of other proteins from colostrum is essential to ensure a correct PIT (Smith and Foster 2007; Leece 1973). These special conditions for intact colostrum protein absorption decrease along the first 48 h after birth. Consequently, it is crucial to feed newborn ruminants and piglets with colostrum during this early time frame (Hernández-Castellano et al. 2015a; Leece 1973; Moore et al. 2005). Besides PIT, colostrum intake also promotes intestinal epithelial cell growth reflected by the increased villus length and width, crypt depth, and mucosal thickness in the duodenum, jejunum, and ileum, which improve nutrient absorption ability and intestinal function development in calves (Blum and Baumrucker 2008; Guilloteau et al. 1997; Yang et al. 2015). Similar effects on the intestinal development have been also observed in newborn small ruminants and piglets as a consequence of the colostrum intake (Kelly et al. 1993; Moretti et al. 2010, 2014). All these facts make colostrum-deprived animals more susceptible to diseases and increase mortality rates during the first hours of life (Nowak and Poindron 2006; Boyd 1972; Declerck et al. 2016).

Several proteomics tools have been used in order to increase the knowledge about the ongoing processes during colostrum absorption and maturation of the immune and digestive system of newborn ruminants and piglets. In newborn calves, Herosimczyk et al. (2013) using a 2D MALDI-TOF/TOF MS approach observed increased amounts of APO A-1, A-IV, and FIB γ chain in blood plasma 48 h after colostrum intake. A similar study using a 2-DIGE MALDI-TOF/TOF MS approach was performed in newborn lambs by Hernández-Castellano et al. (2014b) who found four increased proteins in blood plasma 10 h after the first colostrum feeding. Specifically, APO A-IV, which participates during the enterocyte lipid transport in fatty foods such as colostrum (Stan et al. 2003), was increased. The other three increased proteins (fibrinogen (FIB) γ chain, plasminogen, and SAA) participate through different pathways in the migration of different immune cells such as neutrophils, monocytes, and T lymphocytes to the infection site (Renckens et al. 2006; Badolato et al. 1994; Xu et al. 1995; Kuhns et al. 2001). Those results were confirmed by Hernández-Castellano et al. (2015b) using two more sensitive proteomics approaches; one for detecting the presence/absence of colostrum proteins in

blood plasma from lambs (EASY-nLC TripleTOF 5600 MS) and another for quantifying protein changes in blood plasma from lambs 10 h after the first colostrum intake (iTRAQ LC-MS/MS). Besides the eight proteins increased in plasma by the colostrum intake (APO A-IV, B-100 and E, FIB α chain, trypsin inhibitor, tetranectin, ceruloplasmin precursor, and immunoglobulin μ-chain partial), 26 proteins were exclusively identified in colostrum, whereas they were not identified in the blood plasma (several Ig fractions, caseins, α-lactalbumin, nucleobindin-1, peroxiredoxin-1, and junction plakoglobin, among others). Using a 2D-LC-MS/MS approach, Danielsen et al. (2011) studied the colostrum protein uptake in the porcine gut as well as the colostrum proteome changes during the first 24 h after birth. In this study, different proteomes (sow colostrum and stomach and small intestine tissue from piglets) were studied in order to understand how the individual colostrum proteins pass through the gastrointestinal tract of the newborn piglet. Those results revealed that Igs seem to be efficiently transported, acquitted, and retained in the small intestine. In addition, β-lactoglobulin and caseins were detected at the stomach level, although they were downregulated at the intestinal level. This indicates that either low amount is taken up by the intestine or that those proteins are rapidly digested into peptides when they reach the small intestine (Danielsen et al. 2011). As previously described in this chapter, peptides derived from those two proteins have been shown to actively participate during the immune protection against several pathogens. Besides those two components, other colostrum proteins such as α-lactalbumin, prophenin-1, lactotransferrin, and peptidoglycan recognition protein, all well known for their antimicrobial properties (Zimecki and Kruzel 2007), were identified in the piglet intestine. Milk fat globule-EGF 8, which has been suggested to protect newborn piglets against *E. coli* infections (Shahriar et al. 2006) and osteopontin, a protein related to bone and immune development (Nagatomo et al. 2004; Standal et al. 2004), was also increased in the piglet intestine after colostrum intake. According to the studies of Danielsen et al. (2011) and Hernández-Castellano et al. (2015b), it seems that colostrum contains several compounds that may have an active function in the lumen of the gastrointestinal tract, whereas others may have functions in the intestinal tissue, in the bloodstream, or even in other organs. Further studies will be necessary in order to fully understand colostrum functions and properties. In this way, proteomics can provide a complementary overview of the involved pathways during processes such as colostrogenesis or colostrum absorption in the newborn animal.

5 Conclusion and Future Perspectives

Over the years, colostrum has been the focus of several studies which aimed to describe, explain, and understand the importance of colostrum in the ungulate species as well as the biological processes that take place in the newborn animal after feeding colostrum. The use of proteomics in this field has enormously increased the general knowledge of the proteins present in colostrum and their

relation to passive immune transfer and the immune system development in the newborn animal. The continuous development of new proteomic techniques will increase sensitivity and will allow the identification of very low abundance protein populations in the colostrum, mammary gland, placenta, intestine, and blood plasma. All this information will provide better and clearer overview of the diverse processes that take place during colostrogenesis, colostrum secretion, intestinal colostrum absorption, and, finally, on the newborn immune system development.

References

Actor JK, Hwang SA, Kruzel ML (2009) Lactoferrin as a natural immune modulator. Curr Pharm Design 15:1956–1973

Al-Busadah KA (2007) Efficacy of feeding bovine and caprine colostrum to neonatal camel. J Anim Vet Adv 6:5–7

Almeida AM, Bassols A, Bendixen E et al (2015) Animal board invited review: advances in proteomics for animal and food sciences. Animal 9:1–17

Argüello A, Castro N, Capote J et al (2004) Effect of colostrum administration practices on serum IgG in goat kids. Livest Prod Sci 90:235–239

Badolato R, Wang JM, Murphy WJ et al (1994) Serum amyloid-A is a chemoattractant - Induction of migration, adhesion, and tissue infiltration of monocytes and polymorphonuclear leukocytes. J Exp Med 180:203–209

Barrington GM, Parish SM (2001) Bovine neonatal immunology. Vet Clin N Am Food Anim Pract 17:463–476

Baumrucker CR, Bruckmaier RM (2014) Colostrogenesis: IgG1 transcytosis mechanisms. J Mammary Gland Biol Neoplasia 19:103–117

Beddek AJ, Rawson P, Peng L et al (2008) Profiling the metabolic proteome of bovine mammary tissue. Proteomics 8:1502–1515

Bendixen E, Danielsen M, Hollung K et al (2011) Farm animal proteomics - a review. J Proteomics 74:282–293

Blum JW, Baumrucker CR (2002) Colostral and milk insulin-like growth factors and related substances: mammary gland and neonatal (intestinal and systemic) targets. Domestic Anim Endocrinol 23:101–110

Blum JW, Baumrucker CR (2008) Insulin-like growth factors (IGFs), IGF binding proteins, and other endocrine factors in milk: role in the newborn. Adv Exp Med Biol 606:397–422

Blum JW, Hammon H (2000) Colostrum effects on the gastrointestinal tract, and on nutritional, endocrine and metabolic parameters in neonatal calves. Livest Prod Sci 66:151–159

Boyd JW (1972) The relationship between serum immune globulin deficiency and disease in calves: a farm survey. Vet Rec 90:645–649

Braun JP, Tainturier D, Bezille P et al (1983) Transfer of Ggt from mother colostrum to serum in newborn goats and foals as compared to calves. Clin Biochem 16:A29–A29

Brignole TJ, Stott GH (1979) Colostrum feeding after closure, a protection for agammaglobulinemic and hypogammaglobulinemic calves. J Dairy Sci 62:93–93

Bullen JJ, Rogers HJ, Leigh L (1972) Iron-binding proteins in milk and resistance to Escherichia coli infection in infants. Br Med J 1:69–75

Castro N, Capote J, Bruckmaier RM et al (2011) Management effects on colostrogenesis in small ruminants: a review. J Appl Anim Res 39:85–93

Cebo C (2012) Milk fat globule membrane proteomics: a ‘snapshot’ of mammary epithelial cell biology. Food Technol Biotechnol 50:306–314

Chan TF, Tsai YC, Wu CH et al (2011) The positive correlation between cord serum retinol-binding protein 4 concentrations and fetal growth. Gynecol Obstet Invest 72:98–102

Chmurzynska A (2006) The multigene family of fatty acid-binding proteins (FABPs): function, structure and polymorphism. J Appl Genetics 47:39–48

Clare DA, Swaisgood HE (2000) Bioactive milk peptides: a prospectus. J Dairy Sci 83:1187–1195

Cugno G, Parreira JR, Ferlizza E et al (2016) The goat (Capra hircus) mammary gland mitochondrial proteome: a study on the effect of weight loss using Blue-Native PAGE and two-dimensional gel electrophoresis. PLoS One 11:e0151599

Danielsen M, Pedersen LJ, Bendixen E (2011) An in vivo characterization of colostrum protein uptake in porcine gut during early lactation. J Proteomics 74:101–109

Davies CR, Morris JS, Griffiths MR et al (2006) Proteomic analysis of the mouse mammary gland is a powerful tool to identify novel proteins that are differentially expressed during mammary development. Proteomics 6:5694–5704

De Maio A (1999) Heat shock proteins: facts, thoughts, and dreams. Shock 11:1–12

Declerck I, Dewulf J, Sarrazin S et al (2016) Long-term effects of colostrum intake in piglet mortality and performance. J Anim Sci 94:1633–1643

Desrivieres S, Prinz T, Laria NCP et al (2003) Comparative proteomic analysis of proliferating and functionally differentiated mammary epithelial cells. Mol Cell Proteomics 2:1039–1054

Ebrahim F, Shankaranarayanan JS, Kanwar JR et al (2014) Identification of unprecedented anticancer properties of high molecular weight biomacromolecular complex containing bovine lactoferrin (HMW-bLf). PLoS One 9:e106568

Elfstrand L, Lindmark-Mansson H, Paulsson M et al (2002) Immunoglobulins, growth factors and growth hormone in bovine colostrum and the effects of processing. Int Dairy J 12:879–887

El-Zahar K, Sitohy M, Choiset Y et al (2004) Antimicrobial activity of ovine whey protein and their peptic hydrolysates. Milchwissenschaft 59:653–656

Erhard MH, Luft C, Remler HP et al (2001) Assessment of colostral transfer and systemic availability of immunoglobulin G in new-born foals using a newly developed enzyme-linked immunosorbent assay (ELISA) system. J Anim Physiol Anim Nutr 85:164–173

Farmer C (2013) Review: Mammary development in swine: effects of hormonal status, nutrition and management. Can J Anim Sci 93:1–7

Ford C, Chevalier S (1995) DNA-Replication - ALMOST licensed. Curr Biol 5:1009–1012

Gevaert K, Vandekerckhove J (2011) Gel-free proteomics methods and protocols. Humana Press, New York, NY

Guilloteau P, Huerou-Luron IL, Chayvialle JA et al (1997) Gut regulatory peptides in young cattle and sheep. Zentralbl Veterinarmed A 44:1–23

Hadorn U, Hammon H, Bruckmaier RM et al (1997) Delaying colostrum intake by one day has important effects on metabolic traits and on gastrointestinal and metabolic hormones in neonatal calves. J Nutr 127:2011–2023

Harris P, Johannessen KM, Smolenski G et al (2010) Characterisation of the anti-microbial activity of bovine milk ribonuclease4 and ribonuclease5 (angiogenin). Int Dairy J 20:400–407

Hernández-Castellano LE, Almeida AM, Castro N et al (2014a) The colostrum proteome, ruminant nutrition and immunity: a review. Curr Protein Pept Sci 15:64–74

Hernández-Castellano LE, Almeida A, Ventosa M et al (2014b) The effect of colostrum intake on blood plasma proteome profile in newborn lambs: low abundance proteins. BMC Vet Res 10:85

Hernández-Castellano LE, Morales-delaNuez A, Sanchez-Macias D et al (2015a) The effect of colostrum source (goat vs. sheep) and timing of the first colostrum feeding (2h vs. 14h after birth) on body weight and immune status of artificially reared newborn lambs. J Dairy Sci 98:204–210

Hernández-Castellano LE, Arguello A, Almeida AM et al (2015b) Colostrum protein uptake in neonatal lambs examined by descriptive and quantitative liquid chromatography-tandem mass spectrometry. J Dairy Sci 98:135–147

Hernández-Castellano LE, Suarez-Trujillo A, Martell-Jaizme D et al (2015c) The effect of colostrum period management on BW and immune system in lambs: from birth to weaning. Animal 9:1672–1679

Hernández-Castellano LE, Almeida AM, Renaut J et al (2016a) A proteomics study of colostrum and milk from the two major small ruminant dairy breeds from the Canary Islands: a bovine milk comparison perspective. J Dairy Res 83:366–374

Hernández-Castellano LE, Ferreira AM, Nanni P et al (2016b) The goat (Capra hircus) mammary gland secretory tissue proteome as influenced by weight loss: a study using label free proteomics. J Proteomics 145:60–69

Hernandez-Ledesma B, Recio I, Amigo L (2008) Beta-lactoglobulin as source of bioactive peptides. Amino Acids 35:257–265

Herosimczyk A, Lepczynski A, Ozgo M et al (2013) Blood plasma protein and lipid profile changes in calves during the first week of life. Pol J Vet Sci 16:425–434

Hesselager MO, Codrea MC, Sun Z et al (2016) The pig peptide atlas: a resource for systems biology in animal production and biomedicine. Proteomics 16:634–644

Horsley SW, Knight SJ, Nixon J et al (1998) Del(18p) shown to be a cryptic translocation using a multiprobe FISH assay for subtelomeric chromosome rearrangements. J Med Genet 35:722–726

Jena MK, Janjanam J, Naru J et al (2015) DIGE based proteome analysis of mammary gland tissue in water buffalo (Bubalus bubalis): lactating vis-a-vis heifer. J Proteomics 119:100–111

Kanai Y, Miura K, Uehara T et al (1993) Natural occurrence of nuc in the sera of autoimmune-prone Mrl Lpr mice. Biochem Biophys Res Commun 196:729–736

Keenan TW, Dylewski DP (1995) Intracellular origin of milk lipid globules and the nature and structure of the milk lipid globule membrane. In: Advanced dairy chemistry. Lipids, vol 2. Chapman & Hall, New York, pp 89–130

Kelly D, King TP, McFadyen M et al (1993) Effect of preclosure colostrum intake on the development of the intestinal epithelium of artificially reared piglets. Biol Neonate 64:235–244

Kirby DR, Bradbury S (1965) The hemo-chorial mouse placenta. Anat Rec 152:279–281

Klobasa F, Werhahn E, Butler JE (1987) Composition of sow milk during lactation. J Anim Sci 64:1458–1466

Kolb E, Kaskous S (2003) Constituents of the colostrum and the milk of goats and their significance for the health of kids (a review). Tierarztl Umschau 58:140–146

Koldovsky O (1980) Hormones in milk. Life Sci 26:1833–1836

Koldovsky O (1989) Search for role of milk-borne biologically active peptides for the suckling. J Nutr 119:1543–1551

Korhonen H, Pihlanto A (2007) Technological options for the production of health-promoting proteins and peptides derived from milk and colostrum. Curr Pharm Design 13:829–843

Kubota T, Miyauchi M, Miura K et al (1998) Upregulation of nucleobindin expression in human-activated lymphocytes and non-Hodgkin's lymphoma. Pathol Int 48:22–28

Kuhns DB, Nelson EL, Alvord WG et al (2001) Fibrinogen induces IL-8 synthesis in human neutrophils stimulated with formyl-methionyl-leucyl-phenylalanine or leukotriene B(4). J Immunol 167:2869–2878

Kurosaki M, Zanotta S, Calzi ML et al (1996) Expression of xanthine oxidoreductase in mouse mammary epithelium during pregnancy and lactation: regulation of gene expression by glucocorticoids and prolactin. Biochem J 319:801–810

Lappas M, Hiden U, Desoye G et al (2011) The role of oxidative stress in the pathophysiology of gestational diabetes mellitus. Antioxid Redox Signal 15:3061–3100

Le A, Barton LD, Sanders JT et al (2011) Exploration of bovine milk proteome in colostral and mature whey using an ion-exchange approach. J Proteome Res 10:692–704

Lee DG, Nam J, Kim SW et al (2015) Proteomic analysis of reproduction proteins involved in litter size from porcine placenta. Biosci Biotech Bioch 79:1414–1421

Leece JG (1973) Effect of dietary regimen on cessation of uptake of macromolecules by piglet intestinal epithelium (closure) and transport to the blood. J Nutr 103:751–756

Lerias JR, Hernández-Castellano LE, Suárez-Trujillo A et al (2014) The mammary gland in small ruminants: major morphological and functional events underlying milk production—a review. J Dairy Res 81:304–318

Lopez C (2011) Milk fat globules enveloped by their biological membrane: unique colloidal assemblies with a specific composition and structure. Curr Opin Colloid Interface Sci 16:391–404

Lu J, Liu L, Pang XY et al (2016) Comparative proteomics of milk fat globule membrane in goat colostrum and mature milk. Food Chem 209:10–16

Mahley RW, Innerarity TL, Rall SC et al (1984) Plasma-lipoproteins - Apolipoprotein Structure and gunction. J Lipid Res 25:1277–1294

Marnila P, Korhonen H (2011) Milk | colostrum. In: Fuquay FW, Fox PF, McSweeney PLH (eds) Encyclopedia of dairy sciences, 2nd edn. Academic Press, San Diego, pp 591–597

Mather IH (2000) A review and proposed nomenclature for major proteins of the milk-fat globule membrane. J Dairy Sci 83:203–247

Moore M, Tyler JW, Chigerwe M et al (2005) Effect of delayed colostrum collection on colostral IgG concentration in dairy cows. J Am Vet Med Assoc 226:1375–1377

Moretti DB, Pauletti P, Kindlein L et al (2010) Enteric cell proliferation in newborn lambs fed bovine and ovine colostrum. Livest Sci 127:262–266

Moretti DB, Nordi WM, Lima AL et al (2014) Enteric, hepatic and muscle tissue development of goat kids fed with lyophilized bovine colostrum. J Anim Physiol Anim Nutr 98:201–208

Murata M, Wakabayashi H, Yamauchi K et al (2013) Identification of milk proteins enhancing the antimicrobial activity of lactoferrin and lactoferricin. J Dairy Sci 96:4891–4898

Nagatomo T, Ohga S, Takada H et al (2004) Microarray analysis of human milk cells: persistent high expression of osteopontin during the lactation period. Clin Exp Immunol 138:47–53

Nguyen TV, Yuan L, Azevedo MSP et al (2007) Transfer of maternal cytokines to suckling piglets: in vivo and in vitro models with implications for immunomodulation of neonatal immunity. Vet Immunol Immunopathol 117:236–248

Nissen A, Bendixen E, Ingvartsen KL et al (2012) In-depth analysis of low abundant proteins in bovine colostrum using different fractionation techniques. Proteomics 12:2866–2878

Nowak R, Poindron P (2006) From birth to colostrum: early steps leading to lamb survival. Reprod Nutr Dev 46:431–446

Ontsouka CE, Bruckmaier RM, Blum JW (2003) Fractionized milk composition during removal of colostrum and mature milk. J Dairy Sci 86:2005–2011

Osada T, Watanabe G, Kondo S et al (2001) Male reproductive defects caused by puromycin-sensitive aminopeptidase deficiency in mice. Mol Endocrinol 15:960–971

Padua MB, Kowalski AA, Canas MY et al (2010) The molecular phylogeny of uterine serpins and its relationship to evolution of placentation. Faseb J 24:526–537

Pakkanen R, Aalto J (1997) Growth factors and antimicrobial factors of bovine colostrum. Int Dairy J 7:285–297

Parks DA, Granger DN (1986) Xanthine-Oxidase - Biochemistry, distribution and Physiology. Acta Physiol Scand 126:87–99

Penchev-Georgiev I (2008) Effect of colostrum insulin-like growth factors on growth and development of neonatal calves. Bulg J Vet Med 11:75–88

Peterson JA, Patton S, Hamosh M (1998) Glycoproteins of the human milk fat globule in the protection of the breast-fed infant against infections. Biol Neonate 74:143–162

Pihlanto-Leppala A, Marnila P, Hubert L et al (1999) The effect of alpha-lactalbumin and beta-lactoglobulin hydrolysates on the metabolic activity of Escherichia coli JM103. J Appl Microbiol 87:540–545

Porter P (1973) Studies of porcine secretory IgA and its component chains in relation to intestinal absorption of colostral immunoglobulins by the neonatal pig. Immunology 24:173–176

Prystowsky H (1958) Fetal blood studies .9. Some aspects of oxygen transfer across the hemochorial placenta of the human in postmature pregnancy. Obstet Gynecol 12:164–167

Purup S, Vestergaard M, OPedersen L et al (2007) Biological activity of bovine milk on proliferation of human intestinal cells. J Dairy Res 74:58–65
Reinhardt TA, Lippolis JD (2008) Developmental changes in the milk fat globule membrane proteome during the transition from colostrum to milk. J Dairy Sci 91:2307–2318
Renckens R, Roelofs JJTH, Florquin S et al (2006) Urokinase-type plasminogen activator receptor plays a role in neutrophil migration during lipopolysaccharide-induced peritoneal inflammation but not during Escherichia coli-induced peritonitis. J Infect Dis 193:522–530
Sasaki M, Davis CL, Larson BL (1976) Production and turnover of IgG1 and IgG2 immunoglobulins in the bovine around oarturition1. J Dairy Sci 59:2046–2055
Scammel AW (2001) Production and uses of colostrum. Aust J Dairy Technol 56:74–82
Schroten H, Hanisch FG, Plogmann R et al (1992) Inhibition of adhesion of S-fimbriated Escherichia coli to buccal epithelial cells by human milk fat globule membrane components - a novel aspect of the protective function of mucins in the nonimmunoglobulin fraction. Infect Immun 60:2893–2899
Scoot A, Owen BD, Agar JL (1972) Influence of orally administered porcine immunoglobulins on survival and performance of newborn colostrum-deprived pigs. J Anim Sci 35:1201–1205
Senda A, Fukuda K, Ishii T et al (2011) Changes in the bovine whey proteome during the early lactation period. Anim Sci J 82:698–706
Sgarbieri VC (2004) Propriedades fisiológicas-funcionais das proteínas do soro de leite. Revista de Nutrição 17:397–409
Shahriar F, Ngeleka M, Gordon JR et al (2006) Identification by mass spectroscopy of F4ac-fimbrial-binding proteins in porcine milk and characterization of lactadherin as an inhibitor of F4ac-positive Escherichia coli attachment to intestinal villi in vitro. Dev Comp Immunol 30:723–734
Silanikove N, Shapiro F, Shamay A et al (2005) Role of xanthine oxidase, lactoperoxidase, and NO in the innate immune system of mammary secretion during active involution in dairy cows: manipulation with casein hydrolyzates. Free Radic Biol Med 38:1139–1151
Silva SV, Malcata FX (2005) Caseins as source of bioactive peptides. Int Dairy J 15:1–15
Smith GW, Foster DM (2007) Short communication: absorption of protein and immunoglobulin G in calves fed a colostrum replacer. J Dairy Sci 90:2905–2908
Smolenski G, Haines S, Kwan FYS et al (2007) Characterisation of host defence proteins in milk using a proteomic approach. J Proteome Res 6:207–215
Stan S, Delvin E, Lambert M et al (2003) Apo A-IV: an update on regulation and physiologic functions. Biochim Biophys Acta 1631:177–187
Standal T, Borset M, Sundan A (2004) Role of osteopontin in adhesion, migration, cell survival and bone remodeling. Exp Oncol 26:179–184
Stelwagen K, Carpenter E, Haigh B et al (2009) Immune components of bovine colostrum and milk. J Anim Sci 87:3–9
Talbot NC, Powell AM, Caperna TJ et al (2010) Proteomic analysis of the major cellular proteins of bovine trophectoderm cell lines derived from IVP, parthenogenetic and nuclear transfer embryos: reduced expression of annexins I and II in nuclear transfer-derived cell lines. Anim Reprod Sci 120:187–202
Theil PK, Hurley WL (2016) The protein component of sow colostrum and milk. In: Milk proteins - from structure to biological properties and health aspects. InTech, Rijeka, Croatia, pp 183–198
Tiaden AN, Breiden M, Mirsaidi A et al (2012) Human serine protease HTRA1 positively regulates osteogenesis of human bone marrow-derived mesenchymal stem cells and mineralization of differentiating bone-forming cells through the modulation of extracellular matrix protein. Stem Cells 30:2271–2282
Tyler JW, Hancock DD, Thorne JG et al (1999) Partitioning the mortality risk associated with inadequate passive transfer of colostral immunoglobulins in dairy calves. J Vet Intern Med 13:335–337

Vaerman JP, Arbuckle JB, Heremans JF (1970) Immunoglobulin A in the pig. II. Sow's colostral and milk IgA: quantitative studies and molecular size estimation. Int Arch Allergy Immunol 39:323–333

Westermeier R, Naven T (2002) Expression proteomics. In: Proteomics in practice. Wiley-VCH Verlag GmbH, Weinheim, pp 11–160

Wooding P, Burton G (2008) Synepitheliochorial placentation: ruminants (ewe and cow). In: Comparative placentation: structures, functions and evolution. Springer, Berlin, pp 133–167.

Wooding FB, Flint AP, Heap RB et al (1986) Control of binucleate cell migration in the placenta of sheep and goats. J Reprod Fertil 76:499–512

Xu L, Badolato R, Murphy WJ et al (1995) A novel biologic function of serum amyloid A. Induction of T lymphocyte migration and adhesion. J Immunol 155:1184–1190

Yang M, Zou Y, Wu ZH et al (2015) Colostrum quality affects immune system establishment and intestinal development of neonatal calves. J Dairy Sci 98:7153–7163

Yang M, Cong M, Peng XM et al (2016) Quantitative proteomic analysis of milk fat globule membrane (MFGM) proteins in human and bovine colostrum and mature milk samples through iTRAQ labeling. Food Funct 7:2438–2450

Zhang L, Boeren S, Hageman JA et al (2015) Bovine milk proteome in the first 9 days: protein interactions in maturation of the immune and digestive system of the newborn. PLoS One 10: e0116710

Zimecki M, Kruzel ML (2007) Milk-derived proteins and peptides of potential therapeutic and nutritive value. J Exp Ther Oncol 6:89–106

Proteomics in Milk and Dairy Products

Alessio Soggiu, Paola Roncada, and Cristian Piras

Abstract Milk from different animal species is a valuable source of proteins and other nutrients for human consumption. Its dynamic range of proteoforms among animal species provides different nutritional profiles. Milk can be used plain or after transformation processes such as fermentation, caseification, ripening, and many others. This chapter will describe the actual role of proteomics to study and address safety, quality, and transformation processes.

The large-scale study of proteins (proteomics) is able to provide information about different protein profiles, characteristics of milk, nutrients, lactation stage, and the relative health status of the animal. More extensively, proteomics is useful for the assessment of safety and quality of both milk and dairy products.

The aim of this chapter is to summarize the most recent applications of proteomics, including the methods for sample preparation and analysis, and to make them available and applicable for both research and analytical purposes.

Keywords Milk • Dairy products • Proteomics • Peptides • Cheese • Ripening • Quality • Microbiota

List of Abbreviations

1-DE	One-dimensional electrophoresis
2-D DIGE	Two-dimensional difference gel electrophoresis
2-DE	Two-dimensional electrophoresis
CE	Capillary electrophoresis
CEX	Cation exchange
CFU	Colony-forming units
DGGE	Denaturing gradient gel electrophoresis

A. Soggiu (✉) • C. Piras
Department of Veterinary Medicine, University of Milan, Milan, Italy
e-mail: alessio.soggiu@unimi.it

P. Roncada
Istituto Sperimentale Italiano "Lazzaro Spallanzani", Milan, Italy

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_9

DNA	Deoxyribonucleic acid
ESI	Electrospray ionization
GO	Gene ontology
HIC	Hydrophobic interaction chromatography
HPLC	High-performance liquid chromatography
IEF	Isoelectric focusing
iTRAQ	Isobaric tags for relative and absolute quantitation
KEGG	Kyoto Encyclopedia of Genes and Genomes
LAB	Lactic acid bacteria
LC	Liquid chromatography
MALDI-TOF MS	Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MFGM	Milk fat globule membrane
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MW	Molecular weight
NGS	Next-generation sequencing
NSLAB	Nonstarter lactic acid bacteria
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PTM	Posttranslational modification
QqQ-MS	Triple quadrupole mass spectrometry
qTOF	Quadrupole time of flight
RNA	Ribonucleic acid
RP	Reverse phase
rRNA	Ribosomal RNA
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SRM	Selected reaction monitoring
SSCP	Single-strand conformation polymorphism
TCA	Trichloroacetic acid
UHT	Ultrahigh temperature processing
UPLC	Ultra performance liquid chromatography
αs1-CN	Alpha s1 casein
αs2-CN	Alpha s2 casein
β-CN	Beta casein

Dairy products include all foods that are made from milk and are used for human nutrition. Among the main species used as source of milk there are cows, water buffaloes, sheep, goats, donkeys and camels. Milk from these animals can be used raw or after transformation processes such as fermentation, caseification, ripening, and many others. This chapter describes the actual role of proteomics to study and to address safety, quality, and transformation processes.

1 Analytical Techniques for Milk and Dairy Products

Milk from different animal species represents an extremely valuable source of proteins and other nutrients in human nutrition.

It is composed by a dynamic range of proteoforms different in all animal species and responsible for several specific nutritional profiles.

The large-scale study of proteins (proteomics) is able to provide information about different protein profiles, characteristics of milk, nutrients, lactation stage, and the health status of the animal during this period (mastitis or other bacterial infections) (Li et al. 2017).

Because of its importance in the analysis of milk proteoforms, one of the fastest and easiest ways to analyze the milk proteome is 2D electrophoresis. This method represents a top-down approach able to study the entire proteins, and it requires protein fragmentation just for the identification process (MALDI-TOF MS) (Wang et al. 2006). On the other hand, as bottom-up approach, LC-MS/MS represents a technique able to deeply analyze milk proteome despite some loss of information about the different proteoforms (Zhang et al. 2015; Nissen et al. 2012).

In order to deeply analyze the milk proteome and to maximize protein identification, it is necessary to use casein depletion (Soggiu et al. 2016) or other combined fractionation techniques (Hogarth et al. 2004; Mollé et al. 2009).

The protein components of milk play a key role in the transformation process and are responsible for the characteristics of the final product. Protein fractions of milk can be divided into two main groups: caseins (CN) and whey proteins. The main protein constituents are caseins (alpha, beta, gamma, and kappa), which account for about 80% of the total protein. The remaining 20% of the protein fraction consist of the whey proteins, represented mostly by β-lactoglobulins, α-lactalbumins, immunoglobulins, serum albumin, and proteose peptones. Besides, there are other constituents like free amino acids, urea, and creatine, among others (Fig. 1).

The protein content of cheese is defined as the content of total nitrogen (protein and nonprotein nitrogen). The protein nitrogen originates primarily from denatured casein deriving from acid or enzymatic coagulation and from a small part of whey proteins that, while being soluble in serum, are still incorporated in the casein matrix and from insoluble peptides. The nonprotein nitrogen instead comes from medium-soluble peptides and small-weight molecules like amino acids. Routinely, total nitrogen is determined by the reference methods according to Kjeldahl (Lynch and Barbano 1999) or Dumas (Simonne et al. 1997), but both methods are susceptible to adulterations. The Kjeldahl method in particular is a tedious and time-consuming procedure requiring disposal of hazardous chemicals.

On the other hand, direct protein can be quantitated using different strategies. The dye-binding methods utilize the dyes Coomassie Brilliant Blue G-250 (according to Bradford), Amido Black 10B, or Orange G that specifically bind to proteins. The protein content is determined by measuring the intensity of the dye color, which correlates to the protein concentration. Nowadays these methods are

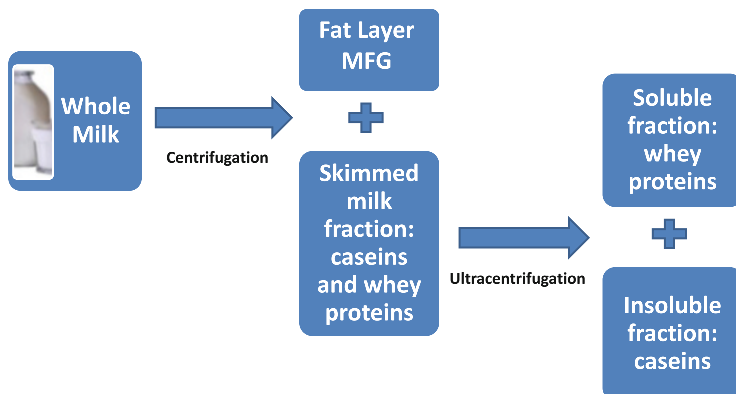

Fig. 1 Schematic representation of the fractionation of milk components and their protein fractions

used exclusively in the laboratory and have been replaced in food industry by infrared spectroscopy that measures directly the protein and peptide content, together with lipids and sugars (Fox et al. 2015).

The chemistry of proteins, in the past, and now proteomics, allowed the analysis of specific protein components in dairy products by using different and complementary analytical techniques with different selectivity and specificity. Electrophoretic and chromatographic techniques, often in combination with mass spectrometry, are widely used to qualitatively and quantitatively analyze proteins, peptides, and amino acids in dairy products. Several analytical applications related to those techniques will be summarized in Sects. 1.2 and 2.1.

1.1 Dynamic Range of Proteins in Milk and Cheese and Depletion Strategies for Main Components

Mature cow's milk contains on average 35 g/l of proteins; of these about 80% are caseins, and about 20% consist of whey proteins. Different fractions of milk caseins exist. The most abundant of these is α-casein (a heterogeneous protein, with numerous variants, such as α-s1 and α-s2), followed by β-casein, κ-casein, and γ-casein (resulting from the degradation of β-casein). The three major serum proteins are β-lactoglobulin, α-lactalbumin, and serum albumin; they represent 50, 20, and 10%, respectively, of the whey protein fraction. The remaining 20% is composed of immunoglobulins and proteins with enzymatic or antibacterial activity. In sheep, the casein fraction represents around 80%, 50% in horse and <50% in human. Whey proteins are proportionally higher in human and nonruminants (horse, donkey) but lower in ruminants. A qualitative and quantitative description

of the protein fractions of mature milk from different species was reported (Claeys et al. 2014).

A comprehensive analysis made on unfractionated whole milk obtained from five different species has been made using 1-DE and 2-DE coupled with MS (Hinz et al. 2012). Fractionation of milk, intended as the separation of the casein fraction from the whey fraction, can be obtained principally using three different techniques: chymosin-induced separation, isoelectric precipitation, and ultracentrifugation. An investigation has been made to evaluate the protein composition obtained from different fractionation strategies using 2-DE and tandem mass spectrometry (Jensen et al. 2012). Other approaches in dairy animals have been directed to the characterization of specific fractions like whey (Yang et al. 2013), casein (Wang et al. 2017), or MFGM (Yang et al. 2015) using specific LC-MS/MS approaches.

Cheese is the product obtained from the coagulation of casein in milk (curd). Major differences in domestic animals (cow, water buffalo, goat, and sheep, listed according to their importance in milk production) are the concentration and types of caseins and their ratio to the whey proteins in milk. Diversity in milk composition, especially in casein profiles, produces differences in cheese proteome profile, because whey proteins do not precipitate in the curd (or in very low amounts) during milk transformation. The processing techniques and the ripening substantially affect the protein content of dairy products: in fresh cheese (Mozzarella, cottage cheese) and soft cheese, the protein content lies between 18 and 25 g/100 g, while it can be 30 g/100 g or more in hard cheese (Grana Padano and Parmigiano-Reggiano cheese). In whey products like Ricotta, the protein content ranges from 8.8 to 11.3 g/100 g. Other dairy products are butter (0.9 g/100 g), yogurt (3.5 g/100 g), fermented milks (2–5.6 g/100 g), and dry milk (26–36 g/100 g) (Muehlhoff et al. 2013).

While the cheese caseins are insoluble (insoluble nitrogen) in most aqueous solvents, medium-sized and short peptides as degradation products can be dissolved, and their concentration increases with the progress of proteolysis. A large number of analytical techniques for fractionation of the protein component in cheese take advantage of these solubilization properties. The need to separate the highly abundant components of the cheese (caseins) from the less abundant (whey proteins, proteins released by bacteria, peptides) requires the application of specific fractionation and depletion techniques. These techniques are fundamental complements to the proteomic analysis in order to characterize the various protein components in cheese. All caseins in Cheddar, for example, can be easily separated from the soluble fraction using an extraction with water at a ratio of 2:1 (water/grated cheese) at 40 °C for 1 h. In this way it is possible to recover up to 90% of the nitrogen soluble in the aqueous phase and caseins in the pellet after homogenization and centrifugation of the resulting mixture (Kuchroo and Fox 1982). This technique can be easily adapted to other kinds of cheese like Grana Padano (Soggiu et al. 2016), a hard-cooked cheese, or Emmenthal (Gagnaire et al. 2004).

Quantitative casein depletion and isolation from cheese can be obtained using an acid or so-called isoelectric precipitation of the aqueous extracts at pH 4.6 with 1M HCl. At this pH that corresponds roughly to their isoelectric point, casein

coagulation and aggregation can be easily obtained at room temperature, and aggregates can be sedimented by low-speed centrifugation (Molina et al. 2000; Fox et al. 2015). If the scope of the investigation is proteomic analysis of caseins, the pellet has to be washed with acidified water to remove salts, and a wash with an organic solvent is required to remove lipids (Molina et al. 2000). Most of the caseins and the deriving high molecular weight peptides are soluble in aqueous 4–8 molar urea solution. This characteristic can be used in the preparation of samples to be subjected to subsequent separation techniques (Fox et al. 2015).

Hydraulic extraction with a press can be used to obtain an aqueous phase from the entire cheese. In this way it is possible to process in one experiment a relatively high sample amount (2–2.5 kg) obtaining a "juice" enriched in soluble proteins (such as lactoferrin, bovine serum albumin, β-lactoglobulin, and α-lactalbumin as well as bacterial proteins). Using the same "juice" or samples with similar composition, it is possible to separate caseins and casein peptides from whey proteins and bacterial proteins and simultaneously desalt the sample using size-exclusion chromatography on Superdex 75 in ammonium acetate buffer (Gagnaire et al. 2004).

The soluble nitrogen fraction (medium to low molecular weight peptides and whey proteins) can be obtained also by extraction and subsequent precipitation with suitable acids on already depleted samples.

Trichloroacetic acid (TCA) is widely used as a precipitating agent of proteins and peptides as well as trifluoroacetic acid or phosphotungstic acid. A further fractionation of cheese peptides before precipitation can be obtained using solid phase extraction on C-18 resin in combination with ultrafiltration using membranes with diverse molecular weight cutoffs (Fox et al. 2015; O'Sullivan and Fox 1990; Singh et al. 1994).

A schematic summary of the possible processing strategies of milk and dairy products is represented in Fig. 2. As will be described below, besides pre-fractionation techniques, there are many other valuable ways to properly perform a proteomic analysis of milk and dairy products, such as the choice among a gel-based and a gel-free approach.

1.2 Gel-Based and Gel-Free Approaches: Selected Applications

Milk is a biological fluid produced by the mammary gland, and it is important for the nourishment of the offspring. Its nutraceutical properties are eminent, as milk is an essential source of proteins, amino acids, and nutraceutical peptides (Verma and Ambatipudi 2016). In order to analyze the complex protein composition of milk and dairy products, beyond depletion strategies, several different proteomic approaches may be used which may benefit from protein separation in gels and from electrophoresis steps. Also in mass spectrometric-based analyses, additional electrophoretic steps prior to MS may be of value as a pre-fractionation step to reduce sample

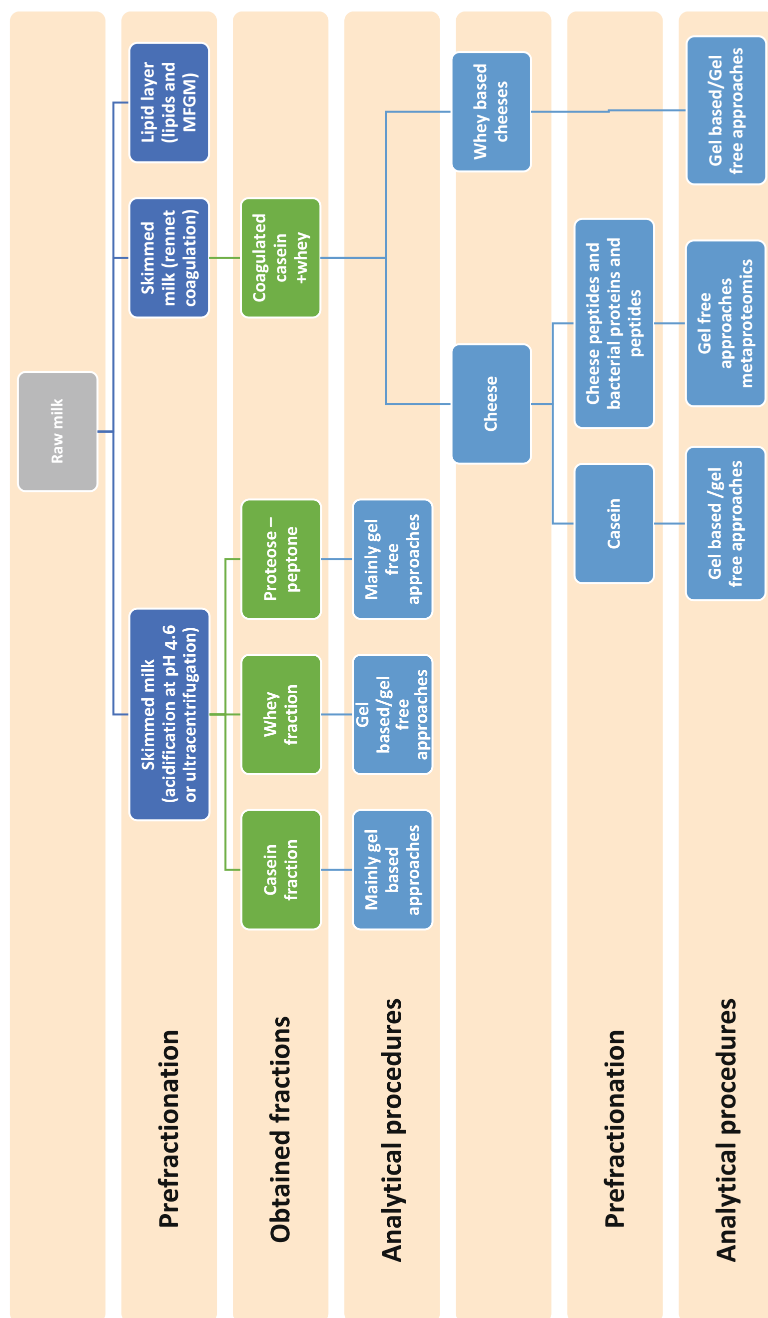

Fig. 2 Image of the workflow summarizing the techniques to be applied to milk and dairy products processing in order to properly analyze the different protein fractions

complexity. This separation can be performed according to the molecular weight of the proteins, their isoelectric points, or both. In either way it is possible to provide a first important classification in the study of these complex protein matrices.

1.2.1 Gel-Based Approaches

As previously mentioned, gel-based approaches are able to quickly provide a visualization of milk protein content in their entire form without a previous tryptic digestion.

Figure 3 shows a representative 1D electrophoresis (Fig. 3a) and a 2D electrophoresis (Fig. 3b) of skimmed cow milk. Main milk proteins identified by mass spectrometry are indicated in 1-DE and 2-DE.

For example, lately, this technique has been successfully applied for the study of consequences of heat treatment on camel milk. Authors used 2-D DIGE to evaluate the changes to the proteomics pattern in the sample at room temperature and after thermal treatment at 63 and 98 °C (Benabdelkamel et al. 2017).

Bacterial infection could as well be responsible of changing the proteome profile of milk. About this topic, 2-DE and LC-MS/MS showed that *S. agalactiae* subclinical mastitis altered the protein profiles of cow milk (Pongthaisong et al. 2016).

Gel-based approaches have been as well widely used in several applications for the characterization of dairy products. One example of successful application is

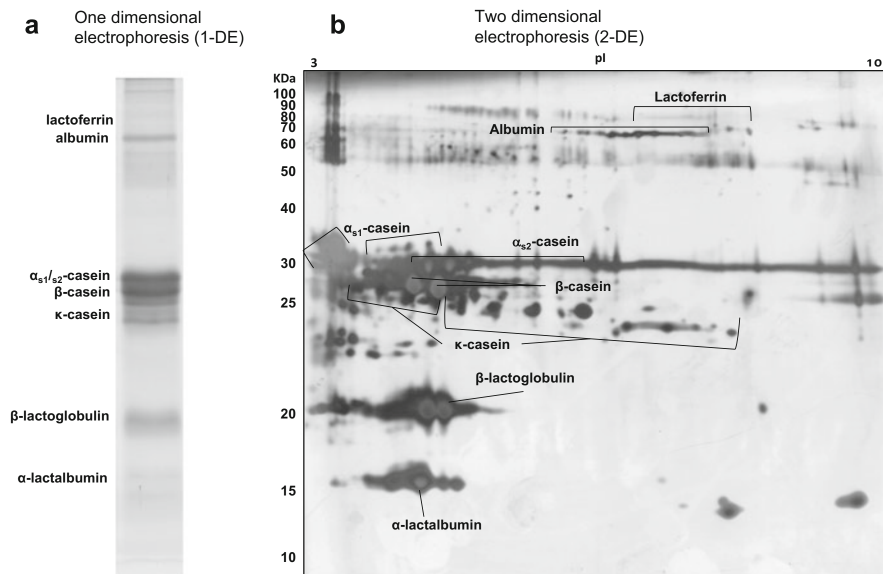

Fig. 3 Representative 1D electrophoresis (**a**) and 2D electrophoresis (**b**) of skimmed cow milk. Main milk proteins identified by mass spectrometry are indicated in 1-DE and 2-DE

related to the study of adulterations, for instance, in Mozzarella cheese. More precisely, one-dimensional electrophoretic techniques like sodium dodecyl sulfate-PAGE and alkaline urea-PAGE were successfully employed to the evaluation of freshness of the raw materials used in the manufacturing of buffalo Mozzarella cheese (Petrella et al. 2015).

Similarly, one-dimensional electrophoretic approaches were successful in the characterization of artisanal cheese as "Coalho" cheese. The authors performed SDS-PAGE of the aqueous extracts of "Coalho" cheese and were able to attribute the protein bands of lactoferrin, β-lactoglobulin, β-lactoglobulin (dimer), α-lactalbumin, bovine serum albumin, α-casein, β-casein, κ-casein, and para-κ-casein. The identified proteomic patterns can be used for the characterization of this artisanal product (Silva et al. 2016).

From another point of view, gel-based approaches like 2D electrophoresis can be used for the thorough proteomic evaluation of microorganisms involved in cheese manufacturing or spoilage as *Mucor* species. In this case, 2D electrophoresis has been successfully applied to the study of species-specific proteins responsible for the adaptive process of *Mucor* to cheese environment (Morin-Sardin et al. 2017).

1.2.2 Gel-Free Approaches

Since bovine milk and related products are extensively used for human nutrition purposes, it is consistent to analyze the differences of bovine colostrum and mature milk in comparison to human. Among the last examples, an advanced iTRAQ approach coupled with KEGG and GO analysis was used and underlined the differences among these specimens (Yang et al. 2017a). Interestingly, many of the differentially represented functions were related to immune and regulatory functions. The differences between bovine colostrum and milk have been investigated by applying two-dimensional liquid chromatography MS/MS in order to investigate which proteins could be most relevant to neonatal calf health. This important study highlighted the presence of 50 proteins specific for colostrum and 13 proteins specific for mature milk (Nissen et al. 2017).

Another iTRAQ proteomic approach made it possible to identify and quantify 920 milk exosome proteins. This study describes the results of a differential proteomic analysis performed among human and bovine exosomes and represents a good example of the successful application of this method to study this milk fraction (Yang et al. 2017b).

The concern about food fraud is increasing especially in dairy products, because protein changes induced by their transformation processes make analyses of the products more difficult. One of the most common frauds is related to the addition of cow milk to goat and sheep milk in dairy products for economic reasons. An interesting approach to solve this problem utilizes UPLC-QqQ-MS for the presence/absence of the tryptic peptides LSFNPTQLEEQCHI and LAFNPTQLEGQCHV to differentiate between cow, sheep, and goat milks. This approach using two peptides

has been proven to be more reliable and did not require a previous electrophoresis step for sample preparation (Chen et al. 2016).

Another interesting approach to discern milks of different species in cheese manufacturing describes the use of HPLC-tandem mass spectrometry. In this case, a multiple reaction monitoring method has been successfully applied for the detection of the four species most important in Italian dairy industry (cow, buffalo, sheep, and goat) (Bernardi et al. 2015).

Bacterial consortia responsible for cheese manufacturing and ripening represent another important field to be studied through proteomic techniques. The study of bacterial communities through proteomics has been made possible because of the recent advances in metaproteomics. This technique based on two-dimensional chromatography coupled with tandem mass spectrometry can be successfully applied in the fields of both food safety and food traceability. Its help made it possible to characterize the differential microbiota composition responsible for the late blowing of Grana Padano cheese (Soggiu et al. 2016).

The described approaches reflect the most recent advances in proteomic-based food science and the study of dairy products and are useful for the investigation of food quality and safety.

2 Proteomics to Investigate Milk and Dairy Products' Safety, Authenticity, and Quality

Safety and quality of milk represent essential features both for its consumption before and after processing. One of the major concerns about milk hygiene and quality is related to the presence of mastitis (Coulona et al. 2002). Milk with a high number of somatic cells is, with high probability, collected from cows affected by mastitis and shows different concentrations of non-casein fractions, salts, free fatty acids, casein content, and lactose (Ogola et al. 2007). The quality of this raw ingredient is therefore really important for the quality of all the derived dairy products.

For the analysis of these traits of food, it is important to analyze not only protein patterns, but even posttranslational modifications and, if required, the presence of allergens. Figure 4 presents a workflow of proteomics and data analysis able to provide information about all characteristics. All the samples to be analyzed must undergo through protein extraction and protein separation (usually gel based) or peptide separation (LC) before being analyzed through MS. After protein separation is possible to investigate posttranslational modifications using Western blot, specific staining, or mass spectrometry and database analysis (Fig. 4).

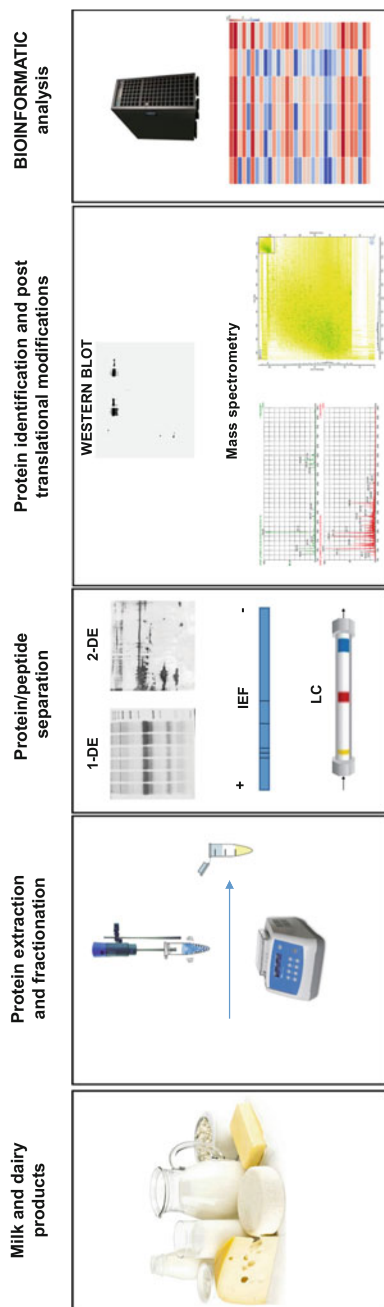

Fig. 4 Proteomics workflow for milk and dairy products analysis (modified from Piras et al. 2016)

2.1 Safety: Detection of Pathogens in Milk and Dairy Samples

Investigating the presence of bacteria responsible of spoilage and of hampering the food safety of milk represents an important task to be achieved. Actual methods based on sequencing are quite costly and time-consuming (Pongthaisong et al. 2016), and the proteomic community is putting a lot of efforts to provide tools to enhance pathogen detection.

One example of a technique developed to overcome this problem describes enhanced detection of bacteria through the use of MALDI-TOF MS coupled with a web tool for hierarchical clustering of peptide mass spectra (SPECLUST) (Vithanage et al. 2017). This method is based on the analysis of MALDI-TOF MS dataset matching against a consensus peak list of potential biomarker peaks representing ribosomal proteins of different bacteria. This approach made it possible to minimize the error due to variables such as culture conditions or sample preparation and supported the assignment of milk isolates up to the species and strain level.

As previously stated, mastitis and subclinical mastitis could represent a burden for the quality of milk. *Streptococcus agalactiae* often represents one of the causes of this complication, and several proteomic studies were performed to evaluate the physiology of this infection and its consequences on cow milk. The milk proteome of subclinical mastitis caused by *Streptococcus agalactiae* has been successfully investigated and, in milk of infected animals, highlighted the presence of the antibacterial protein cathelicidin-1 that positively correlated with the milk somatic cell count ($r = 0.616$) (Pongthaisong et al. 2016).

Bacteria represent key players in the transformation and ripening of dairy products. However, it has to be taken into account that bacteria could also represent hazards in case of pathogen growth, especially if considering the increasing problems of antibiotic resistance.

Most of the times, cheese is considered as safe food; however, some microbial risks are present and can represent a hazard for human health. Among them, *Listeria monocytogenes*, *Staphylococcus aureus*, and *Escherichia coli* can be responsible for outbreaks or food poisoning. For example, *S. aureus* and *L. monocytogenes* have been isolated from unpasteurized soft cheese on sale in California (Choi et al. 2016). An interesting study published in 2014 highlighted that, in the United States in the period from 1998 to 2011, 42% of outbreaks originated from cheese made with unpasteurized milk and 49% from cheese made with pasteurized milk. This data suggests the relevance of strict sanitation and microbiologic monitoring in cheese-making facilities (Gould et al. 2014).

Cheese commerce is strictly regulated, for example, regulations in the European Union require for cheese with a shelf life of more than 5 days <100 CFU/g in at least five samples (Regulation 2007).

An even more important issue is the development and use of methods capable of detecting reliably and fast probable hazards in finished foods. In this field, the

recent advances in proteomic techniques provide an important contribution. One example is related to the detection of *Listeria monocytogenes*, which can take from 4 to 5 days with the cultivation method. In order to overcome this problem, a MALDI-TOF MS method has been proposed (Barbuddhe et al. 2008) both for identification and typing of *Listeria* species. Fructose 1,6-bisphosphate aldolase has been as well proposed as a novel immunogenic surface protein of *Listeria* species useful for its detection. In this case, authors successfully applied MALDI-TOF MS and MS/MS for the characterization of fructose 1,6-bisphosphate aldolase from Coomassie-stained SDS-PAGE; prerequisite has been band identification through immunoblotting (Mendonca et al. 2016).

One of the key features that make *Listeria monocytogenes* capable to survive in various environmental conditions is due to its adaptive capacity at several different conditions, such as low pH, high salinity, oxygen content, and various temperatures. Using a proteomic approach, authors discovered that many of the *Listeria* proteins involved in the protection against cold stress and osmotic stress were involved in the maintenance of cell wall and cellular processes (Pittman et al. 2014).

This described evidence summarizes the most recent advances of proteomics in the detection of pathogens in cheese.

2.2 Protein and Peptide Modifications Related to Authenticity

Milk thermal treatment is mandatory before consumption to reduce bacterial load and to address microbiological safety. The most used ways in food industry are, in order of effectiveness on bacterial load, high-temperature short-time (HTST) pasteurization, ultrahigh temperature (UHT) treatment, and in-container sterilization (Lewis 1994). The drawback of these heating procedures is the vitamin degradation, the reduced bioavailability of some amino acids, and the formation of Maillard reaction products (Oamen et al. 1989; Kilshaw et al. 1982; Pischetsrieder and Henle 2012). For all these reasons, the heating process of milk should be kept as low as possible to reach the standards of safety, and it is therefore necessary to develop analytical methods to monitor the milk heating process in milk industry. One example of the application of proteomics to solve this problem has been provided by Ebner and colleagues (Ebner et al. 2016) who successfully identified 16 different milk peptides differentially modulated by industrial processes. In order to achieve this result, the authors used MALDI-TOF MS after stage-tip extraction and relative quantification using an internal reference peptide.

Cross-species adulteration of milk represents another field of application of proteomics. For example, donkey's milk is considered a good choice for newborn feeding in case of allergy to cow milk (Salimei and Fantuz 2012). In order to avoid allergenicity, it is strictly necessary to avoid any kind of possible adulteration of non-cow milk with the less costly cow milk. It is now possible to detect the presence of cow milk added to donkey milk and goat milk using a sensitive and robust MALDI-TOF MS profiling. Authors successfully applied MALDI-TOF MS

profiling between 2000 and 25,000 kDa able to detect the contamination with cow milk up to 5% (Di Girolamo et al. 2014).

Proteins can be used as markers for properties of many foods, and the evaluation of their profiles can provide useful information about composition and processing (Ortea et al. 2016). Among dairy products, the addition of cow or sheep milk to goat cheese making represents a common fraud (Guarino et al. 2010). To overcome this problem, it is possible to analyze the casein fraction after a simple extraction step and after digestion with plasmin. The hydrolytic products can be analyzed through LC-ESI-MS/MS to specifically detect a fragment of sheep's casein (Guarino et al. 2010).

Generally speaking, the fraudulent use of less expensive cow milk for ewe and goat cheese making represents a common problem in the matter of food authenticity. The reference method for the identification of cow γ2- and γ3-caseins is IEF, and it has been proven to be reliable for cow milk detection (Mayer 2005; Mikulec et al. 2013). However, this technique has some drawbacks as being unable to detect cow milk in amounts lower than 5% and the occurrence of some false positives (Mikulec et al. 2013; Mayer 2005). For example, in terms of allergies, it would be better to be able to detect lower amounts of fraudulent addition of cow milk.

Recent advances described the successful use of MALDI-TOF protein profiling in the detection of bovine milk added to ewe and buffalo milk or in the adulteration of Mozzarella cheese (Cozzolino et al. 2001, 2002). This method can be used as well for the detection of amounts up to 0.5% of bovine milk in milk of other species as cow, ewe, and buffalo (Di Girolamo et al. 2014), and it can be applied as well for the identification of thermal treatments (Sassi et al. 2015). More precisely, it is possible to detect up to 10% of UHT milk added to raw or pasteurized milk (Pinto et al. 2012).

Tracing the protein β-lactoglobulin through ESI-MS and HPLC-ESI-MS, it is possible to detect as well amounts of bovine milk in goat and caprine or ovine milk in a concentration up to 5% (Chen et al. 2004; Műller et al. 2008).

About cheese, a described LC-MS/MS method can be used for the detection of sheep milk in cow and goat cheese. The method targets through SRM a specific peptide originated from plasminolytic digestion of caseins and made it possible to detect up to 2% of sheep milk in cheese made of milk from goat and cow (Guarino et al. 2010).

Powdered milk represents another type of adulteration, and it has been successfully detected using 2-DE coupled with MALDI MS for the identification of lactosylated peptides specific of powdered milk. With this method 1% of powdered milk is still traceable (Calvano et al. 2012).

These described methods provide relevant proteomic tools to be applied in the field of the evaluation of authenticity among dairy products.

2.3 Protein and Peptide Modifications Related to Quality

Milk quality can be influenced by several factors going from animal welfare to heating process or other treatments.

In the first place, milk quality is different among mammalian species, and, as stated before, it is important to underline the main differences among all milks (Fig. 3). These differences then, more than providing different valuable nutrients, are important to avoid the possibility of allergic reactions in patients allergic to cow milk (D'auria et al. 2005). Casein polymorphisms among mammalian species have been extensively investigated both through gel-based and gel-free methods, thoroughly differentiating cow, goat, sheep, buffalo, and mare according to the concentrations of their isoforms (Roncada et al. 2012).

It has been demonstrated that thermal treatment of milk induces protein carbonylation that is directly related to the harshness of the process. Authors discovered 53 unique carbonylated peptides derived from 15 different proteins; the highest number of these was present in infant formula milk and the lowest in raw milk. This study was performed by successfully applying a derivatization process with *O*-(biotinylcarbazoylmethyl) hydroxylamine of reactive carbonyls, an enrichment through avidin-biotin affinity chromatography, and an analytical process by nano-RP-UPLC-ESI-MS (Milkovska-Stamenova et al. 2017).

Among functional milk compartments, besides MFGM, there are also the exosomes. Their presence has been documented, but there was no evidence of differences between human and bovine exosomes in colostrum and mature milk. This gap has been filled by Yang and colleagues who discovered and quantified a total of 920 milk exosome proteins providing important knowledge about their different functions among bovine and human milk (Yang et al. 2017a).

There are numerous studies and publications on the quality of dairy products deriving from domestic animals: a group of foods that stand out for the value of the nutrient content. Many molecules make dairy products "naturally functional foods," for instance, essential amino acids, vitamins, lipids, and other organic molecules that classical chemical approaches have investigated and classified. Nowadays, proteomics-based study of low MW peptides and associated PTM led to the discovery and classification of a multitude of proteolytic peptides with specific biological activity deriving from multiple enzymatic fermentation processes over time. Moreover, the quality of a dairy product depends almost totally on the quality of the milk used during the production, and technologies used in the proteomic field are the essential tools to detect fraud or technological problems during all the processes related to fermentation and cheese making.

As reported in Sect. 2.1, several techniques can be used to analyze dairy samples, and many of them can be applied for the evaluation of the production quality. Using HIC-HPLC without a previous pre-fractionation, αs1-, αs2-, β-, and κ-caseins of fresh and ripened cheese from cows, ewes, and goats can be separated and determined (Bramanti et al. 2003). Similar approaches to analyze the quality have been undertaken using RP-HPLC on β-lactoglobulins (Ferreira and Caçote 2003), CE on caseins (Rodriguez-Nogales and Vázquez 2007), and IEF coupled with CEX-HPLC (Mayer et al. 1997). Also electrophoretic (Manzo et al. 2008) and immunological (Hurley et al. 2004) techniques applied on specific protein species can provide information about adulteration and related quality. This kind of approaches can provide qualitative and quantitative results at the protein level that can be used to

evaluate the quality of the samples. However, the exclusive application of electrophoretic and chromatographic techniques cannot provide an exhaustive view on quality including information on peptides. To circumvent those issues, several state-of-the-art proteomic approaches have been proposed to deeply characterize and evaluate quality in terms of sensitivity. A combined untargeted and targeted approach by monitoring specific transitions provided selective detection of buffalo Mozzarella adulteration three orders of magnitude lower than the official method (Russo et al. 2012). Other investigations demonstrated the validity of MALDI-TOF and LC-MS/MS in detecting adulteration in Ricotta (Camerini et al. 2016; Russo et al. 2016), Mozzarella (Cozzolino et al. 2002), goat's and cow's cheese (Guarino et al. 2010), raw cow milk (Fontenele et al. 2017), and skimmed milk powder (Cordewener et al. 2009).

3 Proteomics and Dairy Product Transformation Processes

Raw milk has a complex microbial population or "microbiota" which includes technologically important bacteria, such as lactic acid bacteria (LAB), but also microorganisms responsible for unwanted negative effects on the quality of the milk and derivatives and sometimes pathogens. Traditional methods to determine the nature of the milk microbiota are based on the isolation and culturing of the microorganisms on defined media prior to their identification by molecular methods.

However, approaches that require a cultivation step are often inaccurate because the species present in low concentration can be masked by the more abundant ones or may not be cultivable in the laboratory. For this reason culture-independent approaches have been introduced analyzing directly nucleic acids (DNA or RNA) that allow determining the microbial composition of a food, without the need of an initial culture passage. Such techniques are more sensitive and rapid compared to conventional culture-dependent methods of analysis, which, although still presenting some limits, have been recognized as suitable for the detection and identification of microorganisms present in dairy products. Several molecular techniques from PCR and SSCP to DGGE have been used in the past to detect differences in microbial communities from milk and cheese (Quigley et al. 2011). Nowadays next-generation DNA sequencing technologies have greatly enhanced our knowledge in the area of fermented food (Ercolini 2013; Kergourlay et al. 2015), including that of raw milk cheese (Quigley et al. 2012). The targets most commonly used for the identification of species are the genes encoding ribosomal RNA (rRNA) 16S and 26S, for bacteria and eukaryotes, respectively. These regions were chosen because they possess both conserved domains that facilitate the use of universal PCR primers to amplify gene portions and hypervariable domains that allow the identification of the corresponding microorganisms. However, despite the capacity to classify hundreds of bacterial genera and make a taxonomic analysis, such an approach fails to investigate the molecular functions of each

microorganism during ripening. Combining metatranscriptomics with metagenomics in a mixed multi-omic approach has shown the potential to filter out microbial populations that are metabolically active during the ripening process (Dugat-Bony et al. 2015).

3.1 Proteomic Analysis Applied to Cheese Maturation

The ripening of cheese is a fundamental process with the purpose to shape product consistency and appearance and, at the same time, to bring out particular flavors and aromas through complex and sometimes long chemical-physical processes. The maturation of cheese is essentially an enzymatic process coupled with physical transformations (pH, free water, humidity, etc.). A multitude of biochemical processes take place for sugars, lipids, and proteins in this period, in particular in the raw material and in the caseins. Among the main causes of slow changes in the cheese (glycolysis, proteolysis, lipolysis) over time are microorganisms which play a pivotal role. The microflora of raw milk and of starter and those important for production and aging environments together with enzymes from rennet have to be considered responsible of the cheese-making process and the development of the organoleptic characteristics of the cheese.

From a proteomic point of view, during ripening, slow proteolysis of casein takes place with the production of polypeptides, peptides, peptones, and free amino acids, which are further transformed to other compounds. The different production technologies, by selecting and mixing specific milk and starters, influence this proteolysis, determining the quality of the various types of cheese.

In cheese the ripening time can range from a period of 3 weeks (Mozzarella) up to 2 years or more for Parmigiano-Reggiano and extra-mature Cheddar, leading to primary changes (proteolysis) and secondary changes (amino acid catabolism). Proteolysis is the most important biochemical event in cheese made from milk of different species and can vary from limited (Mozzarella) to moderate (Cheddar, Grana Padano, Parmigiano-Reggiano) to very extensive (blue cheese, Gorgonzola).

Several proteomic techniques can be applied to understand biochemical changes that occur at the level of cheese proteins during ripening and to detect specific markers of maturation at different time points to be used also in evaluation of quality and authenticity. Electrophoretic techniques (1-DE, 2-DE, urea-PAGE, IEF) also in combination with antibody-based assays are the most appropriate to monitor primary changes like casein proteolysis in large fragments. Medium to short peptides and water-soluble amino acids can be analyzed by HPLC and CE, and, depending on the resolution and sensitivity required, these fractionation methods may be coupled with mass spectrometry.

In cheese with short ripening time and limited proteolysis like Mozzarella, the proteomic approach has to provide information about molecular markers of freshness and to monitor early changes in casein proteolysis. A gel-based approach consisting of SDS-PAGE, alkaline-PAGE with urea, and 2-DE together with

protein identification by LC-ESI qTOF MS and multivariate analysis has been applied to water buffalo Mozzarella. The combination of different electrophoretic techniques has proven suitable to reveal diagnostic casein fragments deriving from plasmin and chymosin activity as well as fragments from highly proteolyzed samples to be used as molecular markers of quality and correct ripening (Petrella et al. 2015).

When analyzing cheeses with short or medium ripening time (up to 1 year) and at multiple time points, electrophoretic and mass spectrometric techniques have to be used in combination to detect early proteolytic changes like high to medium molecular weight fragments of caseins and low MW peptides at the later stage of maturation. An investigation made on ripening dynamics in Ragusano cheese analyzed by means of urea-PAGE and LC-MS/MS samples at 0, 4, and 7 months. Urea-PAGE highlighted extended hydrolysis of β- and αs1-caseins at 4 months, and >90% of both caseins were hydrolyzed after 7 months of aging. Using LC-MS/MS it was possible to identify all the main cleavage sites for β-, αs1-, and αs2-caseins that arose from the enzymatic activity of plasmin and residual chymosin or cathepsin D and also identified the global extensive hydrolysis as marker of ripening of that pasta filata cheese (Gagnaire et al. 2011). A similar approach has been applied to water-soluble extract (WSE) of a low-ripened 120-day aged Teleme (white-brined) cheese made from sheep, goat, or cow milk. Moreover, the simultaneous use of qTOF-MS with N-terminal sequencing and LC/MS provided not only information about ripening but also on the milk source used (Pappa et al. 2008). Approaches only based on CE can also be used to analyze specific enzymatic activities and ripening in cheese made from one (Irigoyen et al. 2000; Strickland et al. 1996) or different milks (Molina et al. 2000).

Many investigations on long-ripened hard cheese have been made coupling chromatography with mass spectrometry, especially to detect short peptides and amino acids deriving from degradation of medium-length peptides by starter LAB and nonstarter LAB (NSLAB). A detailed investigation made on Parmigiano-Reggiano reported the ripening analyzed during 24 months at 14 different time points, starting from the curd to the end of maturation. Another variable, when analyzing cheese rind, is represented by the sampling procedure. As reported, sampling has been made from internal and external parts of the wheels to detect eventual differences in proteolytic events. Peptides have been separated from caseins by acid precipitation and centrifugation, the resulting aqueous fraction has been filtered with cutoff 10 kDa, and semi-quantification with an appropriate internal standard was carried out by reversed-phase LC-ESI-MS. The application of such a strategy enabled the identification of more than 100 different peptides ranging from about 200 to more than 7000 Da and the semi-quantification of 57 peptides. The analysis was also able to distinguish non-proteolytic aminoacyl derivatives (*N*-γ-glutamyl-amino acids and *N*-lactoyl-amino acids) and the proteolytic ones from αs1- and β-CN. Several specific peptides, from bioactive to modified ones, have been highlighted as time-point-specific fingerprints of ripening dynamics over time (Sforza et al. 2012). Similar investigations have been

performed on other well-known long-ripened hard cheese types like Grana Padano (Sforza et al. 2003; Masotti et al. 2010).

Proteolytic changes can be rapidly detected with high accuracy without chromatographic separation, even by MALDI-TOF MS (Fedele et al. 1999; Piraino et al. 2007) on the soluble aqueous fraction. An investigation made on ten different sheep, cow, and goat cheeses of different ripening time and fermentation typologies highlighted diagnostic small peptides deriving from chymosin, plasmin, and enzymes from starter, nonstarter, or secondary microflora in the cheese (Piraino et al. 2007). Moreover, MALDI-TOF MS and other mass spectrometric techniques were able to detect early proteolytic changes in dairy products other than cheese. Investigations have been made on yogurt using MALDI in linear mode on extracts to analyze the protein and peptide profile up to 55,000 Da (Fedele et al. 1999) to study the effects of bacterial strains over time. Other fermented products like kefir have been analyzed using MALDI-TOF and Orbitrap MS that allowed to detect and build a detailed dataset for the identification of bioactive components (Ebner et al. 2015; Dallas et al. 2016).

3.2 *Metaproteomics for Microbiome Analysis of Raw Milk and Dairy Products*

With the appearance of new techniques and especially culture-independent methods based on DNA and 16S-rRNA sequencing, it was possible to deeply characterize the microbiota of raw milk in humans and animals in terms of composition (Quigley et al. 2013). Nevertheless NGS techniques and in particular 16S-rRNA offer only the taxonomic profile of the complex microbial ecosystem (the microbiota) that is present in milk and dairy products, especially those made from raw milk. Due to the complexity of such ecosystems, other complementary techniques are needed to exhaustively describe not only the population inside milk and dairy product but also the dynamics of metabolic functions of each taxon inside the ecosystem (Ferrocino and Cocolin 2017). The analysis of the functional metatranscriptome in milk has been applied very few times, for example, during cheese production to monitor the dynamics of microbiota during ripening (De Filippis et al. 2016).

The exclusive use of NGS techniques offers only a partial view of the complex microbial ecosystem (the microbiome) that is present in dairy products, especially those made from raw milk. Due to the complexity of such ecosystems, several techniques are needed to exhaustively describe not only the population inside a dairy product but also the dynamics of metabolic functions of each taxon inside the ecosystem (Ferrocino and Cocolin 2017). In this context, approaches based on the investigation of the proteome within the dairy product (metaproteome) can deliver simultaneously information about microbial taxonomies and metabolic functions as to provide a complete picture of the microbiome at different time points during ripening.

A pioneering investigation made on Emmenthal cheese on the pre-fractionated aqueous phase by size-exclusion chromatography applied 2D-PAGE in combination with mass spectrometry to investigate the microbial population of the cheese. One hundred fifty spots have been separated by 2D-PAGE, and 62 were unambiguously identified by peptide mass fingerprinting and a de novo sequencing by LC-ESI-MS/MS and led to the identification of several unidentified proteins. The microbiota analysis detected mostly bacterial proteins associated to the starter microflora and several metabolic functions ranging from stress response to glycolytic processes (Gagnaire et al. 2004).

Despite the proteomic strategy adopted, this approach failed to cover the entire complexity of the microbiota due to the technological limits of that time. A quantitative approach based on iTRAQ on experimental Swiss-type cheese at three ripening times investigated the release of bacterial enzymes into the cheese as a time course. This strategy coupled with a nano-LC ESI-qTOF permitted to quantify 30 proteins both from bacterial and bovine origin and to functionally classify them (Jardin et al. 2012).

A metaproteomic approach based on multidimensional chromatography coupled with tandem mass spectrometry shed light on the microbiota of Grana Padano, a raw milk cheese, overcoming the limits of an exclusive gel-based approach. In that investigation a tailored bioinformatics pipeline coupled with a depletion strategy and a 2D-LC-MS/MS permitted to investigate differences in microbial taxonomies and molecular functions and to correlate them to the use of lysozyme and the quantity of clostridial spore during the ripening (Soggiu et al. 2016).

4 Conclusions and Future Perspectives

Nowadays proteomics and related technologies can provide deep knowledge about the protein and peptide composition of raw milk and dairy products thanks to the simultaneous detection and quantitation of hundreds proteoforms and peptides related to technological and ripening processes. The state-of-the-art of described research tools can be complementary and reinforce analytical methods officially used to ensure quality, authenticity, and biodiversity intrinsically linked to each dairy product.

References

Barbuddhe SB, Maier T, Schwarz G, Kostrzewa M, Hof H, Domann E, Chakraborty T, Hain T (2008) Rapid identification and typing of listeria species by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Appl Environ Microbiol 74(17):5402–5407

Benabdelkamel H, Masood A, Alanazi IO, Alzahrani DA, Alrabiah DK, AlYahya SA, Alfadda AA (2017) Proteomic profiling comparing the effects of different heat treatments on camel (*Camelus dromedarius*) milk whey proteins. Int J Mol Sci 18(4):721

Bernardi N, Benetti G, Haouet NM, Sergi M, Grotta L, Marchetti S, Castellani F, Martino G (2015) A rapid high-performance liquid chromatography-tandem mass spectrometry assay for unambiguous detection of different milk species employed in cheese manufacturing. J Dairy Sci 98 (12):8405–8413

Bramanti E, Sortino C, Onor M, Beni F, Raspi G (2003) Separation and determination of denatured α s1-, α s2-, β-and κ-caseins by hydrophobic interaction chromatography in cows', ewes' and goats' milk, milk mixtures and cheeses. J Chromatogr A 994(1):59–74

Calvano CD, Monopoli A, Loizzo P, Faccia M, Zambonin C (2012) Proteomic approach based on MALDI-TOF MS to detect powdered milk in fresh cow's milk. J Agric Food Chem 61 (8):1609–1617

Camerini S, Montepeloso E, Casella M, Crescenzi M, Marianella RM, Fuselli F (2016) Mass spectrometry detection of fraudulent use of cow whey in water buffalo, sheep, or goat Italian ricotta cheese. Food Chem 197(Part B):1240–1248

Chen RK, Chang LW, Chung YY, Lee MH, Ling YC (2004) Quantification of cow milk adulteration in goat milk using high-performance liquid chromatography with electrospray ionization mass spectrometry. Rapid Commun Mass Spectrom 18(10):1167–1171

Chen Q, Ke X, Zhang J, Lai S, Fang F, Mo W, Ren Y (2016) Proteomics method to quantify the percentage of cow, goat, and sheep milks in raw materials for dairy products. J Dairy Sci 99 (12):9483–9492

Choi K-H, Lee H, Lee S, Kim S, Yoon Y (2016) Cheese microbial risk assessments—a review. Asian Australas J Anim Sci 29(3):307

Claeys WL, Verraes C, Cardoen S, De Block J, Huyghebaert A, Raes K, Dewettinck K, Herman L (2014) Consumption of raw or heated milk from different species: an evaluation of the nutritional and potential health benefits. Food Control 42:188–201

Cordewener JHG, Luykx DMAM, Frankhuizen R, Bremer MGEG, Hooijerink H, America AHP (2009) Untargeted LC-Q-TOF mass spectrometry method for the detection of adulterations in skimmed-milk powder. J Sep Sci 32(8):1216–1223

Coulona J-B, Gasquib P, Barnouin J, Ollier A, Pradel P, Pomiès D (2002) Effect of mastitis and related-germ on milk yield and composition during naturally-occurring udder infections in dairy cows. Anim Res 51(05):383–393

Cozzolino R, Passalacqua S, Salemi S, Malvagna P, Spina E, Garozzo D (2001) Identification of adulteration in milk by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Mass Spectrom 36(9):1031–1037

Cozzolino R, Passalacqua S, Salemi S, Garozzo D (2002) Identification of adulteration in water buffalo mozzarella and in ewe cheese by using whey proteins as biomarkers and matrix-assisted laser desorption/ionization mass spectrometry. J Mass Spectrom 37(9):985–991

D'auria E, Agostoni C, Giovannini M, Riva E, Zetterström R, Fortin R, Franco Greppi G, Bonizzi L, Roncada P (2005) Proteomic evaluation of milk from different mammalian species as a substitute for breast milk. Acta Paediatr 94(12):1708–1713

Dallas DC, Citerne F, Tian T, Silva VL, Kalanetra KM, Frese SA, Robinson RC, Mills DA, Barile D (2016) Peptidomic analysis reveals proteolytic activity of kefir microorganisms on bovine milk proteins. Food Chem 197:273–284

De Filippis F, Genovese A, Ferranti P, Gilbert JA, Ercolini D (2016) Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. Sci Rep 6:21871

Di Girolamo F, Masotti A, Salvatori G, Scapaticci M, Muraca M, Putignani L (2014) A sensitive and effective proteomic approach to identify she-donkey's and goat's milk adulterations by MALDI-TOF MS fingerprinting. Int J Mol Sci 15(8):13697–13719

Dugat-Bony E, Straub C, Teissandier A, Onésime D, Loux V, Monnet C, Irlinger F, Landaud S, Leclercq-Perlat M-N, Bento P, Fraud S, Gibrat J-F, Aubert J, Fer F, Guédon E, Pons N, Kennedy S, Beckerich J-M, Swennen D, Bonnarme P (2015) Overview of a surface-ripened cheese community functioning by meta-omics analyses. PLoS One 10(4):e0124360

Ebner J, Arslan AA, Fedorova M, Hoffmann R, Küçükçetin A, Pischetsrieder M (2015) Peptide profiling of bovine kefir reveals 236 unique peptides released from caseins during its production by starter culture or kefir grains. J Proteomics 117:41–57

Ebner J, Baum F, Pischetsrieder M (2016) Identification of sixteen peptides reflecting heat and/or storage induced processes by profiling of commercial milk samples. J Proteomics 147:66–75

Ercolini D (2013) High-throughput sequencing and metagenomics: moving forward in the culture-independent analysis of food microbial ecology. Appl Environ Microbiol 79(10):3148–3155

Fedele L, Seraglia R, Battistotti B, Pinelli C, Traldi P (1999) Matrix-assisted laser desorption/ionization mass spectrometry for monitoring bacterial protein digestion in yogurt production. J Mass Spectrom 34(12):1338–1345

Ferreira IM, Caçote H (2003) Detection and quantification of bovine, ovine and caprine milk percentages in protected denomination of origin cheeses by reversed-phase high-performance liquid chromatography of beta-lactoglobulins. J Chromatogr A 1015(1):111–118

Ferrocino I, Cocolin L (2017) Current perspectives in food-based studies exploiting multi-omics approaches. Curr Opin Food Sci 13:10–15

Fontenele MA, Bastos MSR, dos Santos KMO, Bemquerer MP, do Egito AS (2017) Peptide profile of Coalho cheese: a contribution for protected designation of origin (PDO). Food Chem 219:382–390

Fox P, Uniacke-Lowe T, McSweeney PL, O'Mahony JA (2015) Dairy chemistry and biochemistry. Springer International Publishing Switzerland

Gagnaire V, Piot M, Camier B, Vissers JP, Jan G, Léonil J (2004) Survey of bacterial proteins released in cheese: a proteomic approach. Int J Food Microbiol 94(2):185–201

Gagnaire V, Carpino S, Pediliggieri C, Jardin J, Lortal S, Licitra G (2011) Uncommonly thorough hydrolysis of peptides during ripening of Ragusano cheese revealed by tandem mass spectrometry. J Agric Food Chem 59(23):12443–12452

Gould LH, Mungai E, Barton Behravesh C (2014) Outbreaks attributed to cheese: differences between outbreaks caused by unpasteurized and pasteurized dairy products, United States, 1998–2011. Foodborne Pathog Dis 11(7):545–551

Guarino C, Fuselli F, La Mantia A, Longo L, Faberi A, Marianella RM (2010) Peptidomic approach, based on liquid chromatography/electrospray ionization tandem mass spectrometry, for detecting sheep's milk in goat's and cow's cheeses. Rapid Commun Mass Spectrom 24 (6):705–713

Hinz K, O'Connor PM, Huppertz T, Ross RP, Kelly AL (2012) Comparison of the principal proteins in bovine, caprine, buffalo, equine and camel milk. J Dairy Res 79(2):185–191

Hogarth CJ, Fitzpatrick JL, Nolan AM, Young FJ, Pitt A, Eckersall PD (2004) Differential protein composition of bovine whey: a comparison of whey from healthy animals and from those with clinical mastitis. Proteomics 4(7):2094–2100

Hurley IP, Elyse Ireland H, Coleman RC, Williams JH (2004) Application of immunological methods for the detection of species adulteration in dairy products. Int J Food Sci Technol 39 (8):873–878

Irigoyen A, Izco JM, Ibáñez FC, Torre P (2000) Evaluation of the effect of rennet type on casein proteolysis in an ovine milk cheese by means of capillary electrophoresis. J Chromatogr A 881 (1–2):59–67

Jardin J, Mollé D, Piot M, Lortal S, Gagnaire V (2012) Quantitative proteomic analysis of bacterial enzymes released in cheese during ripening. Int J Food Microbiol 155(1):19–28

Jensen HB, Poulsen NA, Moller HS, Stensballe A, Larsen LB (2012) Comparative proteomic analysis of casein and whey as prepared by chymosin-induced separation, isoelectric precipitation or ultracentrifugation. J Dairy Res 79(4):451–458

Kergourlay G, Taminiau B, Daube G, Champomier Vergès M-C (2015) Metagenomic insights into the dynamics of microbial communities in food. Int J Food Microbiol 213:31–39

Kilshaw P, Heppell L, Ford J (1982) Effects of heat treatment of cow's milk and whey on the nutritional quality and antigenic properties. Arch Dis Child 57(11):842–847

Kuchroo C, Fox P (1982) Soluble nitrogen in Cheddar cheese: comparison of extraction procedures. Milchwissenschaft Milk Sci Int 37:331–335

Lewis M (1994) Heat treatment of milk. In: Robinson (ed) Modern dairy technology. Springer, Dordrecht, pp 1–60

Li S, Wang Q, Lin X, Jin X, Liu L, Wang C, Chen Q, Liu J, Liu H (2017) The use of "omics" in lactation research in dairy cows. Int J Mol Sci 18(5):983

Lynch JM, Barbano DM (1999) Kjeldahl nitrogen analysis as a reference method for protein determination in dairy products. J AOAC Int 82(6):1389–1398

Manzo C, Pizzano R, Addeo F (2008) Detection of pH 4.6 insoluble β-lactoglobulin in heat-treated milk and Mozzarella cheese. J Agric Food Chem 56(17):7929–7933

Masotti F, Hogenboom JA, Rosi V, De Noni I, Pellegrino L (2010) Proteolysis indices related to cheese ripening and typicalness in PDO Grana Padano cheese. Int Dairy J 20(5):352–359

Mayer H (2005) Milk species identification in cheese varieties using electrophoretic, chromatographic and PCR techniques. Int Dairy J 15(6):595–604

Mayer H, Heidler D, Rockenbauer C (1997) Determination of the percentages of cows', ewes' and goats' milk in cheese by isoelectric focusing and cation-exchange HPLC of γ-and para-κ-caseins. Int Dairy J 7(10):619–628

Mendonca M, Moreira GMSG, Conceição FR, Hust M, Mendonca KS, Moreira ÂN, França RC, da Silva WP, Bhunia AK, Aleixo JAG (2016) Fructose 1, 6-bisphosphate aldolase, a novel immunogenic surface protein on Listeria species. PLoS One 11(8):e0160544

Mikulec N, Plavljanić D, Radeljević B, Havranek J, Antunac N (2013) Proving the adulteration of ewe and goat cheeses with cow milk using the reference method of isoelectric focusing of γ-casein. Mljekarstvo 63(3):115–121

Milkovska-Stamenova S, Mnatsakanyan R, Hoffmann R (2017) Protein carbonylation sites in bovine raw milk and processed milk products. Food Chem 229:417–424

Molina E, De Frutos M, Ramos M (2000) Capillary electrophoresis characterization of the casein fraction of cheeses made from cows', ewes' and goats' milks. J Dairy Res 67(02):209–216

Mollé D, Jardin J, Piot M, Pasco M, Léonil J, Gagnaire V (2009) Comparison of electrospray and matrix-assisted laser desorption ionization on the same hybrid quadrupole time-of-flight tandem mass spectrometer: application to bidimensional liquid chromatography of proteins from bovine milk fraction. J Chromatogr A 1216(12):2424–2432

Morin-Sardin S, Jany J-L, Artigaud S, Pichereau V, Bernay B, Coton E, Madec S (2017) Proteomic analysis of the adaptative response of *Mucor* spp. to cheese environment. J Proteomics 154:30–39

Muehlhoff E, Bennett A, McMahon D (2013) Milk and dairy products in human nutrition. Food and Agriculture Organization of the United Nations (FAO), Rome

Műller L, Barták P, Bednář P, Fryšová I, Ševčík J, Lemr K (2008) Capillary electrophoresis-mass spectrometry—a fast and reliable tool for the monitoring of milk adulteration. Electrophoresis 29(10):2088–2093

Nissen A, Bendixen E, Ingvartsen KL, Røntved CM (2012) In-depth analysis of low abundant proteins in bovine colostrum using different fractionation techniques. Proteomics 12 (18):2866–2878

Nissen A, Andersen PH, Bendixen E, Ingvartsen KL, Røntved CM (2017) Colostrum and milk protein rankings and ratios of importance to neonatal calf health using a proteomics approach. J Dairy Sci 100(4):2711–2728

O'Sullivan M, Fox PF (1990) A scheme for the partial fractionation of cheese peptides. J Dairy Res 57(01):135–139

Oamen E, Hansen A, Swartzel K (1989) Effect of ultra-high temperature steam injection processing and aseptic storage on labile water-soluble vitamins in milk. J Dairy Sci 72 (3):614–619

Ogola H, Shitandi A, Nanua J (2007) Effect of mastitis on raw milk compositional quality. J Vet Sci 8(3):237–242

Ortea I, O'Connor G, Maquet A (2016) Review on proteomics for food authentication. J Proteomics 147:212–225

Pappa E, Robertson J, Rigby N, Mellon F, Kandarakis I, Mills E (2008) Application of proteomic techniques to protein and peptide profiling of Teleme cheese made from different types of milk. Int Dairy J 18(6):605–614

Petrella G, Pati S, Gagliardi R, Rizzuti A, Mastrorilli P, La Gatta B, Di Luccia A (2015) Study of proteolysis in river buffalo mozzarella cheese using a proteomics approach. J Dairy Sci 98 (11):7560–7572

Pinto G, Caira S, Cuollo M, Fierro O, Nicolai MA, Chianese L, Addeo F (2012) Lactosylated casein phosphopeptides as specific indicators of heated milks. Anal Bioanal Chem 402 (5):1961–1972

Piraino P, Upadhyay VK, Rossano R, Riccio P, Parente E, Kelly AL, McSweeney PLH (2007) Use of mass spectrometry to characterize proteolysis in cheese. Food Chem 101(3):964–972

Piras C, Roncada P, Rodrigues PM, Bonizzi L, Soggiu A (2016) Proteomics in food: quality, safety, microbes, and allergens. Proteomics 16(5):799–815

Pischetsrieder M, Henle T (2012) Glycation products in infant formulas: chemical, analytical and physiological aspects. Amino Acids 42(4):1111–1118

Pittman JR, Buntyn JO, Posadas G, Nanduri B, Pendarvis K, Donaldson JR (2014) Proteomic analysis of cross protection provided between cold and osmotic stress in *Listeria monocytogenes*. J Proteome Res 13(4):1896–1904

Pongthaisong P, Katawatin S, Thamrongyoswittayakul C, Roytrakul S (2016) Milk protein profiles in response to *Streptococcus agalactiae* subclinical mastitis in dairy cows. Anim Sci J 87 (1):92–98

Quigley L, O'Sullivan O, Beresford TP, Ross RP, Fitzgerald GF, Cotter PD (2011) Molecular approaches to analysing the microbial composition of raw milk and raw milk cheese. Int J Food Microbiol 150(2):81–94

Quigley L, O'Sullivan O, Beresford TP, Ross RP, Fitzgerald GF, Cotter PD (2012) High-throughput sequencing for detection of subpopulations of bacteria not previously associated with artisanal cheeses. Appl Environ Microbiol 78(16):5717–5723

Quigley L, O'sullivan O, Stanton C, Beresford TP, Ross RP, Fitzgerald GF, Cotter PD (2013) The complex microbiota of raw milk. FEMS Microbiol Rev 37(5):664–698

Regulation E (2007) Commission regulation on microbiological criteria for foodstuffs. Off J Eur Union 50:12–29

Rodriguez-Nogales JM, Vázquez F (2007) Application of electrophoretic and chemometric analysis to predict the bovine, ovine and caprine milk percentages in Panela cheese, an unripened cheese. Food Control 18(5):580–586

Roncada P, Piras C, Soggiu A, Turk R, Urbani A, Bonizzi L (2012) Farm animal milk proteomics. J Proteomics 75(14):4259–4274

Russo R, Severino V, Mendez A, Lliberia J, Parente A, Chambery A (2012) Detection of buffalo mozzarella adulteration by an ultra-high performance liquid chromatography tandem mass spectrometry methodology. J Mass Spectrom 47(11):1407–1414

Russo R, Rega C, Chambery A (2016) Rapid detection of water buffalo ricotta adulteration or contamination by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 30(4):497–503

Salimei E, Fantuz F (2012) Equid milk for human consumption. Int Dairy J 24(2):130–142

Sassi M, Arena S, Scaloni A (2015) MALDI-TOF-MS platform for integrated proteomic and peptidomic profiling of milk samples allows rapid detection of food adulterations. J Agric Food Chem 63(27):6157–6171

Sforza S, Ferroni L, Galaverna G, Dossena A, Marchelli R (2003) Extraction, semi-quantification, and fast on-line identification of oligopeptides in Grana Padano cheese by HPLC-MS. J Agric Food Chem 51(8):2130–2135

Sforza S, Cavatorta V, Lambertini F, Galaverna G, Dossena A, Marchelli R (2012) Cheese peptidomics: a detailed study on the evolution of the oligopeptide fraction in Parmigiano-Reggiano cheese from curd to 24 months of aging. J Dairy Sci 95(7):3514–3526

Silva RA, Bezerra VS, Pimentel MdCB, Porto ALF, Cavalcanti MTH, José Luiz Filho L (2016) Proteomic and peptidomic profiling of Brazilian artisanal 'Coalho' cheese. J Sci Food Agric 96 (13):4337–4344

Simonne AH, Simonne EH, Eitenmiller RR, Mills HA, Cresman CP (1997) Could the Dumas method replace the Kjeldahl digestion for nitrogen and crude protein determinations in foods? J Sci Food Agric 73(1):39–45

Singh T, Fox P, Højrup P, Healy A (1994) A scheme for the fractionation of cheese nitrogen and identification of principal peptides. Int Dairy J 4(2):111–122

Soggiu A, Piras C, Mortera SL, Alloggio I, Urbani A, Bonizzi L, Roncada P (2016) Unravelling the effect of clostridia spores and lysozyme on microbiota dynamics in Grana Padano cheese: a metaproteomics approach. J Proteomics 147:21–27

Strickland M, Weimer BC, Broadbent JR (1996) Capillary electrophoresis of Cheddar cheese. J Chromatogr A 731(1):305–313. https://doi.org/10.1016/0021-9673(95)01160-9

Verma A, Ambatipudi K (2016) Challenges and opportunities of bovine milk analysis by mass spectrometry. Clin Proteomics 13(1):8

Vithanage NR, Bhongir J, Jadhav SR, Ranadheera CS, Palombo EA, Yeager TR, Datta N (2017) Species-level discrimination of psychrotrophic pathogenic and spoilage gram-negative raw milk isolates using a combined MALDI-TOF MS proteomics–bioinformatics-based approach. J Proteome Res 16(6):2188–2203

Wang J, Li D, Dangott LJ, Wu G (2006) Proteomics and its role in nutrition research. J Nutr 136 (7):1759–1762

Wang X, Zhao X, Huang D, Pan X, Qi Y, Yang Y, Zhao H, Cheng G (2017) Proteomic analysis and cross species comparison of casein fractions from the milk of dairy animals. Sci Rep 7:43020

Yang Y, Bu D, Zhao X, Sun P, Wang J, Zhou L (2013) Proteomic analysis of cow, yak, buffalo, goat and camel milk whey proteins: quantitative differential expression patterns. J Proteome Res 12(4):1660–1667

Yang Y, Zheng N, Zhao X, Zhang Y, Han R, Ma L, Zhao S, Li S, Guo T, Wang J (2015) Proteomic characterization and comparison of mammalian milk fat globule proteomes by iTRAQ analysis. J Proteomics 116:34–43

Yang M, Cao X, Wu R, Liu B, Ye W, Yue X, Wu J (2017a) Comparative proteomic exploration of whey proteins in human and bovine colostrum and mature milk using iTRAQ-coupled LC-MS/MS. Int J Food Sci Nutr 68(6):671–681

Yang M, Song D, Cao X, Wu R, Liu B, Ye W, Wu J, Yue X (2017b) Comparative proteomic analysis of milk-derived exosomes in human and bovine colostrum and mature milk samples by iTRAQ-coupled LC-MS/MS. Food Res Int 92:17–25

Zhang L, Boeren S, Hageman JA, van Hooijdonk T, Vervoort J, Hettinga K (2015) Bovine milk proteome in the first 9 days: protein interactions in maturation of the immune and digestive system of the newborn. PLoS One 10(2):e0116710

Proteomics in Skeletal Muscle Research

Eva Veiseth-Kent, André M. de Almeida, Brigitte Picard, and Kristin Hollung

Abstract The chapter revolves around the use of proteomics in skeletal muscle research in farm animals. The importance of muscle proteome characterisation is highlighted, together with the use of proteomics in studies related to animal production and meat quality traits. Specifically related to animal production, the chapter gives an overview of proteomics studies related to muscle development and growth during foetal life, muscle hypertrophy, genetics and nutrition. For the understanding of meat quality development, the chapter summarises studies related to postmortem proteome changes as well as studies focusing on meat quality traits such as tenderness, colour, pH and water holding capacity.

Keywords Muscle growth • Genetics • Nutrition • Postmortem changes • Meat quality

1 Characterisation of Muscle Proteomes

Proteomics studies are dependent on how well the proteins in a cell, tissue, organ, organism or population can be separated and quantified and if the individual proteins can be identified. Proteomics studies are therefore dependent on the robustness of protein databases (Soares et al. 2012). Accordingly, particularly in less studied organisms or tissues, pioneer proteomic work involves characterisation of the proteome, i.e. an exhaustive description of the proteins present in a given proteome. Such characterisation is particularly important when proteomics is based on two-dimensional gel electrophoresis (2DE). In fact, the mapping of the proteins

E. Veiseth-Kent (✉) • K. Hollung
Nofima AS, Aas, Norway
e-mail: eva.veiseth-kent@nofima.no

A.M. de Almeida
Instituto Superior de Agronomia, Universidade de Lisboa Tapada da Ajuda, Lisboa, Portugal

B. Picard
INRA—UMRH, Centre Auvergne Rhône Alpes, Saint-Genès-Champanelle, France

© Springer International Publishing AG 2018
A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_10

in a gel, highlighting the location of specific proteins, is a very important tool for other researchers working with similar proteomes and research subjects, by providing an idea on the identification of a given spot or spots of interest. During recent years, and with the advent of shotgun proteomics, such protein mapping has been rapidly converging towards a thorough proteome characterisation and division of proteins according to different categories, typically the biochemical pathways they are involved in or their cellular location. It is noteworthy to mention that a similar exercise was also conducted in the classical 2DE mapping. However, the broadness of proteome analysis allowed by approaches like label-free proteomics is such that the characterisation of the proteome can almost be considered as a "by-product" of a gel-free differential proteomics study.

In farm animals, to the best of our knowledge, the first 2DE mapping of muscle proteome was conducted by Bouley and co-workers in 2004, where a 2DE-based approach was used to characterise the proteome of the bovine semitendinosus (ST) muscle. They were able to identify 129 different spots corresponding to 75 different gene products and classified these proteins according to their molecular role (metabolism, cell structure, cell defence and contractile apparatus). This interesting work also allowed identification on the gel of different isoforms of several proteins such as troponin T. In this work, an immobilised pH 4–7 gradient was used in the first dimension, and thus the whole alkaline range was not included. Two years later, the same team mapped alkaline range (pH 7–11) proteins in a study with 60 spots and 32 identified proteins in total (Chaze et al. 2006). In recent years, the focus on bovine skeletal muscle mapping or characterisation has been focusing on gel-free methods. In a recent study with an emphasis on bovine development in Qinchuan cattle, He and co-workers (2016) characterised the longissimus muscle proteome, identifying over 1300 proteins and at the same time combining transcriptomics and proteomic data. Finally, the bovine longissimus thoracis (LT) muscle proteome has been characterised in a recent paper focusing on the growth path changes in the Alentejana breed (Almeida et al. 2017). They identified over 500 proteins in this study and characterised them according to the different biochemical processes in which they are involved. These studies reflect important advances in proteomics technology, particularly the instrumentation and databases for model/well-studied organisms. A similar example has been seen for the mapping of the rabbit skeletal muscle proteome (Almeida 2013). An earlier 2DE-based muscle mapping in 2009 obtained 45 different identifications classified as either metabolism, contractile apparatus, cell structure, cell defence or an additional "others" category (Almeida et al. 2009). A later shotgun-based analysis by Liu and co-workers (2013) led to the identification of almost 500 proteins.

With the ovine (*Ovis aries*) skeletal muscle, to the best of our knowledge, there is only one in-depth proteome characterisation describing the longissimus lumborum (LL) muscle proteome (Yu et al. 2015). The authors conducted a very interesting approach by fractioning the whole proteome into sarcoplasmic, myofibrillar and insoluble fractions and were able to identify over 380 proteins and classify them according to 23 different classes based on gene ontology. They were also able to assign the majority of the proteins to one of the above-mentioned

fractions with only 18 proteins in common among them. Using another ruminant species, the water buffalo (*Bubalus bubalis*), and with a strategy based on one-dimensional electrophoresis, Ferranti and co-workers (2007) studied the effect of ageing on buffalo muscle proteomic profiles and contributed to the first proteome characterisation in this animal. In pigs (*Sus scrofa*), the muscle phosphoproteome has been characterised by Huang and co-workers (2014) identifying 305 unique proteins, including 160 phosphoproteins and 784 phosphorylation sites.

2 Muscle Development and Growth during Foetal Life

In farm animals, understanding the mechanisms regulating foetal muscle fibre development and growth is of great importance, since the muscle fibre number at birth is known to be important for muscle mass and meat quality. Some of the initial proteomics analysis of in vivo myogenesis was performed on bovine muscle by Chaze and co-workers (2008, 2009). Using 2DE, they described the evolution of around 500 protein spots at 5 key stages of foetal life. In their analysis, they distinguished two periods of foetal life. Firstly, they considered the period from 60 to 180 days postconception (dpc), characterised by cell proliferation and by an increase of the total number of fibres which is fixed at around 180 dpc in bovine muscle (for review, see Picard et al. 2010). Secondly, they analysed the period from 180 to 260 dpc, corresponding to the maturation of fibres. On 60 dpc, the most abundant proteins such as heterogeneous nuclear ribonucleoprotein K (hnRNPK), proliferation-associated protein 2G4 (PA2G4), septin 11 and ARHGDIA (Rho GDP dissociation inhibitor alpha) are involved in embryonic development, migration, proliferation and regulation of cell cycle. The 110-dpc stage was characterised mainly by high abundance of structural protein such as myosin light chain, desmin and cellular organelle organisation (chloride intracellular protein 1, galectin-1). The 180-dpc stage was characterised by large modifications in the abundances of many proteins. The most abundant proteins at this stage were structural and contractile proteins (several isoforms of myosin light chains and of troponin T), proteins involved in energy metabolism (activator of protein ATPase, prostaglandin reductase, prostaglandin F synthase) and proteins involved in the regulation of apoptosis [dimethylaminohydrolase, serum albumin, annexinA4, heat shock protein 27 (Hsp27)]. Some proteins such as CLIC4 (chloride intracellular channel gene) peaked at 60–110 dpc and then decreased in abundance. This protein is a known marker of proliferation and is an effector of apoptosis through the activation of caspases (Ronnov-Jessen et al. 2002). The abundance of other proteins involved in the negative regulation of proliferation decreased after 180 dpc. The last trimester of gestation was mainly characterised by an increase in the abundance of proteins related to energy metabolism and structural organisation. This period was also characterised by several changes in protein isoforms associated with the acquisition of the contractile and metabolic properties of muscle fibres, for example, for troponin T, myosin heavy chains and enolase (Chaze et al. 2009).

Recently, a bioinformatics analysis of the dataset of cattle proteome using ProteINSIDE web service allowed the completion of our current knowledge on cattle myogenesis (Kaspric et al. 2015a, b). This tool revealed eight proteins secreted using a signal peptide: APOA1, P4HB, GSN, ADIPOQ, AFP, ALB, PDIA3 and SERPINA1. Eleven other proteins secreted by other secretory pathways were also revealed. Knowledge of these secreted proteins is very useful to better understand the mechanisms involved in the crosstalk between tissues, for example, between muscle and adipose tissues, which is important for meat production. In addition, Chan and co-workers (2011) analysed the secretome of C2C12 cells by SILAC and shotgun proteomics and proposed three secreted proteins—osteoglycin (OGN), peroxiredoxin 1 (Prx1) and cytokine-induced apoptosis inhibitor 1 (CIAPIN1)—as extracellular factors having an important role in myogenesis.

Moreover, the use of ProteINSIDE also allowed the construction of a network of 171 interactions between the proteins of the bovine foetal muscle dataset. This revealed five central proteins in this network: YWHAZ and YWHAE (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation proteins) which mediate signal transduction by binding to phosphoserine-containing proteins, MYC which is an activator of the transcription of growth-related genes and Hsp27 and Hsp70 known for their role in the protection of structural proteins and as anti-apoptotic factors (Kaspric et al. 2015a). These examples illustrate the power of proteomics associated with bioinformatics to highlight new knowledge about the biological process of myogenesis. This also allows the proposition of new hypotheses such as the secretion of adiponectin by the foetal muscle and the putative role of autophagy during myogenesis (Kaspric et al. 2015a).

Proteome changes during myogenesis have also been described in pig and bird muscles (Doherty et al. 2004; Teltathum and Mekchay 2009; Picard et al. 2010). Comparative biology of the data obtained in these different species (cattle, pig, chicken) revealed similarities (Picard et al. 2010). This showed that proteins such as alpha-fetoprotein, albumin, fetuin and transferrin are likely to play a major role in early myogenesis in mammals and birds as they were more abundant at the early stages of myogenesis and then decreased in abundance. The shift from enolase 1 (ENO1) expressed in the early stages of myogenesis, to expression of ENO3 during the late stages of myogenesis, was observed in all species. In all three species, Hsp27 showed a peak of abundance at stages characterised by a transition between proliferation of myoblasts and differentiation and maturation of muscle fibres. Its role could be through its anti-apoptotic or chaperone action. Another interesting finding was that apolipoprotein A1 (APOA1, high-density lipoprotein cholesterol) was highly abundant during early myogenesis and then decreased in all three species. This protein is known to be expressed predominantly in the liver and intestine, and its presence in the foetal skeletal muscle was confirmed by immunohistochemistry (Picard et al. 2010). An elevation of APOA1 was also reported in the skeletal muscle of newborn chickens (Doherty et al. 2004; Teltathum and Mekchay 2009), and it was hypothesised that APOA1 could act as a local transporter in early post-hatching development.

In pigs, the analysis of piglets with different birth weights showed that the expression of proteins involved in protein synthesis, antioxidant function, catecholamine degradation and fat metabolism was decreased in intrauterine growth-restricted piglets. On the contrary, the expression of mitochondrial F1-ATPase and proteasome was increased (Wang et al. 2008). These results demonstrate that maternal nutrition has a high impact on the expression of genes involved in myogenesis. In cattle, a recent study on the effects of nutrient restriction and refeeding from early to mid-gestation on molecular events of myogenesis revealed higher abundances of enolase 1, annexin 5 and tubulin beta 4B and 5B chains in the restricted group. This group had lower abundances of actin 1, creatine kinase M and alpha-fetoprotein. The restricted and refed group had significantly lower abundances of peptidyl-prolyl cis-trans isomerase D (Bonnet et al. 2015).

Postnatally, several studies in different species showed a relation between muscle growth and an increase in abundance of fast glycolytic proteins. For example, in chicken, the abundance of two glycolytic enzymes, phosphoglycerate mutase (PGAM1) and triosephosphate isomerase (TPI1), was positively correlated with chicken ageing (Doherty et al. 2004). An increase of TPI1 was also observed in pig muscle (Hollung et al. 2009). Fatty acid-binding protein 3 (FABP3) was negatively correlated with chicken ageing (Teltathum and Mekchay 2009) but was found to be associated with muscle growth in commercial pig breeds (Hollung et al. 2009). Several heat shock proteins such as Hsp90, Hsp27 and alpha-B crystallin were found to change in abundance during myogenesis, indicating a putative important role of these proteins and of cellular stress during myogenesis (for review, see Picard et al. 2010).

3 Skeletal Muscle Hypertrophy

Muscle hypertrophy is an important mechanism of postnatal growth and involves an increase in size of skeletal muscle through a growth in size of its component cells, that is, the muscle fibres. Several proteomics analyses have revealed that muscle hypertrophy of genetic origin [monogenic: mutation of myostatin gene (MSTN), or polygenic: from divergent selection on muscle growth potential] is characterised by high abundance of proteins representative of fast glycolytic properties. For example, Bouley and co-workers (2005) found higher abundance of glycolytic enzymes such as phosphoglucomutase (PGM), glycogen phosphorylase and enolase 1 and of fast contractile type (myosin heavy chain IIx (MyHC-IIx)), troponin T fast (fTnT) and myosin-binding protein H (MyBP-H) in the ST muscle of homozygote double-muscled (DM) Belgium Blue young bulls comparatively to homozygote non-DM. On the contrary, proteins of oxidative metabolism, such as isocitrate dehydrogenase, succinate dehydrogenase, pyruvate dehydrogenase, fATP synthase F1 b and D chain, ubiquinol-cytochrome C reductase core protein 1, NADH-ubiquinone oxidoreductase and fatty acid-binding protein (H-FABP), and of slow contractile type (troponin T slow and slow isoforms of myosin light chains) were less abundant in

the DM muscle. The muscle of heterozygote DM young bulls showed intermediate abundances of these proteins. This positive relationship between muscle mass and fast glycolytic properties has been reported also in other species. In muscles of Texel sheep with a QTL for muscle hypertrophy, a comparison to Romanov sheep by Hamelin and co-workers (2006) revealed the same differences in contractile and metabolic type. The proteomics analysis of knockout (KO) mice for the MSTN gene demonstrated a shift towards fast glycolytic type in KO mice (Chelh et al. 2009). These data suggest that some fast glycolytic proteins could be proposed as good biomarkers of muscle mass.

Using proteomics analysis, which allows the identification of several isoforms of a protein, it was highlighted for the first time that the expression of fTnT is finely regulated in cattle with muscle hypertrophy (DM and divergently selected) (Bouley et al. 2005; Picard et al. 2005). Using 2DE, these authors revealed 11 spots of fTnT. Among them, five were of higher abundance in DM and in divergently selected cattle, and the other six were not different according to muscle mass. Mass spectrometry analysis revealed that these differences are originating from alternative splicing with two mutually exclusive exons 16 and 17. Only the fTnT from exon 16 is overexpressed in hypertrophied muscles. Consequently, the ratio of fTnT exon16/fTnT exon17 was proposed by these authors as a good marker of muscle mass (Picard et al. 2005). Proteomics also revealed post-translational modifications, for example, the phosphorylation of myosin light chain 2 (MLC2) in hypertrophied muscles (Bouley et al. 2005).

In MSTN KO mice, proteomics analysis revealed higher abundance of several proteins of survival/mortality such as DJ-1, PINK1, 14-3-3 protein, TCTP/GSK-3beta and Tcp1. On the contrary, pro-apoptotic proteins such as caspase-3, caspase-8, caspase-9, BID (BH3-interacting domain death agonist), ID2 (inhibitor of DNA binding 2), Daxx (death domain-associated protein) and mortalin (GRP75) were less abundant in KO mice (Chelh et al. 2009). These differences in the abundance of proteins involved in apoptosis have been confirmed in bovine muscle (Chelh et al. 2011). These data allow the hypothesis to be proposed that muscle hypertrophy could be originated from a balance between cell survival and programmed cell death.

Some proteins involved in calcium homeostasis such as sarcoplasmic reticulum 53 kDa glycoprotein (SR53G), p20 or sarcosine, were differentially abundant between hypertrophied and non-hypertrophied muscles (Bouley et al. 2005). An intermediate abundance was observed in heterozygote DM comparatively to homozygote DM and non-DM. In coherence with these data, alpha1-syntrophin involved in the regulation of calcium influx was more abundant in MSTN KO mice (Chelh et al. 2009). In other studies of muscle hypertrophy induced by muscle growth in different species (chicken, pig, cattle), proteomics analysis revealed changes in contractile, metabolic and chaperone proteins in accordance with the modifications described above (Doherty et al. 2004; Teltathum and Mekchay 2009; Lametsch et al. 2006). Proteomics studies of muscle atrophy induced by disuse or ageing revealed drastic changes in structural and contractile proteins such as myosin, actin and troponins, which was the inverse to that observed in hypertrophied muscles.

Modification in abundances of stress proteins and of metabolic proteins with a shift from fast glycolytic to slow oxidative was also observed (for review, see Ohlendieck 2011).

4 Effects of Genetics on Muscle Protein Profiles

Genetics has a prominent effect on farm animal productivity as well as on qualities and characteristics inherent to farm animal products. Being a protein-rich product, it is not surprising that proteomics in meat would play a major role in studies aiming to differentiate breeds (or genetic groups) of farm animals with different traits and then in turn originate diverse products. Such information can be of great importance, for instance, to understand particular physiology aspects related to such breeds but very importantly as a tool for product authentication.

The pig is undoubtedly the most studied farm animal species when it comes to muscle proteomics. The majority of the work that can be found in the literature typically contrast breeds associated with intensive farming systems (international) with breeds associated with more extensive farming systems (regional or national breeds), with the first tending to have higher muscle mass and the latter fattier carcasses. The first work on this subject was conducted by Hollung and co-workers (2009). They compared the proteomes of the adductor muscle of Norwegian Landrace and Duroc at three different ages using 2DE and determined almost 100 proteins with differential abundance between the two breeds, classifying them according to molecular role. This work demonstrated for the first time that proteomics could be used to differentiate muscles according to pig breed. In the same year, Xu and co-workers (2009) conducted a similar trial on the Large White and the Meishan breeds. Again, results point towards different types of muscle structure in the two breeds that could be specifically associated with meat science. In 2010, Mach and co-workers conducted a similar study, involving a total of five different breeds (Landrace, Duroc, Large White, Pietrain and Belgian Landrace) and the semimembranosus (SM) and longissimus muscles. Similarly to the previous studies, they suggested several protein biomarkers that can be used to differentiate the muscles and the breeds.

A more complete integrated omics approach was used by an Italian team in order to differentiate the Large White and the local Casertana pig breeds. They combined proteomics with transcriptomics (Murgiano et al. 2010) and proteomics with metabolomics (D'Alessandro et al. 2011). The results showed that the autochthonous Casertana breed had an overexpression of glycolytic enzymes, whereas the selected Large White breed showed an overexpression of cell cycle and skeletal muscle growth-related genes and proteins, as well as a very different lipid composition that was determined by metabolomics. These studies showed the importance of using an integrated omics approach to study a biological question in the field of animal science and, above all, show the importance of transcriptomics to complement proteomics results. In the specific scope of this work, they show also how

metabolomics may bring an additional perspective highly complementary to proteomics and transcriptomics, by specifically focusing on the lipid composition. In pigs in recent years, similar studies have been conducted, comparing, for instance, the Yorkshire breed muscle with that of the Chinese Shaziling indigenous breed (Yang et al. 2016). Again, a combined proteomic and transcriptomics approach was used to analyse the longissimus muscle and determined that several biochemical pathways were different according to breed: acid metabolism, the glycolytic pathway and skeletal muscle growth. Yang and co-workers (2016) suggested several proteins as putative markers and conducted the corresponding validation, relating it to meat quality, which was better in the Chinese breed. Finally in 2016, Liu and co-workers used proteomics and transcriptomics to study the proteome of the longissimus muscle in normal and Canadian double-muscled Large White pigs. They conducted an in-depth integrated study of the double-muscle phenotype in pigs and identified several proteins and genes that relate with this trait.

In cattle, several studies have used proteomics to reveal the effect of different genotypes on meat quality traits. The majority of these studies do not necessarily compare different breeds of cattle but instead compare genetic groups within a population differing, for instance, in meat traits such as beef tenderness (Baldassini et al. 2015; Carvalho et al. 2014) or marbling score (Shen et al. 2012). Within this context, Gagaoua and co-workers (2015a) compared the proteomes of the LT and ST muscles in three breeds differing in meat tenderness: Aberdeen Angus, Blonde d'Aquitaine and Limousin. They established markers for meat tenderness in these three breeds and found important links between some of the proteins involved in meat tenderness (e.g. DJ-1 and Prx6; see the meat tenderness section of this chapter for further details). With a similar rationale, but focusing on meat ageing, Marino and co-workers (2014) have also conducted an interesting study comparing the sarcoplasmic proteomes of Friesian, Podolian and Romagnola–Podolian crosses, relating differences in the expression of several muscle structural proteins to different ageing processes and phenotypes in these breeds. In sheep, we could only find one study in which a breed comparison has been made using muscle proteomics. Almeida and co-workers (2016) compared three sheep breeds with different levels of tolerance to seasonal weight loss: Damara, a hardy breed from the Kalahari Desert (Almeida 2011), as well as the Dorper (intermediate tolerance) and the Australian Merino (lower tolerance). They suggested several proteins as molecular markers of tolerance to weight loss such as desmin, troponin T, phosphoglucomutase and the histidine triad nucleotide-binding protein 1 and further related these results to differences in lipid metabolism, in turn related to the fat tail in Damara animals (Alves et al. 2013; van Harten et al. 2016).

In rabbits and poultry, studies have also been conducted relating muscle proteome profiling in different breeds. In rabbits, Almeida and co-workers (2010) used 2DE and peptide mass fingerprinting to study the gastrocnemius muscle in Iberian wild rabbits and in the most important breed used in meat production, the New Zealand White (NZW) rabbit as influenced by seasonal weight loss. Similar to the examples previously described for swine and cattle, important differences were established concerning structural proteins that had higher levels of expression

in NZW having more muscle than the wild rabbits. In poultry, the breast muscle of several Italian chicken breeds (Pépoi, Padovana and Ermellinata di Rovigo) has been characterised using 2DE and mass spectrometry. The proteomics analysis allowed a clear distinction between the three breeds, generating information that can be used, for instance, in product certification (Zanetti et al. 2011).

The above-mentioned examples clearly demonstrate the usefulness of proteomics approaches to study muscle/meat proteomes specifically aiming towards breed differentiation. The rationale in the majority of the studies seems to relate to the establishment of markers that are in turn related to superior or more desirable meat quality traits. Another important aspect in these studies is the definition of proteins that can be used to specifically characterise and differentiate a meat product associated with a certain breed and/or production system and that therefore has a superior quality, showing this way the important role of proteomics as a certification tool.

5 Effects of Nutrition on Muscle Protein Profiles

Nutrition, and particularly nutritional composition, is among the most studied subjects in the context of animal science. This is not surprising, given the influence that nutritional strategies have on the profitability of animal production and on the composition and the characteristics of animal-origin proteins, particularly muscle and meat. Proteomics has seldom been used for cattle nutrition studies. In fact, all the work that is found in the literature about the subject normally uses proteomics as a tool to characterise the effect of different feeding strategies on the composition and proteome of the bovine muscle of high economic value, such as the ST or the longissimus muscles. Frequently such feeding strategies are extremely aiming to differentiate intensive from extensive production systems. Shibata and co-workers (2009) have, for instance, contrasted Japanese black cattle grass fed and fed on a diet with a high percentage of cereals. They determined several proteins that had higher expression in either groups, and results include, for instance, several glycolytic enzymes that were significantly greater in the grazing group, such as beta-enolase 3, fructose 1,6-bisphosphate aldolase A, triosephosphate isomerase and Hsp27. Finally, they related the expression level of some of the proteins with specific aspects of meat quality. In Korean Hanwoo steers, Kim and co-workers (2009) studied the proteome of the longissimus dorsi (LD) muscle at different phases of the fattening process. Proteins such as zinc finger 323 and myosin light chain were highly expressed in the late-fattening stage, while others like triosephosphate isomerase (TPI) and succinate dehydrogenase (SDH) were expressed more in the early-fattening stage. These results were finally related to the intramuscular fat content and meat quality of the muscles. Both above-mentioned strategies were based on 2DE and mass spectrometry. More recently, Almeida and co-workers (2017) used a shotgun proteomics strategy to characterise the LT muscle in the Portuguese Alentejana breed according to different growth paths, one more intensive (18 month) and based on cereal feeding and the other one

more extensive (24 months) and based on a large cycle of pasture feeding. They proposed several proteins (e.g. Myozenin-2, glycolytic enzymes and 14-3-3 protein zeta/delta) as indicators of a more intensive growth path. Myosin-binding protein H had higher abundance in the DG group, suggesting it could be associated with the discontinuous growth path and hence would be of importance as a certification tool in beef production.

In sheep, only one reference could be found concerning the use of proteomics to study the effects of nutritional levels on muscle characteristics. Almeida and co-workers (2016) compared three different sheep breeds fed under maintenance needs contrasting them to controls. They suggested several proteins as putative biomarkers of tolerance to weight loss—desmin, troponin T, phosphoglucomutase and the histidine triad nucleotide-binding protein 1—as we have previously mentioned in the breed comparison section. Given the intrinsically extensive nature of small ruminant production, it is not likely that nutritional management, contrasting intensive with extensive, will be available in the near future. It is however an important subject that would deserve more focus in the future, particularly at the level of contrasting certified and non-certified products.

In pigs, the effects of nutrition and of nutritional supplements on muscle proteome profiles have been extensively studied. Some of the studies have a physiological approach, while others are specifically related to animal production. The effect of intrauterine growth restriction (IUGR) in pigs has been known to lead to alterations in protein expression in the skeletal muscle (Wang et al. 2008, 2013), with proteolysis (proteasome alpha-5 and alpha-1 subunits) and oxidative stress-responsive (scavenger-receptor protein and alpha-1 acid glycoprotein) proteins being related to IUGR. A similar physiological approach has been conducted by Liu and co-workers (2014) to demonstrate that birth weight has an important influence on how pigs react to high-fat diets both on meat traits and at the muscle proteome levels. The results are particularly interesting as high-fat diets lead to increased expression not just of energy metabolism proteins but also of an important number of stress-responsive proteins. More recently, Pires and co-workers (2016) have studied the effect of low-protein diets in finishing pigs on the LL muscle protein profiles as mediated by lysine restriction. Authors combined iTRAQ and label-free proteomics to study this topic that was ultimately related to animal production indicators and meat quality traits. They identified 35 proteins differentially expressed that were highlighting relevant biological networks and proteins associated with muscle physiology and meat quality. Categories such as muscle contraction and structural constituents of cytoskeleton were the most significantly up-regulated proteins in the muscle from pigs fed reduced-protein diets, whereas animals fed normal-protein diets had an overexpression of proteins involved in the regulation of energy metabolism.

Dietary supplements are an important aspect of pig research, specifically aiming to increase growth performance. Non-starch polysaccharide enzymes (NSPEs), for instance, have long been used in the feed production to degrade non-starch polysaccharide to oligosaccharides increasing the nutritive value of these compounds. The effect of NSPE supplementation has been studied at the level of the longissimus

muscle in growing pigs (Zhang et al. 2015). Feeding NSPEs to pigs led to an increased abundance of proteins related to energy production, protein synthesis, muscular differentiation, immunity, oxidation resistance and detoxification and a decreased abundance of proteins related to inflammation, highlighting the importance of the supplement for rapid muscular growth in pigs. A similar approach was conducted by Costa-Lima and co-workers (2015) to study the effect of ractopamine (a leaning agent) on the sarcoplasmic muscle proteome of pork. Several proteins were found with higher abundance in the ractopamine-treated group. These included, for instance, glyceraldehyde-3-phosphate dehydrogenase and phosphoglucomutase-1, finally suggesting that ractopamine influences the abundance of enzymes involved in glycolytic metabolism, and the differential abundance of glycolytic enzymes could potentially influence the conversion of muscle to meat. Finally, and also focusing on the effect of dietary supplements but this time in poultry, Zheng and co-workers (2014) used 2DE to study the effect of probiotic *Enterococcus faecium* feeding on carcass traits and the pectoral muscle proteome. The study revealed that improved meat quality of broilers fed probiotics is related to proteome expression changes, particularly glycolytic proteins, and further contributes to the importance of the use of probiotics to improve poultry production.

6 Postmortem Changes in Muscle Proteomes

Immediately after slaughter, the muscle starts on the postmortem journey converting muscle to meat, facing restrictions in nutrient and oxygen supply, combined with a drop in temperature. These new conditions trigger the onset of several protecting mechanisms aiming at keeping the muscle functioning, at least to a minimum, for a while. However, prolonged nutrient and oxygen deficiency will eventually induce apoptosis and muscle cell death and a gradual decomposition of the muscle protein structures. This in turn is part of the tenderisation process described in more detail below.

Since the first publication in 2002 by Lametsch and co-workers, a number of studies have been performed aiming at describing the proteome changes in different muscles and species postmortem. The majority of these studies are based on the separation of proteins by 2DE followed by the identification of changing protein spots by mass spectrometry. This approach usually provides a limited number of identifications, but still some trends are appearing in the proteome patterns.

Studies from pork have demonstrated a shift in enzymes involved in energy metabolism and degradation of structural proteins (Lametsch et al. 2002; Morzel et al. 2004; Hwang et al. 2005). Among these are several fragments of actin, demonstrating a successive degradation of this structural protein through a proteolytic process. Laville and co-workers (2009) described the postmortem degradation of actin in cattle LT muscle with differences in tough and tender LT muscle, while Ouali and co-workers (2013) proposed α-actin as a good marker of postmortem apoptosis.

A shift in energy status immediately after slaughter is supported by increased abundance of enzymes involved in the glycolytic pathway and the TCA cycle observed in bovine LT muscle (Jia et al. 2006). Another observation is the shift in protein solubility during the first postmortem period. This has been studied by comparing the protein fractions in the soluble vs. the insoluble protein fractions from the same samples (Jia et al. 2006; Bjarnadottir et al. 2010). Several of the soluble proteins were still observed in the insoluble fraction probably caused by precipitation or aggregation, thus going from a soluble to an insoluble state. The mechanism for this change in solubility of proteins and enzymes is still unclear, but the pH decline caused by increased glycolysis and the release of lactate postmortem could be one explanation, as protein solubility is strongly related to the isoelectric properties. Another explanation could be due to modifications or proteolytic cleavage that also may alter the isoelectric properties.

7 Meat Quality: pH

The pH level in meat is a crucial factor that can have great impact on several important meat quality traits, including colour, postmortem proteolysis, tenderness, water holding capacity (WHC) and shelf life. Several authors studied the proteome muscle in relation to postmortem pH decline in several species (Wu et al. 2014; Huang et al. 2014; Zhang et al. 2014). Gagaoua and co-workers (2015b) analysed the relationships between 20 protein biomarkers of tenderness (Picard et al. 2014) with the pH of the LT muscle of Blonde d'Aquitaine bulls. Muscle pH was measured at three time points postmortem: 45 min (pH 45 min), 3 h (pH 3 h) and 24 h (pHu; ultimate pH). They showed that the abundance of three proteins (actin, Prx6 and MDH1) was negatively related to pH 45 min. The pH 3 h was negatively related to more proteins: glycolytic proteins (ENO3 and LDHB), heat shock proteins (Hsp20, GRP75), proteins involved in oxidative stress (Prdx6, DJ-1) and one structural protein (MyBP-H). The pHu was negatively related to structural proteins (actin, CapZ-beta), GRP75 and oxidative stress proteins (Prx6, DJ-1) and positively related to the structural protein MyHC-IIx/b. The combined analysis of the 3 pH measurements demonstrated that small Hsps and glycolytic proteins were not correlated with pHu. This indicates that glycolytic enzymes and small Hsps play an important role in the pH drop early postmortem. In accordance with these results, Lomiwes and co-workers (2014) showed that a decrease of pH rapidly induced a binding of the small Hsps to the myofibrils. This is in accordance with the implication of apoptosis early postmortem in the conversion of muscle into meat (Ouali et al. 2006, 2013). Pulford and co-workers (2008) showed that small Hsps are very sensitive to low pH values with limited expression after 22 h postmortem. The implication of glycolytic proteins in pH decline during the first hours postmortem is coherent with the high glycolytic activity in hypoxic conditions postmortem to provide energy for preserving cells from death. This is in accordance with the postmortem modifications of proteins described by Jia and

co-workers (2006). Pancholi (2001) proposed also that ENO3, in hypoxic situations, could protect cells by increasing anaerobic metabolism.

Prx6 was the only protein among the 20 studied by Gagaoua and co-workers (2015b) to be correlated with the 3 pH measurements (45 min, 3 h and pHu). This suggests an important role of oxidative process during pH decline. Prx6 is a bifunctional protein with two activities: glutathione peroxidase (GPx, antioxidative) and phospholipase A2 (PLA2). Gagaoua and co-workers (2015b) hypothesised that the action of Prx6 on pH decline could be through its PLA2 activity.

8 Meat Quality: Colour

Meat colour is an important meat quality trait, since meat-purchasing decisions are greatly influenced by colour. The reason for this is that consumers use meat colour as an indicator of freshness and overall quality. Several authors reported by proteomics analysis proteins associated with colour parameters (lightness, *L**; redness, *a**; yellowness, *b**) and colour stability in different species. The reported proteins are mainly involved in muscle contraction, energy metabolism, signalling pathways, chaperone and apoptosis functions (Sayd et al. 2006; Joseph et al. 2012, 2015; Desai et al. 2014; Suman et al. 2014; Canto et al. 2015; Wu et al. 2015). Peroxiredoxin-2, dihydropteridine reductase and Hsp27 were reported to be related to surface colour stability (Joseph et al. 2012). These authors also found several Hsp70 (Hsp70-1A/B, Hsp70-8) related to beef colour stability. Gagaoua and co-workers (2015b) found that these two proteins were negatively correlated with *L** values and positively with *b**, respectively.

Kwasiborski and co-workers (2008) found negative correlations between Hsp72 and Hsp70-1A/B and *L** in pork LT muscle. Moreover, Gagaoua and co-workers (2015b) found a positive correlation between *L** values and alpha-actin and MyHC IIa in LT muscle of Blonde d'Aquitaine bulls. Alpha-actin was also found related to *L** values (for review, see Picard et al. 2010).

In the LT of Blonde d'Aquitaine cattle, Prx6 and MDH1 were positively correlated only with *a** (Gagaoua et al. 2015b). Glycolytic enzymes such as PGM, GAPDH and ENO3 were positively correlated with *a** values in beef (Joseph et al. 2012; Canto et al. 2015). This is coherent with the role of glycolytic pathway in the production of NADH, an important cofactor for metmyoglobin formation.

Zhang and co-workers (2014) reported negative correlations between Hsp90 and *L** and *b** values but no relation with *a**. Similarly, Gagaoua and co-workers (2015b) found seven proteins correlated both with *L** and *b** parameters (alpha-B crystallin, Hsp70-1B, Hsp70-8, MyHCI, μ-calpain, ENO 3, MyHBP-H). In addition, Hsp40 was positively related to both *a** and *b** values, while ENO3 and MyBP-H were positively correlated with *L** and negatively correlated with *b** in LT muscle of cattle. In the study, the authors found that the two proteins Hsp70-1A/B and μ-calpain were correlated with the three colour parameters: negatively with

L^* and positively with a^* and b^*. These two proteins could influence meat colour interactively.

From all these data, Gagaoua and co-workers (2015b) proposed that postmortem reactive oxygen species (ROS) damage the sarcoplasmic reticulum inducing a liberation of calcium ions, which will activate μ-calpain involved in muscle proteolysis. Oxidative stress through lipid peroxidation will produce 4-hydroxy-2-nonetal (4-HNE) which induces carbonylation of Hsp70. Carbonylated Hsp70 is a substrate of μ-calpain. This hypothesis is coherent with the results of Suman and co-workers (2006), indicating that 4-HNE could affect meat colour by the formation of several adducts with muscle proteins which will modify the protein functionality.

Desai and co-workers (2014) compared the muscle proteomes of normal and reddish fillets of catfish. The beta subunit of haemoglobin was more abundant in reddish fillets. No differences were observed for myofibrillar proteins between the two groups of fillets, indicating that the red colour defect originated from a higher abundance of haemoglobin and not from differences in myofibrillar proteins abundances. Later, Desai and co-workers (2016) compared the breast meat proteome of two groups of broilers, pale, soft and exudative (PSE) broiler and normal, characterised by pH 24 and L^*24 as normal (pH 24 5.8–6.2, L^*24 45–55) or PSE-like (pH 24 5.4–5.7, L^*24 55–65). The whole muscle proteome analysis revealed 15 differentially abundant proteins in normal and PSE-like broiler breast samples. Actin alpha, myosin heavy chain, phosphoglycerate kinase, creatine kinase M type, beta-enolase, carbonic anhydrase 2, proteasome subunit alpha, pyruvate kinase and malate dehydrogenase were overabundant in PSE-like broiler breast. On the contrary, phosphoglycerate mutase-1, alpha-enolase, ATP-dependent 6-phosphofructokinase and fructose 1,6-bisphosphatase were more abundant in normal meat.

Nair and co-workers (2016) analysed the sarcoplasmic proteome of SM muscles from carcasses colour-labile inside (ISM) and colour-stable outside (OSM) beef steaks. Glycolytic enzymes (fructose-bisphosphate aldolase A, phosphoglycerate mutase 2 and beta-enolase) and phosphatidylethanolamine-binding protein 1 were more abundant in ISM steaks. These authors proposed that rapid postmortem glycolysis in ISM could lead to rapid pH decline during the early postmortem period, which in turn could alter its colour stability. Wu and co-workers (2016) analysed the relationship between meat colour parameters and the sarcoplasmic proteome of LL and psoas major (PM) muscles from Chinese Luxi yellow cattle during postmortem storage. The colour stability was better for LL than for PM muscle. These authors proposed glycerol-3-phosphate dehydrogenase, fructose-bisphosphate aldolase A isoform, glycogen phosphorylase, Prx2, phosphoglucomutase-1, superoxide dismutase [Cu-Zn] and heat shock cognate protein (71 kDa) as good candidate predictors of meat colour stability during postmortem storage. They found more glycolytic proteins in relation with meat colour in LL muscles, suggesting that the glycolytic metabolism contributes to the postmortem stability of meat.

9 Meat Quality: Water Holding Capacity

Water holding capacity (WHC) is defined as the ability meat has to retain water during postmortem storage and is one of the most important characteristics of meat, both in terms of quality and economy. Poor WHC may cause significantly reduced water content in muscles during storage and processing, and this will affect yield, technological quality and sensory attributes (e.g. juiciness and tenderness) of the meat. Thus, there is a strong interest in understanding the underlying mechanisms for variation in WHC, particularly in pork, where the problem is most pronounced.

Over the last decade, several proteomics studies have been performed with the aim of increasing our understanding of the underlying mechanisms governing this trait and to identify potential protein biomarkers for WHC. In almost all studies, structural proteins and proteins involved in energy metabolism and stress protection have been found to vary depending on WHC. For proteins related to energy metabolism, increased levels of both full-length proteins and fragments have been found to coincide with high drip loss. Hwang and co-workers (2005) found a greater degradation of adenylate kinase in pork with high drip loss, while van de Wiel and Zhang (2007) observed higher levels of full-length creatine kinase in high drip loss pork. The link between increased level of creatine kinase and high drip loss is thought to relate to a more rapid pH decline and muscle contraction in these muscles (van de Wiel and Zhang 2007). Moreover, enzymes involved in glycolysis, lactate production and the TCA cycle have been reported to vary depending on the extent of drip loss in pork (Di Luca et al. 2013; te Pas et al. 2013). Of the structural proteins, changes in both myofibrillar and cytoskeletal proteins have been found to differ between groups with varying WHC (Hwang et al. 2005; Di Luca et al. 2013; te Pas et al. 2013). Another important protein group that has been indicated to play a role in WHC is the stress-related proteins, particularly Hsp70 (Di Luca et al. 2011, 2013). Di Luca and co-workers (2013) found proteins from metabolic pathways, stress response and structural proteins, as well as transport proteins to be associated with WHC, and many of these proteins were also found to discriminate between different levels of drip loss in a follow-up study, where a multivariate data analysis approach was applied (Di Luca et al. 2016).

A proteomics approach has also been used to study the importance of protein denaturation and oxidation in relation to WHC of pork (Bernevic et al. 2011). Specifically, Bernevic and co-workers found more denaturation of myofibrillar proteins in the high drip loss group, while there was a higher abundance of oxidation of contractile proteins in the low drip loss group. Recently, an integrated multi-omics approach has also been used to reveal functional pathways and candidate gene markers for drip loss in pigs (Welzenbach et al. 2016). This study confirmed the importance of the energy-related metabolic processes and showed in particular that the sphingolipid metabolism may be important for drip loss.

10 Meat Quality: Tenderness

Tenderness is the most important meat quality parameter, especially for beef, as judged by consumers (Miller et al. 2001). Protein proteolysis caused by the calpain proteases is the main contributor to tenderness development in meat. However, several other factors are also involved, but the molecular mechanisms are still not fully understood. Over the last years, several studies have focused on unravelling protein markers for tenderness in different cattle breeds and muscles. However, most studies are done with the longissimus muscle. From these proteome studies, proteins related to different cellular functions play a role.

Proteins associated with oxidative stress and apoptosis have been suggested as markers of tenderness in several studies. Peroxiredoxin 6 (Prx6) is an antioxidant enzyme important for the detoxification of reactive oxygen species (ROS) in the cells. Higher levels of Prx6 were observed in the tender muscles; thus a lower extent of ROS can be expected (Jia et al. 2009; Polati et al. 2012; Guillemin et al. 2011b). In line with this, it has also been shown that ROS inhibits the function of μ-calpain, one of the important tenderising enzymes in muscle. Thus, the increased abundance of Prx6 can reduce the level of ROS which will enhance the function of μ-calpain and eventually lead to more tender meat. The correlation between Prx6 and μ-calpain is further supported in a recent study by Gagaoua and co-workers (2015a).

Several members of the heat shock family of proteins have anti-apoptotic properties and may contribute to delay the postmortem onset of apoptosis and thus also influence on the tenderness development and the rate of conversion of muscle to meat (Ouali et al. 2006, 2013). Changes in abundance of heat shock proteins, such as alpha-B crystallin, Hsp20, Hsp27, Hsp40 and Hsp70, are demonstrated in several proteome studies comparing tender and tough meat (Hwang et al. 2005; Morzel et al. 2008; Jia et al. 2009; Bjarnadottir et al. 2012; Polati et al. 2012; Gagaoua et al. 2015a; Picard et al. 2014; Kim et al. 2008; Hwang et al. 2005). Together these studies strongly support that apoptosis is an important mechanism during the early phase of the tenderisation process postmortem.

Loosening of the myofibrillar structure of the muscle is probably the most critical part of tenderness development. Thus, it is expected to find fragments of myofibrillar proteins in the proteome studies of muscle and meat postmortem and to a various extent according to the level of tenderness. This was first described in 2003 by Lametsch and co-workers who found changes in abundance of the major myofibrillar proteins such as actin, myosin heavy chain, myosin light chain and titin, suggesting a relationship between degradation of these proteins postmortem and tenderness in pork. These findings are further supported in numerous studies in several species and muscles. Fragments of alpha-actin are correlated with tenderness in three French beef breeds and are suggested to be biomarker of tenderness (Chaze et al. 2013). This is further supported by observations in other pig and cattle breeds postmortem (Morzel et al. 2008; Zapata et al. 2009).

Glycolytic enzymes have also been associated with tenderness development in several studies. Higher abundance of glycolytic enzymes was, for instance, found in tender LD muscle of Chianina cattle (D'Alessandro et al. 2012). Similarly, Picard and co-workers (2014) compared glycolytic, fast oxido-glycolytic and oxidative muscles and found that the muscles with a greater proportion of fast glycolytic fibres produced the most tender meat and also that oxidative muscles with a greater proportion of slow oxidative fibres will be more tender.

Efforts have been made to combine and compare the different studies aiming at finding protein markers of tenderness in different breeds and species over the years. Bioinformatics was used in a study by Guillemin and co-workers (2011a) to create functional networks of 24 proteins published as potential markers in the ST and LD muscles of beef. The study confirmed that cellular pathways involved in apoptosis, oxidative stress and Hsps are important for tenderness. Four proteins, μ-calpain, Hsp70-1A/B, Hsp70-8 and Hsp27, were related to a high number of other proteins and were used to make four principal networks. The study unravelled both similar and distinct pathways between the two muscles in the study. These results are further strengthened in the study by Gagaoua and co-workers (2015a). By a similar approach making correlation networks of 18 proteins which had been previously suggested as tenderness markers in the literature, they were able to demonstrate robust relationships between proteins belonging to similar or different biological pathways. Some proteins were correlated regardless of muscle or breed, such as the correlation between DJ-1 and Hsp20 and Prx6 and μ-calpain.

In conclusion, the studies published so far suggest that proteins involved in mitochondrial function, oxidative stress or proteins that have cell-protective functions play key roles in tenderness development of muscle and meat. This also supports the hypothesis that apoptosis is a critical event during early postmortem and that a delay of the apoptotic process is beneficial to obtain tender meat.

11 Conclusion

The application of proteomics within the field of skeletal muscle research in farm animals has been of major importance in increasing our understanding of various biological processes as well as the revealing the underlying mechanisms governing many of the important meat quality traits. First, the initial work related to characterisation of muscle proteomes in the various farm animal species has been of great importance. Second, proteomics has given insight into how both genetics and animal management such as feeding have manifested their biological or phenotypic effects on animals. Last, since many meat quality traits have relatively low heritability but are highly influenced by environment and treatment, proteomics has been a great tool to gain more knowledge regarding this influence and to search for potential biomarkers.

References

Almeida AM (2011) The Damara in the context of Southern Africa fat-tailed sheep breeds. Tropl Anim Health Prod 43(7):1427–1441. https://doi.org/10.1007/s11250-011-9868-3

Almeida AM (2013) Rabbit muscle proteomics: a great leap forward. Proteomics 13(15):2225–2226. https://doi.org/10.1002/pmic.201300205

Almeida AM, Campos A, van Harten S, Cardoso LA, Coelho AV (2009) Establishment of a proteomic reference map for the gastrocnemius muscle in the rabbit (*Oryctolagus cuniculus*). Res Vet Sci 87(2):196–199. https://doi.org/10.1016/j.rvsc.2009.02.011

Almeida AM, Campos A, Francisco R, van Harten S, Cardoso LA, Coelho AV (2010) Proteomic investigation of the effects of weight loss in the gastrocnemius muscle of wild and NZW rabbits via 2D-electrophoresis and MALDI-TOF MS. Anim Genet 41(3):260–272. https://doi.org/10.1111/j.1365-2052.2009.01994.x

Almeida AM, Palhinhas RG, Kilminster T, Scanlon T, van Harten S, Milton J, Blache D, Greeff J, Oldham C, Coelho AV, Cardoso LA (2016) The effect of weight loss on the muscle proteome in the Damara, Dorper and Australian Merino Ovine Breeds. PLoS One 11(2). https://doi.org/10.1371/journal.pone.0146367

Almeida AM, Nanni P, Ferreira AM, Fortes C, Grossmann J, Bessa RJB, Costa P (2017) The longissimus thoracis muscle proteome in Alentejana bulls as affected by growth path. J Proteomics 152:206–2015. https://doi.org/10.1016/j.jprot.2016.10.020

Alves SP, Bessa RJB, Quaresma MAG, Kilminster T, Scanlon T, Oldham C, Milton J, Greeff J, Almeida AM (2013) Does the fat tailed Damara Ovine Breed have a distinct lipid metabolism leading to a high concentration of branched chain fatty acids in tissues? PLoS One 8(10). https://doi.org/10.1371/journal.pone.0077313

Baldassini WA, Braga CP, Chardulo LAL, Silva J, Malheiros JM, de Albuquerque LG, Fernandes TT, Padilha PD (2015) Bioanalytical methods for the metalloproteomics study of bovine longissimus thoracis muscle tissue with different grades of meat tenderness in the Nellore breed (*Bos indicus*). Food Chem 169:65–72. https://doi.org/10.1016/j.foodchem.2014.07.131

Bernevic B, Petre BA, Galetskiy D, Werner C, Wicke M, Schellander K, Przybylski M (2011) Degradation and oxidation postmortem of myofibrillar proteins in porcine skeleton muscle revealed by high resolution mass spectrometric proteome analysis. Int J Mass Spectrom 305 (2–3):217–227. https://doi.org/10.1016/j.ijms.2010.11.010

Bjarnadottir SG, Hollung K, Faergestad EM, Veiseth-Kent E (2010) Proteome changes in bovine Longissimus thoracis muscle during the first 48 h postmortem: shifts in energy status and myofibrillar stability. J Agric Food Chem 58(12):7408–7414. https://doi.org/10.1021/jf100697h

Bjarnadottir SG, Hollung K, Hoy M, Bendixen E, Codrea MC, Veiseth-Kent E (2012) Changes in protein abundance between tender and tough meat from bovine Longissimus thoracis muscle assessed by isobaric tag for relative and absolute quantitation (iTRAQ) and 2-dimensional gel electrophoresis analysis. J Anim Sci 90(6):2035–2043. https://doi.org/10.2527/jas.2011-4721

Bonnet M, Kaspric N, Vonnahme K, Barboiron C, Viala D, Chadeyron D, Picard B (2015) Maternal nutrient restriction followed by realimentation from early to mid-gestation in beef cows impacts proteomes of fetal muscular and adipose tissues. ICoMST, Clermont-Ferrand, p 111

Bouley J, Chambon C, Picard B (2004) Mapping of bovine skeletal muscle proteins using two-dimensional gel electrophoresis and mass spectrometry. Proteomics 4(6):1811–1824. https://doi.org/10.1002/pmic.200300688

Bouley J, Meunier B, Chambon C, De Smet S, Hocquette JF, Picard B (2005) Proteomic analysis of bovine skeletal muscle hypertrophy. Proteomics 5(2):490–500

Canto A, Suman SP, Nair MN, Li ST, Rentfrow G, Beach CM, Silva TJP, Wheeler TL, Shackelford SD, Grayson A, McKeith RO, King DA (2015) Differential abundance of sarcoplasmic proteome explains animal effect on beef Longissimus lumborum color stability. Meat Sci 102:90–98. https://doi.org/10.1016/j.meatsci.2014.11.011

Carvalho ME, Gasparin G, Poleti MD, Rosa AF, Balieiro JCC, Labate CA, Nassu RT, Tullio RR, Regitano LCD, Mourao GB, Coutinho LL (2014) Heat shock and structural proteins associated with meat tenderness in Nellore beef cattle, a *Bos indicus* breed. Meat Sci 96(3):1318–1324. https://doi.org/10.1016/j.meatsci.2013.11.014

Chan CYX, McDermott JC, Siu KWM (2011) Secretome analysis of skeletal myogenesis using SILAC and shotgun proteomics. Int J Proteomics 2011:13. https://doi.org/10.1155/2011/329467

Chaze T, Bouley J, Chambon C, Barboiron C, Picard B (2006) Mapping of alkaline proteins in bovine skeletal muscle. Proteomics 6(8):2571–2575

Chaze T, Meunier B, Picard B (2008) Description of in vivo bovine myogenesis using proteomic approach gives new insights for muscle development. Archiv Fur Tierzucht Arch Anim Breed 51:51–57

Chaze T, Meunier B, Chambon C, Jurie C, Picard B (2009) Proteome dynamics during contractile and metabolic differentiation of bovine foetal muscle. Animal 3(7):980–1000. https://doi.org/10.1017/s1751731109004315

Chaze T, Hocquette JF, Meunier B, Renand G, Jurie C, Chambon C, Journaux L, Rousset S, Denoyelle C, Lepetit J, Picard B (2013) Biological markers for meat tenderness of the three main French beef breeds using 2-DE and MS approach. In: Toldra F, Nollet LML (eds) Proteomics in foods. Springer US, New York, pp 127–146

Chelh I, Meunier B, Picard B, Reecy MJ, Chevalier C, Hocquette JF, Cassar-Malek I (2009) Molecular profiles of Quadriceps muscle in myostatin-null mice reveal PI3K and apoptotic pathways as myostatin targets. BMC Genomics 10. https://doi.org/10.1186/1471-2164-10-196

Chelh I, Picard B, Hocquette JF, Cassar-Malek I (2011) Myostatin inactivation induces a similar muscle molecular signature in double-muscled cattle as in mice. Animal 5(2):278–286. https://doi.org/10.1017/s1751731110001862

Costa-Lima BRC, Suman SP, Li ST, Beach CM, Silva TJP, Silveira ETF, Bohrer BM, Boler DD (2015) Dietary ractopamine influences sarcoplasmic proteome profile of pork Longissimus thoracis. Meat Sci 103:7–12. https://doi.org/10.1016/j.meatsci.2014.12.008

D'Alessandro A, Marrocco C, Zolla V, D'Andrea M, Zolla L (2011) Meat quality of the longissimus lumborum muscle of Casertana and Large White pigs: metabolomics and proteomics intertwined. J Proteomics 75(2):610–627. https://doi.org/10.1016/j.jprot.2011.08.024

D'Alessandro A, Marrocco C, Rinalducci S, Mirasole C, Failla S, Zolla L (2012) Chianina beef tenderness investigated through integrated Omics. J Proteomics 75(14):4381–4398. https://doi.org/10.1016/j.jprot.2012.03.052

Desai MA, Joseph P, Suman SP, Silva JL, Kim T, Schilling MW (2014) Proteome basis of red color defect in channel catfish (*Ictalurus punctatus*) fillets. LWT Food Sci Technol 57(1):141–148. https://doi.org/10.1016/j.lwt.2014.01.001

Desai MA, Jackson V, Zhai W, Suman SP, Nair MN, Beach CM, Schilling MW (2016) Proteome basis of pale, soft, and exudative-like (PSE-like) broiler breast (Pectoralis major) meat. Poult Sci 95(11):2696–2706. https://doi.org/10.3382/ps/pew213

Di Luca A, Mullen AM, Elia G, Davey G, Hamill RM (2011) Centrifugal drip is an accessible source for protein indicators of pork ageing and water-holding capacity. Meat Sci 88(2):261–270. https://doi.org/10.1016/j.meatsci.2010.12.033

Di Luca A, Elia G, Hamill R, Mullen AM (2013) 2D DIGE proteomic analysis of early post mortem muscle exudate highlights the importance of the stress response for improved water-holding capacity of fresh pork meat. Proteomics 13(9):1528–1544. https://doi.org/10.1002/pmic.201200145

Di Luca A, Hamill RM, Mullen AM, Slavov N, Elia G (2016) Comparative proteomic profiling of divergent phenotypes for water holding capacity across the post mortem ageing period in porcine muscle exudate. PLoS One 11(3). https://doi.org/10.1371/journal.pone.0150605

Doherty MK, McLean L, Hayter JR, Pratt JM, Robertson DHL, El-Shafei A, Gaskell SJ, Beynon RJ (2004) The proteome of chicken skeletal muscle: changes in soluble protein expression during growth in a layer strain. Proteomics 4:2082–2093

Ferranti P, Gualtieri L, Sacchi R, Quarto M, Chianese L (2007) Molecular characterization of water buffalo meat by proteomic techniques. Ital J Anim Sci 6:1182–1186

Gagaoua M, Terlouw EMC, Boudjellal A, Picard B (2015a) Coherent correlation networks among protein biomarkers of beef tenderness: what they reveal. J Proteomics 128:365–374. https://doi.org/10.1016/j.jprot.2015.08.022

Gagaoua M, Terlouw EMC, Micol D, Boudjellal A, Hocquette JF, Picard B (2015b) Understanding early post-mortem biochemical processes underlying meat color and pH decline in the longissimus thoracis muscle of young blond d'Aquitaine bulls using protein biomarkers. J Agric Food Chem 63(30):6799–6809. https://doi.org/10.1021/acs.jafc.5b02615

Guillemin N, Bonnet M, Jurie C, Picard B (2011a) Functional analysis of beef tenderness. J Proteomics 75(2):352–365. https://doi.org/10.1016/j.jprot.2011.07.026

Guillemin N, Jurie C, Cassar-Malek I, Hocquette JF, Renand G, Picard B (2011b) Variations in the abundance of 24 protein biomarkers of beef tenderness according to muscle and animal type. Animal 5(6):885–894. https://doi.org/10.1017/s1751731110002612

Hamelin M, Sayd T, Chambon C, Bouix J, Bibe B, Milenkovic D, Leveziel H, Georges M, Clop A, Marinova P, Laville E (2006) Proteomic analysis of ovine muscle hypertrophy. J Anim Sci 84 (12):3266–3276. https://doi.org/10.2527/jas.2006-162

He H, Chen S, Liang W, Liu X (2016) Genome-wide proteomics analysis on longissimus muscles in Qinchuan beef cattle. Anim Genet. https://doi.org/10.1111/age.12508

Hollung K, Grove H, Faergestad EM, Sidhu MS, Berg P (2009) Comparison of muscle proteome profiles in pure breeds of Norwegian Landrace and Duroc at three different ages. Meat Sci 81 (3):487–492. https://doi.org/10.1016/j.meatsci.2008.10.003

Huang HG, Larsen MR, Palmisano G, Dai J, Lametsch R (2014) Quantitative phosphoproteomic analysis of porcine muscle within 24 h postmortem. J Proteomics 106:125–139. https://doi.org/10.1016/j.jprot.2014.04.020

Hwang IH, Park BY, Kim JH, Cho SH, Lee JM (2005) Assessment of postmortem proteolysis by gel-based proteome analysis and its relationship to meat quality traits in pig longissimus. Meat Sci 69(1):79–91. https://doi.org/10.1016/j.meatsci.2004.06.019

Jia XH, Hildrum KI, Westad F, Kummen E, Aass L, Hollung K (2006) Changes in enzymes associated with energy metabolism during the early post mortem period in Longissimus thoracis bovine muscle analyzed by proteomics. J Proteome Res 5(7):1763–1769

Jia X, Veiseth-Kent E, Grove H, Kuziora P, Aass L, Hildrum KI, Hollung K (2009) Peroxiredoxin-6—a potential protein marker for meat tenderness in bovine longissimus thoracis muscle. J Anim Sci 87(7):2391–2399. https://doi.org/10.2527/jas.2009-1792

Joseph P, Suman SP, Rentfrow G, Li ST, Beach CM (2012) Proteomics of muscle-specific beef color stability. J Agric Food Chem 60(12):3196–3203. https://doi.org/10.1021/jf204188v

Joseph P, Nair MN, Suman SP (2015) Application of proteomics to characterize and improve color and oxidative stability of muscle foods. Food Res Int 76:938–945. https://doi.org/10.1016/j.foodres.2015.05.041

Kaspric N, Picard B, Reichstadt M, Tournayre J, Bonnet M (2015a) ProteINSIDE to easily investigate proteomics data from ruminants: application to mine proteome of adipose and muscle tissues in bovine foetuses. PLoS One 10(5):e0128086. https://doi.org/10.1371/journal.pone.0128086

Kaspric N, Reichstadt M, Picard B, Tournayre J, Bonnet M (2015b) Protein function easily investigated by genomics data mining using the proteINSIDE online tool. Genomics Comput Biol 1(1):e16. 10.18547/gcb.2015.vol1.iss1.e16

Kim NK, Cho S, Lee SH, Park HR, Lee CS, Cho YM, Choy YH, Yoon D, Im SK, Park EW (2008) Proteins in longissimus muscle of Korean native cattle and their relationship to meat quality. Meat Sci 80(4):1068–1073. https://doi.org/10.1016/j.meatsci.2008.04.027

Kim NK, Lee SH, Cho YM, Son ES, Kim KY, Lee CS, Yoon D, Im SK, SJ O, Park EW (2009) Proteome analysis of the M. longissimus dorsi between fattening stages in Hanwoo steer. BMB Rep 42(7):433–438

Kwasiborski A, Sayd T, Chambon C, Sante-Lhoutellier V, Rocha D, Terlouw C (2008) Pig longissimus lumborum proteome: Part II: relationships between protein content and meat quality. Meat Sci 80(4):982–996. https://doi.org/10.1016/j.meatsci.2008.04.032

Lametsch R, Roepstorff P, Bendixen E (2002) Identification of protein degradation during post-mortem storage of pig meat. J Agric Food Chem 50:5508–5512

Lametsch R, Karlsson A, Rosenvold K, Andersen HJ, Roepstorff P, Bendixen E (2003) Postmortem proteome changes of porcine muscle related to tenderness. J Agric Food Chem 51(24):6992–6997

Lametsch R, Kristensen L, Larsen MR, Therkildsen M, Oksbjerg N, Ertbjerg P (2006) Changes in the muscle proteome after compensatory growth in pigs. J Anim Sci 84(4):918–924

Laville E, Sayd T, Morzel M, Blinet S, Chambon C, Lepetit J, Renand G, Hocquette JF (2009) Proteome changes during meat aging in tough and tender beef suggest the importance of apoptosis and protein solubility for beef aging and tenderization. J Agric Food Chem 57(22):10755–10764

Liu ZY, XN D, Yin CC, Chang ZZ (2013) Shotgun proteomic analysis of sarcoplasmic reticulum preparations from rabbit skeletal muscle. Proteomics 13(15):2335–2338. https://doi.org/10.1002/pmic.201200138

Liu JB, He J, Yu J, Mao XB, Zheng P, Huang ZQ, Yu B, Chen DW (2014) Birth weight alters the response to postnatal high-fat diet-induced changes in meat quality traits and skeletal muscle proteome of pigs. Br J Nutr 111(10):1738–1747. https://doi.org/10.1017/s0007114513004431

Liu SQ, Han WP, Jiang SY, Zhao CJ, CX W (2016) Integrative transcriptomics and proteomics analysis of longissimus dorsi muscles of Canadian double-muscled Large White pigs. Gene 577(1):14–23. https://doi.org/10.1016/j.gene.2015.11.016

Lomiwes D, Hurst SM, Dobbie P, Frost DA, Hurst RD, Young OA, Farouk MM (2014) The protection of bovine skeletal myofibrils from proteolytic damage post mortem by small heat shock proteins. Meat Sci 97(4):548–557. https://doi.org/10.1016/j.meatsci.2014.03.016

Mach N, Keuning E, Kruijt L, Hortos M, Arnau J, Pas M (2010) Comparative proteomic profiling of 2 muscles from 5 different pure pig breeds using surface-enhanced laser desorption/ionization time-of-flight proteomics technology. J Anim Sci 88(4):1522–1534. https://doi.org/10.2527/jas.2009-2256

Marino R, Albenzio M, della Malva A, Caroprese M, Santillo A, Sevi A (2014) Changes in meat quality traits and sarcoplasmic proteins during aging in three different cattle breeds. Meat Sci 98(2):178–186. https://doi.org/10.1016/j.meatsci.2014.05.024

Miller MF, Carr MA, Ramsey CB, Crockett KL, Hoover LC (2001) Consumer thresholds for establishing the value of beef tenderness. J Anim Sci 79(12):3062–3068

Morzel M, Chambon C, Hamelin M, Sante-Lhoutellier V, Sayd T, Monin G (2004) Proteome changes during pork meat ageing following use of two different pre-slaughter handling procedures. Meat Sci 67:689–696

Morzel M, Terlouw C, Chambon C, Micol D, Picard B (2008) Muscle proteome and meat eating qualities of longissimus thoracis of "Blonde d'Aquitaine" young bulls: a central role of HSP27 isoforms. Meat Sci 78(3):297–304. https://doi.org/10.1016/j.meatsci.2007.06.016

Murgiano L, D'Alessandro A, Egidi MG, Crisa A, Prosperini G, Timperlo AM, Valentini A, Zolla L (2010) Proteomics and transcriptomics investigation on longissimus muscles in Large White and Casertana Pig breeds. J Proteome Res 9(12):6450–6466. https://doi.org/10.1021/pr100693h

Nair MN, Suman SP, Chatli MK, Li ST, Joseph P, Beach CM, Rentfrow G (2016) Proteome basis for intramuscular variation in color stability of beef semimembranosus. Meat Sci 113:9–16. https://doi.org/10.1016/j.meatsci.2015.11.003

Ohlendieck K (2011) Skeletal muscle proteomics: current approaches, technical challenges and emerging techniques. Skeletal Muscle 1. https://doi.org/10.1186/2044-5040-1-6

Ouali A, Herrera-Mendez CH, Coulis G, Becila S, Boudjellal A, Aubry L, Sentandreu MA (2006) Revisiting the conversion of muscle into meat and the underlying mechanisms. Meat Sci 74(1):44–58. https://doi.org/10.1016/j.meatsci.2006.05.010

Ouali A, Gagaoua M, Boudida Y, Becila S, Boudjellal A, Herrera-Mendez CH, Sentandreu MA (2013) Biomarkers of meat tenderness: present knowledge and perspectives in regards to our current understanding of the mechanisms involved. Meat Sci 95(4):854–870. https://doi.org/10.1016/j.meatsci.2013.05.010

Pancholi V (2001) Multifunctional alpha-enolase: its role in diseases. Cell Mol Life Sci 58 (7):902–920. https://doi.org/10.1007/pl00000910

Picard B, Bouley J, Cassar-Malek I, Bernard C, Renand G, Hocquette JF (2005) Proteomics applied to the analysis of bovine muscle hypertrophy. In: Hocquette JF, Gigli S (eds) Indicators of milk and beef quality. Wageningen Academic Publishers, Wageningen, pp 379–384

Picard B, Berri C, Lefaucheur L, Molette C, Sayd T, Terlouw C (2010) Skeletal muscle proteomics in livestock production. Brief Funct Genomics 9(3):259–278. https://doi.org/10.1093/bfgp/elq005

Picard B, Gagaoua M, Micol D, Cassar-Malek I, Hocquette JF, Terlouw CEM (2014) Inverse relationships between biomarkers and beef tenderness according to contractile and metabolic properties of the muscle. J Agric Food Chem 62(40):9808–9818. https://doi.org/10.1021/jf501528s

Pires VMR, Madeira MS, Dowle AA, Thomas J, Almeida AM, Prates JAM (2016) Increased intramuscular fat induced by reduced dietary protein in finishing pigs: effects on the longissimus lumborum muscle proteome. Mol Biosyst 12(8):2447–2457. https://doi.org/10.1039/c6mb00213g

Polati R, Menini M, Robotti E, Millioni R, Marengo E, Novelli E, Balzan S, Cecconi D (2012) Proteomic changes involved in tenderization of bovine longissimus dorsi muscle during prolonged ageing. Food Chem 135(3):2052–2069. https://doi.org/10.1016/j.foodchem.2012.06.093

Pulford DJ, Vazquez SF, Frost DF, Fraser-Smith E, Dobbie P, Rosenvold K (2008) The intracellular distribution of small heat shock proteins in post-mortem beef is determined by ultimate pH. Meat Sci 79(4):623–630. https://doi.org/10.1016/j.meatsci.2007.10.027

Ronnov-Jessen L, Villadsen R, Edwards JC, Petersen OW (2002) Differential expression of a chloride intracellular channel gene, CLIC4, in transforming growth factor-beta 1-mediated conversion of fibroblasts to myofibroblasts. Am J Pathol 161(2):471–480. https://doi.org/10.1016/s0002-9440(10)64203-4

Sayd T, Morzel M, Chambon C, Franck M, Figwer P, Larzul C, Le Roy P, Monin G, Cherel P, Laville E (2006) Proteome analysis of the sarcoplasmic fraction of pig Semimembranosus muscle: implications on meat color development. J Agric Food Chem 54(7):2732–2737

Shen YN, Kim SH, Yoon DH, Lee HG, Kang HS, Seo KS (2012) Proteome analysis of bovine longissimus dorsi muscle associated with the marbling score. Asian Australas J Anim Sci 25(8):1083–1088. https://doi.org/10.5713/ajas.2012.12127

Shibata M, Matsumoto K, Oe M, Ohnishi-Kameyama M, Ojima K, Nakajima I, Muroya S, Chikuni K (2009) Differential expression of the skeletal muscle proteome in grazed cattle. J Anim Sci 87(8):2700–2708. https://doi.org/10.2527/jas.2008-1486

Soares R, Franco C, Pires E, Ventosa M, Palhinhas R, Koci K, de Almeida AM, Coelho AV (2012) Mass spectrometry and animal science: protein identification strategies and particularities of farm animal species. J Proteomics 75(14):4190–4206. https://doi.org/10.1016/j.jprot.2012.04.009

Suman SP, Faustman C, Stamer SL, Liebler DC (2006) Redox instability induced by 4-hydroxy-2-nonenal in porcine and bovine myoglobins at pH 5.6 and 4 degrees C. J Agric Food Chem 54(9):3402–3408. https://doi.org/10.1021/jf052811y

Suman SP, Rentfrow G, Nair MN, Joseph P (2014) 2013 EARLY CAREER ACHIEVEMENT AWARD—proteomics of muscle- and species-specificity in meat color stability. J Anim Sci 92(3):875–882. https://doi.org/10.2527/jas.2013-7277

te Pas MFW, Kruijt L, Pierzchala M, Crump RE, Boeren S, Keuning E, Hoving-Bolink R, Hortos M, Gispert M, Arnau J, Diestre A, Mulder HA (2013) Identification of proteomic biomarkers in M. longissimus dorsi as potential predictors of pork quality. Meat Sci 95(3):679–687. https://doi.org/10.1016/j.meatsci.2012.12.015

Teltathum T, Mekchay S (2009) Proteome changes in Thai Indigenous chicken muscle during growth period. Int J Biol Sci 5(7):679–685
van de Wiel DFM, Zhang WL (2007) Identification of pork quality parameters by proteomics. Meat Sci 77(1):46–54. https://doi.org/10.1016/j.meatsci.2007.04.017
van Harten S, Kilminster T, Scanlon T, Milton J, Oldham C, Greeff J, Almeida AM (2016) Fatty acid composition of the ovine longissimus dorsi muscle: effect of feed restriction in three breeds of different origin. J Sci Food Agric 96(5):1777–1782. https://doi.org/10.1002/jsfa.7285
Wang JJ, Chen LX, Li DF, Yin YL, Wang XQ, Li P, Dangott LJ, WX H, GY W (2008) Intrauterine growth restriction affects the proteomes of the small intestine, liver, and skeletal muscle in newborn pigs. J Nutr 138(1):60–66
Wang TJ, Liu C, Feng CP, Wang XQ, Lin G, Zhu YH, Yin JD, Li DF, Wang JJ (2013) IUGR alters muscle fiber development and proteome in fetal pigs. Front Biosci Landmark 18:598–607. https://doi.org/10.2741/4123
Welzenbach J, Neuhoff C, Heidt H, Cinar MU, Looft C, Schellander K, Tholen E, Grosse-Brinkhaus C (2016) Integrative analysis of metabolomic, proteomic and genomic data to reveal functional pathways and candidate genes for drip loss in pigs. Int J Mol Sci 17(9). https://doi.org/10.3390/ijms17091426
Wu G, Farouk MM, Clerens S, Rosenvold K (2014) Effect of beef ultimate pH and large structural protein changes with aging on meat tenderness. Meat Sci 98(4):637–645. https://doi.org/10.1016/j.meatsci.2014.06.010
Wu W, Gao XG, Dai Y, Fu Y, Li XM, Dai RT (2015) Post-mortem changes in sarcoplasmic proteome and its relationship to meat color traits in M. semitendinosus of Chinese Luxi yellow cattle. Food Res Int 72:98–105. https://doi.org/10.1016/j.foodres.2015.03.030
Wu W, QQ Y, Fu Y, Tian XJ, Jia F, Li XM, Dai RT (2016) Towards muscle-specific meat color stability of Chinese Luxi yellow cattle: a proteomic insight into post-mortem storage. J Proteomics 147:108–118. https://doi.org/10.1016/j.jprot.2015.10.027
Xu YJ, Jin ML, Wang LJ, Zhang AD, Zuo B, Xu DQ, Ren ZQ, Lei MG, Mo XY, Li FE, Zheng R, Deng CY, Xiong YZ (2009) Differential proteome analysis of porcine skeletal muscles between Meishan and Large White. J Anim Sci 87(8):2519–2527. https://doi.org/10.2527/jas.2008-1708
Yang H, Xu XL, Ma HM, Jiang J (2016) Integrative analysis of transcriptomics and proteomics of skeletal muscles of the Chinese indigenous Shaziling pig compared with the Yorkshire breed. BMC Genet 17. https://doi.org/10.1186/s12863-016-0389-y
Yu T-Y, Morton JD, Clerens S, Dyer JM (2015) In-depth characterisation of the lamb meat proteome from longissimus lumborum. EuPA Open Proteomics 6:28–41. https://doi.org/10.1016/j.euprot.2015.01.001
Zanetti E, Molette C, Chambon C, Pinguet J, Remignon H, Cassandro M (2011) Using 2-DE for the differentiation of local chicken breeds. Proteomics 11(13):2613–2619. https://doi.org/10.1002/pmic.201000639
Zapata I, Zerby HN, Wick M (2009) Functional proteomic analysis predicts beef tenderness and the tenderness differential. J Agric Food Chem 57(11):4956–4963. https://doi.org/10.1021/jf900041j
Zhang MH, Wang DY, Geng ZM, Bian H, Liu F, Zhu YZ, Xu WM (2014) The level of heat shock protein 90 in pig Longissimus dorsi muscle and its relationship with meat pH and quality. Food Chem 165:337–341. https://doi.org/10.1016/j.foodchem.2014.05.111
Zhang JZ, Gao Y, Lu QP, Sa RN, Zhang HF (2015) iTRAQ-based quantitative proteomic analysis of longissimus muscle from growing pigs with dietary supplementation of non-starch polysaccharide enzymes. J Zhejiang Univ Sci B 16(6):465–478. https://doi.org/10.1631/jzus.B1400266
Zheng AJ, Luo JJ, Meng K, Li JK, Zhang S, Li K, Liu GH, Cai HY, Bryden WL, Yao B (2014) Proteome changes underpin improved meat quality and yield of chickens (*Gallus gallus*) fed the probiotic *Enterococcus faecium*. BMC Genomics 15. https://doi.org/10.1186/1471-2164-15-1167

Proteomics and the Characterization of Fatty Liver Metabolism in Early Lactation Dairy Cows

Björn Kuhla and Klaus L. Ingvartsen

Abstract The high-yielding dairy cow faces major adaptions during the transition period from late pregnancy to early lactation where physiological changes occur in support of the dramatically increase in milk yield. The coordinated physiological changes secure mobilization of nutrients and energy from the body tissue that, in part, covers the rapid increase in nutrient needs for milk production in early lactation. Large amounts of energy are released from fat tissue as non-esterified fatty acids (NEFAs) which together with a depressed increase in energy intake are bringing the cow into a state of negative energy balance. Approximately 25% of the NEFA passes through the liver, where it is metabolized or esterified to triglyceride. When the latter becomes excessive, fatty liver or hepatic lipidosis generally as a subclinical disease state will occur. This chapter investigates realizations in the molecular factors potentially causing fatty liver in the transition cow discovered using proteomics.

Keywords Liver • Proteomics • Dairy cow • Negative energy balance • Macronutrient and energy metabolism • Oxidative stress response

1 Introduction

The transition period of dairy cows is commonly defined as the period from 3 weeks before to 3 weeks after calving (Drackley 1999). Within this period, dry matter intake usually decreases by 10–30% during the last 3 weeks before parturition and continuously increases within the 3 weeks after calving (Ingvartsen and Andersen 2000;

B. Kuhla (✉)
Leibniz Institute for Farm Animal Biology (FBN), Institute of Nutritional Physiology 'Oskar Kellner', Dummerstorf, Germany
e-mail: b.kuhla@fbn-dummerstorf.de

K.L. Ingvartsen
Department of Animal Science, Aarhus University, Tjele, Denmark
e-mail: kli@anis.au.dk

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_11

Esposito et al. 2014). However, the maximum level of dry matter intake is only achieved by 6–8 weeks after calving and then amounts to 200–250% of the level at parturition (Hammon et al. 2009; Weber et al. 2013). However, the peak of milk production precedes the peak in dry matter intake, and thus, nutrient losses due to milk production exceed the level of nutrient intake. Consequently, cows face the challenge of a negative nutrient and energy balance (NEB). These nutrient and energetic gaps are met by the mobilization of body reserves. Glycogen and protein stored in the liver and muscle tissue and some minerals in the bones but mainly triglycerides (TG) deposited in distinct adipose tissues are mobilized during early lactation to meet the high metabolic demand of nutrients in the mammary gland in early lactation. NEB and the mobilization of body reserves are highly associated with perturbed immune functions (e.g., Ingvartsen and Moyes 2015) that may increase the risk of infectious diseases such as retained placenta, metritis, and mastitis as well as increase liability of metabolic diseases such as ketosis, ruminal acidosis, milk fever, and fatty liver (Ingvartsen 2006; Esposito et al. 2014). All these factors are critically important to the health, production, and profitability of dairy cows.

The density of nutrients, particularly protein and fat, in colostrum and milk is highest within the first days of lactation. Lipases cleave TG to yield long-chain, non-esterified fatty acids (NEFAs), resulting in dramatic increases in plasma concentrations in early lactation. Released NEFAs are intensively used as substrate for milk fat synthesis in the mammary gland and can contribute to almost half of milk fat (Palmquist 2006). Besides supporting milk fat production, NEFAs also serve as energy-providing substrate for other organs sparing glucose in organs such as the heart, skeletal muscle, and liver. However, the extent of fat mobilization is not only dependent on the energy requirements of the cows. Higher condition dairy cows generally mobilize more and show a greater loss of body condition score (BCS) which is reflected in higher plasma NEFA concentrations compared to lower condition cows (Vallimont et al. 2011). However, cows entering the transition period with a high BCS are able to transfer more NEFA to the mammary gland for milk fat synthesis and to oxidize more fatty acids (Börner et al. 2013). Accordingly, fat cows produce just as much or more fat-corrected milk (FCM) yield and show a lower feed intake, thus entering into a stronger NEB compared to thinner counterparts (Ingvartsen and Andersen 2000; Hammon et al. 2009; Schäff et al. 2012; Weber et al. 2013).

The key nutrient determining the level of milk production is lactose which in turn is produced from glucose by the mammary gland. More than 80% of the glucose turnover is used for lactose synthesis during peak lactation (Bauman and Currie 1980). The liver is the complementary organ for the mammary gland as it provides the majority of glucose for lactose synthesis. Only a minor portion of glucose is directly absorbed from the intestine or the rumen wall, and only a small amount of circulating glucose is used by other organs than the mammary gland, i.e., as energy source in the brain (Lemosquet et al. 2009). Therefore, hepatic gluconeogenesis is of major importance at the beginning and throughout lactation to meet the glucose demand of the mammary gland.

During NEB, the mammary gland also receives amino acids and smaller peptides for the synthesis of milk proteins. These precursors may originate from proteolysis and through suppression of tissue protein synthesis in various organs including the skeletal muscle, liver, skin, gastrointestinal tract, and uterus (Bell et al. 2000; Kuhla et al. 2011). Supplementing amino acids into the abomasum in early lactation may further increase milk yield, indicating that milk production level is not only determined by glucose but also by the level of available amino acids (Larsen et al. 2015). Taken together, the liver plays a pivotal role in early lactation to meet the nutrient and energy requirements of the mammary gland by providing substrates originating from anabolic and catabolic pathways.

2 Lipid Metabolism

The amount of NEFA taken up by the liver depends on the blood flow rate and amounts to 25% of plasma NEFA passing through the hepatic circulation (Reynolds et al. 2003). Once entered into the liver, NEFA can be (1) completely oxidized to carbon dioxide to provide energy for the liver, (2) partially oxidized to produce ketone bodies that are released into the blood and serve as fuel for other tissues, or (3) reconverted to triglycerides (TG). Triglycerides, which are not exported as VLDL, are stored in the cytosol as fat droplets. When the uptake of NEFA exceeds the capacity for synthesizing or exporting VLDL into circulation and when the capacity for NEFA metabolization is exhausted in early lactation, fatty liver syndrome develops (Drackley 1999; Ingvartsen 2006).

For metabolization, NEFAs are transported into the cytosol and via the carnitine acyltransferase system into the mitochondria to be β-oxidized yielding acetyl-CoA. Besides, long-chain fatty acids are also degraded in peroxisomes and microsomes, but in the ruminant liver, these routes are secondary. In the mitochondria, acetyl-CoA may be metabolized with two major fates, complete oxidation through the TCA cycle or incomplete oxidation through ketogenesis. Due to the limited availability of oxaloacetate in early lactation (Bobe et al. 2004)—which in turn is predominantly used for the synthesis of glucose—the majority of acetyl-CoA is used for the formation of ketone bodies such as β-hydroxybutyrate (BHB). Formed ketone bodies are released into the circulation and can be used as a glucose-sparing fuel by extrahepatic organs such as the heart, muscle, and mammary gland. However, some hepatic acetyl-CoA react with oxaloacetate to be oxidized to CO_2, in the tricarboxylic acid (TCA) cycle, resulting in increased fatty acid oxidation but diminished plasma glucose concentrations (Schäff et al. 2012; Weber et al. 2013). Thus, cows with a high extent of fat mobilization have not only greater plasma NEFA and BHB but also lower glucose concentrations (Schäff et al. 2012; Weber et al. 2013).

Proteome analysis in liver biopsies taken throughout the periparturient period of dairy cows contributed to identify metabolic enzymes regulated in response to lipid mobilization. Liver biopsies taken from cows on days −34, −17, +3, +18, and +30

relative to parturition were analyzed for their total lipid content to retrospectively assign the animals to a high (>24.4% fat/g dry matter) or a low (<24% fat/g dry matter) liver fat content group on day 18 postpartum. Liver tissue extracts were applied to two-dimensional gel electrophoresis (2DE), and differentially expressed spots were identified by MALDI-TOF-TOF mass spectrometry (Schäff et al. 2012). The expression of all four mitochondrial β-oxidative enzymes increased in both groups 1.1–3-fold in early lactation. For the medium-chain acyl-CoA dehydrogenase, the extent of upregulation was greater in high- compared to low-mobilizing cows. However, the β-oxidative enzymes located further downstream, short-chain specific acyl-CoA dehydrogenase, enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase 2, and 3-ketoacyl-CoA thiolase as well as 2,4-dienoyl-CoA reductase 1, were lower expressed in cows with the higher liver fat content postpartum (Schäff et al. 2012).

In a similar study, liver triglyceride content was determined in Holstein dairy cows at week 1 postpartum and grouped again to the lowest (81 ± 22 μmol/g) and highest (234 ± 81 μmol/g) wet weight concentrations (Sejersen et al. 2012). Quantification of hepatic proteins by the iTRAQ method and LC tandem mass spectrometry for protein identification revealed that mitochondrial acyl-CoA synthetase long-chain family member 1 (ACSL1) and very-long-chain specific acyl-CoA dehydrogenase (VLCAD) were more abundant, whereas hydroxyacyl-CoA dehydrogenase was less abundant in cows with a high compared to a low liver triglyceride content. Additionally, enzymes involved in the oxidation of unsaturated fatty acids such as 2,4-dienoyl-CoA reductase 1 and dodecenoyl-CoA delta isomerase and 4-trimethylaminobutyraldehyde dehydrogenase involved in carnitine biosynthesis were less expressed in cows with a high liver fat content (Sejersen et al. 2012).

In a further study, cows were distinguished according to their physiological imbalance postpartum involving more hepatic triglycerides in cows with greater physiological imbalance (Moyes et al. 2013). As analyzed by the iTRAQ LC-MS/MS method, imbalanced mid-lactating cows had also a higher expression of carnitine O-palmitoyltransferase 2, acyl-CoA binding protein, acyl-CoA acyltransferase, and alcohol dehydrogenase 4, the latter involved in ketone body synthesis (Moyes et al. 2013). Results of all three studies (Schäff et al. 2012; Sejersen et al. 2012; Moyes et al. 2013) indicate that cows with the higher liver fat content had greater activation but diminished complete β-oxidation of fatty acids, which has also been observed at mRNA (Li et al. 2012) and enzyme activity levels (Murondoti et al. 2004). Incomplete β-oxidation results in less acetyl-CoA but greater ketone body production in the liver (Fig. 1).

The lower expression of mitochondrial β-oxidative enzymes in cows developing the higher liver fat content should confine the production of acetyl-CoA. Failure of complete mitochondrial β-oxidation seems to result in a proportionally increased degradation of fatty acids in peroxisomes and microsomes. Cows with a higher physiological imbalance or liver fat content postpartum had a greater expression of the peroxisomal enoyl-CoA hydratase, catalase, electron transfer flavoprotein, or acyl-CoA oxidase 2 as well as the microsomal aldehyde dehydrogenase 1 (Schäff et al. 2012; Moyes et al. 2013). It seems that activation of peroxisomal β-oxidation and endoplasmic ω-oxidation is a sink for exhausted mitochondrial β-oxidation

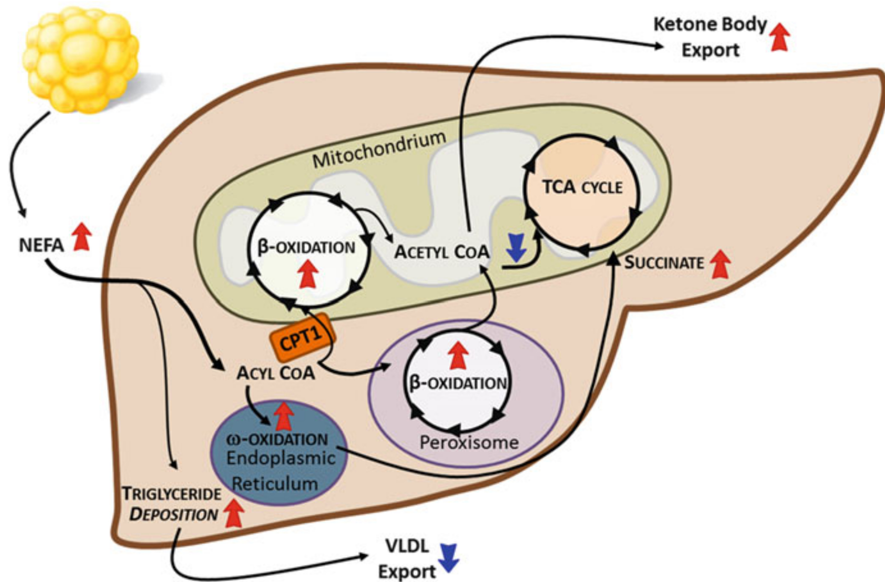

Fig. 1 Hepatic lipid metabolism of dairy cows in early lactation. *Red arrows* indicate activation, whereas *blue arrows* indicate deactivation of pathways in fatty liver as identified by proteome analyses (Schäff et al. 2012; Sejersen et al. 2012)

(Fig. 1). The proteomic data support previous data showing that total (mostly mitochondrial) oxidation of 1-^{14}C-palmitate in liver homogenates increases from parturition until week 5 postpartum, but also contrasts the results on peroxisomal β-oxidation capacity which decreased from week 3 to 12 postpartum (Grum et al. 2002). However, proteome results agree with the finding that cows with a high propensity to mobilize fat have a greater whole-body fat oxidation as compared to leaner counterparts (Börner et al. 2013).

In addition, the content of liver phospholipids and apolipoprotein tended to be lower in cows with a higher liver triglyceride content, suggesting reduced phospholipid and apolipoprotein syntheses or increased degradation of phospholipids or apolipoprotein, respectively, in early lactation. For the latter assumption, the proteome study by Sejersen et al. (2012) revealed a lower abundance of annexin A1, a phospholipase inhibitor, and Schäff et al. showed a reduced apolipoprotein E abundance in cows with pronounced compared to moderate fatty liver. Reduced hepatic phospholipid and apolipoprotein contents may inhibit the secretion of triglycerides in the form of VLDL from the liver into the circulation (Bobe et al. 2004).

3 Carbohydrate Metabolism

In dairy cows, the supply of carbons for hepatic gluconeogenesis is primarily from propionate produced in the rumen, but in early lactation other substrates such as lactate, amino acids, and glycerol gain importance in providing carbons for glucose

production (Larsen and Kristensen 2013). Besides the de novo synthesis of glucose, mobilized hepatic glycogen reserves may contribute to the endogenous glucose production in early lactation. Starting approximately 3 weeks before calving, liver glycogen reserves are continuously degraded until reaching a minimum within the first 4 to 6 weeks of early lactation, while the extent of fat mobilization seemed not to affect the level of glycogen degradation (Schäff et al. 2012; Sejersen et al. 2012). However, cows with a high fat mobilization rate may replenish liver glucose faster than counterparts mobilizing less fat (Sejersen et al. 2012). Accordingly, uridine-5-triphosphate (UTP)-glucose-1-phosphate uridylyltransferase, an enzyme involved in glycogenesis, was more abundant in high- compared to low-mobilizing cows (Schäff et al. 2012; Sejersen et al. 2012). On the other hand, phosphoglucomutase-2, which facilitates the interconversion of glucose 1-phosphate and glucose 6-phosphate, is involved in glycogenolysis and glycogenesis, depending on the energy status of the liver (Sejersen et al. 2012).

Among enzymes involved in gluconeogenesis, propionyl-CoA carboxylase, pyruvate carboxylase, lactate dehydrogenase, and enolase 1 were more abundant in high- compared with the low-mobilizing cows, whereas the amounts of fructose-1,6-bisphosphatase 1 and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) were lower in high-mobilizing cows (Schäff et al. 2012; Sejersen et al. 2012). The direction of expression changes indicates an increased production of oxaloacetate from pyruvate and succinate but reduced gluconeogenesis from phosphoenolpyruvate. Thus, cows with fatty liver syndrome seem to have a greater pyruvate cycling (Fig. 2) and less glucose production as suggested by lower ratio between the rate of glucose appearance and glucose oxidation (Weber et al. 2016). The proteome results are in agreement with investigations of gluconeogenic enzyme activities (Rukkwamsuk et al. 1999). However, previous studies have shown that mRNA expression of gluconeogenic enzymes is not different in livers differing in fat content (Hammon et al. 2009; Weber et al. 2013).

The pentose pathway uses glucose to generate NADPH, pentoses, and ribose 5-phosphate. In contrast to glycolysis, its primary role is anabolic. The abundance of three enzymes of the pentose phosphate pathway, namely, transketolase, triosephosphate isomerase, and GAPDH, was lower in cows developing fatty liver (Schäff et al. 2012), indicating impaired generation of nucleotides, reducing equivalents and aromatic amino acids (also see below).

4 Protein and Amino Acid Metabolism

The liver may not only store energy reserves in form of glycogen but also serve to supply proteins and amino acids in times of negative energy balance. Total hepatic protein content was lower in cows challenged by high fat mobilization in early lactation accompanied with the higher protein abundance of PA28a, promoting hepatic protein degradation, and a lower abundance of EF1δ and calpastatin, both factors activating protein biosynthesis (Schäff et al. 2012). On the other hand,

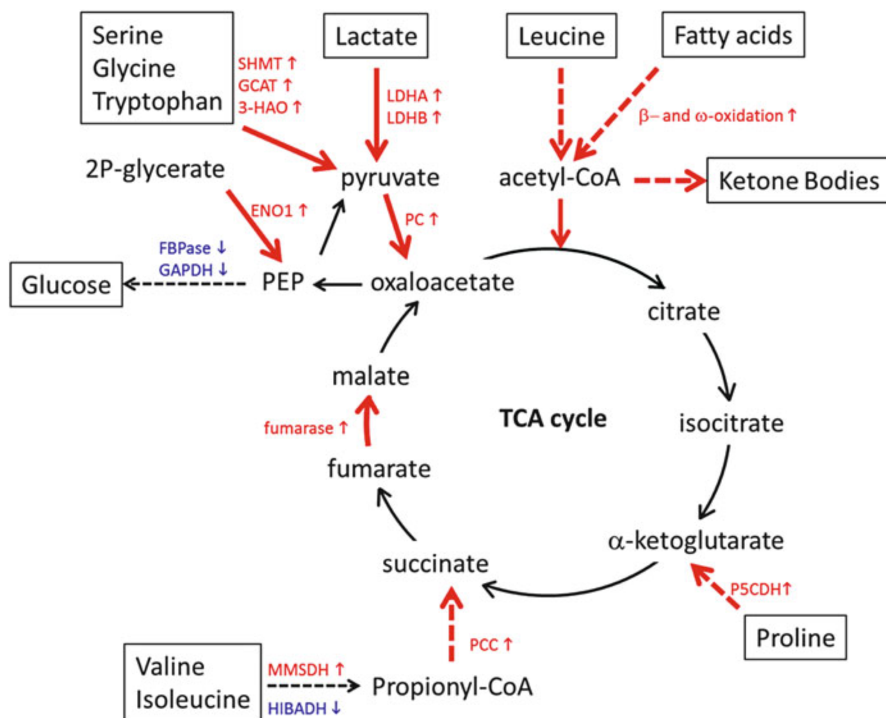

Fig. 2 Hepatic carbohydrate and amino acid catabolism and their converging into the TCA cycle of dairy cows with fatty liver in early lactation. Enzymes and *arrows* in *red* indicate upregulation, whereas enzymes printed in *blue* were downregulated in cows with fatty liver as identified by proteome analyses (Schäff et al. 2012; Sejersen et al. 2012). *3-HAO* 3-Hydroxyanthranilate 3,4-dioxygenase; *HIBADH* 3-hydroxyisobutyrate dehydrogenase; *ENO1* α-enolase; *FBPase* fructose 1,6-bisphosphatase; *GAPDH* glyceraldehyde-3-phosphate dehydrogenase; *GCAT* glycine C-acetyltransferase; *LDHB* L-lactate dehydrogenase B; *MMSDH* methylmalonate-semialdehyde dehydrogenase; *P5CDH* δ-1-pyrroline-5-carboxylate dehydrogenase; *PEP* phosphoenolpyruvate; *PC* pyruvate carboxylase; *PCC* propionyl-CoA carboxylase; *SHMT* serine hydroxymethyltransferase

α-crystallin and 60S ribosomal proteins L7 and L10, all involved in translation and protein folding of newly synthesized protein, are upregulated in cows with fatty liver (Sejersen et al. 2012). Degradation of unused proteins and simultaneous activation of de novo protein synthesis refer to a complex, highly dynamic proteome enabling adaptation of the liver to early lactation and fatty acid load.

Apart from degrading proteins, intermediary amino acid metabolism is closely connected to gluconeogenesis and fatty acid degradation. As stated above, cows that mobilize a lot of fat seem to limit the production of acetyl-CoA derived from β-oxidation. However, acetyl-CoA may also be formed after breakdown of ketogenic amino acids. Indeed, plasma leucine and isoleucine concentrations were found to be significantly lower during early lactation in cows with the high liver fat content compared to cows with a low liver fat content (Schäff et al. 2012),

suggesting that ketone body production in the fatty liver is intensively driven by branched-chain ketogenic amino acids (Fig. 2). Accordingly, protein expression of methylmalonate-semialdehyde dehydrogenase involved in branched-chain amino acid degradation as well as propionate metabolism was found to be higher expressed in mid-lactating cows with physiological imbalance or cows developing fatty liver in early lactation (Moyes et al. 2013; Schäff et al. 2012). Degradation of valine and isoleucine results in the formation of succinate, thereby serving as potential precursors for gluconeogenesis (Fig. 2). On the other hand, valine, leucine, and isoleucine catabolism in the liver were reduced in cows with fatty liver as the hepatic 3-hydroxyisobutyrate dehydrogenase protein was found to be reduced compared to individuals with a low liver fat content (Sejersen et al. 2012).

Based on a 2DE proteome study, serine hydroxymethyltransferase, 3-hydroxyanthranilate 3,4-dioxygenase, and glycine C-acetyltransferase expressions were increased in cows with the higher liver fat content, referring to a prevailing catabolism of serine, glycine, and tryptophan whose carbon skeletons enter the TCA cycle via pyruvate yielding oxaloacetate (Schäff et al. 2012). Furthermore, the protein abundance of δ-1-pyrroline-5-carboxylate dehydrogenase, which participates in the degradation of proline toward α-ketoglutarate, increased from parturition to week 3 of lactation and was greater in high- compared to low-mobilizing cows (Schäff et al. 2012).

In consequence of intensified glucogenic amino acid degradation in cows with fatty liver, one might expect the activation of various amino transferases and urea cycle enzymes. While an increased abundance of arginase 1 and aspartate aminotransferase was described by Sejersen et al. (2012), Schäff et al. (2012) reported a lower abundance of glutamate dehydrogenase, arginase 1, and carbamoyl phosphate synthase 1. The discrepancy between these studies may lie in the level of feed and thus nitrogen intake and the urea cycling between the rumen and the liver. Also, contrasting results may be due to different concentrations of plasma fatty acids in cows with a high versus low liver fat content. Long-chain fatty acids may suppress the induction of urea cycle enzyme genes as reported by Tomomura et al. (1996). However, a reduced abundance of urea cycle enzymes does not exclude a greater entrance of carbon derived from glucogenic amino acids into the TCA cycle.

5 TCA Cycling and Oxidative Phosphorylation

The increased demand for glucose in early lactation withdraws substantial amounts of oxaloacetate from the TCA cycle. This gap needs to be refilled to ensure that acetyl-CoA can be converted to citrate. Results from proteome studies indicate that the process of glycogenolysis and amino acid degradation mainly provide pyruvate and oxaloacetate, succinate, and α-ketoglutarate to the TCA cycle (Fig. 2). Moreover, the higher expression of fumarase, converting fumarate to malate, points to intensified TCA cycling in cows with the higher liver fat content (Schäff et al. 2012). Intensification of TCA cycling is not only to ensure sufficient supply of

oxaloacetate but also to generate enough NADH as this is required among others for gluconeogenesis and ATP production. Enzymes involved in oxidative phosphorylation such as ATP synthase and pyrophosphatase as well as cytochrome c oxidase are increased in early lactation and have higher expression levels in cows with high liver fat content, indicating increased ATP production and mitochondrial respiration in cows with fatty liver (Schäff et al. 2012). Increased oxidative phosphorylation and TCA cycling further suggest that more oxaloacetate remains in the TCA cycle of cows with high compared to low triglyceride accumulation in the liver.

When the oxidative phosphorylation pathway is more active in fatty compared to non-fatty livers, one might assume an increased ATP status and thus a reduced AMP:ATP ratio. The adenosine monophosphate-activated protein kinase (AMPK) is sensitive to the AMP:ATP ratio and is phosphorylated by increased AMP concentrations. In early lactation AMPK phosphorylation (Schäff et al. 2012) and also GTP:AMP phosphotransferase, an enzyme catalyzing the phosphorylation of AMP (Sejersen et al. 2012), were more abundant in animals with fatty liver. These findings argue against an increased hepatic ATP status, but the pAMPK/AMPK ratio was similar between animals with different contents of fat in the liver. Therefore, it seems likely that oxidative phosphorylation is impaired in cows with fatty liver and that upregulation of ATP synthase, pyrophosphatase, and cytochrome c oxidase (Schäff et al. 2012) is a counter-regulatory mechanism to sustain ATP homeostasis (Fig. 3). In human fatty liver, excessive TCA cycling and NADH production have been shown to be a consequence of a damaged electron transport chain and impaired ATP synthesis (Cortez-Pinto et al. 1999; Sunny et al. 2011).

6 Oxidative Stress Response

The high rate of peroxisomal and mitochondrial fatty acid oxidation together with the increase in feed intake is accompanied with high amounts of oxygen consumption in early lactation which confronts the liver with increased oxidative stress levels. The resulting increased H_2O_2 and superoxide radical levels need to be detoxified. Enzymes protecting hepatocytes against increased oxidative stress, such as DJ-1 protein, peroxiredoxin 6, Cu/Zn superoxide dismutase 1, and stress-induced phosphoprotein 1 (Schäff et al. 2012) as well as protein disulfide isomerase, thioredoxin, and peroxiredoxin 2 (Sejersen et al. 2012), were found to be more abundant in cows with the high liver fat content as elaborated by proteomic approaches (Fig. 3). These findings point to an activated oxidative stress defense in these animals. Increased oxidative stress occurs after calving, as indicated, for example, by increased concentrations of plasma reactive oxygen species (ROS), thiobarbituric acid reactive substances (TBARS), or ceruloplasmin concentrations, while the oxygen radical absorbance capacity increases in early lactation (Bernabucci et al. 2005; Osorio et al. 2014). Further, cows with the greater loss of BCS after calving had higher ROS and TBARS plasma concentrations but lower erythrocyte thiol (SH) concentrations and superoxide dismutase activity in the

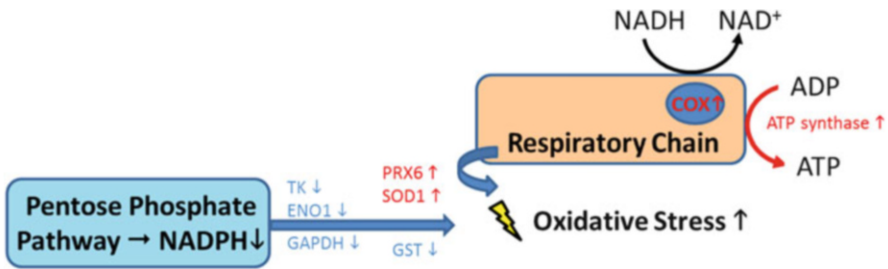

Fig. 3 Regulation of proteins involving increased oxidative phosphorylation and oxidative stress defense, as well as reduced NADPH production in the pentose phosphate pathway leading to increased oxidative stress in cows developing fatty liver

postpartum period as compared to cows with less BCS loss (Bernabucci et al. 2005). Additionally, serum paraoxonase-1 decreased from calving until day 19 after parturition in obese, but not lean cows (Folnožić et al. 2015). All these results suggest that cows with the greater extent of fat mobilization are prone to greater oxidative stress. At least in part, oxidative stress may be counteracted by a greater activation of the anti-oxidative stress machinery in the liver. In serum and blood cells, however, the anti-oxidative status seems to be disturbed in high fat-mobilizing cows.

7 Pentose Phosphate Pathway

Nicotinamide adenine dinucleotide phosphate (NADPH) is required as a reducing agent and is synthesized in the pentose phosphate pathway. To destroy bacteria and combat infections, immune cells use NADPH to reduce O_2 and to form H_2O_2 under the catalysis of NADPH oxidase, a process termed "respiratory burst." Absence of NADPH oxidase in immune cells prevents the formation of ROS and is thus a defense against infections. In the liver, however, NADPH provides the reducing equivalents for biosynthetic regeneration of reduced glutathione which is intensively involved in the protection against the toxicity of ROS. In cows susceptible to develop a high liver fat content, enzymes participating in the pentose phosphate pathway such as transketolase (TK), triosephosphate isomerase (TPI), and GAPDH were lower expressed as compared to cows with less liver fat. From late pregnancy to early lactation, GAPDH and TK decreased to 70–80% in cows with a high liver fat content, before re-increasing by day 40 of lactation (Schäff et al. 2012). Thus, it seems that the NADPH-dependent glutathione system is negatively affected in cows with a high liver fat content, particularly during early lactation. Most interestingly, dietary supplementation of rumen-protected methionine, the first limiting amino acid for glutathione synthesis, has been shown to enhance the de novo glutathione synthesis and increase antioxidant capacity, which was associated with greater assembly or export of very-low-density lipoproteins (Osorio et al. 2014).

8 Conclusions

Proteome analysis in liver biopsies taken throughout the periparturient period of dairy cows has contributed to a better understanding of the regulation of a number of metabolic changes in response to lipid mobilization. The results further indicate that cows with the higher liver fat content had greater activation but diminished complete β-oxidation of fatty acids. Incomplete β-oxidation results in less acetyl-CoA but greater ketone body production in the liver. The lower expression of mitochondrial β-oxidative enzymes in cows developing higher liver fat content probably reduces the production of acetyl-CoA. Failure of complete mitochondrial β-oxidation seems to result in a proportionally augmented degradation of fatty acids in peroxisomes and microsomes. Cows with a high fat mobilization rate appear to replenish liver glucose faster than cows mobilizing less fat, likely by increased production of oxaloacetate from pyruvate and succinate but reduced gluconeogenesis from phosphoenolpyruvate. Thus, cows with fatty liver syndrome seem to have a greater pyruvate cycling.

The abundance of three enzymes of the pentose phosphate pathway was lower in cows developing fatty liver, indicating impaired generation of nucleotides, reducing equivalents and aromatic amino acids. Degradation of unused proteins and simultaneous activation of de novo protein synthesis refer to a complex, highly dynamic proteome enabling adaptation of the liver to early lactation and fatty acid load. Plasma leucine and isoleucine concentrations were found to be significantly lower during early lactation in cows with high liver fat content, suggesting that ketone body production in the fatty liver is intensively driven by branched-chain ketogenic amino acids.

Cows with high liver fat content appear to have an intensified TCA cycling, suggesting that more oxaloacetate remains in the TCA cycle. Further, greater TCA cycling refers to increased NADH production and thereby increased ATP production and mitochondrial respiration. Greater respiration is accompanied by increased production of oxygen-containing radicals. Enzymes protecting hepatocytes against increased oxidative stress levels are more abundant in cows with high liver fat content pointing toward an activated oxidative stress defense in these cows.

The conclusions drawn so far are based on the different abundances of hepatic enzymes controlling macronutrient and energy metabolism and on circulating plasma metabolite concentrations of cows developing different liver fat contents. However, the activity of most of the enzymes is not exclusively determined by their protein abundance and substrate concentration. The presence of cofactors such as vitamins from the B-complex but also and likely more important are the posttranslational modifications of enzymes involved in the above described biochemical pathways. It is well known that predominantly phosphorylation, glycosylation, and acetylation of enzymes control metabolic activity. However, up to now, no phospho-proteomic or glycomic approaches have been performed to elucidate the role of posttranslational modifications of enzymes participating in the adaptation to early lactation and the development of fatty liver of dairy cows. Newly available

antibody and mass spectrometry techniques offer the methodological base to investigate position-specific posttranslational modifications and their quantity relative to their parent proteins. Future work using these approaches will contribute to a better understanding of fatty liver metabolism of transition cows.

References

Bauman DE, Currie WB (1980) Partitioning of nutrients during pregnancy and lactation: a review of mechanisms involving homeostasis and homeorhesis. J Dairy Sci 63(9):1514–1529

Bell AW, Burhans WS, Overton TR (2000) Protein nutrition in late pregnancy, maternal protein reserves and lactation performance in dairy cows. Proc Nutr Soc 59(1):119–126

Bernabucci U, Ronchi B, Lacetera N, Nardone A (2005) Influence of body condition score on relationships between metabolic status and oxidative stress in periparturient dairy cows. J Dairy Sci 88(6):2017–2026

Bobe G, Young JW, Beitz DC (2004) Invited review: pathology, etiology, prevention, and treatment of fatty liver in dairy cows. J Dairy Sci 87:3105–3124

Börner S, Albrecht E, Schäff C, Hacke S, Kautzsch U, Derno M, Hammon HM, Röntgen M, Sauerwein H, Kuhla B (2013) Reduced AgRP activation in the hypothalamus of cows with high extent of fat mobilization after parturition. Gen Comp Endocrinol 193:167–177

Cortez-Pinto H, Chatham J, Chacko VP, Arnold C, Rashid A, Diehl AM (1999) Alterations in liver ATP homeostasis in human nonalcoholic steatohepatitis: a pilot study. JAMA 282:1659–1664

Drackley JK (1999) ADSA Foundation Scholar Award. Biology of dairy cows during the transition period: the final frontier? J Dairy Sci 82(11):2259–2273

Esposito G, Irons PC, Webb EC, Chapwanya A (2014) Interactions between negative energy balance, metabolic diseases, uterine health and immune response in transition dairy cows. Anim Reprod Sci 144(3–4):60–71

Folnožić I, Turk R, Đuričić D, Vince S, Pleadin J, Flegar-Meštrić Z, Valpotić H, Dobranić T, Gračner D, Samardžija M (2015) Influence of body condition on serum metabolic indicators of lipid mobilization and oxidative stress in dairy cows during the transition period. Reprod Domest Anim 50(6):910–917

Grum DE, Drackley JK, Clark JH (2002) Fatty acid metabolism in liver of dairy cows fed supplemental fat and nicotinic acid during an entire lactation. J Dairy Sci 85(11):3026–3034

Hammon HM, Stürmer G, Schneider F, Tuchscherer A, Blum H, Engelhard T, Genzel A, Staufenbiel R, Kanitz W (2009) Performance and metabolic and endocrine changes with emphasis on glucose metabolism in high-yielding dairy cows with high and low fat content in liver after calving. J Dairy Sci 92:1554–1566

Ingvartsen KL (2006) Feeding- and management-related diseases in the transition cow – physiological adaptations around calving and strategies to reduce feeding-related diseases. Anim Feed Sci Technol 126(3–4):175–213

Ingvartsen KL, Andersen JB (2000) Integration of metabolism and intake regulation: a review focusing on periparturient animals. J Dairy Sci 83:1573–1597

Ingvartsen KL, Moyes KM (2015) Factors contributing to immunosuppression in the dairy cow during the periparturient period. Jpn J Vet Res 63(Supp 1):S15–S24

Kuhla B, Nürnberg G, Albrecht D, Görs S, Hammon HM, Metges CC (2011) Involvement of skeletal muscle protein, glycogen, and fat metabolism in the adaptation on early lactation of dairy cows. J Proteome Res 10(9):4252–4262

Larsen M, Kristensen NB (2013) Precursors for liver gluconeogenesis in periparturient dairy cows. Animal 7(10):1640–1650

Larsen M, Galindo C, Ouellet DR, Maxin G, Kristensen NB, Lapierre H (2015) Abomasal amino acid infusion in postpartum dairy cows: effect on whole-body, splanchnic, and mammary amino acid metabolism. J Dairy Sci 98(11):7944–7961

Lemosquet S, Delamaire E, Lapierre H, Blum JW, Peyraud JL (2009) Effects of glucose, propionic acid, and nonessential amino acids on glucose metabolism and milk yield in Holstein dairy cows. J Dairy Sci 92(7):3244–3257

Li P, Li XB, Fu SX, Wu CC, Wang XX, Yu GJ, Long M, Wang Z, Liu GW (2012) Alterations of fatty acid β-oxidation capability in the liver of ketotic cows. J Dairy Sci 95:1759–1766

Moyes KM, Bendixen E, Codrea MC, Ingvartsen KL (2013) Identification of hepatic biomarkers for physiological imbalance of dairy cows in early and mid lactation using proteomic technology. J Dairy Sci 96(6):3599–3610

Murondoti A, Jorritsma R, Beynen AC, Wensing T, Geelen MJ (2004) Unrestricted feed intake during the dry period impairs the postpartum oxidation and synthesis of fatty acids in the liver of dairy cows. J Dairy Sci 87:672–679

Osorio JS, Trevisi E, Ji P, Drackley JK, Luchini D, Bertoni G, Loor JJ (2014) Biomarkers of inflammation, metabolism, and oxidative stress in blood, liver, and milk reveal a better immunometabolic status in peripartal cows supplemented with Smartamine M or MetaSmart. J Dairy Sci 97(12):7437–7450

Palmquist DL (2006) Milk fat: origin of fatty acids and influence of nutritional factors thereon. In: Fox PF, McSweeney PL (eds) Advanced dairy chemistry, volume lipids. Springer, New York, NY, pp 43–92

Reynolds CK, Aikman PC, Lupoli B, Humphries DJ, Beever DE (2003) Splanchnic metabolism of dairy cows during the transition from late gestation through early lactation. J Dairy Sci 86:1201–1217

Rukkwamsuk T, Wensing T, Geelen MJH (1999) Effect of fatty liver on hepatic gluconeogenesis in periparturient dairy cows. J Dairy Sci 82:500–505

Schäff C, Börner S, Hacke S, Kautzsch U, Albrecht D, Hammon HM, Röntgen M, Kuhla B (2012) Increased anaplerosis, TCA cycling, and oxidative phosphorylation in the liver of dairy cows with intensive body fat mobilization during early lactation. J Proteome Res 11:5503–5514

Sejersen H, Sørensen MT, Larsen T, Bendixen E, Ingvartsen KL (2012) Liver protein expression in dairy cows with high liver triglycerides in early lactation. J Dairy Sci 95(5):2409–2421

Sunny NE, Parks EJ, Browning JD, Burgess SC (2011) Excessive hepatic mitochondrial TCA cycle and gluconeogenesis in humans with nonalcoholic fatty liver disease. Cell Metab 14 (6):804–810

Tomomura M, Tomomura A, Dewan MA, Saheki T (1996) Long-chain fatty acids suppress the induction of urea cycle enzyme genes by glucocorticoid action. FEBS Lett 399(3):310–312

Vallimont JE, Dechow CD, Daubert JM, Dekleva MW, Blum JW, Barlieb CM, Liu W, Varga GA, Heinrichs AJ, Baumrucker CR (2011) Heritability of gross feed efficiency and associations with yield, intake, residual intake, body weight, and body condition score in 11 commercial Pennsylvania tie stalls. J Dairy Sci 94:2108–2113

Weber C, Hametner C, Tuchscherer A, Losand B, Kanitz E, Otten W, Singh SP, Bruckmaier RM, Becker F, Kanitz W, Hammon HM (2013) Variation in fat mobilization during early lactation differently affects feed intake, body condition, and lipid and glucose metabolism in high-yielding dairy cows. J Dairy Sci 96:165–180

Weber C, Schäff CT, Kautzsch U, Börner S, Erdmann S, Görs S, Röntgen M, Sauerwein H, Bruckmaier RM, Metges CC, Kuhla B, Hammon HM (2016) Insulin-dependent glucose metabolism in dairy cows with variable fat mobilization around calving. J Dairy Sci 99 (8):6665–6679

Proteomics Research in the Adipose Tissue

Fabrizio Ceciliani, Cristina Lecchi, Jeanne Bazile, and Muriel Bonnet

Abstract Adipose tissue is no longer considered only as a passive fuel reservoir, but a major endocrine organ, distributed in different depots throughout the body and actively involved in complex regulatory processes including appetite, energy expenditure, body weight, inflammation and reproduction. Proteomics has emerged as a valuable technique to characterize both cellular and secreted proteomes from adipose tissues. Adipose tissue is of major importance in farm animals: in dairy animals, it regulates energy metabolism as well as other functions. In cows and pigs, adipose tissue depots' distribution is of fundamental importance for the quality of carcasses. In this chapter, we provide a general overview of adipose tissue functions and its importance in farm animals and summarize the state of art on farm animal adipose tissue proteomics in cattle and pigs but also in chicken and in farmed fish.

Keywords Adipose tissue • Proteomics • Farm animals • Animal physiology

1 Introduction

Adipose tissue is a complex structure composed of several types of cells, including adipocytes and pre-adipocytes, macrophages, endothelial cells, fibroblasts and leucocytes. Adipose tissue historical function is to store the energy in excess, this role being already established for many years. This task is carried out by conserving the heat of the body and controlling the mobilization of lipids (Sethi and Vidal-Puig 2007). When energy is in surplus, it is efficiently deposited as neutral triglycerides in adipose tissue. When energy is required, adipocyte triglycerides are broken down into glycerol and fatty acids. The released glycerol and fatty acids are then

F. Ceciliani (✉) • C. Lecchi
Department of Veterinary Medicine, Università degli studi di Milano, Milano, Italy
e-mail: fabrizio.ceciliani@unimi.it

J. Bazile • M. Bonnet
INRA, VetAgro Sup, UMR Herbivores, Université Clermont Auvergne, Saint-Genès-Champanelle, France

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_12

transported in the blood and subsequently uptaken by the muscles, liver and other organs, orchestrating lipid distribution and whole-body energy balance (Frayn 2002). In farm animal physiology, most of the studies on adipose tissues are focused on specific mechanisms related to animal production, aiming to understand the biochemical and molecular background underlying the mobilization of lipids in lactating animals to fulfil the energy requirements of milk production or the deposition of fat in specific fat depots, to improve the quality of meat. A further role as a major player during systemic metabolic regulation has been increasingly recognized (Luo and Liu 2016). With the discovery of leptin in 1994 (Zhang et al. 1994), adipose tissue was no longer considered as a passive fuel reservoir, but a major endocrine organ, distributed in different depots throughout the body and actively involved in complex regulatory processes including appetite, energy expenditure, body weight, inflammation and reproduction by synthesizing and secreting messenger molecules which are now collectively referred to as "adipokines". Each of these depots plays an essential role in energy homeostasis as well as in endocrine regulation, at both local and systemic levels. The excessive accumulation of fat in adipose tissue causes obesity, a chronic disease that in turn drives obesity-associated metabolic disorders such as diabetes. Although not directly related to obesity, the dysregulation of adipose tissue metabolism represents an issue for farm animals as well. Adipose tissue is involved in the fine-tuning of energy turnover in dairy cows or in carcass and meat qualities in cows and pigs. Therefore, a deeper understanding of the adipose machinery regulation in rearing species can provide tools to overcome metabolic stress-related diseases and improve the quality of carcass and meat.

Among omics disciplines, transcriptomics has provided important advances in understanding the functions of adipose tissue. A major weak point of transcriptomics is that it does not provide any hint about the effective abundance of the proteins, the correlation between the abundances of mRNA and of protein being poor (Griffin et al. 2002). Integration of data derived from gene expression analyses with those from protein expression studies is therefore mandatory to draw a complete and reliable picture of the functions and the activities of adipose tissue. The proteomics of adipose tissue was recently reviewed (Sauerwein et al. 2014). In the present chapter, we will try to provide the reader with a general overview of adipose tissues and summarize the recent advances on adipose tissue proteomics, focusing on farm animals such as cattle and pigs but also chicken and fish.

2 Adipose Tissue: A Bird's-Eye View with a Proteomic Perspective

2.1 The Main Structure and Function of Adipose Tissue: White, Brown and Beige Adipose Tissue

Two main types of adipose tissue exist in mammals, based on differences in morphology, location and functions: white adipose tissue (WAT) and brown

adipose tissue (BAT) (Lizcano and Vargas 2016). WAT is the preferential site to store energy in the form of triacylglycerols during excessive energy disposal and to restore it during fasting periods. Beside this undisputed role, WAT is also involved in systemic metabolic regulation and inflammation. The main role of BAT is to dissipate chemical energy as heat via high levels of uncoupling protein 1 (UCP1), regulating hypothermia by metabolizing glucose and lipids to produce heat participating to non-shivering thermogenesis (Lizcano and Vargas 2016; Louveau et al. 2016).

Recent studies have described another type of thermogenic adipocytes known as beige/brite (brown in white) adipocytes. Beige adipocytes share morphological and functional similarities with brown ones. In BAT, cold stress or β3-adrenoceptor agonists that mimic cold stress (Barbatelli et al. 2010) stimulate UCP1 expression. Therefore, both brown and beige fat fulfil thermogenic roles. In farm animals, the presence of beige or brown-like adipocytes has been suggested in white adipose depots of fattening cattle (Asano et al. 2013) and sheep (Pope et al. 2014).

2.2 Different Depots with Different Proteomes: Comparative Analysis of Subcutaneous and Visceral White Adipose Tissue Proteomics

Adipose tissue develops in various anatomical sites, including the abdominal cavity, the subcutaneous districts and the musculature. They are known to display a different dynamic of growth, which is at the origin of cellular and metabolic features (Bonnet et al. 2010; Hausman et al. 2014; Louveau et al. 2016). Thus, adipose tissue depots respond differentially to rearing practices, which may have implications on the metabolic turnover of lipids and nutrients at the body level for metabolic adaptations, or on carcass qualities and valorization, as visceral fat is discarded, while subcutaneous fat is partly consumed with the muscles. To date, few studies have attempted to investigate the proteome differences among fat depots in ruminants.

A study relying on 2-DGE paired with sequencing mass spectrometry investigated the proteins in undifferentiated and differentiated preadipocytes collected from bovine omental, subcutaneous and intramuscular adipose depots, identifying differentially expressed intracellular proteins during adipogenic conversion (Rajesh et al. 2010). A total of 65 proteins were found to be differentially abundant across the three depots. The preadipocyte differentiation induced a downregulation of many structural proteins, whereas proteins associated with lipid metabolism and metabolic activity, including ubiquinol-cytochrome-c reductase complex core protein I (UQCRC1), ATP synthase D chain, superoxide dismutase (SOD), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), sulfotransferase 1A1 (SULT1A1), carnitine *O*-palmitoyltransferase 2 (CPT2) and heat shock protein beta 1 (HSPB1), were found to be upregulated across the three depots. Among these proteins, five

(tropomyosin alpha-4 chain, rho GDP-dissociation inhibitor 1, purine nucleoside phosphorylase, transgelin, non-muscle caldesmon), one (annexin A1) and four (glyceraldehyde-3-phosphate dehydrogenase, prohibitin, voltage-dependent anion-selective channel protein 2, voltage-dependent anion-selective channel protein 1) proteins were specifically identified during the adipogenic process of omental, subcutaneous and intramuscular cells, respectively. Using a label-free quantification LC-MS/MS, 682 proteins were identified both in subcutaneous and visceral adipose tissues of 15.5-month-old British continental steers regardless of the diet. The abundance of 51% of these proteins was modified depending on the anatomical site. Of these, 240 were known proteins and were assigned to the following top categories of molecular and cellular functions: lipid metabolism, energy production, small molecule biochemistry, post-translational modification, cellular assembly and organization, cell morphology, protein synthesis and cellular function and maintenance (Romao et al. 2013). Notably, proteins related to oxidation of lipids and fatty acids and synthesis of lipids were more abundant in visceral than subcutaneous adipose tissue, in agreement with the repeatedly observed greater metabolic activity in visceral than in subcutaneous adipose tissues. However, any of the identified proteins were reported to be specifically present in one adipose tissue depot. To date, the sole study in ruminants that reports on protein signatures that are specific to the depot location was conducted in the goat (Restelli et al. 2014), in which a LC-MS/MS approach has identified 833 proteins in at least one of the four adipose tissue deposits investigated in 1-month-old goat kids. The tissue coverage was 471 proteins identified in subcutaneous adipose tissue sampled from the base tail, 480 proteins in subcutaneous adipose tissue sampled from the sternum, 587 proteins in omental tissue and 654 proteins in perirenal adipose tissue. These proteins were assigned to four major functional categories, based on similarity of functions, namely, (1) metabolic processes (62.1% of proteins); (2) cell adhesion cytoskeleton, intramuscular transport and membrane integrity (22.2% of proteins); (3) toxic response and folding (9.5% of proteins); and (4) proteins involved in immune and inflammatory response (6.1% of proteins). Of these 39, 30, 72 and 27 proteins were identified in only one adipose depot, the subcutaneous sampled from the base tail and the sternum as well as perirenal and omental depots, respectively. For the purpose of this review, we focused on the lists of proteins specifically identified in subcutaneous, from the base tail and of visceral perirenal adipose tissue, which are the most commonly analysed, in order to identify proteins or pathways related to the depot specificity of fat accumulation. Thus, lists of proteins specifically found in subcutaneous or perirenal adipose tissues were subjected to annotations according to the gene ontology (GO) with ProteINSIDE (Kaspric et al. 2015) in order to provide the biological processes in which they are involved. The 59 (over 72) perirenal and the 35 (over 39) subcutaneous proteins were annotated by 193 and 129 significantly enriched GO terms, respectively. Of these only 16 GO terms were found in common but annotating different proteins with a different *p*-value of enrichment between lists of subcutaneous and perirenal proteins (Fig. 1). This may reveal pathways important for adipose tissue growth in 1-month-old goat kids, however, involving proteins that

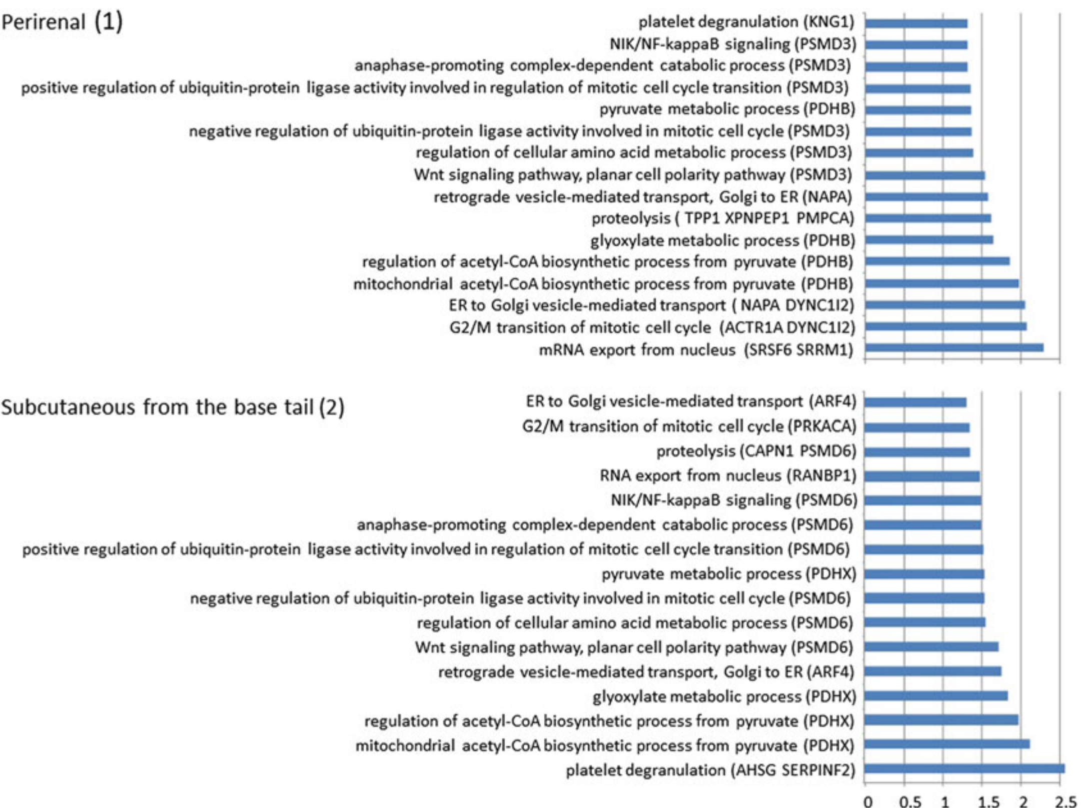

Fig. 1 Identification of 16 shared enriched GO terms (*p-value* < 0.05) of biological process category for the list of 72 proteins from the perirenal adipose tissue (1) and for the list of 39 proteins from the base tail subcutaneous adipose tissue (2) in goat. Results were obtained with ProteINSIDE, and enrichments are expressed as $-\log 10$ (*p-value*) to visually plot them on graphs, which means that a $-\log 10$ (*p-value*) of 3, 2 and 1.3 corresponds to a *p-value* of 0.001, 0.01 and 0.05, respectively. In brackets are the gene names of the proteins annotated by GO terms

differ depending on the anatomical site of adipose tissue. The top ten of the most enriched GO terms that were specifically found for the subcutaneous proteins were related to glucose or lipid metabolism, cell proliferation and extracellular matrix organization, in line with a tissue growth both by hyperplasia and by hypertrophy (Fig. 2). For the perirenal proteins, the top ten of the most enriched GO terms were related to protein translation, rRNA and mRNA processing and DNA repair (Fig. 3). These divergent annotations may reflect the lower maturity of subcutaneous as compared to perirenal adipose tissue, in line with the known depot-specific

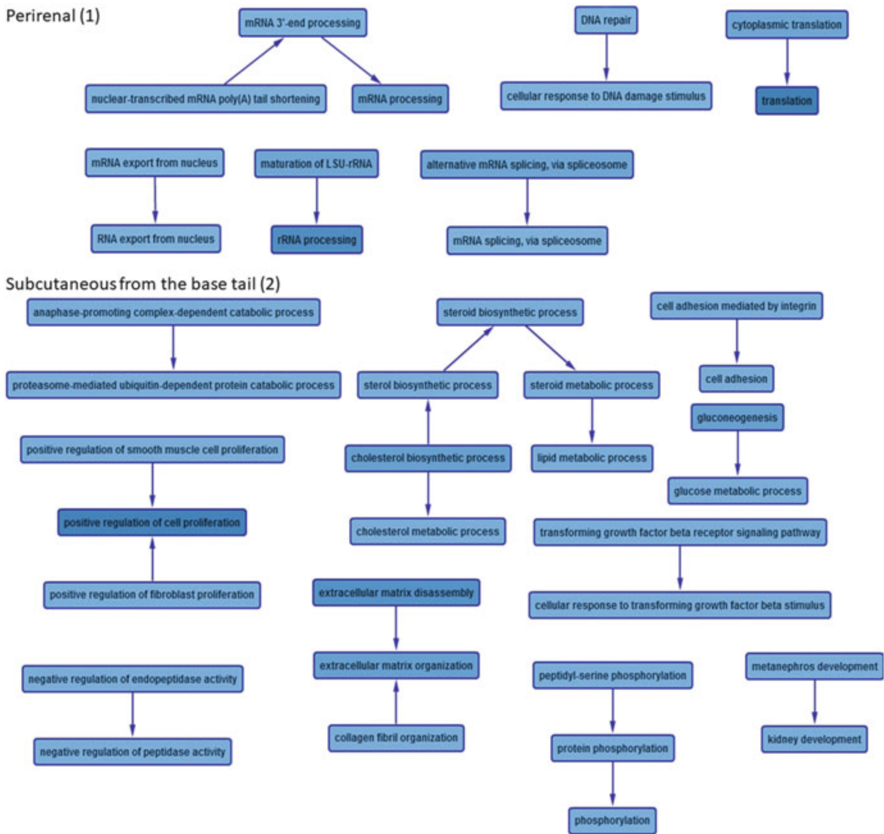

Fig. 2 Relations between the most enriched GO terms were visualized as networks provided by ProteINSIDE. The network links the most enriched GO terms (*p-value* < 0.05) in the biological process category for the list of 72 proteins (59 annotated) from the perirenal adipose tissue (1) and for the list of 39 (35 annotated) proteins from the base tail subcutaneous adipose tissue (2) in goat. The degree of colour saturation is related to the number of proteins annotated by a GO (dark and clear for high and low numbers, respectively)

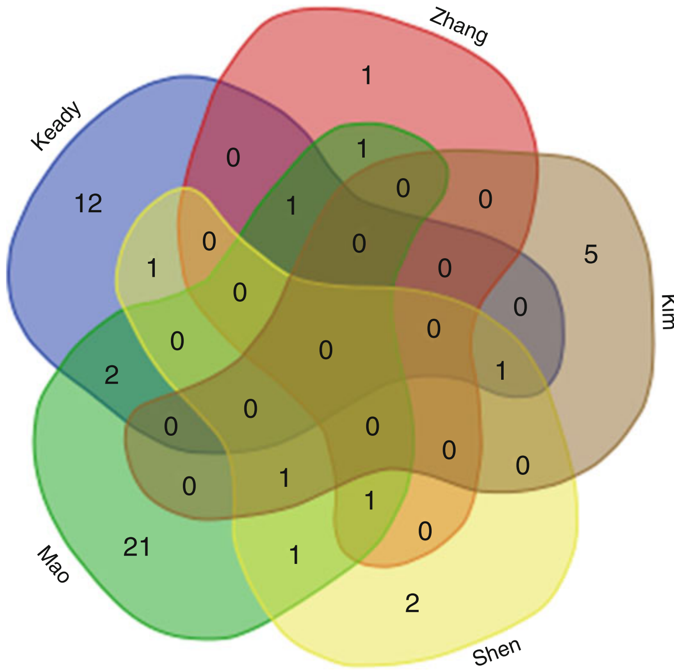

Fig. 3 Distribution and overlap of proteins identified in five proteomic analysis (Kim et al. 2009; Zhang et al. 2010; Shen et al. 2012; Keady et al. 2013; Mao et al. 2016) carried out to provide the molecular basis of marbling and to identify biomarkers. Venn diagram was proceeded with http://bioinformatics.psb.ugent.be/webtools/Venn/

growth patterns reported during the perinatal period in ruminants (Bonnet et al. 2010). Future proteomic studies might contribute to identify depot-specific proteins in order to increment available data and provide the molecular basis of depot specificity of fat accumulation for a robust data mining of the pathways or molecular functions involved.

2.3 *Proteomic Analysis of Brown Adipose Tissue*

Larger mammals such as ruminants are fully developed and able to thermoregulate at birth, thanks to the presence of BAT that disappears after birth and is replaced by WAT. A better knowledge of the molecular features of BAT versus WAT may help in the understanding of the balance between brown and white fat cells and is a prerequisite for a better control of thermogenesis and survival in perinatal period, as well as energy efficiency during the productive life of the ruminants. To the best of our knowledge, only one study has attempted to link the proteome to chemical composition, cellularity, histology, enzyme activities and gene expression during a

time-course analysis of the ontogeny of perirenal adipose tissue in bovine foetuses (Taga et al. 2012). Between 110 and 260 days postconception (dpc, 38 and 90% of gestation length, respectively), the increase in the weight of perirenal adipose tissue resulted from an increase in the adipocyte volume and mainly number and was accompanied by changes in the abundance of 128 proteins among the 143 proteins identified. From these data, an unexpectedly high abundance of the α-subunit of ATP synthase, a member of complex V (ATP synthase) of the respiratory chain which is normally bypassed in BAT, and of aldehyde dehydrogenases ALDH2 and ALDH9A1, has been observed at 180 dpc when compared to other foetal ages. However, these proteins were proposed as hallmark WAT in mouse (Forner et al. 2009). From 210 dpc, an increase in the mRNA abundances of UCP1 and DIO2, which are hallmarks of brown adipocytes, was observed. These results, associated with the presence of numerous unilocular white adipocytes and few brown multilocular adipocytes from 180 dpc, show that foetal ATs up to 260 dpc have molecular and cellular characteristics of WAT in addition to those of BAT (Taga et al. 2012). These results challenged the commonly accepted concept that foetal perirenal AT in cow is brown and highlighted that perirenal AT is a heterogeneous tissue made of brown and white adipocytes up to 260 dpc in bovine foetuses. This study, together with the global gene expression profiling of brown/beige to white adipose tissue during the first month of life in lambs (Basse et al. 2015), provides a useful resource to identify the molecular features of perinatal brown/beige adipose tissue relatively to those of white adipose tissues presented above.

3 Adipose Tissue in Farm Animals: Endocrinology and Immunology

With the discovery of leptin in 1994 (Zhang et al. 1994), adipose tissue has been identified as an endocrine gland secreting messenger molecules that affect a wide variety of bodily functions. From then on, numerous hormones, chemokines, cytokines as well as acute-phase proteins were discovered as adipose products and are now collectively termed as adipokines. Adipokines are involved in the dynamic control of energy metabolism. Being produced by adipose tissue, adipokines are fundamental to deliver to other tissues important information about the nutrient status of the organism, in particular to those responsible for controlling energy intake and expenditure, such as appetite regulation, energy expenditure, insulin sensitivity, glucose metabolism and fatty acid oxidation. Figure 4 presents a short list of adipokines and their main physiological roles. The human adipose tissue secretome determined by proteomic techniques was released in 2007 (Alvarez-Llamas et al. 2007). Products from all cell types present in adipose tissue were included in this first report. In fact, adipose tissue contains not only adipocytes but other cell types as well. A stromal vascular fraction of cells including preadipocytes, fibroblasts, vascular endothelial cells and a variety of immune cells (e.g. macrophages) are also included in adipose tissue depots.

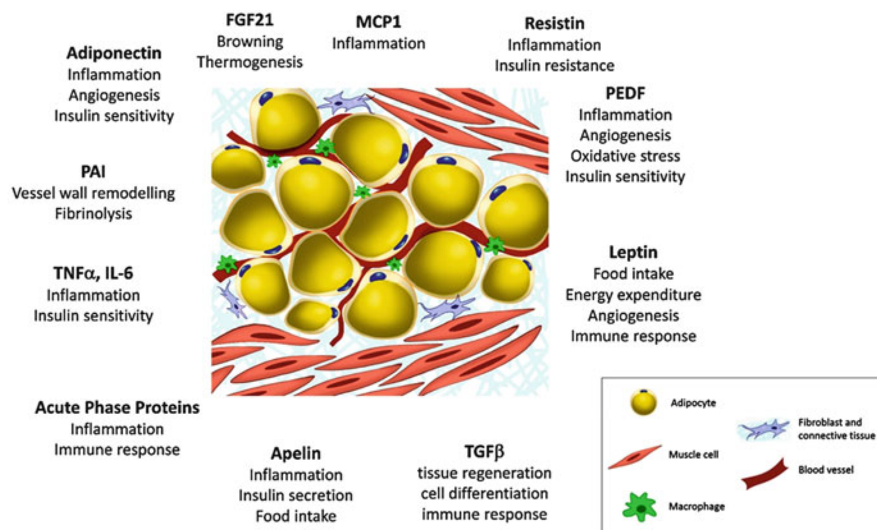

Fig. 4 The complexity of AT and physiological functions of adipokines. Beside adipocytes, adipose tissue also contains preadipocytes; fibroblasts and connective tissues; vascular endothelial cells, providing the capillary network; myocytes; and muscle cells and macrophages that increase when adipose tissue becomes inflamed. In this figure, some of the proteins that are produced by AT, and that can be found in the circulation, the so-called adipokines, are also presented. Some of their main activities are also briefly summarized. *FGF21* fibroblast growth factor 21, *IL-6* interleukin 6, *MCP1* monocyte chemoattractant protein 1, *PAI1* plasminogen activator inhibitor 1, *TNF-α* tumour necrosis factor alpha and *PEDF* pigment epithelium-derived factor. We thank Mrs Helen Arino for drawing the figure

Proteomics analysis carried out on primary human adipocyte cell culture identified 347 proteins of which 263 were predicted to be secreted (Lehr et al. 2012). The list of these which are regarded as adipokines includes not only proteins whose abundance is restricted to adipose tissue but also molecules taken up by adipose tissues that are known to be secreted from other tissues, such as cytokines and acute-phase proteins. Most of adipokines that were initially discovered and considered as adipose specific were then found to be secreted by other tissues such as skeletal muscle and are now classified as myokines. Both adipokines and myokines play major roles in exchanging information between skeletal muscle fibres and adipocytes in the framework of an endocrine cross talk (Trayhurn et al. 2011). Recently, a computational prediction of the large-scale secretome of adipose tissues and muscles in ruminant was achieved by applying algorithms to 24 publications reporting transcriptomics or proteomics results from bovine, ovine and caprine species and eight transcriptomics dataset series. In this way, 1749 proteins were proposed as secreted both by adipose tissues and muscles, and 188 and 357 proteins were predicted to be secreted either by adipose tissues or muscles, respectively (Bonnet et al. 2016).

As a consequence of the cell diversity in adipose tissue, during human obesity, macrophages forming a part of the stromal vascular fraction infiltrate adipose tissue

Table 1 Proteins identified as related to intramuscular fat deposition in at least two studies from five, in which strong differences in marbling resulted from age (Kim et al. 2009[1]; Zhang et al. 2010[2]), breed (Keady et al. 2013[3]) or marbling score at 24 (Shen et al. 2012[4]) or 30 (Mao et al. 2016[5]) months of age in bovine

Gene names	Positive or negative correlations between protein abundance and marbling
CA2	Negative[2,4,5]
MYL1	Negative[3,4] and positive[1]
HSPB1	Negative[5] and positive[2,3]
TPI1	Negative[1,5] and positive[4]
CAPZA2	Positive[3,4]
MYLPF	Negative[3,5]
MYL3	Negative[2,5]
PYGM	Negative[5] and positive[3]
PGK1	Negative[5] and positive[4]

Superscripts refer to the publications as defined in the table caption

in increasing numbers, providing important contribution to the secretory function of adipose tissue, in particular for inflammatory cytokines, such as TNF-α and IL-6 (Weisber et al. 2003). The increase in circulating levels of pro-inflammatory macrophage-derived factors during obesity induces a chronic low-grade inflammatory state leading to the development of insulin resistance and diabetes (Xu et al. 2003). It must be said that immune cells were hardly detectable in both visceral and subcutaneous adipose tissue in dairy cattle. Therefore, macrophages might not be involved in the immunity and metabolism of adipose tissue in nonobese lactating animals (Akter et al. 2012). In the previously described study on goat adipose tissue proteomics (Restelli et al. 2014), it was demonstrated that goat adipose tissue contains 51 proteins that are related to inflammatory and immune responses. The list of these proteins includes acute-phase proteins, such as C-reactive protein, ceruloplasmin and alpha-1-acid glycoprotein and galactoside-binding proteins (LGALS), such as beta-galactoside-binding lectin precursor and galactoside 3. LGALS act as agonists of platelet activation and are pro-apoptotic for immune cells, activate and increase the adhesion of neutrophil and are chemotactic for monocytes (Table 1). Whether these proteins are produced within adipose tissue or are present in adipose tissue because of their uptake from the blood remains to be investigated in ruminants.

4 When Things Turn Wrong: Adipose Tissue in Dairy Animals and Metabolic Imbalance Syndrome

Hepatic lipidosis (also known as “fatty liver disease” or “fat cow syndrome”) is a common production problem of dairy cows occurring during the critical physiologic transition from pregnancy to lactation (Bobe et al. 2004; Hammon et al.

2009). During recent decades, the genetic selection carried out on dairy cows was directed toward an increase of milk production. The result was positive, since milk production reached performance peaks unthinkable even a few years ago. This major gain did not come without serious consequences for cow health. The excessive demand for nutrients frequently ends up in a severe energetic deficit at the beginning of lactation (Mulligan and Doherty 2008; Hammon et al. 2009). The severity of this effect is increased by the need of nutrients for the foetus that is growing exponentially and the fact that feed intake does not equally increase. Animals affected by this major energy deficiency try to counteract by developing insulin resistance (IR) to spare glucose for the foetus (Bauman and Currie 1980) and rapidly mobilizing fat depots, in the attempt of providing nonesterified fatty acids (NEFA) as an energy source. Triacylglycerols are released from hepatocytes as part of lipoproteins, mostly very low-density lipoproteins (VLDL). In ruminants, the secretion of VLDL from the liver is limited as compared with other species. Consequently, the hepatic uptake of NEFA and storage in the form of triacylglycerols exceed their elimination. The storage of excess lipids in hepatocytes leads to liver damage and depressed liver functions (Geelen and Wensing 2006). A morphologic change in liver tissue occurs, characterized by the accumulation of lipid vacuoles within hepatocytes. This metabolic change is called hepatic lipidosis or liver fatty change (Cebra et al. 1997; Imhalsy et al. 2014). Considering this dual role of regulating energy storage by storing and releasing fatty acids, and acting as a major endocrine capable of influencing metabolism by secreting and regulating hormones and adipokines (Kershaw and Flier 2004), adipose tissue metabolism plays an essential role in the development of metabolism syndrome in transition dairy cows. Adipose tissue reacts to the increase of energy demand by increasing lipolysis and by regulating the development of major metabolic changes including, among the other, insulin resistance or sensitivity (Bell and Bauman 1997; Rabe et al. 2008; Loor et al. 2013).

A recent study determined the proteome of transition cow adipose tissue focusing on the relationship between insulin sensitivity (IS) and resistance (IR) of subcutaneous adipose tissue (Zachut 2015). The results showed that a number of 143 proteins out of 586 were differentially abundant in prepartum versus postpartum adipose tissue. The lipid metabolism-related functions that were significantly changed postpartum compared with prepartum include fatty acid metabolism, the esterification of lipids and oxidation of fatty acids. The proteins whose abundance was decreased in postpartum included fatty acid synthase, complement C3, annexin A1 and acyl-CoA desaturase. Finally, the study also addressed proteomic differences between IR and IS adipose tissues. Out of 586 proteins, 111 were differentially abundant between IS and IR cows. Most of them (a number of 106) were increased in IR versus IS adipose, whereas only five were decreased. As expected, the most relevant pathways differentially activated between IR and IS adipose tissue included energy-related pathways, such as gluconeogenesis and glycolysis, 14−3−3-mediated signalling, TCA cycle and ERK/MAPK signalling. The most relevant function in IR as compared to IS adipose tissue was inflammatory response, such as leukocyte migration and proliferation of T lymphocytes,

confirming the relationship between adipose tissue IR and proteins related to inflammation, and organization of actin cytoskeleton.

5 The Marbling Issue: Adipose Tissue in Beef

Understanding muscular adipogenesis and identifying a protein profile associated with intramuscular fat deposition also termed marbling are prerequisites to develop strategies to manipulate marbling in cow and pig in an effort to produce even healthier and tastier meat products for consumers. In order to provide the molecular basis of marbling and to identify biomarkers to predict the animal's ability to deposit intramuscular fat, proteomics of fat accumulation was assayed in the bovine species in five proteomic investigations mainly through 2DE followed by MS/MS analysis. The proteome of *longissimus lumborum* or *longissimus dorsi* was explored in steers, mainly of Asian breeds, showing strong differences in marbling as the result of age (Zhang et al. 2010; Kim et al. 2009), breed (Keady et al. 2013) or marbling score at 24 (Shen et al. 2012) or 30 (Mao et al. 2016) months of age that are among the major drivers of individual variations in the intramuscular fat percentage (Shingfield et al. 2013).

By merging results from these five studies, 50 unique proteins were proposed to be involved in the deposition of intramuscular fat content that are informative of the molecular basis and of the major molecular pathways involved. Among them four (Zhang et al. 2010) and seven (Kim et al. 2009) proteins were found to be related to a high intramuscular fat deposition induced by an increase in age of Korean steers. These proteins were heat shock protein beta 1 (HSPB1, upregulated), ATP synthase D chain mitochondria (ATP5H), carbonic anhydrase II (CA2), myosin light chain 3, slow-twitch muscle [MYL3, all downregulated (Zhang et al. 2010)], as well as myosin light chain 1, slow-twitch muscle A isoform (MYL1), thioredoxin-dependent peroxide reductase (PRDX3, both upregulated), actin, aortic smooth muscle (ACTA2), actin, cytoplasmic 1 (ACTB), succinate dehydrogenase [ubiquinone] flavoprotein subunit (SDHA), triosephosphate isomerase (TPI1) and zinc finger protein 323 [ZSCAN31, downregulated (Kim et al. 2009)]. These proteins are mainly involved in glycolysis/gluconeogenesis (TPI1), the oxidation of glucose and fatty acids (ATP5H) and the regulation of muscle contraction (MYL3, MYL1, ACTA2). Additionally, the most enriched GO terms provided by ProteINSIDE (Kaspric et al. 2015) for these 11 proteins are "muscle filament sliding" (MYL1 MYL3), "muscle contraction" (MYL1 ACTA2), "vascular endothelial growth factor receptor signalling pathway" (ACTB HSPB1) and "negative regulation of apoptotic process" (HSPB1 PRDX3).

Differential abundances of 7 and 28 proteins were found to be related to divergent marbling scores, in Xiangxi yellow × Angus cattle steers (Mao et al. 2016) and in Hanwoo cattles (Shen et al. 2012), respectively, three of them being reported in both breeds. The abundance of carbonic anhydrase II (CA2) was reported to be negatively related to marbling score in both breeds. However,

negative relations were reported between the abundances of triosephosphate isomerase (TPI1), phosphoglycerate kinase 1(PGK1) and the intramuscular fat content in the Xiangxi yellow × Angus cattle steers, while a positive relation was reported in Hanwoo cattle.

Differential abundances of 17 proteins were found to be related to divergent intramuscular fat content in the muscle of Belgian Blue × Holstein Friesian compared to Aberdeen Angus × Holstein Friesian (Keady et al. 2013). The top canonical pathways identified by the authors were glycolysis/gluconeogenesis (glycogen phosphorylase (PYGM), phosphoglycerate mutase 2 (PGAM2) and aldolase A (ALDOA) as greater in abundance in highly marbled Aberdeen Angus), the citric cycle (aconitase 2 (ACO2) and 2-oxoglutarate dehydrogenase (OGDH) as greater in abundance in highly marbled Aberdeen Angus), the protein kinase A signalling pathway (myosin light chain 1 (MYL1), myosin light chain, phosphorylatable (MYLPF), PYGM and troponin I (TNNI2) differing in abundance across breed), and the pentose phosphate pathway [with ALDOA greater and phosphoglucomutase (PGM1) decreased in abundance in Aberdeen Angus compared to Aberdeen Angus].

Altogether these data highlight that the main pathways related to intramuscular fat deposition in beef are metabolic pathways related to glucose, oxidative pathways, the molecular features of muscle fibres (and thus of muscle type) and pathways related to apoptosis. Some of them could be related to previous observations that glucose rather than acetate is a major precursor for lipogenesis, and hence fat accumulation, within intramuscular adipocytes (Smith and Crouse 1984) or to the repeatedly observed higher marbling in skeletal muscle of type I than II (Bonnet et al. 2007).

Of these 50 unique proteins, only 9 were related to intramuscular fat deposition in at least two publications and thus could be potential robust biomarkers to predict meat marbling (Fig. 3 and Table 1). The proteins HSPB1, MYL1, CA2 and TPI1 were declared as differentially abundant according to the intramuscular fat deposition in three studies. Of these, CA2 was always reported to be lower in abundance in muscle with high intramuscular fat content. The proteins CAPZA2, MYLPF, PYGM, MYL3 and PGK1 were declared differentially abundant according to the intramuscular fat deposition in two publications. Among them, CAPZA was always reported to be higher in abundance in muscle with high intramuscular fat content, while MYLPF and MYL3 were reported to be lower (Table 1). The abundance of HSPB1, MYL1, TPI1, PYGM and PGK1 was reported to have a positive or negative relation with marbling depending on studies. The apparent inconsistencies of these results may be explained by an increase in some pathways related to lipogenic capabilities in adipocytes of small volume that are then decreased in large adipocytes, as already observed for proteins involved in lipid synthesis (Bonnet et al. 2007; Romao et al. 2014). A variation in the abundances of HSPB1, MYL1, TPI1, PYGM and PGK1 may partially explain the result of differences depending on breed, age, nutrition and not only of marbling. To sum up, CA2, CAPZA, MYLPF and MYL3 may be proposed as robust biomarkers of

marbling in beef since their abundances were related to intramuscular fat deposition regardless of breed, age or level of marbling.

This literature review lastly shows that proteins related to marbling (Zhang et al. 2010; Kim et al. 2009; Keady et al. 2013; Shen et al. 2012; Mao et al. 2016) are different than those related to back fat thicknesses (Charolaise × Red Angus and Hereford × Angus; Zhao et al. 2010) in steers.

6 Adipose Tissue in Pigs

The pig is a major source of meat for human consumption. Knowledge of the biochemical mechanisms and understanding the biological significance of adipose tissue development in pigs are important for optimal growth efficiency and meat quality. A specific issue of porcine production is the processed meat products. The dynamic of lean and fat growth is of fundamental importance to improve the production of differentiated end products through breeding and feeding strategies. As an example, optimal adipogenesis requirement is different if the final product is fresh versus processed meat products, such as the production of regional and traditionally cured and dried high-value products, like Parma and Serrano ham (Candek-Potokar and Skrlep 2012). Meanwhile, with the background that swine physiology, genomics and nutritional requirements are very similar to that of humans, the pig has become an increasingly important animal model for human metabolic diseases and obesity (De Almeida and Bendixen 2012; Ceciliani et al. 2014). Despite these various interests, proteomic experiments to understand adipogenesis at the protein level in pig are still very limited, as compared to the information made available from genome-wide association, targeted candidate genes and large-scale transcriptomic studies (Cánovas et al. 2010; Ramayo-Caldas et al. 2012; Corominas et al. 2013; Puig-Oliveras et al. 2014; Xing et al. 2015, 2016; Shen et al. 2015; Ros-Freixedes et al. 2016; Zhang et al. 2016). Proteome research has mainly been focused on mapping the muscle proteome with postmortem modifications driving the transformation of muscle to meat, such as meat tenderness, postmortem proteolysis (Pioselli et al. 2011; Mora et al. 2015; Gallego et al. 2016), phosphoproteomics (Huang et al. 2014), water holding capacity (Di Luca et al. 2013, 2016) and meat colour traits (Lomiwes et al. 2014). Very few studies have been carried out on the pig adipose tissue proteome. The available proteomics were carried out in muscle with the aim to identify protein and pathways related to intramuscular fat deposition and in subcutaneous adipose tissue to identify pathways related to body adiposity.

Although intramuscular fat content is an important determinant of meat quality, the information related to deposition of intramuscular fat at protein and proteome levels is scarce. Differences between high and low ability for fat deposition in longissimus dorsi muscle from commercial crosses originating from Pietrain and an industrial cross originating from Duroc, Hampshire and Large White founders were identified in comparative study including both transcriptome and proteome profiles

(Liu et al. 2009). Differences for marbling were correlated to transcriptome and proteome profiles, and 40 muscular genes were identified as differently expressed between high-fat and low-fat groups, either at the mRNA level (29) or encoding proteins (12). Among them, only GSTP1 was found differently abundant at both the mRNA and protein levels. Gene ontology analysis indicated that differentially expressed genes were involved in metabolic processing, cell communication, binding and response to stimulus. The results of this study suggest that interindividual variability in intramuscular fat content might arise essentially from differences in early muscular adipogenesis that may have modified either the volume or the number of intramuscular adipocytes. A recent investigation, pairing transcriptomics with proteomics, identified several genes and proteins involved in fatty acid metabolism and intramuscular fat deposition (Yang et al. 2016). A total of 23 differentially expressed proteins were identified, several of which were potentially associated with fatty acid metabolism. In particular, adipocyte fatty acid-binding protein A, alpha-enolase isoform X1 and beta-enolase isoform X1 were more abundantly found in Chinese indigenous Shaziling pig, a fat-type line with high intramuscular fat (IMF), as compared with the Yorkshire breed, which has a leaner meat ratio. By identifying differences in breed-related protein and transcript abundance patterns between the two breeds, these data provide insights into the mechanisms of growth and development of porcine skeletal muscle and how it might influence the IMF deposition.

The main adipose tissue studied in pig was the subcutaneous. Indeed, as compared to cattle which has a major deposition of fat within visceral and intermuscular adipose tissues, pigs deposit fat mostly within subcutaneous adipose tissue. An excessive growth of subcutaneous adipose tissue decreases the gain production of pig. A characterization of subcutaneous adipose tissue proteomes of young piglets was carried out focusing on the metabolic control and acute-phase response associated with adipogenesis in lean (Large White) as compared to fat (Basque) breeds (Gondret et al. 2012). Mass spectrometry and MS/MS analyses identified 65 proteins, 57 of which were significantly different between subcutaneous adipose tissues from Basque and Large White pigs, and 12 others were expressed with different abundance. The authors demonstrated that several metabolic pathways are differentially expressed in the lean versus fat breeds including, most significantly, the pentose-phosphate pathway, with aldolase C, glucose-6-phosphate dehydrogenase and ribose 5-phosphate isomerase A; the citrate cycle with pyruvate carboxylase, dihydrolipoyl dehydrogenase and aconitase; and the pyruvate metabolism with aldolase C and dihydrolipoyl dehydrogenase; again, malic enzyme; and pyruvate carboxylase proteins. The most significant biological function affected includes lipid metabolism, which was increased in Basque breed as compared to Large White lean one. Carboxylesterase 1 was more abundant in adipose tissue of Basque breed, suggesting an increased activity of triacylglycerol synthesis and degradation aimed to protect cells against fat overload. Adipose tissue of Basque piglets also showed an increased abundance of inflammatory-related proteins, suggesting a low-grade inflammation in “fat” pigs, probably induced by an increased oxidative stress related to adipocyte differentiation. This effect closely

resembles human obesity and inflammatory metabolic syndrome status. The increased abundance of selenium-binding protein, which has ROS scavenging properties, also confirms the need of quenching ROS aiming to protect tissues from oxidative damages. To understand the impact of maternal nutrition on prenatal and offspring metabolism and body adiposity, the proteomes of subcutaneous adipose tissues in piglets born from sows fed on either low, high or normal protein diets were determined. The 2D proteome profiles between the three groups were different (Sarr et al. 2010, 2012). In details, proteins related to fatty acid metabolism and lipid transport were upregulated in piglets born from sows fed on low protein diets. Analogue transcriptome studies have confirmed that foetal programming indeed affects adiposity of pigs (Oster et al. 2011, 2012). Besides increasing our knowledge about the relationship between maternal nutrition and offspring adipose tissue development in human nutrition, these two studies also provide important information that might be implemented in improving rearing and possibly driving the practices affecting the partitioning of fat between subcutaneous and intramuscular adipose tissues.

7 Adipose Tissue in Other Species: Fish and Poultry

7.1 Adipose Tissue Proteomics in Broiler Production

Adipose tissue is regarded as a negative trait in poultry science, although it must be said that genetic pressure for rapid growth had also an increase of fat deposition as side effects (Nones et al. 2006). Proteomic analysis was carried out to obtain insights into the molecular basis of fat deposition in broiler combining 2D electrophoresis with identification by MALDI-TOF (Wang et al. 2009). Approximately 1000 protein spots were identified in adipose tissue, of which 15 proteins were shown to be differentially expressed between lean and fat chickens, most of them being involved in lipid metabolism, including adipocyte FABP, apolipoprotein A-I and acyl-coenzyme A dehydrogenase, and oxidative stress, such as HSP 27, which was downregulated, and glutathione-*S*-transferase-α and *S*-transferase-β, which were upregulated. Furthermore, members of cytoskeleton family were differentially abundant, underlying a different cytoskeleton rearrangement. These results were partially confirmed by a recent study that investigated the differences in abdominal adipose tissue proteomes between broiler lines divergently selected for abdominal fat content (NEAULF breed) (Wu et al. 2016). Thirteen proteins were found to be differentially expressed between lean and fat lines. In particular, the expression of abdominal adipose tissue Apo A-I, PPIase FKBP4 and cytokeratin otokeratin was found significantly higher in lean birds as compared with fat birds. These proteins are mainly involved in lipid metabolism, amino acid metabolism, signal transduction, energy conversion and antioxidant and cytoskeleton as well as in adipose tissue metabolism.

7.2 Fish Adipose Tissue Proteomics

Lipids are the main source of energy for fish (Weil et al. 2013). Adipose tissue is distributed in fat mainly located around the digestive tract in the abdominal cavity (visceral fat). The percentage of fat in fish may be very high (up to 25% of the body weight). Located on the ventral and dorsal area of the fish is the subcutaneous fat, which is located around the body of the fish, prominently in dorsal and ventral zones. The quality of carcass is dependent on the distribution of adipose tissue between subcutaneous and perivisceral depots, given that adipocytes are also located in muscle myosepta, a connective tissue membrane separating muscle sheets.

Only two proteomics studies have been carried out on fish adipose tissue, both of them on rainbow trout (*Oncorhynchus mykiss*). Weil et al. (2009) carried out an in vitro investigation on isolated and cultured adipocytes from subcutaneous (dorsal and ventral) and visceral fat depots. The study investigated the differentiated adipocytes from preadipocytes, while other cells that normally compose the adipose tissue depots, such as connective tissues, endothelial cells or blood cells, were excluded. Proteins were separated by means of 2D electrophoresis, and MALDI-TOF identified nine differentially abundant spots depending on differentiation stage. Of these proteins, beta actin and albumin were associated with visceral adipose tissue, whereas annexin, ATP synthase subunit-β, serum deprivation response protein and heart fatty acid-binding protein (H-FABP) were found to be more abundant in subcutaneous AT. The authors found no differences between adipocytes isolated from visceral and subcutaneous adipose tissue, although the amount of protein expressed was different for some of them. Besides proteins related to cell cultivation procedures, such as bovine serum albumin or serum deprivation response protein, the authors identified actin as the more abundant protein in visceral-derived adipocytes. Actin is involved in cytoskeleton structure, and it was found to be overexpressed during adipogenesis of intramuscular bovine adipocytes (Takenouchi et al. 2004) and is involved in cytoskeletal rearrangement as well. ATP synthase, annexin and H-FABP were found to be more abundant in subcutaneous adipocytes, suggesting that this depot is more metabolically active than visceral ones, which is in contrast with results so far demonstrated for mammalian species. The investigation was integrated with histological parameters, such as adipocyte cell diameter determination, presenting the evidence that the diameter of adipocytes derived from visceral AT was higher as compared to those derived from subcutaneous adipose tissue. A study on transcriptomics and proteomics rainbow trout liver focusing on the differences in fat allocation between visceral and muscle AT was carried out after feeding with high- and low-energy diets (Kolditz et al. 2008). The authors sorted the animals in two groups differing for adipose tissue anatomical distribution, namely, those with mainly visceral adipose tissue (lean muscle line) and mainly muscle adipose tissue (fat muscle line). H-FABP was upregulated in the fat muscle cell line when compared with the lean muscle cell line, confirming its importance in fish adipose metabolism.

GAPDH was upregulated as well, whereas apolipoprotein A-1, which is usually overexpressed in adipose tissue, was on the contrary downregulated in the liver.

8 Conclusions: New Insights into the New Knowledge Contributed by System Biology Approach to Better Understanding of Adipose Tissue in Farm Animals

Postgenomic applications in veterinary medicine, including transcriptomics, proteomics and metabolomics, are increasing exponentially. The number of omics studies carried out in farm animals pales if compared to those carried out in humans and a true system biology approach to the impact of omics in livestock are far from being fully implemented. Animal proteomics is further lagging behind those in human biology (Almeida et al. 2015). Before the full development of the "omics" revolution, the knowledge of the physiological bases of adipose tissue development and IMF deposition was obtained through independent analysis of the activity of each single protein. The study of adipose tissue carried out using proteomic approaches, in particular when applied together with a system biology approach, including also transcriptomics and metabolomics, provides an integrated network of the single elements, the knowledge of which provides, in turn, greater information than the sum of individual parts. Transcriptomics and proteomics are an evident leap forward in our understanding of basic biology of adipose tissue, integrating the information necessary to link, for example, adipose tissue and metabolic stress in dairy cows or optimal IMF deposition in cows and pigs. This information might be readily implemented into the field, on the basis that adipose tissue metabolism may be modified by changing fatty acid content through diet. Omics technologies remain quite expensive, proteomics even more. Yet, technology moves rapidly forward and the costs for omics application are constantly dropping. The goal of a $1000 genome has been almost reached, and it is expected that further drops in omic experiment costs will result in an exponential increase of proteomic studies as well.

Acknowledgements Hereby we acknowledge the support of Dr Helena Arino (he.ar.ba@gmail.com) for the drawing of Fig. 4.

References

Akter SH, Häussler S, Germeroth D et al (2012) Immunohistochemical characterization of phagocytic immune cell infiltration into different adipose tissue depots of dairy cows during early lactation. J Dairy Sci 95(6):3032–3044

Almeida AM, Bassols A, Bendixen E et al (2015) Animal board invited review: advances in proteomics for animal and food sciences. Animal 9(1):1–17

Alvarez-Llamas G, Szalowska E, de Vries MP et al (2007) Characterization of the human visceral adipose tissue secretome. Mol Cell Proteomics 6(4):589–600

Asano H, Yamada T, Hashimoto O et al (2013) Diet-induced changes in Ucp1 expression in bovine adipose tissues. Gen Comp Endocrinol 184:87–92

Barbatelli G, Murano I, Madsen L et al (2010) The emergence of cold-induced brown adipocytes in mouse white fat depots is determined predominantly by white to brown adipocyte transdifferentiation. Am J Physiol Endocrinol Metab 298(6):E1244–E1253

Basse AL, Dixen K, Yadav R et al (2015) Global gene expression profiling of brown to white adipose tissue transformation in sheep reveals novel transcriptional components linked to adipose remodeling. BMC Genomics 16:215

Bauman DE, Currie WB (1980) Partitioning of nutrients during pregnancy and lactation: a review of mechanisms involving homeostasis and homeorhesis. J Dairy Sci 63(9):1514–1529

Bell AW, Bauman DE (1997) Adaptations of glucose metabolism during pregnancy and lactation. J Mammary Gland Biol Neoplasia 2(3):265–278

Bobe G, Young JW, Beitz DC (2004) Invited review: pathology, etiology, prevention, and treatment of fatty liver in dairy cows. J Dairy Sci 87(10):3105–3124

Bonnet M, Faulconnier Y, Leroux C et al (2007) Glucose-6-phosphate dehydrogenase and leptin are related to marbling differences among Limousin and Angus or Japanese Black × Angus steers. J Anim Sci 85:2882–2894

Bonnet M, Cassar-Malek I, Chilliard Y et al (2010) Ontogenesis of muscle and adipose tissues and their interactions in ruminants and other species. Animal 4:1093–1109

Bonnet M, Tournayre J, Cassar-Malek I (2016, August) Integrated data mining of transcriptomic and proteomic datasets to predict the secretome of adipose tissue and muscle in ruminants. Mol BioSyst 12(9):2722–2734. https://doi.org/10.1039/c6mb00224b

Candek-Potokar M, Skrlep M (2012) Factors in pig production that impact the quality of dry-cured ham: a review. Animal 6(2):327–338

Cánovas A, Quintanilla R, Amills M et al (2010) Muscle transcriptomic profiles in pigs with divergent phenotypes for fatness traits. BMC Genomics 11:372

Cebra CK, Garry FB, Getzy DM et al (1997) Hepatic lipidosis in anorectic, lactating holstein cattle: a retrospective study of serum biochemical abnormalities. J Vet Intern Med 11 (4):231–237

Ceciliani F, Restelli L, Lecchi C (2014) Proteomics in farm animals models of human diseases. Proteomics Clin Appl 8(9–10):677–688

Corominas J, Ramayo-Caldas Y, Puig-Oliveras A et al (2013) Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. BMC Genomics 14:843

de Almeida AM, Bendixen E (2012) Pig proteomics: a review of a species in the crossroad between biomedical and food sciences. J Proteome 75(14):4296–4314

Di Luca A, Elia G, Hamill R, Mullen AM (2013) 2D DIGE proteomic analysis of early post mortem muscle exudate highlights the importance of the stress response for improved water-holding capacity of fresh pork meat. Proteomics 13(9):1528–1544

Di Luca A, Hamill RM, Mullen AM et al (2016) Comparative proteomic profiling of divergent phenotypes for water holding capacity across the post mortem ageing period in porcine muscle exudate. PLoS One 11(3):e0150605

Forner F, Kumar C, Luber CA et al (2009) Proteome differences between brown and white fat mitochondria reveal specialized metabolic functions. Cell Metab 10:324–335

Frayn KN (2002) Adipose tissue as a buffer for daily lipid flux. Diabetologia 45(9):1201–1210

Gallego M, Mora L, Concepción Aristoy M et al (2016) The use of label-free mass spectrometry for relative quantification of sarcoplasmic proteins during the processing of dry-cured ham. Food Chem 196:437–444

Geelen MJ, Wensing T (2006) Studies on hepatic lipidosis and coinciding health and fertility problems of high-producing dairy cows using the "Utrecht fatty liver model of dairy cows". A review. Vet Q 28(3):90–104

Gondret F, Guével B, Com E et al (2012) A comparison of subcutaneous adipose tissue proteomes in juvenile piglets with a contrasted adiposity underscored similarities with human obesity. J Proteome 75(3):949–961
Griffin TJ, Gygi SP, Ideker T et al (2002) Complementary profiling of gene expression at the transcriptome and proteome levels in Saccharomyces cerevisiae. Mol Cell Proteomics 1 (4):323–333
Hammon HM, Stürmer G, Schneider F et al (2009) Performance and metabolic and endocrine changes with emphasis on glucose metabolism in high-yielding dairy cows with high and low fat content in liver after calving. J Dairy Sci 92(4):1554–1566
Hausman GJ, Basu U, Wei S et al (2014) Preadipocyte and adipose tissue differentiation in meat animals: influence of species and anatomical location. Annu Rev Anim Biosci 2:323–351
Huang H, Larsen MR, Palmisano G et al (2014) Quantitative phosphoproteomic analysis of porcine muscle within 24 h postmortem. J Proteome 106:125–139
Imhasly S, Naegeli H, Baumann S et al (2014) Metabolomic biomarkers correlating with hepatic lipidosis in dairy cows. BMC Vet Res 10:122
Kaspric N, Picard B, Reichstadt M et al (2015) ProteINSIDE to easily investigate proteomics data from ruminants: application to mine proteome of adipose and muscle tissues in bovine foetuses. PLoS One 10:e0128086
Keady SM, Kenny DA, Ohlendieck K et al (2013) Proteomic profiling of bovine M. longissimus lumborum from crossbred Aberdeen Angus and Belgian Blue sired steers varying in genetic merit for carcass weight. J Anim Sci 91(2):654–665
Kershaw EE, Flier JS (2004) Adipose tissue as an endocrine organ. J Clin Endocrinol Metab 89 (6):2548–2556
Kim NK, Lee SH, Cho YM et al (2009) Proteome analysis of the m. longissimus dorsi between fattening stages in Hanwoo steer. BMB Rep 42:433–438
Kolditz CI, Paboeuf G, Borthaire M et al (2008) Changes induced by dietary energy intake and divergent selection for muscle fat content in rainbow trout (Oncorhynchus mykiss), assessed by transcriptome and proteome analysis of the liver. BMC Genomics 9:506
Lehr S, Hartwig S, Lamers D et al (2012) Identification and validation of novel adipokines released from primary human adipocytes. Mol Cell Proteomics 11(1):M111.010504
Liu J, Damon M, Guitton N, Guisle I, Ecolan P, Vincent A, Cherel P, Gondret F (2009) Differentially-expressed genes in pig Longissimus muscles with contrasting levels of fat, as identified by combined transcriptomic, reverse transcription PCR, and proteomic analyses. J Agric Food Chem 57(9):3808–3817
Lizcano F, Vargas D (2016) Biology of beige adipocyte and possible therapy for type 2 diabetes and obesity. Int J Endocrinol 2016:9542061
Lomiwes D, Farouk MM, Wiklund E, Young OA (2014) Small heat shock proteins and their role in meat tenderness: a review. Meat Sci 96(1):26–40
Loor JJ, Bionaz M, Drackley JK (2013) Systems physiology in dairy cattle: nutritional genomics and beyond. Annu Rev Anim Biosci 1:365–392
Louveau I, Perruchot MH, Bonnet M et al (2016) Invited review: pre- and postnatal adipose tissue development in farm animals: from stem cells to adipocyte physiology. Animal 10:1839–1847
Luo L, Liu M (2016) Adipose tissue in control of metabolism. J Endocrinol 231(3):R77–R99
Mao YW, Hopkins DL, Zhang YM et al (2016) Beef quality with different intramuscular fat content and proteomic analysis using isobaric tag for relative and absolute quantitation of differentially expressed proteins. Meat Sci 118:96–102
Mora L, Gallego M, Escudero E et al (2015) Small peptides hydrolysis in dry-cured meats. Int J Food Microbiol 212:9–15
Mulligan FJ, Doherty ML (2008) Production diseases of the transition cow. Vet J 176(1):3–9
Nones K, Ledur MC, Ruy DC (2006) Mapping QTLs on chicken chromosome 1 for performance and carcass traits in a broiler × layer cross. Anim Genet 37(2):95–100

Oster M, Murani E, Metges CC et al (2011) A high protein diet during pregnancy affects hepatic gene expression of energy sensing pathways along ontogenesis in a porcine model. PLoS One 6 (7):e21691
Oster M, Murani E, Metges CC et al (2012) A low protein diet during pregnancy provokes a lasting shift of hepatic expression of genes related to cell cycle throughout ontogenesis in a porcine model. BMC Genomics 13:93
Pioselli B, Paredi G, Mozzarelli A (2011) Proteomic analysis of pork meat in the production of cooked ham. Mol BioSyst 7(7):2252–2260
Pope M, Budge H, Symonds ME (2014) The developmental transition of ovine adipose tissue through early life. Acta Physiol (Oxford) 210:20–30
Puig-Oliveras A, Ramayo-Caldas Y, Corominas J et al (2014) Differences in muscle transcriptome among pigs phenotypically extreme for fatty acid composition. PLoS One 9(6):e99720
Rabe K, Lehrke M, Parhofer KG et al (2008) Adipokines and insulin resistance. Mol Med 14 (11–12):741–751
Rajesh RV, Heo G, Park M et al (2010) Proteomic analysis of bovine omental, subcutaneous and intramuscular preadipocytes during in vitro adipogenic differentiation. Comp Biochem Physiol Part D Genomics Proteomics 5:234–244
Ramayo-Caldas Y, Mach N, Esteve-Codina A et al (2012) Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. BMC Genomics 13:547
Restelli L, Codrea MC, Savoini G et al (2014) LC-MS/MS analysis of visceral and subcutaneous adipose tissue proteomes in young goats with focus on innate immunity and inflammation related proteins. J Proteome 108:295–305
Romao JM, Jin W, He M et al (2013) Elucidation of molecular mechanisms of physiological variations between bovine subcutaneous and visceral fat depots under different nutritional regimes. PLoS One 8(12):e83211
Romao JM, He ML, McAllister TA et al (2014) Effect of age on bovine subcutaneous fat proteome: Molecular mechanisms of physiological variations during beef cattle growth. J Anim Sci 92:3316–3327
Ros-Freixedes R, Gol S, Pena RN et al (2016) Genome-wide association study singles out SCD and LEPR as the two main loci influencing intramuscular fat content and fatty acid composition in Duroc pigs. PLoS One 11(3):e0152496
Sarr O, Louveau I, Kalbe C et al (2010) Prenatal exposure to maternal low or high protein diets induces modest changes in the adipose tissue proteome of newborn piglets. J Anim Sci 88:1626–1641
Sarr O, Louveau I, Le Huërou-Luron I et al (2012) Adipose tissue proteomes of intrauterine growth-restricted piglets artificially reared on a high-protein neonatal formula. J Nutr Biochem 23(11):1417–1424
Sauerwein H, Bendixen E, Restelli L et al (2014) The adipose tissue in farm animals: a proteomic approach. Curr Protein Pept Sci 15(2):146–155
Sethi JK, Vidal-Puig AJ (2007) Thematic review series: adipocyte biology. Adipose tissue function and plasticity orchestrate nutritional adaptation. J Lipid Res 48(6):1253–1262
Shen YN, Kim SH, Yoon DH et al (2012) Proteome analysis of bovine longissimus dorsi muscle associated with the marbling score. Asian-Australas J Anim Sci 25:1083–1088
Shen L, Luo J, Du J et al (2015) Transcriptome analysis of Liangshan pig muscle development at the growth curve inflection point and asymptotic stages using digital gene expression profiling. PLoS One 10(8):e0135978
Shingfield KJ, Bonnet M, Scollan ND (2013) Recent developments in altering the fatty acid composition of ruminant-derived foods. Animal 7:132–162
Smith SB, Crouse JD (1984) Relative contributions of acetate, lactate and glucose to lipogenesis in bovine intramuscular and subcutaneous adipose tissue. J Nutr 114:792–800
Taga H, Chilliard Y, Meunier B et al (2012) Cellular and molecular large-scale features of fetal adipose tissue: is bovine perirenal adipose tissue brown? J Cell Physiol 227:1688–1700

Takenouchi T, Miyashita N, Ozutsumi K et al (2004) Role of caveolin-1 and cytoskeletal proteins, actin and vimentin, in adipogenesis of bovine intramuscular preadipocyte cells. Cell Biol Int 28 (8–9):615–623

Trayhurn P, Drevon CA, Eckel J (2011) Secreted proteins from adipose tissue and skeletal muscle – adipokines, myokines and adipose/muscle cross-talk. Arch Physiol Biochem 117(2):47–56

Wang D, Wang N, Li N et al (2009) Identification of differentially expressed proteins in adipose tissue of divergently selected broilers. Poult Sci 88(11):2285–2292

Weil C, Sabin N, Bugeon J et al (2009) Differentially expressed proteins in rainbow trout adipocytes isolated from visceral and subcutaneous tissues. Comp Biochem Physiol Part D Genomics Proteomics 4(3):235–241

Weil C, Lefevre F, Bugeon J (2013) Characteristics and metabolism of different adipose tissues in fish. J Rev Fish Biol Fisheries 23:157

Weisber SP, McCann D, Desai M et al (2003) Obesity is associated with macrophage accumulation in adipose tissue. J Clin Invest 112(12):1796–1808

Wu CY, Wu YY, Liu CD et al (2016) Comparative proteome analysis of abdominal adipose tissues between fat and lean broilers. Proteome Sci 14(1):9

Xing K, Zhu F, Zhai L et al (2015) Integration of transcriptome and whole genomic resequencing data to identify key genes affecting swine fat deposition. PLoS One 10(4):e0122396

Xing K, Zhu F, Zhai L et al (2016) Identification of genes for controlling swine adipose deposition by integrating transcriptome, whole-genome resequencing, and quantitative trait loci data. Sci Rep 6:23219

Xu H, Barnes GT, Yang Q et al (2003) Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance. J Clin Invest 112(12):1821–1830

Yang H, XL X, Ma HM, Jiang J (2016) Integrative analysis of transcriptomics and proteomics of skeletal muscles of the Chinese indigenous Shaziling pig compared with the Yorkshire breed. BMC Genet. 17(1):80

Zachut M (2015) Defining the adipose tissue proteome of dairy cows to reveal biomarkers related to peripartum insulin resistance and metabolic status. J Proteome Res 14(7):2863–2871

Zhang Y, Proenca R, Maffei M et al (1994) Positional cloning of the mouse obese gene and its human homologue. Nature 372(6505):425–432

Zhang Q, Lee H-G, Han J-A et al (2010) Differentially expressed proteins during fat accumulation in bovine skeletal muscle. Meat Sci 86:814–820

Zhang J, Ma J, Long K et al (2016) Dynamic gene expression profiles during postnatal development of porcine subcutaneous adipose. Peer J. 4:e1768

Zhao YM, Basu U, Dodson MV et al (2010) Proteome differences associated with fat accumulation in bovine subcutaneous adipose tissues. Proteome Sci 8:14

Proteomics and Mammary Gland Research in Dairy Species

André M. de Almeida and Peter David Eckersall

Abstract Proteomics has over the last decade become a technology that has been used more and more frequently in the study of mammary gland function in health and disease. There is now considerable evidence that the potential of this advanced analytical technology is broadly applied in mammary research and biology and starting to fulfil its potential. Proteomics has widely been used for several purposes. The first studies refer to the characterization of the mammary gland proteome, an objective achieved in the most important dairy species, cow and goat. Proteomics was also used to address the effects of genotype on the mammary gland proteome by contrasting, for instance, different breeds of cattle. Surprisingly, in bovines, proteomics has had limited use when comparing different nutritional levels. Proteomics has also been widely used as a tool to characterize the lactation cycle in cattle with studies related to morphology and milk secretion, among others.

Mastitis is the economically most important disease of dairy cows, and there has been substantial proteomics-based investigation of the change following infection. Both Gram-positive and Gram-negative organisms cause mastitis and milk from udders infected with either species causing substantial change to the milk proteome. High abundance milk protein such as caseins, β-lactoglobin and α-lactalbumin are reduced, while others such as albumin and IgG are increased. Recent advances in proteomic analysis of milk from dairy cows with mastitis have allowed over 290 milk proteins to be identified as they change in concentration following infection. Consistent results across a number of varied investigations have shown that proteins such as haptoglobin, cathelicidins and peptidoglycan recognition protein are among the most up-regulated components in milk during this disease, while fatty acid binding protein was consistently down-regulated.

A.M. de Almeida (✉)
Instituto Superior de Agronomia, Universidade de Lisboa, Lisboa, Portugal
e-mail: aalmeida@isa.ulisboa.pt

P.D. Eckersall
Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK

Faculty of Veterinary Medicine, University of Zagreb, Zagreb, Croatia

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_13

Proteomics has been extensively used in other species including dairy goats, buffalo or yaks and will likely increase in the near future, expanding, for instance, to wild mammal species. Finally, quantitative proteomics in particular is at the cusp of providing substantially greater insight into mammary gland metabolism and phenotypic alterations in the physiology of lactation and in disease processes.

Keywords Proteomics • Mammary gland • Dairy species • Mastitis

List of Abbreviations

1D	One dimension
2D	Two dimension
2DE	Two-dimensional electrophoresis
DIGE	Differential in-gel electrophoresis
ESI	Electrospray ionization
iTRAQ	Isobaric tag for relative and absolute quantitation
RPLC	Reversed phase liquid chromatography
LC	Liquid chromatography
MALDI TOF/TOF	Matrix assisted desorption ionization-time of flight
MS	Mass spectrometry
NMR	Nucleic magnetic resonance
QConCAT	Concatenated peptides
SCC	Somatic cell count
TLC	Teat canal lining

1 Introduction

The production of milk and dairy products such as butter, cream or yoghurt is one of the most important agricultural sectors worldwide. In 2013, the world milk production was estimated by the FAO[1]—Food and Agriculture Organization of the United Nations—to be approximately 770 billion tonnes, the majority being bovine (*Bos taurus* and *Bos indicus*) milk (83%). Other dairy species are the domestic water buffalo (*Bubalus bubalis*), goat (*Capra hircus*), sheep (*Ovis aries*) and camel (*Camelus dromedarius*) with, respectively, 13%, 2%, 1% and 0.3% of the world's fresh milk production. Milk production trends vary enormously from country to country and region to region as does the relative importance of each dairy species previously mentioned. As in all mammals, a very complex organ, the mammary gland, produces milk.

The mammary gland is in fact a modified sudoriferous gland that produces colostrum and milk for the nourishment of the newborn (Frandson et al. 2009). In

[1] http://faostat3.fao.org/download/Q/QL/E (accessed Oct 2016).

ruminants, mammary gland(s) are located in an udder. Udders are very complex organs with different types of cells with various roles according to their function. The udder is made up of a supportive system, an epithelial cell-based secretory system and a duct system, in addition to blood, lymph and nerve systems. The number of mammary glands varies from species to species with cows having four separate mammary glands and small ruminants having only two. The mammary gland secretes colostrum and milk and undergoes a complex and several-month-long lactation cycle with important alterations in size and function. The anatomical, physiological and hormonal aspects that govern colostrum and milk production in ruminants throughout such a complex lactation cycle have been extensively studied over the last few decades. Striking differences may be found in mammary gland physiology, morphology and function according to species (e.g. bovine vs. small ruminant and sheep vs. goat), breed (e.g. beef breeds vs. dairy breeds), stage of the lactation cycle (e.g. early vs. late lactation), nutritional status (e.g. over-conditioned vs. under-conditioned cows) or the presence of mastitis-causing microorganisms.

The last two decades have witnessed a gradual increase in the importance of the so-called post-genomics tools to the study of numerous aspects of farm animal biology, production and health (Eckersall et al. 2012; Almeida et al. 2015). Such tools include different fields like transcriptomics, metabolomics and proteomics and are usually integrated in a systems biology approach that confers a global multi-omics overview of the study subject contributing to an in-depth knowledge and understanding of the multitude of different factors at the molecular level.

This chapter has the objective of reporting on the applications and main achievements of proteomics in the context of mammary gland studies in dairy species. The chapter is divided into five different sections. Each is dedicated to a specific aspect of mammary gland research. After this brief introduction, the first section concerns efforts made to characterize the mammary gland proteome. A second section is about the effects of genotype and nutrition on the mammary gland proteome. The third section discusses changes in the bovine mammary gland proteome according to the different stages of the lactation cycle, and the fourth section is focused on the importance of proteomics to the study of mammary gland health, particularly the occurrence of mastitis. In a fifth section, we provide an emphasis to proteomic studies conducted on dairy species other than cattle. Finally, we end the chapter with a section on the major conclusions and future trends related to the use of proteomics to the study of mammary gland physiology and health.

2 Characterization of the Bovine Mammary Gland Proteome

Proteomic studies are dependent on how well the proteomes can be separated and individual protein abundance determined. These in turn depend on the existence of databases (public or private) that are crucial for the determination of protein abundance and identification (Soares et al. 2012). Accordingly, in less studied tissues,

organs and organisms, the first studies to be undertaken involved the characterization of the proteome. Such proteome mappings had the initial idea of locating protein spots in a 2D gel, typically identifying a few hundred proteins, but as proteomics technology evolved, rapidly moved to large datasets with thousands of proteins listed. Having numerous entries on protein databases and later being a sequenced organism, the *Bos taurus* species always had a quite high number of proteins identified in proteomics experiments. As such, it is not surprising that, to the best of our knowledge, no comprehensive 2DE map was ever created in cattle for the mammary gland, as the majority of the studies were about differential expression (e.g. Zhao et al. 2015) and focused primarily on the identification of proteins with differential abundance. The first large-scale characterization of the bovine mammary gland (and milk) proteomes was reported by Bislev et al. (2012a). In this publication, the authors present the "Bovine PeptideAtlas" which was essentially a first collection of several *Bos taurus* proteome datasets of different origins. The rationale behind this work was to build a source of information for proteomics-based milk production and mammary gland health studies. It has a very high potential and can be effectively used also for other dairy and animal science communities. The atlas comprised a total of 1921 proteins and 8559 distinct peptides.

Specific cell proteomes within the bovine mammary gland have also been characterized in recent years. Janjanam and colleagues (Janjanam et al. 2013) working with *Bos indicus* have characterized the proteome of mammary epithelial cells (milk-secreting cuboidal cells located in the innermost layer of alveoli) and investigated their role on lactation. Using a 2DE MALDI-TOF/TOF MS- and 1D-Gel-LC-MS/MS-based approach and mammary epithelial cells isolated from milk using immunomagnetic beads, the authors were able to identify almost 500 different proteins in a total of 28 different pathways that were described as involved in the lactation function. Furthermore, the study also contributed to the annotation of the bovine genome. In 2015, Smolenski and co-workers (2015) studied the proteome of the bovine teat canal lining (TCL). This structure provides the first line of defence against pathogenic bacteria infecting the mammary gland. The authors used 2DE, Western blotting and fluorescence immunohistochemistry to identify the abundance and location of selected proteins in this structure. The authors identified two dominant clusters of proteins that included proteins in the keratin and S100 families and determined their location in the teat canal keratinocytes and the cornified outermost layer of the teat canal epithelium. Significant between-animal variation in the abundance of the S100 proteins in the TCL was demonstrated. An additional function for some of these proteins (antimicrobial activity) reinforced the role of the structure in the context of udder health.

Proteome characterization has also been focused on specific organelles and cell structures. Peng and co-workers (2008) carried out a proteomic analysis of mammary tissue microsomes from lactating bovine mammary tissue using one-dimensional SDS-PAGE with RPLC-ESI-MS/MS. The authors identified over 700 proteins that were classified according to their subcellular localizations and biological functions, with a particular emphasis on lipid metabolism. Yan and co-workers (2015) investigated the proteome map of plasma membrane fractions of

mammary gland in lactating cows using 1D-Gel-LC-MS/MS. The authors identified almost 900 non-redundant proteins with 16% having an undetermined or unknown function and with 215 specifically associated with the plasma membrane. Molecular roles included binding, transport and catalytic proteins.

The above-mentioned studies reflect the different types of tissues and a multitude of functions that contribute to the highly complex nature of the mammary gland. Research focus seems to be directed more and more at very specific aspects within the organ: cells and structures or even organelles and not to the whole organ. It is likely that in the coming years, other proteomes, within the different structures of the mammary gland, will be characterized. These studies will be the basis of proteomic research on the functionality of the mammary gland in animals subjected to different production or sanitary conditions, such as different feeding strategies or the occurrence of mastitis, to name but a few examples.

3 The Bovine Mammary Gland Proteome: Effects of Genotype and Nutrition

Genotype and nutrition are the two most important factors conditioning the success of a dairy operation. In the majority of the Western countries, bovine milk production is based on one breed, the Holstein-Friesland, which has been selected for high yields and that, accordingly, has to be fed on high-energy rations for the system to be profitable. Nevertheless, systems that are more extensive can also be implemented using natural pastures and lower yield animals or breeds such as the Ayrshire or the Jersey cows. With such a context, it is not surprising that the majority of the bovine mammary gland proteomics literature is focused on the Holstein breed. In fact and to the best of our knowledge, no study has ever been conducted to compare the Holstein mammary gland to that of other bovine breeds. Nevertheless, such comparisons were made for the milk proteome (Murgiano et al. 2009) and for the liver (Timperio et al. 2009) of Holstein versus the Italian beef breed Chianina, in a study that aimed essentially at investigating differences between dairy and beef breeds. More recently similar comparisons have also been made by contrasting the milk proteome of Holstein and Jersey cows (Tacoma et al. 2016) and in the milk proteome of cows with the DGAT1 K232A (diglyceride *O*-acyltransferase 1) genotype (Lu et al. 2015). Given the differences between the milk proteomes of these breeds, it is likely that in the near future these comparisons will be extended to the mammary gland or to specific tissues (e.g. secretory tissue) or organelles. It would be particularly interesting to contrast the mammary gland proteome of different breeds of cattle that have been selected for different milk yields and milk composition.

Only one paper was found on the subject of nutritional effects. Daniels and co-workers (2006) used a proteomic approach to characterize the effects of peripubertal feeding on the heifer mammary (parenchymal tissue) development.

Using an approach based on 2DE, the authors were able to detect 820 different spots of which 131 had differential abundance according to diet (restricted vs. elevated) and 108 were influenced by body weight. This study reveals the importance of nutrition on the protein profiles of the mammary gland in heifers and clearly suggests the pertinence of similar studies in animals subjected to different feeding systems. It would be interesting, for instance, to compare the mammary gland proteome of cows under intensive and pasture-based feeding systems. Such studies, particularly if conducted together with milk proteome studies, would be of interest to the characterization of the different feeding systems and would highlight the effects of intensive, concentrate-based feeding system on the udder milk production machinery. Such types of study would have implications to animal welfare in milk production.

4 The Bovine Mammary Gland Proteome Throughout the Lactation Cycle

A normal healthy mammary gland undergoes significant morphological and physiological changes throughout the lactation cycle affecting the size and functionality of the gland. This is a very complex process that has been extensively reviewed in ruminants (Collier et al. 2012; Lerias et al. 2014). Proteomics has been used to characterize several aspects of the process in dairy cows. In 2012, Rawson and co-workers (2012) used 2D-DIGE to compare the enzyme abundances between the liver and mammary gland of lactating Friesian cows concluding that enzymes catalysing gluconeogenesis and β-oxidation were most prominent in the liver and enzyme abundances in mammary tissue were consistent with fat synthesis rather than β-oxidation. This work was essentially focused on the metabolic changes that these two organs undergo during the lactation period, particularly the energy metabolism.

Janjanam and co-workers (2014) also used 2D-DIGE to characterize protein changes in mammary epithelial cells at different stages of lactation (early, peak and late) comparing low and high yield dairy cows. The authors identified 41 proteins that showed changes in abundance related to lactation stages. Proteins up-regulated during late stage of lactation are associated with stress-induced signalling pathways coinciding with the shutdown of lactopoiesis and the preparation for the dry period. The authors also identified 22 proteins and several signalling pathways (Akt, PI3K and p38/MAPK) that are related to higher milk yields. Examples of the first group of proteins included, for instance, annexin 5, lamin B1 or CALR protein, whereas the second comparison filtered out proteins such as moesin, transaldolase or vimentin. The authors created a rather complex network analysis of the differentially expressed proteins for both comparisons highlighting the complex regulation of lactation and different yields. A similar rationale may be found in the works of Zhang and co-workers (2015) that studied milk proteins throughout the lactation

curve of four cows. The authors identified over 200 differentially regulated proteins. Major affected pathways included lipid synthesis (fatty acid-binding protein, perilipin-2, butyrophilin) that were found to be up-regulated, whereas proteins related to cholesterol transport such as apolipoprotein E were down-regulated. These changes were in accordance with the long-described changes in milk fat yield throughout lactation, indicating the increase of de novo mammary fatty acid synthesis as lactation advances. The authors also found a high abundance of immune-related proteins in early lactation, which indicates a strong role of early milk secretions in the immune system development of calves. These two works reveal the high complementarity of milk and mammary gland proteomics and how the analysis of organ plus secretion can contribute significantly to understanding lactation regulation. Proteomics has also been used to characterize the changes in mammary gland epithelial cells treated with oestrogen, given the importance of this hormone on mammary gland development and function (Huang et al. 2012). The authors treated bovine mammary epithelial cells with oestrogen and used 2DE and MALDI TOF/TOF to determine the effect of this hormone on nuclear phosphorylated proteins. Seven proteins were identified differentially up-regulated in mammary epithelial cells upon 24-h oestrogen exposure. Proteins included glycyl-tRNA synthetase, of importance in milk protein synthesis, as well as other proteins involved in various cellular functions, such as translation initiation factors, GTP-binding nuclear proteins and heat shock proteins, contributing to understanding the molecular mechanisms behind milk synthesis.

The above-mentioned studies made a significant contribution to understand the changes occurring in the mammary gland during the different phases of lactation, revealing the adaptation of the cellular machinery as well as the changes in organ morphology and function that are inherent to lactation. They also reveal how the whole process is related and governed by protein synthesis and the numerous and complex changes that affect the mammary gland proteome. The selected references mentioned above also finally highlight the importance of combining mammary gland proteomic studies with milk proteomics in order to obtain a comprehensive view on the lactation process and how it may be better controlled in a production context.

5 Proteomics and Bovine Mastitis

Considering animal health and welfare and in the context of mammary gland biology, mastitis is the most important disease being caused by the infection of the gland by bacterial or viral pathogens and leads to substantial economic loss to dairy farming (Hogeveen et al. 2011). During the disease process, metabolic changes take place in the mammary gland in order to defend it against the infection resulting in major alteration to the components of the secreted milk. Protein is the major organic component of milk (Singh et al. 2014), and as would be expected significant changes occur in the composition of milk proteins after the onset of

mastitis, continuing until recovery is established. Over the last decade, there has been increasing application of proteomics in animal sciences (Almeida et al. 2015), and the proteins of milk have been examined by these methods in increasing depth. This topic has been the subject of several reviews (Roncada et al. 2012; Abd El-Salam 2014). In relation to the analysis of milk proteome, the modified composition of the total milk protein during mastitis has been one of the most explored areas, not only to study the pathophysiology of infection on this easily accessible biological fluid but also to identify and assess the value of protein biomarkers of mastitis in milk. It could have a major impact on the health of dairy cows if a biomarker for early detection of especially subclinical mastitis could be identified and developed. It has been recognized that—among current technological advances in mastitis detection and monitoring—proteomics has a leading role to play. The analysis of milk protein in response to infection has accelerated over the last decade with the application of increasingly sensitive proteomic methodology and instrumentation (Wheeler et al. 2012; Boehmer 2011; Duarte et al. 2015; Verma and Ambatipudi 2016). Thus advances in proteomic technology have been mirrored by increasingly informative investigations of milk, stretching from early investigations where 2DE and MALDI-TOF were used to separate and identify the high abundance proteins of milk (Galvani et al. 2001). Presently, the changes that take place following mammary gland infection (Hogarth et al. 2004) represent 292 proteins that have abundance changes at particular time points of the process (Mudaliar et al. 2016).

However, if repeatability is taken as the essence of valid results, then a combination of results from publications describing the proteome of milk from quarters with mastitis is a valuable exercise and reveals that 66 proteins have been described in 3 publications or more as having increased or decreased during the disease (Table 1). This includes known changes that occur in the high abundance proteins of milk with reductions in caseins, β-lactoglobulin and α-lactalbumin and increases in albumin, lactoferrin and immunoglobulins. While the alteration due to infection in these well-known milk proteins is widely accepted, it is surprising that there are few studies in which it has been demonstrated by electrophoretic separation. In a recent study (Thomas et al. 2016b), the reduction in concentration of these high abundance proteins of normal, healthy milk has been demonstrated to occur 30 h postinfection in experimentally induced *Streptococcus uberis* mastitis (Fig. 1); subsequently albumin starts to increase at 36 h and lactoferrin at 57 h after infection.

From these publications there have been consistent findings in change in lower abundance proteins with a number being observed to increase in six or more publications (Table 1) and which can be categorized as acute phase protein (fibrinogen, haptoglobin, serum amyloid A, complement C), somatic cell-derived host defence protein (cathelicidin 1, peptidoglycan recognition peptide, calgranulin) and serum or tissue proteins (apolipoprotein, enolase). In contrast, of the low abundance proteins, only fatty acid binding protein has been shown to decrease in six publications. The number of publications that report a change in concentration of a specific protein is a relatively crude means to validate the findings as these

Table 1 Proteins recognized as either increasing or decreasing in three or more proteomic investigations of mastitis

Number	Milk protein	Reference reporting increase in mastitis	Reference reporting decrease in mastitis
1	14-3-3 protein zeta chain, epsilon	2, 10, 16	
2	Actin, cytoplasmic-1	2, 3, 7, 14, 16	
3	Albumin	1, 2, 3, 4, 5, 6, 7, 10, 11	
4	Alpha-1-acid glycoprotein	3, 7, 10, 12, 13, 16	
5	Alpha-1-antitrypsin	3, 5, 6, 7	
6	Alpha-2-glycoprotein 1, zinc-binding	10	6, 9, 14, 16
7	Alpha-2-macroglobulin	5, 6, 7, 10, 16	
8	Alpha-lactalbumin		1, 3, 4, 6, 11, 14, 16
9	Annexin A1	2, 5, 10, 16	
10	Apolipoproteins	2, 3, 4, 5, 6, 7, 10, 16	
11	Bactenecin 5 (cathelicidin-2)	3, 5, 7, 9, 10, 12, 13, 16	3
12	Beta-lactoglobulin		1, 3, 4, 5, 6, 8, 11, 14, 15, 16
13	Butyrophilin subfamily 1 member A1		6, 13, 16
14	Calgranulin B (protein S100-A9)	2, 5, 7, 9, 16	
15	Caseins		1, 3, 11, 13
16	Cathelicidin (general)	3, 5, 6, 9, 13, 10	3, 6
17	Cathelicidin-4	9, 13, 14	
18	Cathelicidin-7	9, 10, 16	
19	Chitinase-3-like protein-1	7, 6, 9, 16	
20	Clusterin	5, 7, 14	
21	Cofilin 1 (nonmuscle)	6, 7, 16	12
22	Complement C3	3, 5, 7, 9, 10, 14	15
23	Complement C4	3, 5, 7, 16	
24	Complement component 3	5, 6, 9, 10	
25	Complement factor B	6, 7, 10, 16	
26	Complement factor H	6, 10, 16	
27	Coronin, actin binding protein, 1A	2, 6, 10, 16	
28	Cyclic dodecapeptides (cathelicidin-1)	2, 3, 4, 5, 7, 9, 10, 13, 14, 15, 16	
29	Endopin 2B	5, 9	
30	Endopin-1	5, 6, 9, 13	
31	Enolase 1 (alpha)	2, 6, 7, 9, 13, 16	
32	Fatty acid-binding protein		3, 6, 13, 14, 15, 16
33	Fetuin (α-2-HS-glycoprotein)	2, 3, 5, 6, 7, 16	
35	Fibrinogen	2, 3, 4, 5, 6, 7, 9, 10, 13, 16	3
36	Gelsolin	6, 7, 16	
37	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	2, 7, 9, 13, 16	

(continued)

Table 1 (continued)

Number	Milk protein	Reference reporting increase in mastitis	Reference reporting decrease in mastitis
38	Glycosylation-dependent cell adhesion molecule-1 (Glycam-1) (lactophorin)	4, 5, 8, 14	6, 16
39	Haptoglobin	5, 6, 7, 9, 10, 13, 16	
40	Heat shock protein	2, 6, 10, 16	
41	Hemopexin	7, 9, 10, 16	
42	Histone	2, 5, 10, 16	
43	Immunoglobulin heavy chain	5, 13, 14, 15	
44	Immunoglobulin lambda light chain variable region	6, 14, 15	13
45	Indolicidin (cathelicidin-4)	3, 5, 7, 10, 16	
46	Inter-alpha-trypsin inhibitor heavy chain (4 and H1)	5, 7, 10, 14, 16	
47	Kininogen 1	5, 7, 10	
48	Lactadherin (milk fat globule EGF factor 8 protein)	5, 7	3, 6, 14
49	Lactoperoxidase	7, 10	6, 15, 16
50	Lactotransferrin (lactoferrin)	4, 5, 6, 9, 10, 11, 14, 16	15
51	Leukocyte elastase inhibitor (LEI) (Serpin B1)	9, 10, 16	
52	L-lactate dehydrogenase	6, 9, 16	
53	Nucleobindin-1	7	6, 14, 16
54	Osteopontin	7, 10, 14	
55	Peptidoglycan recognition protein	2, 4, 5, 6, 7, 9, 10, 16	
56	Peroxiredoxin-5, mitochondrial	9, 10, 16	
57	Plasminogen	5, 6, 7, 16	
58	Platelet glycoprotein 4		10, 13, 16
59	Polymeric immunoglobulin receptor		3, 6, 15, 16
60	Profilin-1	6, 7, 16	
61	Protein S100-A12 (calcium-binding protein in amniotic fluid 1) (CAAF1)	2, 3, 9, 13, 16	
62	Protein S100-A8 (calgranulin A)	7, 6, 9, 10, 13, 14, 16	
63	Serotransferrin	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 14, 16	
64	Serpin A3-1	10, 13, 16	
65	Serum amyloid A 3	5, 6, 7, 10, 12, 13 16	
66	Xanthine dehydrogenase/oxidase	7	6, 14, 16

(1) Hogarth et al. (2004), (2) Smolenski et al. (2007), (3) Boehmer et al. (2008), (4) Boehmer et al. (2010b), (5) Danielsen et al. (2010), (6) Ibeagha-Awemu et al. (2010), (7) Boehmer et al. (2010a), (8) Kim et al. (2011), (9) Alonso-Fauste et al. (2012), (10) Reinhardt et al. (2013), (11) Li et al. (2014), (12) Yang et al. (2012), (13) Zhao et al. (2015), (14) Zhang et al. (2015), (15) Pongthaisong et al. (2016), (16) Mudaliar et al. (2016)

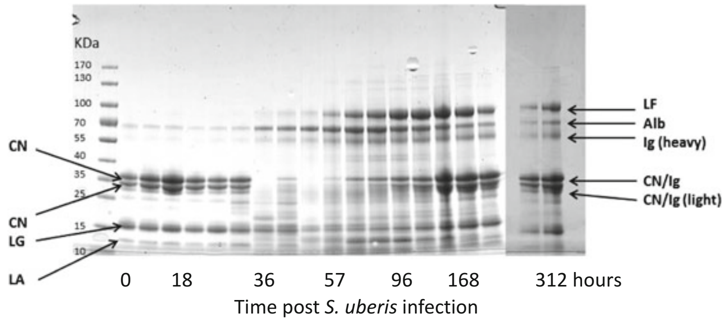

Fig. 1 One-dimensional gel showing high abundance proteins from a mammary quarter challenged with *Streptococcus uberis* (from left to right: size marker with band size in kDa, 0, 6, 12, 18, 24, 30, 36, 42, 48, 57, 72, 81, 96, 120, 144, 68, 192, 240 and 312 h post challenge. *LF* lactoferrin, *Alb* albumin, *Ig* Immunoglobulin, *CN* casein, *LG* β-lactoglobulin, *LA* α-lactalbumin. Reproduced from Thomas et al. (2016) Molecular Biosystems 12 2735–2747 with permission from the Royal Society of Chemistry

publications cover considerable differences in experimental technique and sensitivity. Still, it does provide substantial support to the identification of molecules of importance in the host response to the mastitis and also in identification of putative biomarkers of the disease. Indeed these findings would validate empirically designed investigation of mastitis biomarkers that have proposed acute phase proteins such as haptoglobin (Thomas et al. 2015) or host defence proteins such as cathelicidins (Addis et al. 2016) as markers of the disease and could be valuable information for multiplex assay development with many target analytes, as has been proposed (Bislev et al. 2012b).

In the early proteomic investigations of mastitis, the identification of protein change was the aim, but as experience has been gained, the rationale for undertaking studies with this advanced analytical approach has become more defined. In the development of the proteomic analysis of milk from mastitis, electrophoresis on 2D gels has been used to characterize the increased diversity of proteins in milk in natural infections causing mastitis with, for example, 62 (Alonso-Fauste et al. 2012) or 95 distinct gene products being identified (Smolenski et al. 2007). These methods were able to show that significant proteins which increased in the presence of mastitis are lactoferrin, cathelicidins and the acute phase proteins such as haptoglobin and serum amyloid A. Further investigations using 2DE and MS/MS have examined proteome changes in cows with mastitis during the transition period in which haptoglobin, SAA and also α_1 acid glycoprotein (AGP) were again recognized as being increased (Yang et al. 2012). Furthermore, in naturally acquired subclinical mastitis caused by *Streptococcus agalactiae*, cathelicidin-1 was highly correlated to somatic cell count (SCC), the conventional test used to detect mastitis (Pongthaisong et al. 2016) revealing an interesting biomarker putative role. However, increasingly non-gel-based proteomics has been used for the study of the milk

proteomes in mastitis, and a number of reports have focused on characterization of the response to mastitis caused by specific bacterial infection.

Gram-positive bacteria such as *Staphylococcus aureus* and *Streptococcus uberis* are major causative pathogens of mastitis. Proteome investigations have focused on determining the effects of these bacteria on host reactions. An extended study of the proteins in subfractions of milk from cows experimentally infected with *S. aureus* has examined the milk exosomes, the milk fat globule membranes and the whey proteins identifying 300 proteins related to host defence, 94 of which were significantly altered in response to mastitis (Reinhardt et al. 2013). Of interest was that proteins related to neutrophil extracellular traps were particularly high in the milk fat globule membrane, demonstrating that examination of milk subfractions can extend the potential for unlocking the pathophysiology of an infected udder. Investigation of natural cases of mastitis caused by *S. aureus* infection has shown that not only milk is affected but that the mammary tissue also undergoes extensive change in response to pathogenic bacteria. A study using iTRAQ quantitative proteomics (Huang et al. 2014) has shown that 36 tissue proteins were up-regulated in tissue from infected cows, including fibrinogen, collagen, proteins S100-A12 (calgranulin C) and inter-alpha trypsin inhibitor H4. In contrast, 19 proteins were down-regulated including casein, apolipoprotein A4 and butyrophilin with a conclusion that such changes related to the tissue damage and subsequent repair being undertaken by the mammary tissue in response to infection.

Lipoteichoic acid is a component of the cell wall of Gram-positive bacteria and is known to stimulate the host reaction to infections. Investigation of the response of the proteome of milk from udder quarters individually inoculated with lipoteichoic acid extracted from *S. aureus* has demonstrated that this bacterial factor does indeed stimulate major host responses (Larsen et al. 2010). The focus of this investigation was on the changes occurring in the smaller proteins and peptides in milk and revealed that the lipoteichoic acid treatment induced significant proteolysis of normal milk proteins. Twenty different peptides were found and on identification were demonstrated to be derived from proteolytic breakdown of α_{s1}- and β-casein. It was also determined that plasmin, cathepsins, elastase and amino/carboxylpeptidases were responsible for the degradation of the normal milk proteins.

Proteome investigation of the responses to *S. uberis*, initially by 2DE and LC-MS, have revealed 68 host defence-associated proteins having antimicrobial, pathogen-recognition and immune or inflammatory regulatory functions (Smolenski et al. 2014).

Recently, experimental infection with *S. uberis* has been used as the model for an extensive investigation of the response of the milk proteome in mastitis and has been integrated into an extended omics investigation including peptidomics, quantitative proteomics and metabolomics, in an attempt to define a mastitomic (mastitis-omics) overview with a systems biology approach. Investigation of the changes in the peptidome of milk from udder quarters over the 13 days following inoculation with *S. uberis* used ultrafiltration to remove proteins of greater than 25 kDa but omitted a trypsinization step. This allowed capillary electrophoresis

linked to MS to identify a biomarker panel of 77 peptides that, combined together, could be used in the detection, diagnosis and monitoring of mastitis (Thomas et al. 2016b). This extended a previous investigation of the peptides present in milk from natural cases of mastitis where a similar approach using a peptide panel was not only able to detect mastitis but showed potential in the differentiation of Gram-negative from Gram-positive bacteria as being the cause of the infection (Mansor et al. 2013). The mastitomic investigation included a label-free quantitative proteomic approach for the analysis of change in the proteome after the *S. uberis* infection (Mudaliar et al. 2016). While a peak in bacteriology was seen at 12 h postinfection, the major changes in the proteome were evident at 57 and 81 h post challenge. Peptidoglycan recognition protein and cathelicidins were among the proteins to be increased in abundance by the highest fold expression at the earliest time points investigated (57 h post challenge), but at later stages (81 h) haptoglobin had the highest fold increase in regulation. Ingenuity pathway analysis of the data from the study revealed that the acute phase protein pathways were the most up-regulated during the infection. The mastitomics investigation also included metabolomic analysis of the same milk samples (Thomas et al. 2016a) and revealed significant change in carbohydrate, lipid and nucleic acid pathways. Increases in free amino acids, di-, tri- and oligopeptides, from milk protein degradation were also found by metabolomics with an unexpected increase in pathways related to bile acid metabolism. The combination of omics technologies in a systems biology approach to investigating mastitis shows great promise to provide the greatest insight into the multitude of host responses to the infection of the mammary gland.

Mammary tissue from dairy cows infected with *S. uberis* has been used in development of a targeted quantitative proteomic approach for multiplexed assay of proteins for mastitis investigation and diagnosis (Bislev et al. 2012b). Standard QconCAT peptides of 20 proteins related to the bovine mammary gland's response to *S. uberis* infection were generated and used as the basis for single reaction monitoring for assessment of the concentration of the proteins in a single sample. In the final delivery of the system, QconCAT peptides for 12 of the proteins provided sufficient sensitivity to be used to estimate changes in tissue samples. The target proteins included acute phase proteins such as haptoglobin and serum amyloid A and immune-derived host factors such as interleukin-6 and calgranulin B and C.

Mastitis caused by Gram-negative bacteria, especially *Escherichia coli*, has also been investigated with proteomic approaches. Experimental infusion of udder quarters with lipopolysaccharide (LPS) endotoxin derived from *E. coli* cell walls showed that after 12 h low abundance proteins were detectable including serotransferrin, protein S100-A12 (calgranulin C) and cathelicidins (Hinz et al. 2012). This investigation also assessed the change in peptides produced by the lipopolysaccharide infusion revealing that peptides derived from α_s- and β-casein were present in increased amounts due to the action of plasmin and cathepsin D. The effects of *E. coli* on mammary tissue has been assessed using a 2DE and label-free approaches (Zhao et al. 2015) and revealed 56 tissue proteins affected by the natural *E. coli* infection of the gland. They were related to the stress,

transport and localization pathways with vimentin and α-enolase being central to the functional hubs in the network. A more specialized investigation utilizing glycoproteomics (Yang et al. 2014) has examined changes in the *N*-glycosylated mammary gland protein, revealing that differences were apparent in *E. coli* mastitis and in particular for the lysosome and *O*-glycan biosynthesis pathways. A further alternative use of proteomics in relation to mastitis biology has been the investigation of the proteome of different *E. coli* strains with an emphasis on comparing strains causing persistent and transient mastitis (Lippolis et al. 2014). An iTRAQ shotgun analysis found that 28 bacterial proteins were associated with expression changes that correlated to disease with bacterial motility being a key phenotype in determining the extent of persistence or transience of the infection.

Although the milk and mammary tissue proteomes have been assessed in their response to *S. aureus*, *S. uberis* and *E. coli*, there are few comparative data between host responses to the different infections. However, a recent investigation did compare treatment for mastitis with silver ion as an antibacterial agent against both *E. coli* and *S. aureus* infection, but did not include extensive analysis of the difference in host response to the bacteria (Kang et al. 2016). The investigations of the differing pathogens have been made by different groups with differing methodologies and instrumentation. It would be of great value if proteomic and peptidomic investigations could be made to investigate host responses to different pathogens with a consistent protocol to allow direct comparison, not only to understand differences in pathophysiology but also to determine if protein biomarker assays could have the potential to discriminate between the causative agents. This could have tremendous therapeutic value such that antibacterial treatment could be targeted more precisely. This would reduce excessive waste of antibiotics and contribute to reducing antimicrobial resistance.

The host defence proteome of milk from a bulk tank sample has also been compared to that of human milk with high concentrations of cathelicidins and lactoperoxidase found in the pooled samples from dairy cows (Hettinga et al. 2011). The level of somatic cell count (SCC) in milk is indicative of the presence and extent of mastitis. A study of the milk proteome in milk samples from dairy cows with elevated SCC, using a dimethyl labelling and LC-MS/MS approach, identified increases in immune-related and other biological functions such as protease inhibition (Zhang et al. 2015).

Mastitis may not be the only disease of dairy cows that can be assessed by analysis of the milk proteome as it has been investigated to assess whether putative biomarkers of disease resistance can be identified and utilized by examination of the proteins in milk (van Altena et al. 2016). An investigation using NanoLC-MS/MS identified 13 proteins which were more abundant in low-disease-resistant cows rather than in high-disease-resistant cows. Of these, lactoferrin was identified as the single most valuable biomarker in milk for determination of low-disease-resistant cows such that dairy cows with higher levels of lactoferrin had a higher risk of being culled in the following 12 months than cows with low lactoferrin.

Peptidomics, the analysis of low molecular weight proteins and peptides, has not only been investigated for the increases found in peptide content of milk following

mastitis (Mansor et al. 2013; Thomas et al. 2016b) but has been used to characterize the peptides present in healthy bovine milk. However, a similar conclusion was made that, at peak lactation when peptides were detected, they were identified as the degradation products of β-casein, α_{s1}-casein and α_{s2}-casein caused by plasmin, cathepsin and elastase activity (Dallas et al. 2014). In contrast, no peptides from the breakdown of β-lactoglobulin, lactoferrin or secretory IgA were found. It appears that even in milk from healthy dairy cows there is ongoing breakdown of caseins which is greatly increased during mastitis, probably due to the influx of endogenous proteases released from the neutrophils among the somatic cells. Examination of the milk peptidome is of increasing importance in the dairy industry for assessing the quality of milk and during its processing. It is also of special value in current efforts to identify bioactive peptides which may have health benefits to consumers as well as to the calves of the dairy cows on farm (Sanchez-Rivera et al. 2014).

Overall, the investigation of milk and the mammary gland in response to mastitis using proteomic approaches has gathered pace over the last decade and has kept pace with the technological innovations made in instrumentation. It is likely that this will accelerate as the potential for detailed analysis of the milk phenotype by the use of proteomics gains further appreciation. As well as providing targets for development of disease biomarkers, the real value of proteome analysis will come when the technology is integrated with a full systems biology approach. This will answer questions not only of importance in animal husbandry and dairy production but also in understanding the host's multifarious responses and identifying where therapeutic, immunological, biologic or nutritional interventions will have the most effective and efficient mechanism of restoring full health and welfare to infected dairy cows.

6 Proteomics and the Mammary Gland in Other Dairy Species

In animal science mammary gland proteomics, the majority of the works conducted so far has focused on the bovine species because of its importance and economic value in worldwide dairy production. In fact, mammary gland proteomic studies in other species are extremely limited, and frequently, results obtained in cattle may be extrapolated to other dairy species. However, huge differences exist in milk composition and proteome (Hernandez-Castellano et al. 2016a; Yang et al. 2016), mammary gland structure and lactation lengths (Lerias et al. 2014; Suarez and Alvarez-Morujo Suarez 1982) of these species, and extrapolations and generalizations should be considered very carefully.

In the domestic water buffalo (*Bubalus bubalis*) mammary gland, two works may be found, both by the same group based in India. The geographic origin of this work is not surprising, given that in that country a large proportion of milk

produced originates from this species. In 2015, Jena et al. (2015) compared the proteomes of lactating buffalo females with that of heifer's using 2D-DIGE and mass spectrometry. The authors identified over 20 proteins with increased abundance in the lactating buffalo mammary gland, while 8 were up-regulated in the heifer's mammary gland, out of a total of 41 with differential abundance/regulation. Bioinformatic analyses showed that the majority of the proteins are involved in different metabolic processes of varied importance in the context of buffalo's lactation. This very interesting work showed for the first time the existing differences between the mammary gland of lactating and pre-lactating buffalo heifers. To the best of our knowledge, no similar work has been conducted for cattle. However, it would be of the utmost interest to conduct a parallel study in *Bos taurus* or *Bos indicus* cattle to ascertain the differences between species. As a follow-up study, the same group (Anand et al. 2016) studied aspects related to MGP-40, a chitinase-like protein which is overexpressed during mammary gland involution, trying to unravel its physiological function in the mammary gland. The authors used a cell line designed to overexpress this gene and have determined that the gene enhanced the proliferation of buffalo mammary epithelial cell and protected the cells from apoptosis. Furthermore, it reduced dome formation, acinar polarization and casein synthesis in buffalo mammary epithelial cells in the presence of lactogenic hormones. This work contributed significantly to the understanding of the apoptosis mechanisms and mammary gland drying process in water buffalo and again it would be very interesting to contrast it with similar studies in the bovine mammary gland. Although a totally different approach (transcriptomics based), it would be interesting to mention the work of Kapila et al. (2016) on the impact of heat stress on the adaptation of mammary epithelial cells. Overall, heat stress resulted in a decrease in cell viability and proliferation inducing cellular apoptosis and necrosis. The transcriptomic profile of heat-stressed cells led to 153 genes found to be up-regulated, while 8 genes were down-regulated as a consequence of heat stress. The authors finally identified several genes from different functional classes and biological pathways that could be termed as heat responsive in buffalo mammary gland epithelial cells. It would be interesting to contrast these results with a similar or parallel proteomic approach integrating proteomics and transcriptomics results. Finally, and on a different bovid species, it is interesting to mention the works of Li et al. (2014) that compared milk whey from yak (*Bos grunniens*) with clinical mastitis to whey from healthy animals. The authors identified 13 proteins with differential concentration; several of them are of blood origin such as lactoferrin that were related to the mastitic process in addition to regular whey proteins that had different levels according to the presence or absence of mastitis.

In small ruminants (sheep and goat), proteomic studies on the mammary gland are very rare. In sheep (*Ovis aries*), mammary gland omics research seems to be rather focused on transcriptomic research with the majority of the works originating from France and Italy. In fact, several works may be found in the literature including ovine and bovine comparisons (Singh et al. 2013), milking ability in dairy sheep (Dhorne-Pollet et al. 2012) and dairy ewe mammary gland

transcriptome comparison according to breed (Churra vs. Assaf) (Suarez-Vega et al. 2015, 2016). These studies include interesting aspects of the ovine lactation and, using the same samples and experimental design, could easily be replicated in a proteomic approach and further combined to a thorough integrated systems biology analysis.

Finally and focusing on the ovine mammary gland epithelial cells, it is worth mentioning the work by Addis and co-workers (Addis et al. 2013). The authors investigated the lactating sheep mammary gland response to infection with the mastitis-associated *Streptococcus uberis*. The proteomic investigation was focused on the milk fat globule but seen from a mammary gland perspective. The authors determined that several proteins involved in inflammation, chemotaxis of immune cells and antimicrobial defence were overexpressed in animals subjected to mastitis and finally contributed to the development of more sensitive tools for monitoring mastitis in sheep. In goats a study of the same nature has also been conducted (Olumee-Shabon et al. 2013) with goats being challenged with a mastitis-causing lipopolysaccharide (LPS). The authors analysed the milk at several collection points upon the administration of LPS with several proteins related to inflammatory response being up-regulated. In other studies, the ovine milk from ewes with *Staphylococcus chromogenes* mastitis has been examined by 2DE and MS (Chiaradia et al. 2013), and it has been shown that there were quantitative milk protein changes between non-infected samples and those from ewe udder with high levels of somatic cell count. Using 2DE and GeLC-MS/MS, the milk fat globule of ewes naturally infected with *Mycoplasma agalactiae* has been shown to contain proteins which increase during the infection, causing mastitis (Addis et al. 2011). Classes of proteins shown to increase include inflammation and host defence, heat shock proteins and those related to oxidative stress, while proteins related to normal milk fat globule function and metabolism were decreased.

The goat (*Capra hircus*) mammary gland proteome has been extensively studied in two goat breeds from the Canary Islands (Spain) with different levels of tolerance to seasonal weight loss (*Palmera*, susceptible, and *Majorera*, tolerant), the most important limitation to animal production in the tropics and the Mediterranean. The two breeds have been subjected to experimentally induced weight loss/feed restriction and their weight changes and milk production (Lerias et al. 2013) and blood biochemical profiles (Lerias et al. 2015) determined. Both breeds showed similar levels of weight loss and milk production as a consequence of the restricted diet and at the same time differences in blood creatinine, urea, non-esterified fatty acids (NEFAs), cholesterol, IGF-1 and T3 as a result of feed restriction finally highlighting different breed responses to undernutrition. The authors subsequently conducted a whole proteomic analysis of the mammary gland secretory tissue (Hernandez-Castellano et al. 2016b), a mitochondrial proteome analysis (Cugno et al. 2016) and an NMR-based metabolomic analysis (Palma et al. 2016). The secretory tissue label-free proteomic analysis led to the identification of over 1000 proteins, a first characterization of the caprine mammary gland proteome. A total of 96 proteins showed differential regulation between two of the groups within the studied comparisons, with the Majorera (tolerant) breed showing higher abundance

of immune system-related proteins and the Palmera (susceptible) breed displaying higher concentration of proteins related to apoptosis. This analysis confirmed once more that the two goat breeds have a distinct metabolism reaction to seasonal weight loss and the authors suggested several proteins of the immune system and apoptosis pathway (cadherin-13, collagen alpha-1, nidogen-2, clusterin and protein s100-A8) as possible markers of tolerance to seasonal weight loss. The mammary gland mitochondrial proteome was analysed using standard Coomassie-stained 2DE gels and also Blue-Native-PAGE (BN) gels. The proteomic analysis of the mitochondria enabled the resolution of a total of 277 proteins, and 148 (53%) were identified by MALDI-TOF/TOF mass spectrometry. Some of the proteins were identified as subunits of the glutamate dehydrogenase complex and the respiratory complexes I, II, IV and V from mitochondria as well as numerous other proteins with functions in metabolism, development, localization, cellular organization and biogenesis, biological regulation and response to stimulus, among others, which were mapped in both BN and 2DE gels. The comparative proteomic analysis allowed the identification of several proteins: NADH-ubiquinone oxidoreductase 75 kDa subunit and lamin B1 mitochondrial (up-regulated in the *Palmera* breed), guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2 (up-regulated in the *Majorera* breed), mitochondrial cytochrome b-c1 complex subunit 1 and chain D, bovine F1-C8 sub-complex of ATP synthase (down-regulated in the *Majorera* breed) related to the specific differential response to seasonal weight loss of the two breeds. In the metabolomic analysis, NMR was used to compare the metabolome of an aqueous fraction of the mammary gland and milk serum from both breeds. $^{(1)}$H NMR spectra were collected from the aqueous extract of the mammary gland biopsies and milk serum. Profiling analysis led to the identification of 46 metabolites in the aqueous extract of the mammary gland, the most extensive mammary gland metabolome characterization established so far, with lactose, glutamate, glycine and lactate being the most abundant. The analysis of milk serum allowed the identification of 50 metabolites, the most abundant being lactose, citrate and creatine. Significant differences were observed, in mammary gland biopsies and milk serum, between control and restricted-fed groups in both breeds, albeit with no differences between the breeds. The authors linked the variations to metabolism adaptation to the low-energy diet and regarded them indicative of breed-specific microflora. Finally and also very importantly, the authors found that milk serum showed more metabolites varying between control and restricted groups than the mammary gland. These studies, which include also mammary gland fatty acid and transcriptomics profiling (Almeida, unpublished data), finally highlight the importance of conducting multi-omics integrated studies to fully enlighten complex molecular mechanisms such as those underlying differential breed response to seasonal weight loss.

While bovine mastitis has been the subject of the vast majority of proteomic investigation of milk and mammary responses to mastitis, there have been investigations into other species where mastitis causes animal health and welfare issues.

The above-mentioned examples demonstrate how the animal science proteomics field is expanding in species other than bovine. In fact, the relatively low but

increasing number of mammary-gland-related proteomic studies conducted in buffalo, yak, sheep or goat seems to demonstrate that there is a growing interest in the field and the possibilities allowed by the technology, either through a generalized use of the technology or together with other omics such as transcriptomics and metabolomics. Given the less-universal nature of dairy production and the relatively restricted areas where dairy production of these species has an important economic interest, it is likely that in the future research on the proteome of the mammary gland in buffalo and small ruminants will continue depending on the research budgets and the specific problems these countries or regions wish to address. Questions and research topics are therefore likely to address issues such as certification and animal welfare.

7 Conclusions

It is clear from the extent of current investigations that there is currently a great potential for further uses of proteomic approaches in the study of mammary gland biology in health and disease. As far as the investigations into the proteome of the mammary gland go, it is clear that the examination of specific compartments of the gland should be undertaken; thus the teat canals and the cistern should be characterized in both health and diseased tissue and also under different production conditions. Advances in proteomics where specialized technology is developing, for example, to examine the glycoproteins, the phosphoproteins and the lipoprotein subfractions, would yield further knowledge on the physiology and pathophysiology of the mammary gland and lactation. The potential for molecular imaging, a proteomic method combining histological examination of tissue sections with mass spectrometry could have an especially useful role in the future.

There is a timely move in genomics investigation to associate gene changes to the resultant phenotype, and proteomics along with metabolomics are key technologies that will enable this to proceed. In fact, investigations into developing the nutrition of dairy cows leading to high-quality milk are becoming a critical area, where assessing of the mammary gland physiology using a proteomic approach is starting to unveil the potential of the technology. The effects of contrasting diets on mammary gland homeostasis and protein production as estimated by quantitative proteomics will have substantial benefit to the understanding of diet-related adaptations. This could be used to assess the effects of intensive, concentrate-based feeding systems compared to pasture rearing on the udder milk production machinery. Such types of study would have implications to animal health and welfare.

Proteomics has a major part to play in fully understanding the lactation cycle in dairy cows and of special interest will be the quantitative proteomic analysis at critical times such as the initial stimulation with production of colostrum and the drying off period where milk production is halted. Investigations should be made into the mechanism of adaptation of the cellular machinery and the changes in organ morphology and function in these key periods.

The extensive results that have now been generated by proteomic investigations of mastitis have provided the basis for biomarker development that are likely to have significant impact on the diagnosis of the disease and especially for the difficult to detect subclinical form of the disease. In addition, the investigations have enabled a theory on the mechanism of the alteration of protein in milk to be developed. Thus the increases seen in proteins in milk during mastitis can derive from different sources as illustrated in Fig. 2. Proteins such as albumin or IgG may transfer from serum through a more permeable blood-mammary barrier, protein such as mammary-associated serum amyloid A may also be synthesized in the mammary epithelium in response to cytokine released in response to bacterial infection, or proteins such as cathelicidins may be released from neutrophil derived somatic cells in the milk. Interestingly, haptoglobin and possibly other proteins may be derived from any or all of these sources and clarification of the proportion of the total haptoglobin from each route may be revealed, for instance, by glycoproteomic analysis. The mechanism for loss of the high abundance proteins in healthy milk is, for β-lactoglobulin and α-lactalbumin, located at a reduction in synthesis rate level, while for caseins the action of proteases also contributes to the reduction of this protein in milk from quarters with mastitis and the appearance of peptides as degradation products of casein breakdown.

Although the vast majority of investigations of livestock mammary gland have been in dairy cows, due to their importance in providing edible and nutritious protein for the increasing human population, the use of proteomics in examination of mammary biology in other farm species has started and should gather pace as the technology becomes more accessible. The studies conducted on buffalo, yak, sheep or goat mammary or milk proteome demonstrate that there is great potential for the growing interest in the application of this and other omics technologies. Indeed for

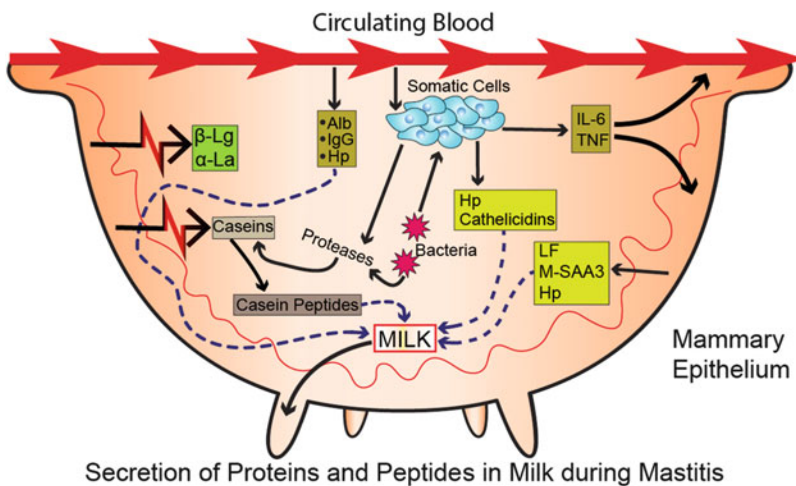

Fig. 2 A diagram representing the change in milk protein and peptides following bacterial infection of an udder quarter of the mammary gland

the study on host responses to the bacterial causes of mastitis, investigators could use small ruminants such as sheep or goats instead of dairy cows as experimental models of the disease in order to reduce resource requirements. Indeed the use of such experimental models has been a neglected tool in investigation of the host responses to bacterial infection under in vivo conditions. The mammary gland is an ideal incubator for bacterial growth providing nutrient and temperature control, but the growth is largely curtailed by the innate and acquired immune systems. No other in vivo system can provide the volume of sample that can be obtained by the milking regime allowing multiple investigations to be undertaken, while no in vitro system can provide the native host responses which can be allied to therapeutic interventions. Proteomic investigation of the mammary gland and milk has now been established, and it awaits exploitation by future researchers in mammary biology, animal science, biomarker development and host responses to bacterial infection.

Acknowledgements Author AM Almeida acknowledges funding from *Fundação para a Ciência e Tecnologia* (FCT, Lisbon, Portugal) through the research project PTDC/CVT/116499/2010—Lactation and milk production in Goat (*Capra hircus*): identifying molecular markers underlying adaptation to seasonal weight loss. The authors are members of COST actions FA1002—Proteomics in Farm Animals and FA1308—Dairycare to whom network funding is acknowledged.

References

Abd El-Salam MH (2014) Application of proteomics to the areas of milk production, processing and quality control – a review. Int J Dairy Technol 67(2):153–166. https://doi.org/10.1111/1471-0307.12116

Addis MF, Pisanu S, Ghisaura S, Pagnozzi D, Marogna G, Tanca A, Biosa G, Cacciotto C, Alberti A, Pittau M, Roggio T, Uzzau S (2011) Proteomics and pathway analyses of the milk fat globule in sheep naturally infected by Mycoplasma agalactiae provide indications of the in vivo response of the mammary epithelium to bacterial infection. Infect Immun 79 (9):3833–3845. https://doi.org/10.1128/iai.00040-11

Addis MF, Pisanu S, Marogna G, Cubeddu T, Pagnozzi D, Cacciotto C, Campesi F, Schianchi G, Rocca S, Uzzau S (2013) Production and release of antimicrobial and immune defense proteins by mammary epithelial cells following Streptococcus uberis infection of sheep. Infect Immun 81(9):3182–3197. https://doi.org/10.1128/iai.00291-13

Addis MF, Tedde V, Puggioni GMG, Pisanu S, Casula A, Locatelli C, Rota N, Bronzo V, Moroni P, Uzzau S (2016) Evaluation of milk cathelicidin for detection of bovine mastitis. J Dairy Sci 99(10):8250–8258. https://doi.org/10.3168/jds.2016-11407

Almeida A, Bassols A, Bendixen E, Bhide M, Ceciliani F, Cristobal S, Eckersall P, Hollung K, Lisacek F, Mazzucchelli G, McLaughlin M, Miller I, Nally J, Plowman J, Renaut J, Rodrigues P, Roncada P, Staric J, Turk R (2015) Animal board invited review: advances in proteomics for animal and food sciences. Animal 9(1):1–17

Alonso-Fauste I, Andres M, Iturralde M, Lampreave F, Gallart J, Alava MA (2012) Proteomic characterization by 2-DE in bovine serum and whey from healthy and mastitis affected farm animals. J Proteome 75(10):3015–3030

Anand V, Jaswal S, Singh S, Kumar S, Jena MK, Verma AK, Yadav ML, Janjanam J, Lotfan M, Malakar D, Dang AK, Mohanty TK, Kaushik JK, Mohanty AK (2016) Functional characterization of mammary gland protein-40, a chitinase-like glycoprotein expressed during mammary gland apoptosis. Apoptosis 21(2):209–224. https://doi.org/10.1007/s10495-015-1196-z

Bislev SL, Deutsch EW, Sun Z, Farrah T, Aebersold R, Moritz RL, Bendixen E, Codrea MC (2012a) A bovine eptideAtlas of milk and mammary gland proteomes. Proteomics 12 (18):2895–2899

Bislev SL, Kusebauch U, Codrea MC, Beynon RJ, Harman VM, Rontved CM, Aebersold R, Moritz RL, Bendixen E (2012b) Quantotypic properties of QconCAT peptides targeting bovine host response to Streptococcus uberis. J Proteome Res 11(3):1832–1843. https://doi.org/10.1021/pr201064g

Boehmer JL (2011) Proteomic analyses of host and pathogen responses during bovine mastitis. J Mammary Gland Biol Neoplasia 16(4):323–338

Boehmer JL, Bannerman DD, Shefcheck K, Ward JL (2008) Proteomic analysis of differentially expressed proteins in bovine milk during experimentally induced Escherichia coli mastitis. J Dairy Sci 91(11):4206–4218. https://doi.org/10.3168/jds.2008-1297

Boehmer JL, DeGrasse JA, McFarland MA, Tall EA, Shefcheck KJ, Ward JL, Bannerman DD (2010a) The proteomic advantage: label-free quantification of proteins expressed in bovine milk during experimentally induced coliform mastitis. Vet Immunol Immunopathol 138(4):252–266. https://doi.org/10.1016/j.vetimm.2010.10.004

Boehmer JL, Ward JL, Peters RR, Shefcheck KJ, McFarland MA, Bannerman DD (2010b) Proteomic analysis of the temporal expression of bovine milk proteins during coliform mastitis and label-free relative quantification. J Dairy Sci 93(2):593–603. https://doi.org/10.3168/jds.2009-2526

Chiaradia E, Valiani A, Tartaglia M, Scoppetta F, Renzone G, Arena S, Avellini L, Benda S, Gaiti A, Scaloni A (2013) Ovine subclinical mastitis: proteomic analysis of whey and milk fat globules unveils putative diagnostic biomarkers in milk. J Proteome 83:144–159. https://doi.org/10.1016/j.jprot.2013.03.017

Collier RJ, Annen-Dawson EL, Pezeshki A (2012) Effects of continuous lactation and short dry periods on mammary function and animal health. Animal 6(3):403–414. https://doi.org/10.1017/s1751731111002461

Cugno G, Parreira JR, Ferlizza E, Hernandez-Castellano LE, Carneiro M, Renaut J, Castro N, Arguello A, Capote J, Campos AMO, Almeida AM (2016) The goat (Capra hircus) mammary gland mitochondrial proteome: a study on the effect of weight loss using blue-native PAGE and two-dimensional gel electrophoresis. PLoS One 11(3). https://doi.org/10.1371/journal.pone.0151599

Dallas DC, Guerrero A, Parker EA, Garay LA, Bhandari A, Lebrilla CB, Barile D, German JB (2014) Peptidomic profile of milk of Holstein cows at peak lactation. J Agric Food Chem 62 (1):58–65. https://doi.org/10.1021/jf4040964

Daniels KM, Webb KE, McGilliard ML, Meyer MJ, Van Amburgh ME, Akers RM (2006) Effects of body weight and nutrition on mammary protein expression profiles in Holstein heifers. J Dairy Sci 89(11):4276–4288

Danielsen M, Codrea MC, Ingvartsen KL, Friggens NC, Bendixen E, Rontved CM (2010) Quantitative milk proteomics—host responses to lipopolysaccharide-mediated inflammation of bovine mammary gland. Proteomics 10(12):2240–2249. https://doi.org/10.1002/pmic.200900771

Dhorne-Pollet S, Robert-Granie C, Aurel MR, Marie-Etancelin C (2012) A functional genomic approach to the study of the milking ability in dairy sheep. Anim Genet 43(2):199–209. https://doi.org/10.1111/j.1365-2052.2011.02237.x

Duarte CM, Freitas PP, Bexiga R (2015) Technological advances in bovine mastitis diagnosis: an overview. J Vet Diagn Investig 27(6):665–672. https://doi.org/10.1177/1040638715603087

Eckersall PD, de Almeida AM, Miller I (2012) Proteomics, a new tool for farm animal science. J Proteome 75(14):4187–4189

Frandson RD, Wilke WL, Falls AD (2009) Anatomy and physiology of farm animals. Wiley-Blackwell, London

Galvani M, Hamdan M, Righetti PG (2001) Two-dimensional gel electrophoresis/matrix-assisted laser desorption/ionisation mass spectrometry of commercial bovine milk. Rapid Commun Mass Spectrom 15:258–264

Hernandez-Castellano LE, Almeida AM, Renaut J, Arguello A, Castro N (2016a) A proteomics study of colostrum and milk from the two major small ruminant dairy breeds from the Canary Islands: a bovine milk comparison perspective. J Dairy Res 83(3):366–374. https://doi.org/10.1017/s0022029916000273

Hernandez-Castellano LE, Ferreira AM, Nanni P, Grossmann J, Arguello A, Capote J, Cai GH, Lippolis J, Castro N, de Almeida AM (2016b) The goat (Capra hircus) mammary gland secretory tissue proteome as influenced by weight loss: a study using label free proteomics. J Proteome 145:60–69. https://doi.org/10.1016/j.jprot.2016.03.030

Hettinga K, van Valenberg H, de Vries S, Boeren S, van Hooijdonk T, van Arendonk J, Vervoort J (2011) The host defense proteome of human and bovine milk. PLoS One 6(4):e19433. https://doi.org/10.1371/journal.pone.0019433

Hinz K, Larsen LB, Wellnitz O, Bruckmaier RM, Kelly AL (2012) Proteolytic and proteomic changes in milk at quarter level following infusion with Escherichia coli lipopolysaccharide. J Dairy Sci 95(4):1655–1666. https://doi.org/10.3168/jds.2011-4813

Hogarth CJ, Fitzpatrick JL, Nolan A, Young FJ, Pitt A, Eckersall PD (2004) Differential protein composition of bovine whey: a comparison of whey from healthy animals and from those with clinical mastitis. Proteomics 4:2094–2100

Hogeveen H, Huijps K, Lam TJGM (2011) Economic aspects of mastitis: new developments. N Z Vet J 59(1):16–23. https://doi.org/10.1080/00480169.2011.547165

Huang JG, Gao XJ, Li QZ, LM L, Liu R, Luo CC, Wang JL, Bin Q, Jin X (2012) Proteomic analysis of the nuclear phosphorylated proteins in dairy cow mammary epithelial cells treated with estrogen. In Vitro Cell Dev Biol Anim 48(7):449–457. https://doi.org/10.1007/s11626-012-9531-y

Huang JM, Luo GJ, Zhang ZJ, Wang XG, Ju ZH, Qi C, Zhang Y, Wang CF, Li RL, Li JB, Yin WJ, Xu YX, Moisa SJ, Loor JJ, Zhong JF (2014) iTRAQ-proteomics and bioinformatics analyses of mammary tissue from cows with clinical mastitis due to natural infection with Staphylococci aureus. BMC Genomics 15:839. https://doi.org/10.1186/1471-2164-15-839

Ibeagha-Awemu EM, Ibeagha AE, Messier S, Zhao X (2010) Proteomics, genomics, and pathway analyses of Escherichia coli and Staphylococcus aureus infected milk whey reveal molecular pathways and networks involved in mastitis. J Proteome Res 9(9):4604–4619. https://doi.org/10.1021/pr100336e

Janjanam J, Jamwal M, Singh S, Kumar S, Panigrahi AK, Hariprasad G, Jena MK, Anand V, Kumar S, Kaushik JK, Dang AK, Mukesh M, Mishra BP, Srinivasan A, Reddy VS, Mohanty AK (2013) Proteome analysis of functionally differentiated bovine (Bos indicus) mammary epithelial cells isolated from milk. Proteomics 13(21):3189–3204. https://doi.org/10.1002/pmic.201300031

Janjanam J, Singh S, Jena MK, Varshney N, Kola S, Kumar S, Kaushik JK, Grover S, Dang AK, Mukesh M, Prakash BS, Mohanty AK (2014) Comparative 2D-DIGE proteomic analysis of bovine mammary epithelial cells during lactation reveals protein signatures for lactation persistency and milk yield. PLoS One 9(8):e102515. https://doi.org/10.1371/journal.pone.0102515

Jena MK, Janjanam J, Naru J, Kumar S, Kumar S, Singh S, Mohapatra SK, Kola S, Anand V, Jaswal S, Verma AK, Malakar D, Dang AK, Kaushik JK, Reddy VS, Mohanty AK (2015) DIGE based proteome analysis of mammary gland tissue in water buffalo (Bubalus bubalis): lactating vis-a-vis heifer. J Proteome 119:100–111. https://doi.org/10.1016/j.jprot.2015.01.018

Kang SJ, Il Cho Y, Kim KH, Cho ES (2016) Proteomic analysis to elucidate the antibacterial action of silver ions against bovine mastitis pathogens. Biol Trace Elem Res 171(1):101–106. https://doi.org/10.1007/s12011-015-0510-5

Kapila N, Sharma A, Kishore A, Sodhi M, Tripathi PK, Mohanty AK, Mukesh M (2016) Impact of heat stress on cellular and transcriptional adaptation of mammary epithelial cells in riverine buffalo (Bubalus Bubalis). PLoS One 11(9):e0157237. https://doi.org/10.1371/journal.pone.0157237

Kim Y, Atalla H, Mallard B, Robert C, Karrow N (2011) Changes in Holstein cow milk and serum proteins during intramammary infection with three different strains of Staphylococcus aureus. BMC Vet Res 7:51. https://doi.org/10.1186/1746-6148-7-51

Larsen LB, Hinz K, Jorgensen ALW, Moller HS, Wellnitz O, Bruckmaier RM, Kelly AL (2010) Proteomic and peptidomic study of proteolysis in quarter milk after infusion with lipoteichoic acid from Staphylococcus aureus. J Dairy Sci 93(12):5613–5626. https://doi.org/10.3168/jds.2010-3409

Lerias JR, Hernandez-Castellano LE, Morales-delaNuez A, Araujo SS, Castro N, Arguello A, Capote J, Almeida AM (2013) Body live weight and milk production parameters in the Majorera and Palmera goat breeds from the Canary Islands: influence of weight loss. Trop Anim Health Prod 45(8):1731–1736. https://doi.org/10.1007/s11250-013-0423-2

Lerias JR, Hernandez-Castellano LE, Suarez-Trujillo A, Castro N, Pourlis A, Almeida AM (2014) The mammary gland in small ruminants: major morphological and functional events underlying milk production – a review. J Dairy Res 81(3):304–318. https://doi.org/10.1017/s0022029914000235

Lerias JR, Pena R, Hernandez-Castellano LE, Capote J, Castro N, Arguello A, Araujo SS, Saco Y, Bassols A, Almeida AM (2015) Establishment of the biochemical and endocrine blood profiles in the Majorera and Palmera dairy goat breeds: the effect of feed restriction. J Dairy Res 82(4):416–425. https://doi.org/10.1017/s0022029915000412

Li X, Ding XZ, Wan YL, Liu YM, GZ D (2014) Comparative proteomic changes of differentially expressed whey proteins in clinical mastitis and healthy yak cows. Genet Mol Res 13(3):6593–6601. https://doi.org/10.4238/2014.August.28.4

Lippolis JD, Brunelle BW, Reinhardt TA, Sacco RE, Nonnecke BJ, Dogan B, Simpson K, Schukken YH (2014) Proteomic analysis reveals protein expression differences in Escherichia coli strains associated with persistent versus transient mastitis. J Proteome 108:373–381. https://doi.org/10.1016/j.jprot.2014.06.008

Lu J, Boeren S, van Hooijdonk T, Vervoort J, Hettinga K (2015) Effect of the DGAT1 K232A genotype of dairy cows on the milk metabolome and proteome. J Dairy Sci 98(5):3460–3469. https://doi.org/10.3168/jds.2014-8872

Mansor R, Mullen W, Albalat A, Zerefos P, Mischak H, Barrett DC, Biggs A, Eckersall PD (2013) A peptidomic approach to biomarker discovery for bovine mastitis. J Proteome 85:89–98

Mudaliar M, Tassi R, Thomas FC, McNeilly TN, Weidt SK, McLaughlin M, Wilson D, Burchmore R, Herzyk P, Eckersall PD, Zadoks RN (2016) Mastitomics, the integrated omics of bovine milk in an experimental model of Streptococcus uberis mastitis: 2. Label-free relative quantitative proteomics. Mol BioSyst 12(9):2748–2761. https://doi.org/10.1039/c6mb00290k

Murgiano L, Timperio AM, Zolla L, Bongiorni S, Valentini A, Pariset L (2009) Comparison of milk fat globule membrane (MFGM) proteins of Chianina and Holstein cattle breed milk samples through proteomics methods. Nutrients 1(2):302–315. https://doi.org/10.3390/nu1020302

Olumee-Shabon Z, Swain T, Smith EA, Tall E, Boehmer JL (2013) Proteomic analysis of differentially expressed proteins in caprine milk during experimentally induced endotoxin mastitis. J Dairy Sci 96(5):2903–2912. https://doi.org/10.3168/jds.2012-5956

Palma M, Hernandez-Castellano LE, Castro N, Arguello A, Capote J, Matzapetakis M, de Almeida AM (2016) NMR-metabolomics profiling of mammary gland secretory tissue and milk serum in two goat breeds with different levels of tolerance to seasonal weight loss. Mol BioSyst 12(7):2094–2107. https://doi.org/10.1039/c5mb00851d

Peng LF, Rawson P, McLauchlan D, Lehnert K, Snell R, Lordan TW (2008) Proteomic analysis of microsomes from lactating bovine mammary gland. J Proteome Res 7(4):1427–1432. https://doi.org/10.1021/pr700819b

Pongthaisong P, Katawatin S, Thamrongyoswittayakul C, Roytrakul S (2016) Milk protein profiles in response to Streptococcus agalactiae subclinical mastitis in dairy cows. Anim Sci J 87(1):92–98. https://doi.org/10.1111/asj.12391

Rawson P, Stockum C, Peng LF, Manivannan B, Lehnert K, Ward HE, Berry SD, Davis SR, Snell RG, McLauchlan D, Jordan TW (2012) Metabolic proteomics of the liver and mammary gland during lactation. J Proteome 75(14):4429–4435. https://doi.org/10.1016/j.jprot.2012.04.019

Reinhardt TA, Sacco RE, Nonnecke BJ, Lippolis JD (2013) Bovine milk proteome: quantitative changes in normal milk exosomes, milk fat globule membranes and whey proteomes resulting from Staphylococcus aureus mastitis. J Proteome 82:141–154

Roncada P, Piras C, Soggiu A, Turk R, Urbani A, Bonizzi L (2012) Farm animal milk proteomics. J Proteome 75(14):4259–4274

Sanchez-Rivera L, Martinez-Maqueda D, Cruz-Huerta E, Miralles B, Recio I (2014) Peptidomics for discovery, bioavailability and monitoring of dairy bioactive peptides. Food Res Int 63:170–181. https://doi.org/10.1016/j.foodres.2014.01.069

Singh M, Thomson PC, Sheehy PA, Raadsma HW (2013) Comparative transcriptome analyses reveal conserved and distinct mechanisms in ovine and bovine lactation. Funct Integr Genomics 13(1):115–131. https://doi.org/10.1007/s10142-012-0307-y

Singh H, Boland M, Thompson A (2014) Milk proteins from expression to food. Academic, Amsterdam

Smolenski G, Haines S, Kwan FYS, Bond J, Farr V, Davis SR, Stelwagen K, Wheeler TT (2007) Characterisation of host defence proteins in milk using a proteomic approach. J Proteome Res 6(1):207–215

Smolenski GA, Broadhurst MK, Stelwagen K, Haigh BJ, Wheeler TT (2014) Host defence related responses in bovine milk during an experimentally induced Streptococcus uberis infection. Proteome Sci 12:19. https://doi.org/10.1186/1477-5956-12-19

Smolenski GA, Cursons RT, Hine BC, Wheeler TT (2015) Keratin and S100 calcium-binding proteins are major constituents of the bovine teat canal lining. Vet Res 46:113. https://doi.org/10.1186/s13567-015-0227-7

Soares R, Franco C, Pires E, Ventosa M, Palhinhas R, Koci K, de Almeida AM, Coelho AV (2012) Mass spectrometry and animal science: protein identification strategies and particularities of farm animal species. J Proteome 75(14):4190–4206

Suarez AJ, Alvarez-Morujo Suarez AJ (1982) Morphological variations of mammary gland alveoli. Anat Histol Embryol 11:11

Suarez-Vega A, Gutierrez-Gil B, Klopp C, Robert-Granie C, Tosser-Klopp G, Arranz JJ (2015) Characterization and comparative analysis of the milk transcriptome in two dairy sheep breeds using RNA sequencing. Sci Rep 5:18399. https://doi.org/10.1038/srep18399

Suarez-Vega A, Gutierrez-Gil B, Klopp C, Tosser-Klopp G, Arranz JJ (2016) Comprehensive RNA-Seq profiling to evaluate lactating sheep mammary gland transcriptome. Sci Data 3:160051. https://doi.org/10.1038/sdata.2016.51

Tacoma R, Fields J, Ebenstein DB, Lam YW, Greenwood SL (2016) Characterization of the bovine milk proteome in early-lactation Holstein and Jersey breeds of dairy cows. J Proteome 130:200–210. https://doi.org/10.1016/j.jprot.2015.09.024

Thomas FC, Waterston M, Hastie P, Parkin T, Haining H, Eckersall PD (2015) The major acute phase proteins of bovine milk in a commercial dairy herd. BMC Vet Res 11:207

Thomas FC, Mudaliar M, Tassi R, McNeilly TN, Burchmore R, Burgess K, Herzyk P, Zadoks RN, Eckersall PD (2016a) Mastitomics, the integrated omics of bovine milk in an experimental model of Streptococcus uberis mastitis: 3. Untargeted metabolomics. Mol BioSyst 12 (9):2762–2769. https://doi.org/10.1039/c6mb00289g

Thomas FC, Mullen W, Tassi R, Ramirez-Torres A, Mudaliar M, McNeilly TN, Zadoks RN, Burchmore R, Eckersall PD (2016b) Mastitomics, the integrated omics of bovine milk in an experimental model of Streptococcus uberis mastitis: 1. High abundance proteins, acute phase proteins and peptidomics. Mol BioSyst 12(9):2735–2747. https://doi.org/10.1039/c6mb00239k

Timperio AM, D'Alessandro A, Pariset L, D'Amici GM, Valentini A, Zolla L (2009) Comparative proteomics and transcriptomics analyses of livers from two different Bos taurus breeds: "Chianina and Holstein Friesian". J Proteome 73(2):309–322. https://doi.org/10.1016/j.jprot.2009.09.015

van Altena SEC, de Klerk B, Hettinga KA, van Neerven RJJ, Boeren S, Savelkoul HFJ, Tijhaar EJ (2016) A proteomics-based identification of putative biomarkers for disease in bovine milk. Vet Immunol Immunopathol 174:11–18. https://doi.org/10.1016/j.vetimm.2016.04.005

Verma A, Ambatipudi K (2016) Challenges and opportunities of bovine milk analysis by mass spectrometry. Clin Proteomics 13:8. https://doi.org/10.1186/s12014-016-9110-4

Wheeler TT, Smolenski GA, Harris DP, Gupta SK, Haigh BJ, Broadhurst MK, Molenaar AJ, Stelwagen K (2012) Host-defence-related proteins in cows' milk. Animal 6(3):415–422. https://doi.org/10.1017/s1751731111002151

Yan QX, Tang SX, Tan ZL, Han XF, Zhou CS, Kong JH, Wang M (2015) Proteomic analysis of isolated plasma membrane fractions from the mammary gland in lactating cows. J Agric Food Chem 63(33):7388–7398. https://doi.org/10.1021/acs.jafc.5b02231

Yang YX, Wang JQ, DP B, Li SS, Yuan TJ, Zhou LY, Yang JH, Sun P (2012) Comparative proteomics analysis of plasma proteins during the transition period in dairy cows with or without subclinical mastitis after calving. Czeh J Anim Sci 57(10):481–489

Yang YX, Shen WJ, Zhao XW, Zhao HL, Huang DW, Cheng GL (2014) Proteomics and pathway analysis of N-glycosylated mammary gland proteins in response to Escherichia coli mastitis in cattle. Vet J 200(3):420–425. https://doi.org/10.1016/j.tvjl.2014.03.026

Yang YX, Zheng N, Zhao XW, Zhang YD, Han RW, Yang JH, Zhao SG, Li SL, Guo TJ, Zang CJ, Wang JQ (2016) Metabolomic biomarkers identify differences in milk produced by Holstein cows and other minor dairy animals. J Proteome 136:174–182. https://doi.org/10.1016/j.jprot.2015.12.031

Zhang L, Boeren S, van Hooijdonk ACM, Vervoort JM, Hettinga KA (2015) A proteomic perspective on the changes in milk proteins due to high somatic cell count. J Dairy Sci 98(8):5339–5351. https://doi.org/10.3168/jds.2014-9279

Zhao XW, Yang YX, Huang DW, Cheng GL, Zhao HL (2015) Comparative proteomic analysis of proteins expression changes in the mammary tissue of cows infected with Escherichia coli mastitis. J Vet Sci 16(3):253–263. https://doi.org/10.4142/jvs.2015.16.3.253

Proteomics in Wool and Fibre Research

Jeffrey E. Plowman

Abstract Wool is almost entirely composed of proteins, the major component of which are the keratins and the keratin-associated proteins (KAPs). Though they were first identified in 1934, it is only comparatively recently that they have been subject to study using modern proteomic techniques. Using a variety of approaches both gel and gel-free proteomics, many new keratins and KAPs have been identified and characterised in the mature fibre and its various subcomponents as well as through the various stages of keratinisation of wool follicle. Preliminary studies have also revealed distinctive differences both within and between breeds. Proteomic approaches have also allowed investigations to be extended into examining the effect of feed restriction on protein composition as well as modifications to the proteins caused by either environmental or process damage.

Keywords Electrophoresis • Mass spectrometry • Keratins • KAPs • Follicles

Abbreviations

%T	Total percentage of acrylamide
2DE	Two-dimensional electrophoresis
ESI	Electro-spray ionisation
HGTP	High-glycine-tyrosine protein
HSP	High-sulphur protein
KAP	Keratin-associated protein
LC	Liquid chromatography
MALDI	Matrix-assisted adsorption/desorption ionisation
MS	Mass spectrometry
TOF	Time of flight
UHSP	Ultrahigh-sulphur protein

J.E. Plowman (✉)
AgResearch Ltd, Christchurch, New Zealand
e-mail: jeff.plowman@agresearch.co.nz

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_14

1 Introduction

Wool is the mammalian fibre grown on sheep, its properties being highly dependent on the breed it is obtained from. Domestic breeds of sheep were thought to have originated some 8–10,000 years ago in southwestern and central Asia (Zeuner 1963). Apart from the obvious difference of colour, the wool from these sheep was characterised by an outer coat of coarse hair and an undercoat of short fine wool. Domestication of wild sheep dates back to around 3000 BC from such breeds as the European mouflon (*Ovis musimon*), steppe or urial (*Ovis vignei*) and argali (*Ovis ammon*). From these have arisen the plethora of breeds that exist today, a process that has also seen the disappearance of the coarse guard hairs and colour in the major wool-producing ones (Cottle 1991).

The wool fibres themselves originate from follicles in the dermis of the skin and undergo a programmed apoptosis that, in the cells of the central cortex, reduces their internal content to structures known as macrofibrils which define the physical characteristics of the fibre (Marshall et al. 1991). First identified in 1934 (Goddard and Michaelis 1934), the main structural component of these macrofibrils are keratins with their α-helical core. They are further divided into two types, the acidic Type Is and the neutral-basic Type II, these two associating to form a coiled-coil heterodimer, the primary building block of the macrofibril (Parry and Steinert 1995). The other major class of proteins in the wool fibre are the keratin-associated proteins (KAPs), which have an influence on the mechanical properties of the fibres through the interactions with the keratins. These have been artificially subdivided into three different subclasses based on the relative proportion of three amino acids: the high-sulphur proteins (HSPs) with cysteine contents of less than 30 moles%, the ultrahigh-sulphur proteins (UHSPs) with cysteine contents greater than 30 moles% and the high-glycine-tyrosine proteins (HGTPs) which can have anything from 10 to 30 moles% of these two amino acids (Powell 1996). More recent research has revealed new KAP families that have relatively low cysteine contents but are high in other amino acids such as serine or have intermediate to high levels of glycine and proline (Zhou et al. 2012; Gong et al. 2011).

Proteomics has a lot to offer in the study of wool. While genomics provides information on the genotype of the wool of the animal and the level of expression of these proteins during the early stages of follicle development, proteomics deals with the phenotype of the wool fibre, i.e. which proteins are actually present and in what relative amounts. More importantly proteomics can also reveal information on post-translational modifications as well as modifications that have occurred as a result of weathering or processing (Herbert et al. 1997, Paton et al. 2008). Proteomics has also been used to examine the interaction between the keratins and their associated KAPs by progressive reduction followed by alkylation (Deb-Choudhury et al. 2015). In all of this the techniques of two-dimensional electrophoresis (2DE) and mass spectrometry (MS), be it either liquid chromatography (LC) MS or matrix assisted adsorption/desorption ionisation (MALDI), are a critical part of the process of unlocking this information.

2 Mapping the Wool Cortex Proteome by 2DE

Some of the earliest studies into mapping keratin proteins of whole wool fibres by two-dimensional electrophoresis (2DE) involved separation in the first dimension by a non-equilibrium approach. This involved labelling the proteins with ^{14}C-iodoacetic acid, which gave the proteins an overall negative charge, allowing the separation to be carried out at pH 8.9. The proteins were then separated in the second dimension on the basis of their molecular mass in polyacrylamide gels. Because of the radiolabelling, the resulting protein pattern could then be visualised by autoradiography. Using this approach most of the proteins appeared along a line of spots and bands running along a diagonal in the gel, though the keratins and some HGTPs appeared off the diagonal (Marshall 1981). Separation in the first dimension at pH 3 was also investigated and found to lead to a better dispersion of protein spots in the non-equilibrium 2DE map (Marshall and Ley 1986). The issue these early proteomic approaches faced was identification of the proteins in the map. Both the keratins and HGTPs were easy to locate based on their respective molecular weights but more difficulty was experienced identifying the HSPs and UHSPs. One approach taken with one HSP family and the cuticle UHSPs involved in vitro expression of these proteins and then separation by non-equilibrium 2DE (MacKinnon et al. 1990).

While this approach was useful, it had its limitations, largely because of the poor resolution between the proteins in the clusters observed on the map. The next step in the process of 2DE mapping wool proteins involved combining isoelectric focusing in the first dimension with molecular weight separation in the second dimension (Herbert and Woods 1994; Herbert et al. 1996). Using this approach ^{14}C-iodoacetamide-alkylated wool proteins were focused between pH 4 and 7 in the first dimension separating the keratins into two distinct regions, the acidic Type I keratins being found in a tight cluster at low pH and the Type II keratins appearing in a long string at higher pH and molecular mass. After electrophoresis individual protein spots were excised from the gel, eluted onto polyvinylidene fluoride membranes and analysed by amino acid analysis (Herbert et al. 1997). Identification of keratins was by the way of amino acid compositional matching via international sequence databases, though the number of amino acid sequences available limited this approach. At that stage, only two full sequences were known for Type I and two for Type II keratins in SWISS-PROT (http://www.expasy.org/) or NCBI (http://www.ncbi.nlm.nih.gov/) databases. The substitution of dithiothreitol by the tertiary phosphine, tributylphosphine, led to further increases in the number of minor proteins visualised on the gel, especially when silver staining was used (Herbert et al. 1998).

Improvements in the 2DE technique saw the resolution of the Type I keratin region into four distinctive chains of spots, rather than long smears. At the same time, MALDI time-of-flight (TOF) mass spectrometry replaced amino acid analysis as a way of identifying the proteins in those spots. At this stage proteins were matched to known sequences using peptide mass fingerprinting with search engines such as MASCOT (http://www.matrixscience.com/search_form_select.html)

(Plowman et al. 2000). Nevertheless, there was still the same limitation of sequence availability, and with only two known Type I and two Type II keratin sequences in these databases, the four strings of spots in the Type II region could only be matched to two proteins (Plowman et al. 2002). One thing that also became evident in this approach was the problem of high sequence homology within members of the various keratin and KAP families. In the case of the KAP1 HSP, this was compounded by the limited number of basic residues in the proteins resulting in unique identification of proteins in this family being dependent on only four tryptic peptides. The presence of proline next to some of these residues placed further limitations on this as uniquely identifying the various members of this family was highly dependent on some low abundance tryptic peptides.

Further progress in the identification of proteins in the wool protein 2DE map became possible with access to an in-house expressed sequence tag (EST) database, which offered many new sequences of both keratins and KAPs. From this it was apparent that there is a considerable homology between individual keratin proteins such as the four Type I keratins, K31, K33a, K33b and K34, where the number of identical residues is around 92%, while sequence homology can be as high as 96%. This is also true for the major Type II keratins with the number of identical residues being as high as 78% with homologies as high as 89%. Under such circumstances, identifying proteins in each spot became highly dependent on locating the relatively few unique peptides for that protein. To investigate this keratin, proteins were separated between pH 4 and 7 and then on 8% T gels in the second dimension to maximise the resolution between protein spots on the gels. Both MS and MS/MS data were collected with both a MALDI-Q-TOF MS and a tandem quadrupole TOF MS (Deb-Choudhury et al. 2010). The peak lists generated by both approaches were searched against the NCBInr database augmented by an in-house EST database of sheep sequences using an in-house MASCOT server. To determine the distribution of these highly homologous proteins in the spots on the 2DE map, the data was filtered for unique masses for Type I and II keratins. Then, all unique peptides were tabulated, and from these, the number of unique peptides masses per keratin per spot was determined. By these means the predominant keratin in each spot was determined. Thus, the four major strings of spots in the Type I keratin region that had previously been assigned to two proteins (Plowman et al. 2002) could now be assigned to four keratins, K31, K33a, K33b and K34 (Fig. 1), though there appeared

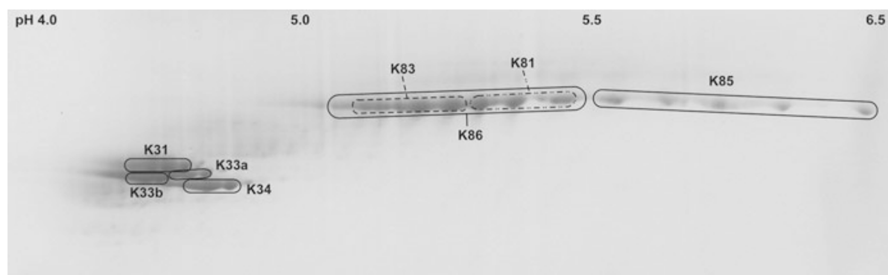

Fig. 1 The 2DE map of the keratin region of the mature wool fibre, separation being carried out between pH 4 and 7 in the first dimension and on 8% T gels in the second dimension

to be a considerable overlap between K33a and K33b. One interesting finding of this study was the detection of at least one unique peptide for K38 with the K31 string and unique peptides for K35 in the K33a/K33b strings, MALDI imaging suggesting that K35 may also be found within the K31 string of spots. In the Type II keratin region, unique peptides for K81, K83 and K86 were detected in the low pH region of the string. Furthermore, K81 tended to be found at the higher pH end of this cluster of spots, while K83 was at the lower pH end. Finally, unique peptides for K85 were predominantly found in the six spots at the high pH end of the Type II keratin string.

One problem with the keratins is that they tend to dominate the wool protein 2DE map. However, by modifying the extraction solutions through lowering either the urea or DTT concentrations or lowering the pH, it was possible to reduce the amount of keratin extracted and thus increase the protein load of KAPs on 2DE gels (Plowman et al. 2010). The KAPs were found to fall into two major groups, depending on whether they focused at acidic or alkaline pHs (Fig. 2). Two protein families were found at acidic pH, the KAP1 HSP family of four members between 20 and 30 kDa and the KAP3 family between 10 and 15 kDa. Of the KAP1 family three of the spots have been identified by finding unique peptides to KAP1.1, KAP1.3 and KAP1.4 (Plowman et al. 2002), while KAP3.2 and KAP3.4 have been positively identified in the lower molecular weight string. The alkaline KAPs can be further divided into two major groups, the HSPs which separate

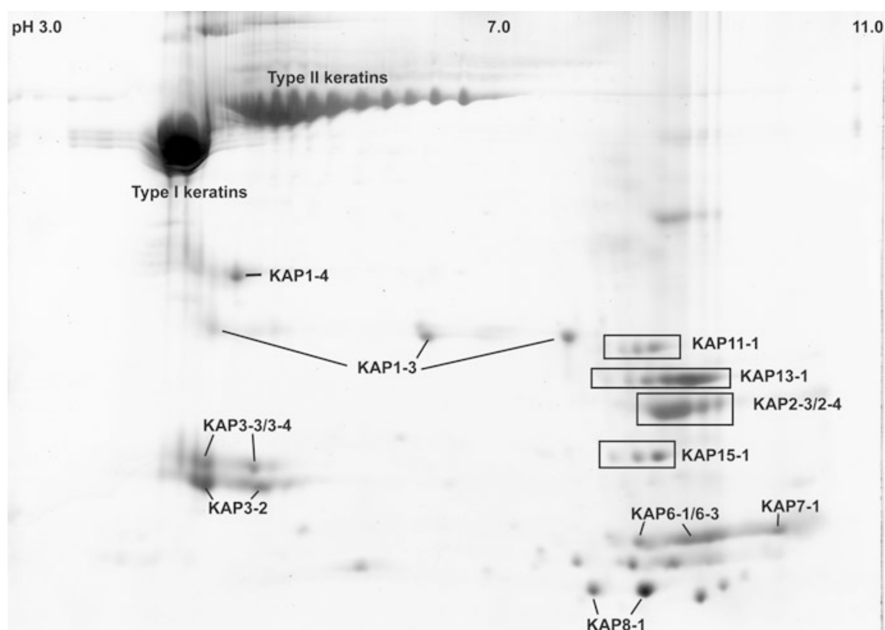

Fig. 2 The 2DE map of keratin and KAP of the mature wool fibre, separation being carried out between pH 3 and 11 in the first dimension and on 15% T gels in the second dimension

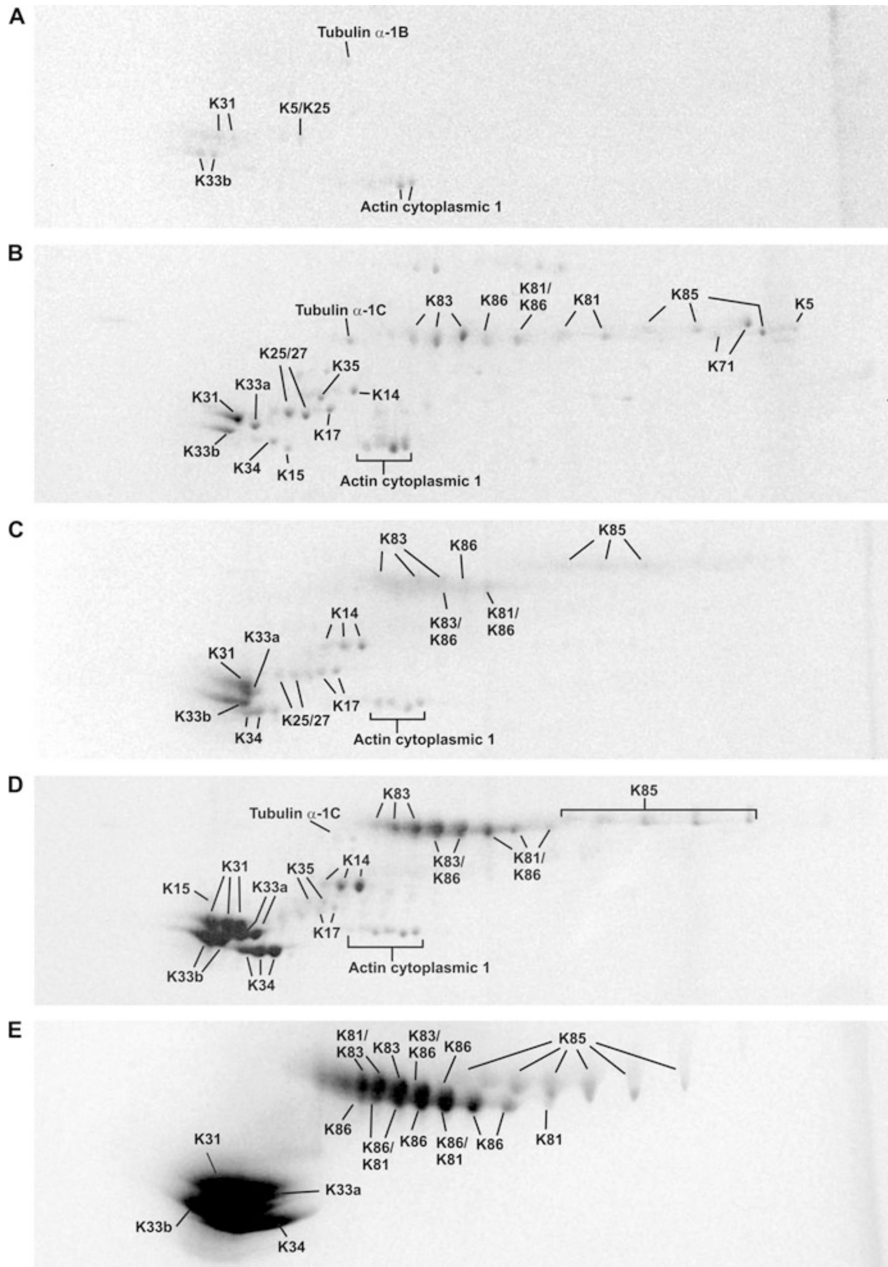

Fig. 3 2DE maps of the keratin region of the various regions of the wool follicle and fibre: (**a**) the bulb portion, (**b**) the elongation portion, (**c**) the keratogenous portion, (**d**) the keratinisation portion, (**e**) the mature fibre

between 15 and 20 kDa and the HGTPs below 10 kDa. Of the HSPs unique peptides for three members of the KAP2 family have been found in a broad spot, while KAP11.1 and KAP13.1 resolve above them and KAP15.1 below. Among the HGTPs unique peptides for both KAP6.1 and KAP6.3 have been identified in one string, along with KAP7.1 at higher pH, and below them, two spots are consistent with KAP8.1 (Fig. 3).

3 Mining the Proteome of Wool by MS

Early studies into the protein composition of wool suggested that there were a maximum of four Type I and four Type II keratins in wool, and including the proteins from the nine KAP families, the total keratin-related proteins in wool was thought to be 23 (Powell 1996; Powell and Rogers 1997). However, recent advances in hair research have seen the identification of 11 and 6 Type I and II keratins, respectively (Langbein et al. 1999, 2001, 2007), while 25 human hair KAP families have now been identified with a total of 88 individual KAPs (Rogers et al. 2001, 2002, 2008). In contrast a total of 9 Type I and 7 Type II keratins have been identified in wool (Yu et al. 2011), but the total number of KAPs was still unknown.

As the sheep genome had not been fully characterised around the time that these other KAPs had been identified in human hair, mass spectrometry was used to drill deeper into the wool proteome with the view to identifying new proteins and thus update the wool protein database (Clerens et al. 2010). After tryptic digestion wool protein extracts were analysed by a variety of mass spectral techniques, including LC-MS/MS, LC-MALDI and 2D-LC-MS/MS. To extend the coverage of the wool proteome separation of the extracts, prior to mass spectrometry, was carried by one-dimensional gel electrophoresis off-line or on-line via a variety of LC techniques including reverse phase, strong cation exchange, strong anion exchange and hydroxy-apatite chromatography. This resulted in the production of 167 data sets which were separated individually to an in-house Mascot server for identification against an in-house sheep expressed sequence tag (EST) database restricting taxonomy to *Ovis aries*. Searches were then extended to other even-toed ungulates, excluding *Bos taurus*, with combinations of peptides identifying proteins in *Bovidae*-w/o-*Bos* but not in *Ovis* considered to represent genuine proteins not described for the *Ovis* proteome. These proteins were added to the database, with names assigned based on their homology with existing sequences in other species. The search was extended to the *Laurasiatheria* taxonomy (placental mammals) and the same approach followed, any new proteins identified being added to the database. This database was then validated by importing MS/MS peak lists into ProteinScape (Bruker Daltonics) and searching these with the database augmented with the new *Ovis* sequences and excluding all positives by rejecting those with Mascot scores of less than 25, while proteins with a combined score of 40 were always rejected. This led to an initial short list of 115 proteins, which was reduced to 104 after the removal of false positives. Subsequent BLAST homology searches led to the removal of two duplicate entries,

but the discovery of some unique peptides gave rise to 11 additional proteins bringing the total number identified to 113.

This approach resulted in the identification of 15 keratins including 2 polymorphic variants, 13 epithelial keratins and 36 KAPs, including 6 polymorphic variants, as well as trichohyalin. This means that now 23 KAP families and 74 individual KAPs have been identified in sheep. Also identified were keratin-anchoring proteins such as desmoglein, desmoplakin and plakophillin, and other filamentous proteins such as lamin and tubulin. Of the remaining proteins there were a number of nuclear proteins, such as histones, and enzymes, and this is, no doubt, largely because the keratinisation process does not completely remove all other protein material, leaving behind what is termed the cytoplasmic remnant in the centre of the cell.

4 Differential Proteomics of the Wool Fibre

Interest in the protein composition of the individual components of the wool fibre has led to efforts to separate and analyse these individually. In one approach the cuticle and cortical cells were separated by mechanical disruption in either detergents or formic acid (Marshall and Ley 1986). Subsequent separation by non-equilibrium 2DE, when separation was carried out in the first dimension at pH 3, resulted in a significantly different pattern of proteins from that observed in the cortex, the cuticle HSPs appearing to be of higher molecular weight than cortical HSPs. Mechanical disruption of wool snippets treated with trypsin was used to release individual orthocortical cells from the fibre leaving behind a core of paracortical material (Dowling et al. 1990). After extraction from the fibre, the keratin fraction was precipitated at pH 4.4 with potassium chloride and analysed separately from the KAP fraction by non-equilibrium 2DE. From this approach it was apparent that the eight main components of the keratin fraction were expressed in the same relative amounts in the two cell types, but the paracortex was found to contain a higher concentration of UHSPs than the orthocortex, though one HSP component was observed to be absent in the paracortex. Orthocortical cells were also released from wool fibre snippets by disruption with an ultrasound probe after treatment with Pronase E to leave a fibre remnant consisting mostly of paracortical cells (Plowman et al. 2009). Subsequent separation by 2DE, when proteins were separated in the first dimension between pH 4 and 7, revealed that the KAP3 HSP family was found to be in higher concentrations in the paracortex.

The most easily extractable part of the wool fibre are the macrofibrillar components of the cortical cells. Extraction solutions composed of high concentrations of urea at alkaline pHs in conjunction with reductants 2-mercaptoethanol or dithiothreitol will normally only solubilise these proteins, leaving behind the cell membrane complex (the remnant of the cell wall), intermacrofibrillar material, the cytoplasmic remnant (including the remains of the cell nucleus) and the cuticle. However, substitution of dithiothreitol with tris-(2-carboxyethyl)phosphine has

been found to extract proteins from the exocuticle and part of the endocuticle, in addition to those in the macrofibrils (Bringans et al. 2007). Alternatively, digestion of the whole wool fibre with enzymes such as Pronase E leads to the removal of the cell membrane complex, intermacrofibrillar material, the cytoplasmic remnant and the endocuticle of the cuticle cells. When these two processes are combined, extracting first with tris-(2-carboxyethyl)phosphine, followed by digestion with Pronase E a fraction that is highly enriched in the outermost layer of the exocuticle, the a-layer, is all that is left behind. Subsequent digestion of this with 2-nitro-5-thiocyano-benzoic acid led to the identification of peptides from the UHSP KAP5.5, as well as peptides with strong homologies to human HSPs and UHSPs. In extending this work to the whole cuticle, a further study revealed that the cuticle is composed of a diverse range of proteins (Koehn et al. 2010). While keratins and KAPs make up 13% and 27%, respectively, of the total, the remaining 60% includes cytoskeletal proteins; cell adhesion proteins (including desmosomes); binding and transport proteins; transcription, translation and regulation proteins; protein modification and metabolism proteins; binding/transport proteins; and immune response proteins.

5 The Wool Follicle

The procedure of in situ hybridisation has been used to determine the expression pattern of keratins and two KAPs in both the cuticle and cortex of the wool follicle (Yu et al. 2009, 2011). These studies have revealed that the first keratins to be expressed in the follicle are the Type I K35 and Type II K85, followed shortly after by the K31, with the other keratins appearing further up the follicle shaft, an observation that is consistent with that observed for human hair (Langbein et al. 1999, 2001). The expression pattern of two KAPs was also mapped by this process, both showing differential expression, KAP4.3, a UHSP, being localised on the paracortical side of the follicle, and KAP6.1, a HGTP, on the orthocortical side (Yu et al. 2011). While ISH is useful for indicating the start and end of keratin expression in the follicle, it does not provide any information on the changes to the proteins as they progress up the follicle shaft, something possible with proteomics.

Using a gel-free proteomic approach, ten whole follicles were dissected from the skin, then extracted with lysis buffer and analysed by mass spectrometry (Plowman et al. 2015). From this approach, 87 proteins were identified in the follicle, including 10 trichocyte keratins, 15 epithelial keratins and trichohyalin, a KAP related more with the inner root sheath of the follicle. Gel-free proteomic analysis of 30 follicles subdivided into four zones by dissection provided further confirmation that K31, K35 and K85 are the first trichocyte keratins expressed in the follicle, being found in the region from the base of the follicle to widest part of the bulb (Table 1). Expression of the other keratins was observed to start shortly afterwards in the elongation portion and they were also present in the keratogenous and keratinisation portions of the follicle. The first KAP to appear was trichohyalin in

Table 1 Protein composition of the fibre and the four follicle portions, determined by gel-free proteomic analysis

Type of protein	Follicle portion				Fibre
	Bulb	Elongation	Keratogenous	Keratinisation	
Trichocyte keratins	K31, K35, K85	K31, K33a, K33b, K34, K35, K83, K84, K85, K86	K31, K33a, K33b, K34, K35, K81, K83, K85, K86	K31, K33a, K33b, K34, K35, K81, K83, K85, K86	K31, K33a, K33b, K34, K35, K81, K83, K85, K86
KAPs	Trichohyalin	Trichohyalin KAP11.1	Trichohyalin, KAP2.3, KAP11.1, KAP13.1	Trichohyalin, KAP3.2, KAP4.7, KAP11.1, KAP19.6	KAP2.3, KAP3.2, KAP4.2, KAP4.7, KAP4.9, KAP4.12, KAP6.1, KAP7.1, KAP11.1, KAP13.1, KAP19.5
Epithelial Keratins	K1, K5, K10, K14, K17, K25, K27, K28, K71, K73	K1, K5, K6C, K8, K10, K14, K17, K19, K25, K27, K28, K71, K73	K1, K6C, K8, K14, K17, K19, K25, K27, K28, K71, K72, K73	K6C, K7, K8, K14, K17, K25, K27, K28, K71, K72, K74	
Keratin and IF-anchor proteins	Desmoplakin, vimentin	Desmoplakin, vimentin	Desmoplakin, vimentin	Vimentin	
Collagens	Collagen α-1(XII), collagen α-1(VI), collagen α-2(VI), collagen α-3(VI)	Collagen α-1(XII), collagen α-1(VI), collagen α-2(VI), collagen α-3(VI)	Collagen α-1(VI), collagen α-3(VI)	Collagen α-1(VI), collagen α-3(VI)	
Ribosomal proteins	40S ribosomal protein (S16, S10-like, S4-like, S6-like, nucleolin, ribosomal protein S25) 60S ribosomal proteins (S14,P2, L7a, L18, acidic P0)	40S ribosomal protein (S9-like, S10-like, S4-like, S2-isoform 2, S16) 60S ribosomal proteins (L34, L6-like, P2, L7a, L7-like, L18, acidic P0)	40S ribosomal protein (S7, S10-like, S2-like, S16) 60S ribosomal proteins (L14, P2, L18, L7-like, acidic P0)	60S ribosomal protein P2	
Nuclear proteins	Prelamin-A/C, histones: H1.3, H3.2, H2A Type I	Histones: H1.3, H2A, H3A	Histones: H1.3, H2A, laminin	Histones: H2A, 3A	

the bulb region, followed by KAP11.1 in the elongation portion. This was followed by the appearance of KAP2.3 and KAP13.1 in the keratogenous zone and KAP3.2, KAP4.7 and KAP19.6 in the keratogenous portion, an observation that was consistent with what has been observed in human hair (Rogers et al. 2002).

Epithelial keratins were also prominent in the four follicle zones but not the fibre itself, most of them coming from either the outer route sheath (K5, K7, K14, K19) or the inner root sheath (K25, K27, K28, K71, K74). Ribosomal proteins were found to decline in number moving up from the follicle bulb, an observation consistent with electron micrographs of the follicle in these regions. Other proteins found included nuclear proteins (histones) and keratin anchor proteins such as desmoplakin and vimentin, actin and tubulin.

A two-dimensional electrophoretic (2DE) study revealed that the trichocyte keratins undergo a considerable change as they progress up the fibre 3. Around the upper part of the follicle bulb, the Type Is formed a tight group at low pH and the Type IIs appeared in a long string at higher pH, while K35 and a number epithelial keratins were found in a cluster between these two families. As they moved into the keratogenous zone in the follicle, they gradually assumed the classic pattern of spots found in the mature fibre, the epithelial keratins finally becoming less apparent than in the mature fibre. The Type I keratins were notable for the subtle change they underwent during the keratinisation process. Initially they appeared as a single spot with a narrow streak extending towards lower pH in the elongation portion. This was observed to increase in intensity and length in the keratogenous portion, until it finally coalesced into distinct strings of spots, something that was characteristic of the mature keratinised fibre.

6 Breed Diversity

Early proteomic research has shown that there is considerable diversity in sheep based on the composition of members of the KAP1 HSP family. This family of four HSPs is characterised by a repeat decapeptide sequence in which KAP1.1 has four such repeats, KAP1.2 three, KAP1.3 two and KAP1.4 five (Powell 1996; Powell and Rogers 1997). A study of Merino, Romney and Corriedale sheep by 2DE has shown that the patterns of spots of this protein family differ not only between the three breeds but within the same breed (Flanagan et al. 2002). Of the four proteins in this family, KAP1.3 and KAP1.4 appear to be expressed in every breed. The expression of other two, however, is variable. In Merino three distinct populations have been seen, those with both of these KAPs present or those in which one or the other are present. In contrast, only KAP1.1, KAP1.3 and KAP1.4 are expressed in Romney and Corriedale and the pattern of spots differ significantly from that of Merino. Though there was no relationship between the expression of these proteins and staple tenacity in Merino or Romney sheep, principal component analysis pointed to a relationship between the protein expression and curvature in Corriedale sheep.

7 Post-translational, Environmental and Process Modifications of Keratins

Phosphorylation is known to be important in the assembly process of intermediate filaments such as the lamins (Peter et al. 1991). As such it has been of some interest to know whether the keratins in wool are phosphorylated, as it was thought that it might also explain the multiple spots observed in wool 2DE maps. Some of the earliest evidence for this came from a phosphoamino acid analysis study which claimed to have found phosphoserine in seven of the Type II keratin spots from a 2DE map along with trace amounts of phosphotyrosine (Herbert et al. 1997). However, a more recent investigation applying specific phosphoprotein stains to 2DE wool maps found no evidence for phosphoproteins in wool (Paton et al. 2008). Instead, when individual spots were eluted from the gel and then re-run by 2DE, new spots were formed on both the acidic and basic side of the original spot. In fact when several eluted spots were run together, the whole pattern could be recreated. From this, it was concluded that the cause of the charge heterogeneity in the 2DE map was due to the presence of conformational equilibria between a number of different forms of the same protein in the rehydration solution.

The development of cross-links between keratins and KAPs is an important part of keratinisation, the location, pattern and accessibility of them underpinning the properties of the network. Preliminary investigations into the effect of bleaching on wool found that only few specific cysteine residues were being oxidised to cysteic acid by treatment with hydrogen peroxide (Plowman et al. 2003). This led to a study using a process of staged labelling of both keratins and KAPs where the network was progressively exposed by the addition of chaotropes and reductants and the affected cysteine residues labelled with distinguishable alkylation agents (Deb-Choudhury et al. 2015). Labelling of the KAPs was first observed to occur only after the addition of reductant, and the affected cysteine residues were close to the end domains of the protein suggesting that they were easily accessible in the fibre for forming inter-disulphide linkages keratins or other KAPs. In the case of the HSPs and UHSPs, the majority of the cysteines were in close proximity to a proline, which are thought to play a significant role in forming a close loop conformation in cysteine-rich pentapeptide sequences (Parry et al. 2006). In the case of the keratins, some of the accessible cysteines were found on in the rod domains but were positioned on the exposed surface of the α-helix. The rest were found in the head and tail domains of both these proteins, supporting the idea that they are not buried inside the keratins but hence available for interaction with KAPs.

Proteomics also provides a way of studying the environment and industrial processing on wool. To study the effect of ultraviolet-induced photooxidation, enriched wool cuticle and cortex fractions were irradiated with UVB light and then subjected to a comprehensive analysis of the redox-induced modifications to the aromatic residues in the fibre (Dyer et al. 2010). From this work, an oxidative classification system was developed based on the degree of photo-modification of tyrosine and tryptophan, and this was used to examine the differential nature of

photooxidation. One observation from that was, although the cuticle had a higher level of baseline oxidation than the cortex, the cortex exhibited significantly higher levels of photooxidation under UVB radiation, a fact that was supported by the observation that there were significantly higher levels of photoyellowing in the cortex. The effect of elevated temperatures samples of wool in 0.05 M tetraborate at pH 8.5 were subjected to heat treatment at 90 °C for periods ranging from 2 to 8 h (Grosvenor et al. 2012). While many peptides were observed to be degraded by this process, 20 keratin peptides were observed to consistently degrade during hydrothermal exposure to the extent that they could be used as markers of damage. In fact three of them (ATAENEFVALK, ENAELESR and SNHEEEVNTLR) were of particular note because of both their consistency to oxidative damage and their high level of homology to human hair and skin proteins, hence their usefulness in evaluating protein damage in these tissues and materials. Oxidative damage that leads to the discolouration of wool is of particular interest to the wool industry because it lowers the price farmers receive for wool, and while the colour can be removed by bleaching, the chemicals used to achieve this further damage the wool. In this respect proteomics has been of considerable assistance to the industry in identifying and characterising non-scourable chromophores in wool (Dyer et al. 2007). The identification of these led to an examination of the effect of industrially applicable reagents some of which showed promise in improving the colour in canary yellow wool (Grosvenor et al. 2009). Other studies have identified photooxidation products within photoyellowed bleached wool proteins (Dyer et al. 2006a, b) and have been used to evaluate methods to minimise photooxidation and photobleaching of wool products, such as wool carpets.

In another study the effect of seasonal weight loss on wool production was examined (Almeida et al. 2014). Samples of wool were taken from two different groups of sheep, the control group that had been fed a growth diet and a second group fed a restricted diet. Using isobaric labelling to quantitate the protein expression of the two flocks, an increase in the level of the HSP KAP13.1 and three HGTPs, KAP6, KAP6.2 and KAP6.3var1, was observed. For KAP13.1, this would seem to be at odds with other observations of a deficiency of cysteine in wool grown in nutritionally stressed environments, but it was thought that it may be related to the fact that KAP13.1 is a key protein in the early stages of fibre keratinisation. In contrast the increase in the HGTPs may be related to changes in cortical morphology. HGTPs are preferentially expressed in the orthocortex, and an increase of these proteins in the restricted diet animals was thought to be related to an increase in the relative proportion of these cells and hence decreases in crimp frequency.

8 Conclusions

Thus, the versatility of the proteomic approach as applied to wool is clearly evident from these studies. While in the past it has primarily been used to catalogue the protein composition of the fibre and its various components, it has also proved

useful in the study of fibre development from the follicle and in investigating the relationship between breed diversity and protein composition. More recently, it has been used to examine the effect of environmental or process damage and restricted feed on the fibre. Proteomics has proved useful in not only studying both oxidative and photooxidative damage that leads to the discolouration of wool (such as canary-yellowed wool) but in the evaluation of industrially applicable approaches to mitigate some of this damage or preventing the development of photobleaching and photodegradation of products like carpet wool.

References

Almeida AM, Plowman JE, Harland DP, Thomas A, Kilminster T, Scanlon T, Milton J, Greeff J, Oldham C, Clerens S (2014) Influence of feed restriction on the wool proteome: a combined iTRAQ and fiber structural study. J Proteomics 103:170–177. https://doi.org/10.1016/j.jprot.2014.03.036

Bringans SD, Plowman JE, Dyer JM, Clerens S, Vernon JA, Bryson WG (2007) Characterization of the exocuticle a-layer proteins of wool. Exp Dermatol 16:951–960. https://doi.org/10.1111/j.1600-0625.2007.00610.x

Clerens S, Cornellison CD, Deb-Choudhury S, Thomas A, Plowman JE, Dyer JM (2010) Developing the wool proteome. J Proteomics 73:1722. https://doi.org/10.1016/j.jprot.2010.05.005

Cottle DJ (1991) Australian sheep and wool handbook, 1st edn. Inkata Press, Melbourne

Deb-Choudhury S, Plowman JE, Rao K, Lee E, van Koten C, Clerens S, Dyer JM, Harland DP (2015) Mapping the accessibility of the disulfide crosslink network in the wool fiber cortex. Proteins 83:224–234. https://doi.org/10.1002/prot.24727

Deb-Choudhury S, Plowman JE, Thomas A, Krsinic GL, Dyer JM, Clerens S (2010) Electrophoretic mapping of highly homologous keratins: a novel marker peptide approach. Electrophoresis 31:2894–2902. https://doi.org/10.1002/elps.201000134

Dowling LM, Ley KF, Pearce AM (1990) The protein composition of cells in the wool cortex. In: Proceedings of the 8th International Wool Textile Research Conference. pp 205–214

Dyer JM, Bringans SD, Aitken GD, Joyce NI, Bryson WG (2007) Extraction and characterisation of non-scourable chromophores from discoloured fleece wool. Color Technol 123:54–58. https://doi.org/10.1111/j.1478-4408.2006.00060.x

Dyer JM, Bringans SD, Bryson WG (2006a) Characterisation of photo-oxidation products within photoyellowed wool proteins: tryptophan and tyrosine derived chromophores. Photochem Photobiol Sci 5:698–706. https://doi.org/10.1039/b603030k

Dyer JM, Bringans SD, Bryson WG (2006b) Determination of photo-oxidation products within photoyellowed bleached wool proteins. Photochem Photobiol 82:551–557. https://doi.org/10.1562/2005-08-29-RA-663

Dyer JM, Plowman J, Krsinic G, Deb-Choudhury S, Koehn H, Millington K, Clerens S (2010) Proteomic evaluation and location of UVB-induced photomodification in wool. Photochem Photobiol B Biol 98:118–127. https://doi.org/10.1016/j.jphotobiol.2009.11.008

Flanagan LM, Plowman JE, Bryson WG (2002) The high sulphur proteins of wool: towards an understanding of sheep breed diversity. Proteomics 2:1240–1246. https://doi.org/10.1002/1615-9861(200209)2:9<1240::AID-PROT1240>3.0.CO;2-#

Goddard DR, Michaelis L (1934) A study on keratin. J Biol Chem 106:605–614

Gong H, Zhou H, Dyer J, Hickford J (2011) Identification of the ovine KAP11-1 gene (KRTAP11-1) and genetic variation in its coding sequence. Mol Biol Rep 38:5429–5433. https://doi.org/10.1007/s11033-011-0697-2

Grosvenor AJ, Morton JD, Dyer JM (2012) Determination and validation of markers for heat-induced damage in wool proteins. Am J Anal Chem 3:431–436. https://doi.org/10.4236/ajac.2012.36056

Grosvenor AJ, Plowman JE, Dyer JM (2009) Targeted mitigation of non-scourable wool discolouration. Int J Sheep Wool Sci 57:Article 9

Herbert BR, Chapman AL, Rankin DA (1996) Investigation of wool protein heterogeneity using two-dimensional electrophoresis with immobilised pH gradients. Electrophoresis 17:239–243. https://doi.org/10.1002/elps.1150170141

Herbert BR, Molloy MP, Gooley AA, Walsh BJ, Bryson WG, Williams KL (1998) Improved protein solubility in two-dimensional electrophoresis using tributyl phosphine as reducing agent. Electrophoresis 19:845–851

Herbert BR, Molloy MP, Yan JX, Gooley AA, Bryson WG, Williams KL (1997) Characterisation of wool intermediate filament proteins separated by micropreparative two-dimensional electrophoresis. Electrophoresis 18:568–572. https://doi.org/10.1002/elps.1150180339

Herbert BR, Woods JL (1994) Immobilised pH gradient isoelectric focusing of wool proteins. Electrophoresis 15:972–976. https://doi.org/10.1002/elps.11501501142

Koehn H, Clerens S, Deb-Choudhury S, Morton JD, Dyer JM, Plowman JE (2010) The proteome of the wool cuticle. J Proteome Res 9:2920–2928. https://doi.org/10.1021/pr901106m

Langbein L, Rogers MA, Praetzel-Wunder S, Böckler D, Schirmacher P, Schweizer J (2007) Novel type I hair keratins K39 and K40 are the last to be expressed in differentiation of the hair: completion of the human hair keratin catalogue. J Invest Dermatol 127:1532–1535

Langbein L, Rogers MA, Winter H, Praetzel S, Beckhaus U, Rackwitz H-R, Schweizer J (1999) The catalog of human hair keratins. I. Expression of the nine type I members in the hair follicle. J Biol Chem 274:19874–19884. https://doi.org/10.1074/jbc.274.28.19874

Langbein L, Rogers MA, Winter H, Praetzel S, Schweizer J (2001) The catalog of human hair keratins. II. Expression of the six type II members in the hair follicle and the combined catalog of human type I and II keratins. J Biol Chem 276:35123–35132. https://doi.org/10.1074/jbc.M103305200

MacKinnon PJ, Powell BC, Rogers GE (1990) Structure and expression of genes for a class of cysteine-rich proteins of the cuticle layers of differentiating wool and hair follicles. J Cell Biol 111:2587–2600. https://doi.org/10.1083/jcb.111.6.2587

Marshall RC (1981) Analysis of the proteins from single wool fibers by two-dimensional polyacrylamide gel electrophoresis. Text Res J 51:106–108

Marshall RC, Ley KF (1986) Examination of proteins from wool cuticle by two-dimensional gel electrophoresis. Text Res J 56:772–774

Marshall RC, Orwin DFG, Gillespie JM (1991) Structure and biochemistry of mammalian hard keratin. Electron Microsc Rev 4:47–83

Parry DAD, Smith TA, Rogers MA, Schweizer J (2006) Human hair keratin-associated proteins: sequence regularities and structural implications. J Struct Biol 155:361–369. https://doi.org/10.1016/j.jsb.2006.03.018

Parry DAD, Steinert PM (1995) Intermediate filament structure. Springer, Heidelberg, Germany

Paton LN, Gerrard JA, Bryson WG (2008) Investigations into charge heterogeneity of wool intermediate filament proteins. J Proteomics 71:513–529. https://doi.org/10.1016/j.jprot.2008.08.001

Peter M, Heitlinger E, Häner M, Aebi U, Nigg EA (1991) Disassembly of *in vitro* formed lamin head-to-tail polymers by CDC2 kinase. EMBO J 10:1535–1544

Plowman JE, Bryson WG, Flanagan LM, Jordan TW (2002) Problems associated with the identification of proteins in homologous families: the wool keratin family as a case study. Anal Biochem 300:221–229. https://doi.org/10.1006/abio.2001.5459

Plowman JE, Bryson WG, Jordan TW (2000) Application of proteomics for determining protein markers for wool quality traits. Electrophoresis 21:1899–1906. https://doi.org/10.1002/(SICI)1522-2683(20000501)21:9<1899::AID-ELPS1899>3.0.CO;2-R

Plowman JE, Deb-Choudhury S, Bryson WG, Clerens S, Dyer JM (2009) Protein expression in orthocortical and paracortical cells of merino wool fibers. J Agric Food Chem 57:2174–2180. https://doi.org/10.1021/jf803290h

Plowman JE, Deb-Choudhury S, Thomas A, Clerens S, Cornellison CD, Grosvenor AJ, Dyer JM (2010) Characterisation of low abundance wool proteins through novel differential extraction techniques. Electrophoresis 31:1937–1946. https://doi.org/10.1002/elps.201000053

Plowman JE, Flanagan LM, Paton LN, Fitzgerald AC, Joyce NI, Bryson WG (2003) The effect of oxidation or alkylation on the separation of wool keratin proteins by two-dimensional gel electrophoresis. Proteomics 3:942–950. https://doi.org/10.1002/pmic.200300419

Plowman JE, Harland DP, Ganeshan S, Woods JL, van Shaijik B, Deb-Choudhury S, Thomas A, Clerens S, Scobie DR (2015) The proteomics of wool fibre morphogenesis. J Struct Biol 191:341–351. https://doi.org/10.1016/j.jsb.2015.07.005

Powell BC (1996) The keratin proteins and genes of wool and hair. Wool Tech Sheep Breed 44:100–118

Powell BC, Rogers GE (1997) The role of keratin proteins and their genes in the growth, structure and properties of hair. In: Formation and structure of human hair. Birkhäuser, Basel, pp 59–148

Rogers MA, Langbein L, Praetzel-Wunder S, Giehl K (2008) Characterization and expression analysis of the hair keratin associated protein KAP26.1. Br J Dermatol 159:725–729. https://doi.org/10.1111/j.1365-2133.2008.08743.x

Rogers MA, Langbein L, Winter H, Ehmann C, Praetzel S, Korn B, Schweizer J (2001) Characterization of a cluster of human high/ultrahigh sulfur keratin-associated protein genes embedded in the type I keratin gene domain on chromosome 17q12-21. J Biol Chem 276:19440–19451. https://doi.org/10.1074/jbc.M100657200

Rogers MA, Langbein L, Winter H, Ehmann C, Praetzel S, Schweizer J (2002) Characterization of a first domain of human high glycine-tyrosine and high sulfur keratin-associated protein (KAP) genes on chromosome 21q22.1. J Biol Chem 277:48993–49002. https://doi.org/10.1074/jbc.M206422200

Yu Z, Gordon SW, Nixon AJ, Bawden CS, Rogers MA, Wildermoth JE, Maqbool NJ, Pearson AJ (2009) Expression patterns of keratin intermediate filament and keratin associated protein genes in wool follicles. Differentiation 77:307–316. https://doi.org/10.1016/j.diff.2008.10.009

Yu Z, Wildermoth JE, Wallace OAM, Gordon SW, Maqbool NJ, MacLean P, Nixon AJ, Pearson AJ (2011) Annotations of sheep keratin intermediate filament genes and their patterns of expression. Exp Dermatol 20:582–588. https://doi.org/10.1111/j.1600-0625.2011.01274.x

Zeuner FE (1963) A history of domesticated animals. Hutchinson, London

Zhou H, Gong H, Yan W, Luo Y, Hickford JG (2012) Identification and sequence analysis of the keratin-associated protein 24-1 (KAP24-1) gene homologue in sheep. Gene 511:62–65

Proteomics of Meat Products

Gianluca Paredi, Federica Mori, and Andrea Mozzarelli

Abstract Processed meat, representing about 22% of per capita meat consumption, undergoes technological transformations, such as salting, drying, cooking, or frying. Proteomic methods have been applied to monitor the effects of the processing on meat from cow, yak, pig, lamb, goat, and chicken with the aim of correlating protein pattern with meat quality as well as of identifying adulterations.

Keywords Meat • Processed meat products • Farm animal proteomics • 2DE • LC-MS/MS • Meat adulteration • Meat quality marker

Abbreviations

2DE	Two-dimensional electrophoresis
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GST	Glutathione S-transferase
OECD	Economic Co-operation and Development
LC-MS/MS	Liquid chromatography mass spectrometry
LL	Longissimus lumborum
LTOH	Longtime ohmic cooking
MALDI TOF/TOF	Matrix-assisted laser desorption ionization time of flight
Mb	Myoglobin
MRM	Multiple reaction monitoring
MS	Mass spectrometry
NMR	Nuclear magnetic resonance
PCR	Polymerase chain reaction
PGAM	Phosphoglucomutase
PYGM	Glycogen phosphorylase

G. Paredi • F. Mori • A. Mozzarelli (✉)
Department of Food and Drug, University of Parma, Parma, Italy

Interdepartmental Center SITEIA.PARMA, University of Parma, Parma, Italy
e-mail: andrea.mozzarelli@unipr.it

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_15

PM	Psoas major
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
STOH	Short-time ohmic cooking
SRM	Selected reaction monitoring
TMT	Tandem mass tag
WB	Water bath
WHC	Water-holding capacity

1 Introduction

Meat consumption in the world has increased from 24.2 kg per capita in 1964/1966 to 42.9 kg in 2012, with a predicted small increase to 45.3 kg in 2030 (FAO 2014). Not surprisingly, there are significant differences among countries. In industrialized countries meat consumption has increased from 61.5 kg per capita in 1964/1966 to 76.2 kg in 2012, whereas in developing countries meat consumption, in the same time window, increased from 10.2 to 33.5 kg. If China and Brazil are excluded, the increase in developing countries is only about twofold. Considering year 2012, worldwide, cattle and buffalo account for 22% of meat source, sheep and goat for 4.6%, pork for 36.3%, and poultry for 35.2%. The bovine meat source remains constant or slightly decreases, whereas poultry meat and pork have both increased significantly. Meat is used either fresh or after being processed (Fig. 1). In the United States, the amount of red meat consumption from processed source is 22% (Daniel et al. 2011), with similar values for Europe (Linseisen et al. 2002). Meat products can be classified depending on the applied processing technology, according to Fig. 1 (Heinz and Hautzinger 2007).

Therefore, there is a large variety of processed meat products, some of them manufactured worldwide, whereas others manufactured only in specific areas. Depending on the processing technology, meat proteins undergo extensive modifications that have a deep impact on meat quality, shelf-life, nutrition, and health impact. Meat processing for prolonged storage is a strategy applied long before the industrial revolution as the only procedure allowing meat to be a source for the diet over extended periods.

In this chapter, we report proteomic studies on several meat products from different animal sources. In spite of the relevance of protein fingerprinting for the characterization of processed meat products (Paredi et al. 2012, 2013), for the direction of technological protocols toward higher quality, and for the detection of adulterations, proteomic studies are still very limited. We hope that this chapter, containing some representative examples, will stimulate further activities in this very promising field using powerful proteomic tools.

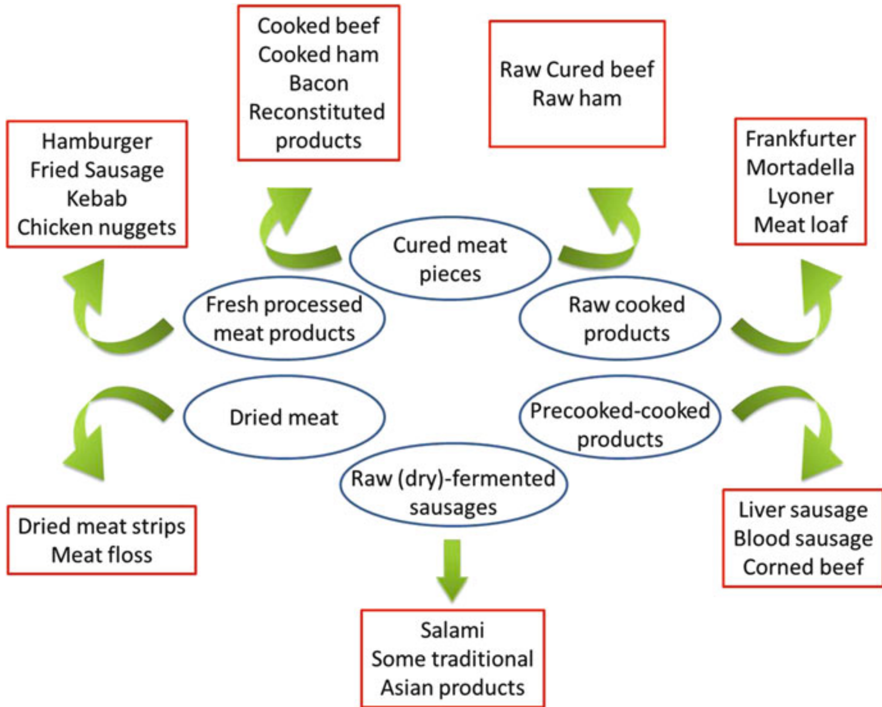

Fig. 1 Classification of processed meat products

2 Beef and Yak Meat

As the main protein source in the human diet, meat plays an essential role in daily lives. However, its production and consumption have significant environmental and health impacts. The Organization for Economic Co-operation and Development (OECD), in its seasonal agricultural report (OECD 2014), describes meat consumption relevance among different countries over the years. Beef, whose consumption was about 24.7 kg per capita in the United States in 2015, shows a consolidated role in the nutritional habits of almost every country. This leads to a natural tendency of improving knowledge and technologies related to processed beef meat products' quality and safety. Proteomics techniques were applied for understanding muscle-specific biochemical pathways correlated to meat quality.

Bresaola, a salted beef typical of Northern Italy, is a good example of a processed beef product. The processing steps involved in its production are (1) careful selection and skillful trimming cuts of meat, (2) dry salting, and (3) slow maturation. The processing can affect the nutrient bioavailability and the nutritional value. A protocol was developed for evaluating the protein bioavailability of Bresaola after in vitro digestion (Bordoni et al. 2014). The underlying assumption is that digestion and absorption, the so-called intestinal bioavailability, are the main

rate-limiting factors of the total bioavailability. The protocol is a combined application of SDS-PAGE and NMR spectroscopy techniques, such as proton nuclear magnetic resonance (H-NMR) and time domain NMR relaxometry. SDS-PAGE allowed characterization of proteins and peptides with molecular mass higher than 5 kDa during digestion, whereas H-NMR spectroscopy allowed the simultaneous detection of free amino acids, peptides, and proteins (even with molecular weight smaller than 5 kDa). Furthermore, time domain NMR provided information on juices accessibility to the components of the food matrix and on the disaggregation of biopolymers. This protocol is a good example of a new foodomics vision that considers the food as a "whole" and can be applied to different protein-rich foods for understanding their nutritional value in a deeper and more complete way.

One of the processing steps that most influences the bioavailability of proteins is cooking. Food-science research focuses on identifying some reliable monitoring parameters to minimize the nutritional loss of food products. Different advanced techniques for characterizing protein modifications caused by hydrothermal treatment were described (Deb-Choudhury et al. 2014). In this study, beef samples (*biceps femoris*) were minced and exposed to boiling water under reflux for different time periods and centrifuged for obtaining a soluble and insoluble fraction to be investigated separately. The efficacy of a LC-MS/MS analysis together with a redox proteomics approach was demonstrated and allowed the identification of oxidative and other heat-induced modifications of proteins. The results showed a general decrease in the level of essential amino acids and a steady increase in the oxidative modification score, both depending on the increased heat exposure time.

It is challenging and stimulating to exploit science innovation for understanding cooking-induced proteome changes and for developing new cooking methods that can be used for guaranteeing a more efficient industrial production and for meeting the consumer's requirement for safer and higher-quality meat products. The proteome changes of beef *M. longissimus dorsi* was investigated under two different cooking techniques, a water bath and an ohmic heating process, both at the same endpoint temperature (Tian et al. 2016). Ohmic cooking is an electric resistance cooking processing, and, besides a shortened cooking time and uniform temperature distribution, it has the advantage of producing cooked meat products with more attractive characteristics such as a more consistent appearance, more stable color, and more acceptable tenderness. The effects of WB (water bath), STOH (short-time ohmic cooking), and LTOH (long-time ohmic cooking), in terms of protein profiles, were investigated using 2D-PAGE. The differentially abundant proteins were identified by MALDI TOF/TOF. Both STOH and LTOH treatments led to an abundance of sarcoplasmic and myofibrillar proteins compared to WB treatment. These proteomic differences potentially explain some of the differential meat properties. For example, the abundance of myosin light chain 1 and myosin heavy chain 1 in the myofibrillar fraction might be associated with beef tenderness, whereas, in the sarcoplasmic fraction, the abundance of chaperone proteins and glycolytic enzymes possibly determines a better meat color. This evidence suggests that ohmic cooking is a great opportunity for the industrial production of meat products.

Food preparation and processing can play a significant role in the development of a novel form of food allergy: red meat allergy. This severe allergic reaction is

characterized by IgE antibodies directed against the carbohydrate galactose-α-1,3-galactose (α-Gal) epitope. This compound can be generated during different food processing, like heating or other treatments. The potential allergenicity of different beef preparations (raw, medium rare, fried, and boiled) was investigated by an immunoproteomics approach (Apostolovic et al. 2014). Protein profiles of the differently cooked beef preparations were determined by 2DE. Furthermore, to identify IgE-binding proteins, a methodology was developed that consists in the application of 2DE followed by immunoblot with sera of meat-allergic patients. The protein spots of interest were analyzed by peptide mass fingerprint, and high-resolution MS/MS analysis allowed the identification of 18 -IgE-binding proteins. A 2D immunoblot developed with anti-α-Gal antibodies was used for identifying α-Gal-containing proteins. The comparison between the two different types of immunoblots indicated that many IgE-binding proteins also contain the α-Gal epitope. This approach, build on immunoproteomics and MS/MS analysis, has allowed the identification of seven novel α-Gal-containing allergens, of which four were stable to heat treatment: creatine kinase M-type, aspartate aminotransferase, β-enolase, and α-enolase. The stability to heat treatment was investigated with the comparison of protein profiles between raw and cooked beef extracts. Aware of the fact that many aspects of red meat allergy remain to be elucidated, this study demonstrates the essential role of proteomics in understanding the biochemical pathways related to safer meat products and its potential role in food product quality control.

Another aspect that can undermine the safety of meat food products is the use of anabolic steroids in their production. In fact, despite the ban by the European Union, growth promoters are being illicitly used to enhance productivity at the expense of human health. The current available analytical techniques for drug residue control have some limits. In particular, they exhibit a great difficulty in the detection of low-dosage or unknown chemical structure anabolic compounds. Proteomics techniques can represent a possible solution for this problem because they allow the screening of indirect biological effects of growth promoter's administration, such as the alteration of protein expression. Therefore, the identification of protein biomarkers associated with steroid treatments represents a stimulating and promising challenge. Targeted mass spectrometry, such as selected reaction monitoring (SRM) also known as multiple reaction monitoring (MRM), has gained more and more importance because it enables precise protein quantification that can be standardized across laboratories. The study carried out by Stella et al. (2015) is a good example of the possible application of proteomics in this field. The aim of this study was to investigate differential muscle (*biceps brachii*) protein expression patterns between bulls treated with an ear implant (Revalor-XS) containing trenbolone acetate and estradiol and untreated animals. First, the building of a list of candidate protein markers of treatment was performed using a shotgun proteomics approach based on a combination of SDS-PAGE and LC-MS/MS. Before the application of mass spectrometry and after the extraction from SDS-PAGE gels, peptide mixtures were labeled with 6-plex TMT (tandem mass tag) reagents to allow the simultaneous quantification of up to six samples in the same LC-MS/MS analysis. It was possible to identify 28 candidate protein markers to be used to

discriminate between treated and untreated animals. Then, to support these results, the hypothetical candidate protein markers were subjected to a targeted proteomics approach using SRM. By applying HPLC coupled to a triple quadrupole mass spectrometer, 18 out of 28 potential protein markers were detected showing significantly different relative concentrations between the two animal groups. This study demonstrates the successful application of a combination of shotgun and targeted proteomics for drug residue analysis.

A highly relevant aspect for food product quality control is the identification of adulterated meat. The fraudulent blending of meat from different species is associated with both economical and ethical aspects. With regard to this specific field, the study of von Bargen and colleagues (2013) represents an extremely effective application of innovative proteomics techniques. The aim was the detection of species-specific biomarker peptides that can help by mass spectrometry the detection of trace amounts of horse and pork in beef meat samples. First, it was necessary to identify species-specific polymorphisms with a nontarget proteomics approach. Subsequently, these novel specific biomarker peptides were utilized for the development of a sensitive MRM-based method. This protocol, which allowed the identification of 12 biomarker peptides specific for pork and horse meat, showed sensitivities that are similar to the most sensitive PCR and ELISA methods available, highlighting the enormous potential of mass spectrometry.

Advanced proteomics techniques have been applied to investigate the basis of beef meat color. Meat color is one of the most relevant quality attributes in influencing the purchase decision of consumers. Therefore, understanding the molecular mechanisms responsible for its stability is a crucial goal for the meat industry from an economic perspective. In particular, proteomics techniques can be applied to correlate the differences in protein patterns with variations in color and biochemical traits influencing meat color. The sarcoplasmic proteome, containing many distinct proteins, including myoglobin and enzymes, controls some of the biochemical processes that influence meat color stability. Many studies have been carried out aimed at identifying differences in the abundance of sarcoplasmic proteins associated with different color stability. In particular, due to their color features and biochemical characteristics, the most representative examples of color-stable and color-labile beef meat are, respectively, beef *Longissimus lumborum* (LL) and *psoas major* (PM). The differential sarcoplasmic proteomes of LL and PM beef muscles were characterized by 2DE and tandem mass spectrometry (Joseph et al. 2012). The results indicated that a higher color stability of LL muscle is associated with the abundance of antioxidant proteins, glycolytic enzymes, and chaperones. These results were later confirmed by the analysis of the sarcoplasmic protein patterns of LL and PM muscles of Chinese Luxi Yellow cattle during postmortem storage (Wu et al. 2015).

Although beef *Longissimus lumborum* muscle shows a greater color stability when compared to other beef muscles, it was demonstrated that animal-to-animal variations can influence this peculiar feature. The biochemical basis of the variation in beef LL muscles color stability was further investigated through 2DE and tandem mass spectrometry (Canto et al. 2015). A significant difference was found between sarcoplasmic protein patterns of color-labile and color-stable LL muscles associated

with an overabundance of glycolytic enzymes in the color-stable meat. The application of proteomics techniques can help to identify breed, meat storage, and processing conditions for preserving an optimal beef color. For example, this can be achieved by the development of the proper muscle-specific processing strategies, including the addition of muscle-specific antioxidants or specific modified atmosphere packaging, and by the identification of potential predictors of meat discoloration.

Yak, an Asiatic member of the Bovidae biological family, is commonly associated with high-quality meat products. Therefore, the preservation of yak meat quality during postmortem aging is a primary goal for the meat industry. A combination of proteomics and bioinformatics was applied to investigate one of these features, the water-holding capacity (WHC) (Zuo et al. 2016). Poor WHC is a common issue related to postmortem storage and to meat food processing. The aim of the study was to obtain the differentially expressed proteins in yak muscle (*longissimus lumborum)* during postmortem storage. Samples, classified according to their specific drip loss capacity, were analyzed by 2DE, which allowed the determination of differences between protein patterns. MALDI TOF/TOF mass spectrometry allowed the identification of several proteins. Three proteins (myosin light chain, heat shock protein 27, and triosephosphate isomerase), for which identification was also validated by Western blot analysis, showed a great potential as protein biomarkers for WHC prediction. This peculiar application of proteomics techniques represents a further evidence of the forefront role of proteomics in meat science.

3 Pork

Sausage, salami, bacon, speck, and cooked and dry-cured hams are the main products obtained from pork. The percentage of processed versus fresh pork consumed per capita is about 20%. Some of the processed pork products are relatively cheap and widely consumed, such as salami, bacon, and sausages, and some other products are expensive and consumed by selective groups of persons, such as dry-cured ham because it is a delicatessen product perceived by consumers as high-quality meat food. Nowadays, considering only the European countries, at least 30 dry-cured hams have some sort of protected geographical indication. These include the Italian Parma and San Daniele hams, the Spanish Jamon Serrano, the French Bayonne, and the Portuguese Presunto. Several factors strongly influence the quality of the final product, the selection of pig breeds, the feeding, and the technological parameters applied in ham production, i.e., salt concentration, ripening time, and temperature and humidity for aging. A central role in the production process is played by salt, which is rubbed on the whole pork legs. Sodium chloride, the only salt used in Italian dry-cured ham processing, has several positive effects like bacteriostatic activity, reduction of water content, and a crucial impact on the sensory and quality parameters of the final product. Regarding the effects of

ripening phase on pork quality, proteolytic activity is retained during all the phases affecting the quality of the final product.

Proteomic methods have been applied to dry-cured ham aimed at characterizing the modifications happening to the pork leg during the ripening phase. The modifications at proteome level were analyzed for Parma and San Daniele dry-cured hams ripened for 6, 10, and 14 months (Di Luccia et al. 2005). A differential extraction of water-soluble and myofibrillar proteins was carried out. The protein extract was analyzed by 2DE in order to improve the identification of less abundant proteins. For the water-soluble fraction, a decrease in abundance was observed for glyceraldehyde-3-phosphate dehydrogenase (GAPDH), enolase, and creatine kinase. These findings are in agreement with previous studies that showed a strong proteolytic activity on these proteins. Moreover, at 14 months of ripening, two spots in the acid region appeared, and they were identified as α- and β-tropomyosin. According to the authors' hypothesis, this finding was due to salt diffusion during the ripening phase which increases the water solubility of these proteins. Data on sarcoplasmic proteins led to the supposition of a reduction of proteolytic activity for this class of proteins. Among myofibrillar proteins, modifications of several spots were observed. Particularly, after 6 months of ripening, spots identified in raw meat as myosin light chain 3 disappeared, and after 10 months several spots identified as myosin light chain fragments were detected. Finally, at 12–14 months, there was no evidence for spots related to the intact forms of α- and β-tropomyosin. These results indicate that the proteolytic activity strongly influences the myofibrillar composition of dry-cured ham during all the ripening period affecting the quality of the product. A LC-MS/MS approach was applied to follow the modification in dry-cured ham proteome during a period of 9 months (Gallego et al. 2016). The quantitative changes were characterized using a relative label-free quantification for 20 sarcoplasmic proteins. All these proteins showed a reduction in protein quantity during dry-cured ham processing due to the proteolytic activity of endopeptidase and exopeptidase toward the sarcoplasmic protein fraction. In particular, in the first 2 months, the abundance reduction was observed for GAPDH, glycogen phosphorylase (PYGM), phosphoglucomutase (PGAM), myoglobin (Mb), and glutathione S-transferase (GST). GAPDH, Mb, and GST were almost undetectable upon the salting phase. After 3.5 months, a reduction in the abundance of glucose 6-phosphate isomerase and triosephosphate isomerase was also observed. Aldolase showed a reduction of 50% in 9 months, whereas enolase, which exhibits a higher stability, showed a reduction of 2.8-fold in 9 months. Proteomic studies are in progress in our laboratory (Paredi et al., 2017) monitoring the effect of a pressure step applied on legs before the salting phase. For the characterization of the molecular events that take place during the first 18 days of ham processing (involving two steps of salting and resting), 2DE analysis and LC-MS were carried out on ham exudates, the protein-rich solution that is formed due to the salt-induced rupture of sarcoplasmic cells. Different protein profiles were obtained as a function of time for both pressed and not pressed exudate samples. Furthermore, protein profile differences were detected comparing pressed and not pressed samples.

Cooked ham is a delicatessen that requires a multistep process for its production: (i) injection of brine, a solution with a salt concentration in the range of 15–45%, containing also spices and casein, (ii) tumbling of meat pieces using different time lengths and distinct temperatures according to the quality of raw meat, (iii) cooking that is carried out on the assembled meat pieces at 68–72 °C to reduce microbial load, and (iv) refrigeration, packaging, and aging for at least 1 month. To date, only two investigations have been carried out aimed at characterizing proteome modifications related to the production process of cooked ham. Barbieri and Rivaldi (2008) applied SDS-PAGE and 2DE to investigate protein variation during all the steps of cooked ham processing. The protein solubility during tumbling was evaluated for a period of 6 h and was found that myofibrillar protein extraction increased during tumbling reaching a plateau after 5 h. The cooking step was monitored in the temperature range between 35 and 75 °C. Differences in protein extraction were observed particularly between 40 and 65 °C with an increase in quantity of actin and desmin and a reduction of myosin heavy chain at 45 °C. A deeper proteomic study was carried out on the first steps of cooked ham processing applying 2DE approach, using IPG strips in the pH range 4–7 (Pioselli et al. 2011). The proteome changes during brine injection and tumbling were determined on exudates released from hams as a function of brine concentrations (15, 30, and 45%), tumbling time (30% brine for either 4, 16, or 20 h of tumbling), and temperature (4 and 10 °C). The comparison of spot distribution showed that each working condition leads to a specific protein pattern with brine concentration strongly influencing the extraction of several proteins. For example, spots identified as myosin light chain 1 and myosin regulatory light chain were more intense in exudates generated with brine at 15 and 30% than 45%, and spots identified as tropomyosin alpha chain were more concentrated in exudate at 30% brine. Tumbling temperature caused changes in extraction of tropomyosin beta chain that showed a twofold increase at 10 °C compared to 4 °C. Finally, the length of tumbling modified the extraction of three proteins, identified as actin alpha skeletal muscle. The results of this proteomic analysis clearly show that each protein profile is associated with a specific processing condition that, in turn, is likely associated with a defined ham texture and quality of the final product.

Proteome modifications triggered by heat were investigated carrying out a proteomic analysis on cooked ham, mortadella, and wüsterl (Di Luccia et al. 2015). A preliminary analysis with SDS-PAGE revealed that the sarcoplasmic fraction of cooked products exhibited a low extractability in weak ionic strength buffer. It was supposed that heat treatment induced aggregation, conformational transitions, and an increase in protein-lipid interactions. 2DE profiles of cooked pork showed that the highest number of spots, 211 and 242, was observed for cooked hams Parma and Prague, respectively, whereas 70 and 68 spots were detected in mortadella and wüsterl that showed missing spots related to several myofibrillar proteins. This finding was explained by assuming that these proteins are endowed by a lower extractability in mortadella and wüsterl. Indeed, it was demonstrated that an ultrasonic step in the extraction protocol led to an increase in the number of observed 2DE spots.

AUT-PAGE/SDS-PAGE was used as an alternative to classic 2DE for the comparison of water-soluble proteins obtained from microbiologically fermented products, such as Naples-type salami; semi-fermented product, such as Italian “coppa”; dry-cured ham; and raw pork (Picariello et al. 2006). Results showed that dry-cured ham, after 12 months of ripening, showed an intense proteolysis of four sarcoplasmic proteins and the appearance of two spots with molecular weight of 37 and 38 kDa. The productive process of “Naples-type” salami requires a ripening phase of only 30 days. The protein pattern showed a reduction in intensity of spots identified as GAPDH, enolase B, and triosephosphate isomerase. Despite the shorter ripening period compared to dry-cured ham, the addition of lactic microflora led to a faster degradation of sarcoplasmic proteins. This is explained by the distribution of bacteria within the meat mixture, a procedure that cannot be achieved in dry-cured ham. Finally, the protein pattern for “coppa” was similar to that observed for raw meat, with only the disappearance of glycogen phosphorylase and the appearance of a spot with estimated molecular weight of 14 kDa.

4 Lamb and Goat Meat

A common method to process meat is by boiling for different lengths of times. Heat causes a variety of protein modifications that affect carbonyl content and cause aromatic residue alterations, formation of glycation end-products, and proteolysis. These effects impact on meat quality and nutrition. A proteomic analysis of lamb meat as well as of the proteins that are released in water was carried out via SDS-PAGE coupled to LC−MS/MS upon digestion and SDS-PAGE coupled to LC−MS/MS without digestion, respectively. The aim of the work was the characterization of the change in proteomic profile and amino acid modifications boiling either for 10 or 240 min (Yu et al. 2015). Not surprisingly, the sample boiled for 240 min allowed the identification of a lower number of proteins, 91, with respect to the samples boiled for 10 min, 156, close to the control, 142. Boiling for a long period might increase the extent of cross-linking among proteins, such as dityrosine, Schiff bases, disulfide bridges, and protein carbonylation, thus leading to lower solubility and increasing aggregation. It was also found that boiling causes the modification of several amino acids, such as oxidation of methionine, acetylation and carboxymethylation of lysine, and deamidation of glutamine.

The alternative method of meat cooking is roasting. A proteomic analysis was carried out for the characterization of the protein profile as well as protein modifications upon roasting lamb meat under control conditions (Yu et al. 2016). The proteomic analysis, carried out via SDS-PAGE coupled to LC-MS/MS, showed that upon heating protein concentration in urea-thiourea extracts decreased tenfold, likely due to myofibrillar protein aggregation and cross-linking. Furthermore, several amino acid modifications were observed, some of them previously observed when lamb meat was cooked by boiling (Yu et al. 2015). In another investigation, the effect of boiling goat muscle at either 100 °C for 30 min or at 121 °C and 15 PSI for 20 min

in an autoclave was investigated by 2DE and MALDI TOF/TOF spectrometry (Sarah et al. 2013). Meat storage at 4 °C was used as a control. The two heating conditions were selected as models for home cooking and industrial cooking. The focus was on the fate of actin as a function of different heating conditions. It was found that there was no significant difference between the two procedures on the amount of actin. Furthermore, analysis of the peptides generated from actin between residue 21 and 374 seems to suggest that these peptides might be exploited as biomarkers for the identification of goat in food products. This investigation was extended to identify the more heat-stable proteins (Sarah et al. 2014). These include myosin light chain, actin, tropomyosin, troponin T, myoglobin, and creatine kinase.

5 Poultry Meat

Chicken is a leading source of meat worldwide. Its consumption is via boiling or roasting with no traditional or industrial strategies for prolonged conservation. The only exception is represented by fried chicken, the so-called chicken nuggets, for which, however, no proteomic investigations have been carried out so far. However, because chicken meat has a lower commercial value than pig and especially beef, it is frequently added to beef or pork meat products as a fraud. Furthermore, because mechanically removed chicken meat has a lower commercial value than hand-removed chicken meat, the first product might be added to the second. Not surprisingly, given the analytical power of proteomics for both frauds, proteomic tools have been exploited for the detection of chicken meat in mixture of pig and beef meat and for the identification of markers of mechanical vs handmade procedure. In order to identify proteins or peptides that are specific of chicken meat with respect to beef and pork, the heat stability of chicken, pig, and beef proteins was determined and compared with fresh meat (Sentandreu et al. 2010). It was found that the peptide DQGTFEDFVGLR generated by trypsin digestion of myosin light chain-3 was chicken-specific. This peptide was always detected in chicken meat extracts and in pork mixture containing as low as 1% chicken meat. Thus, this peptide represents a robust biomarker for chicken meat. A second chicken peptide, the acetylated form of ALGQNPTNAEINK, was identified as a biomarker of chicken meat. It was detected by LC-MS/MS in cooked pork in the presence of 0.5–10% of cooked chicken meat. A similar approach was followed for the identification of peptide biomarkers specific of different classes of fish characterized by distinct commercial value (Carrera et al. 2007) and in the detection of the added soybean proteins in meat (Leitner et al. 2006).

Another study searched for protein biomarkers able to discriminate between chicken meat obtained via mechanic procedures and via manual interventions (Surowiec et al. 2011). Among the few proteins, which were identified by 2DE and LC-MS/MS, differently present in the chicken meat obtained with the two procedures, only hemoglobin was evaluated to be a robust biomarker for mechanically recovered meat.

6 Conclusions and Future Perspectives

Proteomic methods applied to meat have the potential of shaping and directing actions toward the overall goal of achieving high-quality products. This could occur at distinct but strongly interlinked levels: (i) animal selection, (ii) animal growth and feed, (iii) animal slaughtering, (iv) muscle to meat transition, (v) meat technological processing, and (vi) meat adulteration detection. However, somewhat surprisingly, relatively few proteomic studies have been carried out on the last two levels. This calls for intense activities toward the application of proteomic tools for the identification of robust protein markers in beef, broiler, and pork processing.

References

Apostolovic D, Tran TAT, Hamsten C et al (2014) Immunoproteomics of processed beef proteins reveal novel galactose-α-1,3-galactose-containing allergens. Allergy 69(10):1308–1315

Barbieri G, Rivaldi P (2008) The behaviour of the protein complex throughout the technological process in the production of cooked cold meats. Meat Sci 80(4):1132–1137

Bordoni A, Laghi L, Babini E et al (2014) The foodomics approach for the evaluation of protein bioaccessibility in processed meat upon in vitro digestion. Electrophoresis 35(11):1607–1614

Canto ACVCS, Suman SP, Nair MN et al (2015) Differential abundance of sarcoplasmic proteome explains animal effect on beef Longissimus lumborum color stability. Meat Sci 102:90–98

Carrera M, Cañas B, Piñeiro C et al (2007) De novo mass spectrometry sequencing and characterization of Species-Specific Peptides from Nucleoside Diphosphate Kinase B for the classification of commercial fish species belonging to the family Merlucciidae. J Proteome Res 6 (8):3070–3080

Daniel CR, Cross AJ, Koebnick C et al (2011) Trends in meat consumption in the United States. Public Health Nutr 14(4):575–583

Deb-Choudhury S, Haines S, Harland D et al (2014) Effect of cooking on meat proteins: mapping hydrothermal protein modification as a potential indicator of bioavailability. J Agric Food Chem 62(32):8187–8196. https://doi.org/10.1021/jf502668w

Di Luccia A, la Gatta B, Nicastro A et al (2015) Protein modifications in cooked pork products investigated by a proteomic approach. Food Chem 172:447–455

Di Luccia A, Picariello G, Cacace G et al (2005) Proteomic analysis of water soluble and myofibrillar protein changes occurring in dry-cured hams. Meat Sci 69(3):479–491

FAO (2014) Livestock commodities. In: Bruinsma J (ed) World agriculture: towards 2015/2030 an FAO perspective. Earthscan Publications Ltd, London, pp 85–97

Gallego M, Mora L, Concepción Aristoy M et al (2016) The use of label-free mass spectrometry for relative quantification of sarcoplasmic proteins during the processing of dry-cured ham. Food Chem 196:437–444

Heinz G, Hautzinger P (2007) Categories of processed meat products. In: Meat processing technology for small to medium scale producers. Pacific Faaootunrofaat, Bangkok, pp 97–102

Joseph P, Suman SP, Rentfrow G et al (2012) Proteomics of muscle-specific beef color stability. J Agric Food Chem 60(12):3196–3203

Leitner A, Castro-Rubio F, Marina ML et al (2006) Identification of marker proteins for the adulteration of meat products with soybean proteins by multidimensional liquid chromatography−tandem mass spectrometry. J Proteome Res 5(9):2424–2430

Linseisen J, Kesse E, Slimani N et al (2002) Meat consumption in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts: results from 24-hour dietary recalls. Public Health Nutr 5(6b):1243–1258

OECD (2014) Meat consumption

Paredi G, Raboni S, Bendixen E et al (2012) “Muscle to meat” molecular events and technological transformations: the proteomics insight. J Proteome 75(14):4275–4289

Paredi G, Sentandreu M-A, Mozzarelli A et al (2013) Muscle and meat: new horizons and applications for proteomics on a farm to fork perspective. J Proteomics 88:58–82

Paredi G, Benoni R, Pighini G et al (2017) Proteomics of Parma Dry-Cured Ham: analysis of salting exudates. J Agric Food Chem 65(30):6307–6316

Picariello G, De Martino A, Mamone G et al (2006) Proteomic study of muscle sarcoplasmic proteins using AUT-PAGE/SDS-PAGE as two-dimensional gel electrophoresis. J Chromatogr B 833(1):101–108

Pioselli B, Paredi G, Mozzarelli A (2011) Proteomic analysis of pork meat in the production of cooked ham. Mol BioSyst 7(7):2252–2260

Sarah S, Karsani S, Amin I et al (2014) A proteomic based assessment on changes in myofibrillar proteins of goat Longissimus muscle as affected by heat treatments. J Anim Plant Sci 24 (2):406–412

Sarah S, Karsani S, Mokhtar N et al (2013) Differences in thermostable actin profile of goat meat as observed in two-dimensional gel electrophoresis (2DE). Int Food Res J 20(2):897–901

Sentandreu MA, Fraser PD, Halket J et al (2010) A proteomic-based approach for detection of chicken in meat mixes. J Proteome Res 9(7):3374–3383

Stella R, Biancotto G, Arrigoni G et al (2015) Proteomics for the detection of indirect markers of steroids treatment in bovine muscle. Proteomics 15(13):2332–2341

Surowiec I, Koistinen KM, Fraser PD et al (2011) Proteomic approach for the detection of chicken mechanically recovered meat. Meat Sci 89(2):233–237

Tian X, Wu W, Yu Q et al (2016) Quality and proteome changes of beef M. longissimus dorsi cooked using a water bath and ohmic heating process. Innovative Food Sci Emerg Technol 34:259–266

von Bargen C, Dojahn J, Waidelich D et al (2013) New sensitive high-performance liquid chromatography–tandem mass spectrometry method for the detection of horse and pork in halal beef. J Agric Food Chem 61(49):11986–11994

Wu W, Gao X-G, Dai Y et al (2015) Post-mortem changes in sarcoplasmic proteome and its relationship to meat color traits in M. semitendinosus of Chinese Luxi yellow cattle. Food Res Int 72:98–105

Yu T-Y, Morton JD, Clerens S et al (2015) Proteomic investigation of protein profile changes and amino acid residue level modification in cooked lamb meat: the effect of boiling. J Agric Food Chem 63(41):9112–9123

Yu T-Y, Morton JD, Clerens S et al (2016) Proteomic investigation of protein profile changes and amino acid residue-level modification in cooked lamb longissimus thoracis et lumborum: the effect of roasting. Meat Sci 119:80–88

Zuo H, Han L, Yu Q et al (2016) Proteome changes on water-holding capacity of yak longissimus lumborum during postmortem aging. Meat Sci 121:409–419

Proteomics in Fish and Aquaculture Research

Pedro M. Rodrigues, Samuel A.M. Martin, Tomé S. Silva, Surintorn Boonanuntanasarn, Denise Schrama, Márcio Moreira, and Cláudia Raposo

Abstract The demand for animal protein for human consumption is currently on the rise fueled mainly by an exponential increase of the world population. The higher demand of fishery products and capture restrictions as a result of wild fish stock exploitation made aquaculture an extremely important source of protein (mainly fish, shellfish, and algae) available in human diet. Production statistics database from FAO states a value of about 97.2 million tonnes, of which around 70.0 million tonnes of the total food fish and 27.0 million tonnes of aquatic plants. The awareness that nowadays competitiveness is extremely dependent on scientific knowledge and new technologies made the number of manuscripts published in this area to rise almost exponentially. Aquaculture faces many challenges in order to continuously deliver a high-quality farmed fish through a sustainable production system. In order to achieve this goal, new management strategies need to be addressed, and state-of-the-art technologies like proteomics have been applied to study many factors like welfare, safety, nutrition, and diseases, which are directly responsible for the end-product quality.

P.M. Rodrigues (✉) • D. Schrama • C. Raposo
Departamento de Química e Farmácia, Universidade do Algarve, Faro, Portugal

Centro de Ciências do Mar do Algarve – CCMar, Universidade do Algarve, Faro, Portugal
e-mail: pmrodrig@ualg.pt

S.A.M. Martin
Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, UK

T.S. Silva
SPAROS Lda, Área Empresarial de Marim, Olhão, Portugal

S. Boonanuntanasarn
School of Animal Production Technology, Institute of Agricultural Technology, Suranaree University of Technology, Nakhon Ratchasima, Thailand

M. Moreira
IPMA – Portuguese Institute for the Sea and Atmosphere, EPPO – Aquaculture Research Station, Olhão, Portugal

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_16

Keywords Proteomics • Fish • Aquaculture • Fish proteomics • Fish biology • Fish allergens • Fish welfare • Fish diseases • Fish genomics

In this review we will address the latest proteomic studies published in each one of these influencing factors, giving a special importance to welfare since this is seen as a complex interaction of all the other factors. Also a brief review on the actual genomic resources is presented.

1 Proteomics and Fish Welfare

In fish with specific reference to aquaculture, the relationship between fish welfare, stress, and health is a complex interaction of many different variables making welfare in aquaculture a difficult subject to define. For all animals there are some fundamental definitions that should be considered when animal welfare is assessed. These include freedom from hunger, thirst, discomfort, pain, injury, disease, fear and distress, and ability to behave in a "normal way" and are discussed in detail by Ashley (2007). However several of these definitions are clearly difficult to transfer to fish (Berrill et al. 2012; Huntingford and Kadri 2014), but as a starting point, it does give clear lines that can be investigated. It can be hypothesized that when fish are in a good welfare state, they will perform well, have efficient conversion of food to growth, have a well-functioning immune system capable of dealing with immunological challenges, and in general result in a high-quality product. Compromised welfare not only reflects poorly on those who maintain the fish but is also of societal concern; hence there is a demand for high welfare status of farmed fish. Gauging the welfare status of fish is far from being easy, as we cannot easily perceive the emotional well-being of a large number of fish. The basic assumption for fish welfare is that the physiology and behavior do not significantly deviate from what is expected as normal (Prunet et al. 2012). The normal zone of tolerance for many physiological parameters needs to be assessed in relation to stress, nutrition, health, physiology, and behavior, with major changes potentially being indicators of compromised welfare. Developing markers, and how to interpret them, has been a goal in defining welfare in recent years mostly driven by the enormous growth in the aquaculture industry. So two aspects need to be considered: acute welfare issues (such as stress) and chronic welfare issues, which may be ongoing environmental changes such as water quality, environment, and health.

From a proteomic perspective, there are a number of core processes that can be examined and give indications of deviation from the expected normal, with the ambition of developing biomarkers for welfare assessment (Marco-Ramell et al. 2016). This term of "expected normal" requires a baseline to be established for the abundance and presence of proteins or how these proteins are posttranslationally modified. Such approaches have been carried out, and a recent example is reported in carp, where a multi-organ transcriptome and proteome have been assessed

(Kolder et al. 2016). Other papers have examined single tissues such as ovarian fluid (Johnson et al. 2014) and the brain (Gebriel et al. 2014). With the assumption that such analysis is performed on fish in a state of high welfare, the proteome and relationship of abundance of proteins can be useful in comparative investigations.

Acute stress in fish has been extensively studied, and there are clear markers including increased cortisol levels (Ellis et al. 2012), behavioral changes, food intake, and partitioning of energy requirements (Santos et al. 2010). From these observations, it is clear that multiple metabolic processes are changed, and many of these may be related to reduced food intake and not as a direct response to the stressor. Fish handling, netting, and reducing water volume are known to be major acute stressors of fish and are likely to induce perturbations. Repeated handling in Senegalese sole (*Solea senegalensis*) was examined for stress-induced responses in the liver (Cordeiro et al. 2012) where fish were repeatedly handled once a week for 4 weeks. Although the stress events were described as scarce, they could represent handling conditions in the commercial environment. Over 300 proteins were found to be consistently modulated in expression between handled and control fish relating to cellular response to redox stress, and a large number of heat shock proteins (HSPs) were found altered. These results were in line with earlier studies (Alves et al. 2010) where both repeated handling and crowding were used as stressor.

As fish are ectotherms, they have a temperature range where they have maximal performance and when confronted with annual changes in water temperature will acclimate accordingly (Johnston and Dunn 1987). However in culture conditions they cannot move to a more suitable temperature. Incorrect temperature induces stress by increasing oxygen demand, along with overall metabolic rate. Additionally, lower temperatures can also result in an unbalanced fish physiology. Fish behavior and its relation to temperature choice have been shown also to be linked with immune capacity as described for zebra fish (Boltaña et al. 2013). A number of studies have examined the proteome in relation to changing temperatures. Larval sea bream has been examined in relation to warming ocean temperatures (Madeira et al. 2016). Although focused on future climate change, the paper has relevance for aquaculture welfare. Whole animal proteomics showed that larval fish were unable to modify proteins relating to energy metabolism as would have been anticipated with warmer temperature, resulting in other physiological stresses. Although only 15 proteins were identified, some interesting conclusions were drawn. HSPs and protein degradation-related proteins were increased, suggesting dysregulation of protein folding. Other stress-related processes were changed including intracellular transport and porphyrin metabolism indicating reduced oxygen transport. Wild sturgeon larvae exposed to different temperatures (18 and 26 °C) in combination with selenium (Silvestre et al. 2010) were examined for proteomic changes induced by temperature, following 2-DE. Fifteen proteins were identified suggesting processes relating to protein folding, protein turnover (protein synthesis and protein degradation), ATP supply, and structural proteins changed in abundance in response to heat and/or selenium. These possible biomarkers could act as early indicators of dysfunction of larval development. These examples which are based on studies of whole larval fish do not allow for tissue-specific responses to be ascertained and may mask important changes in key tissues.

Natural changes in water temperature can help interpret when fish are stressed or in a state of compromised welfare. The murrel (*Channa striatus*) native to northeast India inhabits streams emerging from hot springs and can live at temperatures up to 38 °C. In this work, fish were acclimated to this temperature and compared with fish at a stable aquaculture temperature of 25 °C. The proteome analysis revealed a panel of HSPs at higher levels in the warmwater fish as well as a number of antioxidant proteins (Mahanty et al. 2016). In temperate latitudes, natural populations of grayling (*Thymallus thymallus*) that naturally inhabit warm or cooler water were examined for muscle proteomics (Mäkinen et al. 2015). Nearly all proteins identified were associated with Gene ontology related to muscle development and are interpreted "as driving the population closer to or to the thermal gene expression optimum." As these fish naturally tolerate and live in such environments, it may be debatable if these can be viewed as markers of welfare or a normal biological response to natural environmental fluctuations.

Although most worries are related to increasing water temperature for fish, as oxygen levels drop and metabolic stress is more likely to occur, warmer water species such as sea bream can have welfare issues at decreased lower winter temperatures. Winter disease occurs when water temperatures drop below about 12 °C, and a combination of organ malfunctions occurs including ionic regulation by gills, poor digestion, and compromised immune function (Castillo et al. 2009). To assess the key hepatic changes during a lowering of temperature in sea bream, fish acclimated to 20 °C were transferred to 8 °C for 10 days before a proteome comparative analysis of fish at the two temperatures was performed. There was a clear shift in the proteome with more proteins being reduced in abundance than increased (Ibarz et al. 2010b). The proteins identified that were changed in abundance suggested that protein and amino acid metabolism was being altered as seen by increased abundance of proteasome components and trypsinogen; secondly, changes in antioxidant activity were evidenced by increase in catalase and glutathione *S*-transferase. Taken together the authors concluded the cold shock resulted in hepatic oxidative damage and might potentially impact on winter disease in sea bream. Diets have been developed (winter feed, WF) to help mitigate these conditions, based on high marine protein and krill oil diets (Silva et al. 2014a) on which the fish grew and performed better. To define at a proteomic level the impacts of the WF and define potential protein markers for improved performance, plasma (Schrama et al. 2016) and liver (Richard et al. 2016) proteins related to protein metabolism, lipid metabolism, and immune function were identified. Authors suggest that fish fed the WF diet had improved oxidative stress capacity and increased amino acid metabolism.

Behavior and emotions are key aspects of behavior with such activities being controlled by brain function. However with the brain being such a complex tissue and often overlooked by aquaculture-related researchers, there is little known regarding the fish brain proteome. An interesting example is the zebra fish being used as a model for sleep disorders where these fish were maintained in continuous light/dark conditions (Purushothaman et al. 2015). These researchers were interested in circadian biology and endogenous daily rhythms controlled by the internal

clock with 78 proteins found altered as a result of changed photoperiod. Several proteins related to γ-aminobutyric acid (GABA)ergic receptors were modified as shown by 2-DE, and further circadian clock genes found modified by real-time PCR. These results could be expanded to the aquaculture environment where fish are often kept on artificial photoperiods for enhanced food intake or for controlling key life history events (Lorgen et al. 2015). Changes in brain proteome have also been examined in carp following anoxia, a species that can survive anoxic conditions, but little knowledge is known on how the brain deals with the lack of oxygen. Smith et al. (2009) found a decrease in abundance of proteins involved in the glycolysis pathway as well as proteins related to repression of neuronal apoptosis and decrease in neuronal degradation, demonstrating coping strategies for this fish species to environmental extremes it can face in natural environment.

Fish diets in aquaculture have changed significantly in recent years, with a move away from wild-sourced marine fish meal and fish oil to terrestrial plants and oils. Such diets can have impacts on fish welfare as they may contain anti-nutritional factors that interfere with digestion and intestinal function (Krogdahl et al. 2015; Król et al. 2016). To assess the impacts on the proteome of fish feeding on such diets, both the liver and intestine have been examined for potential metabolic changes that could indicate changed metabolism and welfare issues. Rainbow trout fed soybean meal rich diets had proteome alterations in the liver (Martin et al. 2003; Vilhelmsson et al. 2004) suggesting changes in lipid-binding proteins and primary energy metabolism. The intestine tissues themselves have received little in the way of proteome analysis; however Vasanth et al. (2015) found that microbial feed additives were able to reduce Atlantic salmon intestinal inflammation and showed five proteins that could be associated with poor intestinal morphology. Interestingly calreticulin, a multifunctional protein involved in extracellular matrix, was also altered in the skin of salmon that were being fed functional feeds associated with reduced sea lice burden (Micallef et al. 2017). Starvation is also directly relevant to welfare in fish; however as in many examples above, the biology of fish species is so plastic in that there is debate when starvation in salmonids becomes a welfare issue. Short-term food withdrawal (2 weeks) has been examined in rainbow trout (Martin et al. 2001) where several enzymes including cathepsin D suggested changes in protein turnover were occurring. More recently the impact of a 4-week food withdrawal was assessed for the intestinal tissue proteome of rainbow trout. In this study several immune-related function proteins and cellular stress showed significant changes (Baumgarner et al. 2013).

2 Proteomics in Fish Nutrition

Nutrition is a central topic in aquaculture research due to its essential role in fish metabolism, growth, health, and welfare. As such, it is not surprising that proteomic techniques have been extensively applied in this field in order to measure biological effects associated with particular dietary treatments or nutritional factors. Given the

wide range of feeding behaviors, digestive physiologies, and nutritional tolerances displayed by different species of fish, as well as the continuous introduction of new alternative ingredients in fish feed formulations, the use of such untargeted approaches can be seen as particularly beneficial by increasing the probability of detecting unforeseen nutritional effects.

Though most proteomic studies in fish nutrition focus on the liver as the target tissue, given its central role in regulating metabolism and adapting to nutritional changes, the muscle seems to be another common target, due to its importance as a peripheral energy-demanding tissue and its role in growth processes. Of particular relevance to nutritional studies is also the gut/intestine, due to its direct contact with bulk digesta and its particular susceptibility to the presence of anti-nutritional factors, as well as its essential role in the immune system and in modulating nutrient intake. Besides these, skin mucus and blood plasma are also seen as attractive targets due to the possibility of sampling them through nonlethal methods, being particularly suited to study the effects of dietary treatments on fish welfare and health. Finally, some studies simply perform protein extraction and analysis of the whole-body proteome, particularly when analyzing larvae or small fish, due to the difficulty in isolating specific tissues.

An important area of research concerns the general physiological effects of feeding (Mente et al. 2017), starvation, and refeeding (Baumgarner et al. 2013; Enyu and Shu-Chien 2011; Martin et al. 2001), as well as the impact that dietary energy intake levels can have on fish nutritional status (Jury 2005; Jury et al. 2008; Kolditz et al. 2008). For example, the works of Martin et al. (2001) and of Enyu and Shu-Chien (2011) show that starvation affects not only energy metabolism (glycolysis, gluconeogenesis, electron transfer chain) and oxidative stress response (peroxiredoxin, catalase, heat shock proteins), as one would expect, but also pathways such as methionine metabolism and lysosomal proteolysis (cathepsin D). Also, some of these works underline the dynamic nature of the hepatic proteome in particular and the need to consider the effect of, e.g., subjecting fish to fasting prior to sampling on proteomic observations. In general, this line of research is essential to assist in the correct interpretation of proteome alterations in fish nutrition studies.

The introduction of alternative ingredients in fish feed formulations (such as plant proteins, vegetable oils, and processed animal proteins) is seen as an important topic in aquaculture, and many proteomic studies focus on this issue, given the potential for unexpected deleterious effects (Ghisaura et al. 2014; Jessen et al. 2012; Kolditz et al. 2007; Kwasek 2012; Martin et al. 2003; Nuez-Ortín et al. 2016; Vilhelmsson et al. 2004; Wulff et al. 2012). Some of the proteins that consistently seem to be affected by the replacement of fish meal by vegetable ingredients include apolipoproteins, fatty acid-binding proteins, heat shock proteins, nitric oxide synthase, homogentisate 1,2-dioxygenase, and methionine/homocysteine metabolism proteins (adenosylhomocysteinase and betaine-homocysteine methyltransferase). Still within this context is supplementation of feeds with amino acids, particularly those displaying low abundance in vegetable ingredients. In this sense, the effect of diets containing variable levels of lysine on the muscle

and whole-body proteome of zebra fish has been characterized (de Vareilles et al. 2012; Gómez-Requeni et al. 2011), showing a high impact not only on structural proteins (actin, myosin, tropomyosin) but also proteins such as apolipoprotein A-I, Pdlim7, and proteins associated to energy metabolism. Another recent concern is the use of genetically modified organisms in fish feeds and its possible impact on fish health and nutritional safety. One study on the effect of genetically modified soy (compared to a near-isogenic non-GM soy) on Atlantic salmon displayed a minimal impact on its hepatic proteome, which suggests this particular strain of GM soy induces no obvious deleterious impact on fish nutrition and health (Sissener 2009; Sissener et al. 2010).

Understanding the effects of dietary micronutrient levels on fish metabolism and health is essential in the context of ever-changing feed formulations, where the possibility of micronutrient deficiencies is not negligible. In this sense, studies of the dietary effects of micronutrient supplementation through proteomic approaches have been undertaken, with works published both on phosphorus (Veiseth-Kent et al. 2013; Ye et al. 2016) and vitamin K (Richard et al. 2014) supplementation.

A particular issue with fish larvae is their high phospholipid requirements, which complicate the formulation of adequate replacements for live feed. Given this, some researchers studied the effect of different levels of soybean lecithin supplementation on the liver proteome of pike perch (Hamza et al. 2010). Results showed growth differences between dietary treatments, which were attributed to observed changes at the level of proteins related to oxidative stress (increased peroxiredoxin and reduced GRP75 and glutathione *S*-transferase with increasing lecithin levels), energy metabolism (changes in the levels of pyruvate carboxylase, phosphoglucomutase, fructose-biphosphate aldolase, and propionyl-CoA carboxylase), and choline metabolism (increased level of sarcosine dehydrogenase with increasing lecithin levels).

There are also studies on the impact of functional feeds, which are formulated or supplemented with particular additives with the purpose of boosting the metabolic and immune status of fish, to help them cope with particularly stressful situations (Richard et al. 2016; Schrama et al. 2016) or ward off infections (Jensen 2015; Jensen et al. 2015; Provan et al. 2013). A particularly strong trend in the field of functional feeds is the use of probiotics (nonpathogenic microorganisms) and bioactive substances derived from microorganisms (e.g., β-glucan), given their putative effects in terms of fish health and even growth performance (Ghaedi et al. 2016; Hosseini et al. 2016; Sveinsdóttir et al. 2009).

Finally, there are studies which focus on the effects of other dietary additives on fish proteomes: ranging from nucleotides (Keyvanshokooh and Tahmasebi-Kohyani 2012) and carbon sources, like α-ketoglutarate (Ibarz et al. 2010a) and glycerol (Silva et al. 2012), to secondary plant metabolites, like maslinic acid (Matos et al. 2013; Rufino-Palomares et al. 2011). These underline the versatility of proteomic approaches as general tools in fish nutrition studies to screen for potential effects at the level of cellular stress and metabolism.

An important detail in proteomic studies of fish nutrition is that the proteomes are intrinsically dynamic and context-dependent, which can make the interpretation

of the results highly challenging. In this sense, improving the design of experiments and data analysis approaches can bring real benefits to fish nutrition studies that leverage proteomic techniques. One of the ways of dealing with this complexity and context-dependence is to include more than one reference (control group), such as a negative control and a positive control. For example, if one is interested in knowing whether a particular feed additive induces nutritional stress, it might make sense to include a positive control diet (i.e., basal diet with an additive known to be stress-inducing) beside the negative control diet (i.e., basal diet). With such approach, we can convert ambiguous questions ("are the treatment samples similar to the negative control samples?") into more objective ones ("are the treatment samples more similar to negative control samples than to positive control samples?"). Following this concept that nutritional effects on proteomes should be interpreted in relative terms compared to reference group(s), rather than in absolute terms, one also should consider, particularly in long-term studies, the possibility of taking and analyzing samples from the start of the experiment and use them as a reference group. Another important detail that can contribute toward correct interpretation of proteomic observations is the co-measurement of complementary information, from easy-to-measure zootechnical parameters (such as fish body weight, body length, condition factor, hepatosomatic index, etc.) to other biological information obtained through the use of high-throughput profiling techniques (metabolomics, transcriptomics). This type of information can be used, on one hand, to isolate the treatment effect from other confounding effects (e.g., when comparing two groups of different mean weight, it is important to ensure that the treatment effects cannot simply be explained by body weight differences) and, on the other hand, to confirm the plausibility and consistency of the interpretation of the results (e.g., if a certain pathway is shown to be affected both at the proteomic and transcriptomic levels, one can be much more certain that the observation is not spurious). With these improvements, and others, related to the technical evolution of higher-throughput gel-free techniques, application of proteomics to the problematics of fish nutrition can provide an invaluable complement to other classical and omics approaches.

2.1 Safety of Aquaculture Products and Fish Allergens

In aquaculture industry, safety is of enormous importance to prevent health hazards, such as biological (bacteria, parasites, and viruses), chemical (heavy metals, dioxins, and aromatic hydrocarbons), and physical (bones, plastic, and glass) hazards (Teklemariam et al. 2015). To control this, the Food and Agriculture Organization (FAO) of the United Nations stabilized a code of practice for fish and fishery products (FAO 2012) where handling of fresh and frozen fish is described following the rules of hazard analysis and critical control points (HACCP). The European Union established Directive EC No 2073/2005 for regulation of microbial contamination (European 2005) and recently published EC No 1379/2013 for labeling and traceability characteristics to control fishery and

aquaculture products (European 2013). Authentication and labeling of fish species these days is very important as the human population is easily mislead by seafood identity substitution, as more than 20000 species of fish and seafood are known to be consumed (Rasmussen and Morrissey 2008).

More recently, proteomics has been emerging in the aquaculture field as a promising approach toward a high-quality end product (Mazzeo and Siciliano 2016). To achieve this goal, these advanced technologies have been used to improve the knowledge regarding potential biomarkers for environmental monitoring, risk assessment, including allergens' detection, traceability, and authenticity (Addis et al. 2010; Mazzeo and Siciliano 2016; Tedesco et al. 2014). Proteomics has been shown in numerous studies to deepen the genomic and transcriptomic approaches since it allows the study of the proteome, which reflects the physiological state of a fish at a given moment, in response to a stimulus. Although the lack of available information at the genome level registered for the majority of the aquaculture species is an enormous obstacle, a deeper proteome coverage of these species was achieved due to complementary studies comprising proteomics, genomics, and transcriptomics (Barbosa et al. 2012; Rodrigues et al. 2012). In case of traceability and authentication, several proteomic-related studies have been performed using fish species such as perch (Berrini et al. 2006), cod, mackerel (Martinez and Jakobsen Friis 2004; Martinez et al. 2007), hake (Pineiro et al. 2001; Carrera et al. 2006), sea bass, sea bream, and tilapia among others (Mazzeo et al. 2008). Muscle samples were used to identify specific proteins, such as parvalbumin, actin, tropomyosin, and myosin light chains. Recently, overexpression of the parvalbumin protein was detected in farmed gilthead sea bream against the wild species using shotgun proteomics (Piovesana et al. 2016). Using 2-DE and MALDI-TOF-MS, species of hake and grenadier were differentiated by the analysis of parvalbumin patterns in white muscle. This differentiation was confirmed using de novo sequencing of nucleoside diphosphate kinase B (Mazzeo and Siciliano 2016). Procedures like 2-DE, MALDI-TOF-MS, and PCR have been contributing in an extensive way for fish authentication (Siciliano et al. 2016; Carrera et al. 2013).

Depending on the research aim, the protein expression levels (comparative proteomics) and the posttranslational modifications (PTMs) can be assessed (Barbosa et al. 2012).

Food allergies are a worldwide issue and it is increasing fast. In 90% of the cases, an allergic reaction is caused due to a food protein of the Big 8, which includes milk, eggs, peanuts, tree nuts, soy, wheat, fish, and shellfish (Ahsan et al. 2016). The majority of food allergic reactions are mediated by immunoglobulin E (IgE). In case of fish allergy, it is estimated to affect up to 2% of adults and up to 7% of infants (Ballmer-Weber et al. 2015) and might cause symptoms like asthma, diarrhea, abdominal pain, or even anaphylaxis (Kuehn et al. 2014). As these symptoms might be severe, it is important to characterize, identify, and quantify all protein allergens (Di Girolamo et al. 2015). Allergens are the proteins used to mediate the allergenicity, and the major fish allergen has been identified as parvalbumin (Kuehn et al. 2014). Proteomics can be an important tool to characterize fish allergens. The major fish allergen can now be detected in less than 2 h using proteomic approaches using selected MS/MS ion monitoring (SMIM) in a

linear ion trap (LIT) mass spectrometer (Swoboda et al. 2002; Carrera et al. 2011, 2012). All these approaches result in a new "omics" era, namely, the allergenomics. After extraction of the proteins and separation by 1-DE or 2-DE, the visualization of fish protein allergens can be performed using immunoblotting with sera of allergic patients, and the N-terminal amino acids can be sequenced after the Edman degradation. The quantification of these allergens can be done by an ELISA using specific antibodies. Characterization and mapping of the IgE epitopes can be performed using liquid chromatography combined with a tandem mass spectrometer after 2-DE separation of the fish proteins (Fig. 1) (Di Girolamo et al. 2015). A different method for absolute allergen quantification has been developed by

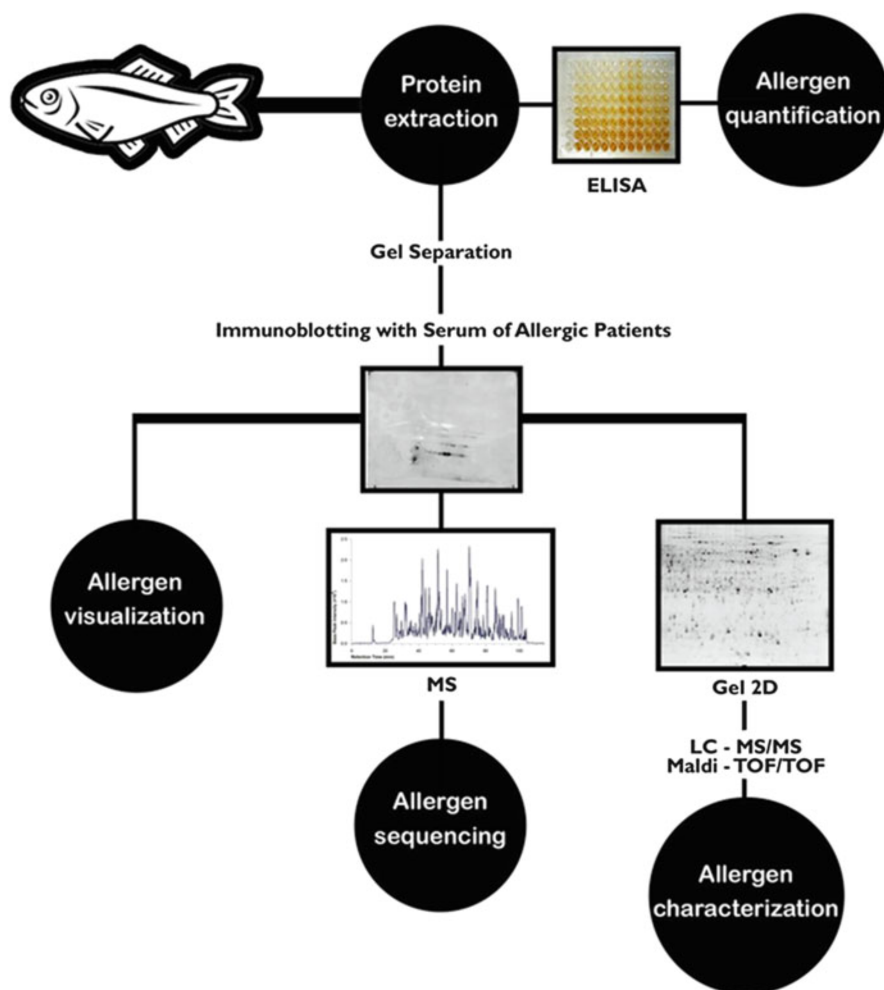

Fig. 1 Fish allergens identification using a proteomic approach

the way of triple quadrupole (QQQ) mass spectrometers using selected reaction monitoring (SRM) in plural multiple reaction monitoring (MRM) (Picotti et al. 2009), but this method is limited to known allergens (Ahsan et al. 2016). More recently Kobayashi and colleagues quantified 22 species of fish by their parvalbumin content using SDS-PAGE. They observed that parvalbumin is present in higher quantities in white muscle and that large-sized translocating species like tuna, swordfish, and salmon show lower quantities of this protein and lower reactivity of IgE from allergic patients (Kobayashi et al. 2016b). They showed in a different study that the amount of parvalbumin determines allergenicity and not the molecular differences of the allergen between species (Kobayashi et al. 2016a).

Different isoforms of parvalbumin have been identified in freshwater carp, and it has been shown that divergent developmental stages may express other isoforms (Brownridge et al. 2009). A commercial antibody against parvalbumin has been used to show its presence in various fish species and also demonstrates that heat treatment of the muscle alters the recognition of the antibody (Saptarshi et al. 2014). A few years earlier, it had been shown that in smoked fish species like salmon, mackerel, and haddock, a novel band of parvalbumin appeared at 30 kDa, and altered immunogenicity was shown on processed cod, salmon, trout, and pickled herring (Sletten et al. 2010). Recent proteomic-based studies identified enolase, tropomyosin, and creatine kinase as novel allergenic proteins, between others (Tomm et al. 2013), and characterized the allergenome of transgenic and non-transgenic fish, showing no difference in expression of parvalbumin and triose-phosphate isomerase, between others (Nakamura et al. 2009). The analysis of parvalbumin allergenicity in different fish species showed that the β-lineage is the most identified, and the International Union of Immunological Societies Allergen Nomenclature Subcommittee (www.allergen.org) contains 21 parvalbumins registered from 12 fish species (Kuehn et al. 2014).

3 Fish Diseases

Farmed fish are susceptible to a wide range of bacterial, viral, parasitic, and fungal infections, and losses through disease not only constitute a serious constraint to this industry, making a significant impact on the quality and volume of the fish produced in Europe and throughout the world (Hill 2005), but also have led people to question the safety of aquaculture (Adams and Thompson 2006).

Several pathogen detection methods (traditional, immunological, molecular) have been extensively used to improve fish health (Parrington and Coward 2002; Burge et al. 2016). And since scientific advances in aquatic health continue to close the gap with clinical and veterinary medicine, new techniques are becoming a reality that offers untold benefits to the aquaculture industry (Adams and Thompson 2006; Oskoueian et al. 2016). Proteomics, still mostly focused on gel-based techniques (Silva et al. 2014b), is one of those new tools and constitutes one of the best approaches for health management in aquaculture (Rodrigues et al. 2012, 2016; Silva et al. 2011) and to better understand fish diseases and epidemiology (Alves et al. 2010).

Fish diseases can be divided in two main areas: infectious fish disease and noninfectious fish diseases.

Infectious fish diseases are caused by pathogens such as virus, bacteria, fungi, and parasites and are the main source of economical loss in farm fish industry (Shinn et al. 2015).

Several proteomic studies related to infectious diseases have been described in the literature in areas like pathogenesis (Park et al. 2012), vaccine development (Lee 2001; Chen et al. 2004), disease diagnosis (Chen et al. 2004), disease resistance (Almeida et al. 2015), physiological response to pathogens (Rodrigues et al. 2012; Peng 2013; Addis et al. 2010), pathogen characterization (Dumpala et al. 2010; Buján et al. 2015; Fernández-Álvarez et al. 2016), immune proteins and immune system characterization and responses (Encinas et al. 2010; Coates and Decker 2016), disease biomarkers (Braceland et al. 2015), and organism response to disease treatment products (Varó et al. 2010).

In Table 1, a summary of some of the proteomic techniques applied in the study of infectious fish diseases is presented. Interestingly the number of proteomic studies in parasites is far lower than the number of studies in virus or bacteria. This is probably due to the availability of more DNA, RNA, and protein information from virus and bacteria in comparison with fish parasites in different databases (Burge et al. 2016).

Noninfectious fish diseases are mostly related to an external stimulus caused for instance by nutrition or the environment. These are normally associated with the production technology and can be the cause of several problems to aquaculture production as malformation, low growth rate, tumors, anorexia, poor quality of the product, or even high death rates (Forné et al. 2010).

The study of the influence of these external factors using proteomics is addressed in several papers such as the ones describing fish response to contaminants like PAH or PCBs (Galland et al. 2015), exposure to heavy metals or radioactive compounds (Hogstrand et al. 2002; Smith et al. 2015; Yadetie et al. 2016), exposure to toxins (Karim et al. 2011), response to stressors (Cordeiro et al. 2012), physical trauma (Wu et al. 2004) or fish development characterization used to reduce malformation incidence (Chicano-Gálvez et al. 2015), characterization of gas bubble disease caused by hyperoxygenation of the tanks (Salas-Leiton et al. 2009), and characterization of fish tumors (Stentiford et al. 2005; Lerebours et al. 2013).

As can be observed in Table 1, most proteomic studies in this field use top-down approaches (mainly 2-DE, followed by mass spectrometry). The major reason for this is related to the use of proteomics in aquaculture being still in its early days and progress in defining fish proteomes is expected to be slower than genome sequencing. Also, datasets from diseased fish and from fish pathogens need to be collected and available on a large scale before this technology can be fully used. In addition, although 2-D electrophoresis is the main technique used for detecting variation in the expression of proteins, this procedure is time-consuming and expensive, and reproducibility is a problem. Even in combination with mass spectrometry, only the more abundant proteins can be detected, thus indicating the need for new technologies (Zhou et al. 2012; Rodrigues et al. 2016).

Table 1 Summary of some of the proteomic techniques applied in the study of infectious fish diseases

Fish species	Aetiological agent	Disease	Tissue	Proteomic technique	Publication
Vírus					
Singapore grouper	Iridovírus	–	Iridovirus envelope proteins	1-DE-MALDI-TOF/TOF-MS/MS and LC-MALDI-TOF/TOF-MS/MS	Zhou et al. (2011)
Common carp	Spring viremia of carp virus	Spring viremia of carp	Epithelioma papulosum cyprini cells	2-DE MALDI-TOF/TOF	Liu et al. (2013)
Atlantic salmon	Salmonid alphavirus subtype 3	PD disease	Serum	2-D nanoflow UHPLC-ESI-MS/MS	Braceland et al. (2013)
Zebra fish	Rhabdovirus	Viral hemorrhagic septicemia	Whole body	2-D-DIGE MALDI-TOF/TOF	Encinas et al. (2010)
Zebra fish	Megalocytivirus	Infectious spleen and kidney necrosis virus	Whole body	2-DE MALDI-TOF/TOF	Xiong et al. (2011)
Bacteria					
Gilthead sea bream	*Moraxella* sp.	Bacteria	Kidney	2-D PAGE	(Addis et al. 2010)
Channel catfish	*Edwardsiella ictaluri*	Bacteria	Head kidney	2-D PMF MALDI-TOF-MS/MS	Booth and Bilodeau-Bourgeois (2009)
Turbot	*Edwardsiella tarda*	Bacteria	Bacterial strains	DIGE MALDI TOF/TOF	Buján et al. (2015)
Various species	*Flavobacterium columnare*	Columnaris disease	*F. columnare*	2-D LC ESI MS/MS and 2-DE MALDI TOF/TOF MS	Dumpala et al. (2010)
Sea bass	*Aeromonas salmonicida* subsp. *salmonicida*	Bacteria	Multiple organs	MALDI-TOF-MS	Fernández-Álvarez et al. (2016)
Various species	*Vibrio anguillarum*	Bacteria	*V. anguillarum* outer membrane proteins	2-DE LC-nano ESI-Q-TOF MS⁄MS	Kao et al. (2009)
Salmonids	*Yersinia ruckeri*	Bacteria	*Y. ruckeri*	Nano LC-ESI	Kumar et al. (2016)
Various species	*F. columnare*	Columnaris disease	*F. columnare* outer membrane proteins	SDS-PAGE RP-HPLC MS/MS	Liu et al. (2008)

(continued)

Table 1 (continued)

Fish species	Aetiological agent	Disease	Tissue	Proteomic technique	Publication
Various species	*Flavobacterium psychrophilum*	Bacteria	*F. psychrophilum*	2-DE LC-MS/MS	Ponnerassery et al. (2007)
Rainbow trout	*Aeromonas salmonicida*	Bacteria	Acute response system to vaccine	2-D-PAGE MALDI-TOF and ESI MS/MS	Russell et al. (2006)
Various species	*Edwardsiella tarda*	Hemorrhagic septicemia	Virulence determinants	2-DE ESI tandem MS	Srinivasa Rao et al. (2004)
Parasites					
Atlantic salmon	Sea lice, *Lepeophtheirus salmonis*	–	Epidermal mucus	LC-MS/MS	Provan et al. (2013)

4 Genomic Resources

Genomic resources provide the bioinformatic tools needed for proteomics. In general, the proteome is dynamic in different cells, organs, growth stages, and environmental conditions, and the differences in the proteome may be affected by a number of factors. For instance, differential splicing of RNA or alternative splicing generates multiple protein translated products produced from a single gene. There are posttranslational processes that result in the modification of protein products. Therefore, proteomic studies that complement genomic information can provide a useful tool to investigate the entire biological, physiological, and metabolic processes in an organism. Genomics is defined as the systematic study of genomes, which refers to the entire genetic material of an organism. A database of animal genome sizes, which have been estimated using haploid DNA contents (C-values, in picograms), has been constructed with genomes available for over 5600 animal species (http://www.genomesize.com/). Recent advances in DNA sequencing technologies and bioinformatics have brought revolutionary advances in genomics for several aquatic animals (Table 2). In addition, more genomic information for

Table 2 Genomic databases of aquatic animals

Species	Website
Zebra fish (*Danio rerio*)	http://zfin.org/cgi-bin/webdriver?MIval=aa-newmrkrselect.apg
Salmon (*Salmo salar*)	http://web.uvic.ca/grasp/ http://www.salmobase.org/
Fugu (*Takifugu rubripes*)	http://www.fugu-sg.org/index.html
Catfish (*Ictalurus* spp.)	http://catfishgenome.org/
Rainbow trout (*Oncorhynchus mykiss*)	https://www.genoscope.cns.fr/trout/
Oyster (*Crassostrea gigas*)	http://www.oysterdb.com/FrontHomeAction.do?method=home
Shrimp (*Litopenaeus vannamei*)	http://www.shrimp.ufscar.br/en/introduction/ http://shrimpgpat.sc.mahidol.ac.th/ShrimpGPATV2/
Marine genomic database	http://mgnew.clemson.edu/
Medaka (*Oryzias latipes*)	http://mepd.cos.uni-heidelberg.de/mepd/ http://utgenome.org/medaka/ http://asia.ensembl.org/Oryzias_latipes/Info/Index http://mbase.nig.ac.jp/mbase/medaka_top.html
Pufferfish (*Tetraodon nigroviridis*)	http://www.genoscope.cns.fr/externe/tetraodon/ https://www.genome.gov/11008305/
Stickleback (*Gasterosteus aculeatus*)	http://sticklebrowser.stanford.edu/cgi-bin/hgGateway?hgsid=21904
Tilapia (*Oreochromis niloticus*, *Astatotilapia burtoni*, *Metriaclima (Maylandia) zebra*, *Pundamila nyererei*, *and Neolamprologus brichardi*)	http://cichlid.umd.edu/CGCindex.html https://www.broadinstitute.org/tilapia/tilapia-genome-project

aquatic animals are expected to be accessible in the near future. The applications of genome technologies have implications for fisheries sciences and aquaculture such as the management of fish genetic resources, improvement of aquaculture productivity for food security, and environmental sustainability of the aquaculture industry (Wenne et al. 2007; Quinn et al. 2012). Genomic research areas include structural genomics, functional genomics, epigenomics, and metagenomics.

Structural genomics describes genome structure, organization, and evolution including genetic map construction, genome sequencing, and the determination of a protein and its three-dimensional structure. Recently, salmonid genomes provide the valuable sources of whole genome duplications, which have been an important landmark for vertebrate evolution (Berthelot et al. 2014; Lien et al. 2016). The National Center for Biotechnology Information (NCBI) genomic information organizes databases on whole genome sequences, maps, assemblies, and annotations of over 80 fishes (https://www.ncbi.nlm.nih.gov/genome/browse/). The genome sequences have been published for a number of aquatic animals (Spaink et al. 2014) (Table 2). In addition, Ensembl (http://asia.ensembl.org/index.html) has been available as a genome browser for supporting the comparative genomic information of vertebrates. With the extensively growing number of genomic databases, whole genome-based selection for aquaculture species is expected to be possible in the near future. The genetic information of the mitochondria of fish has also been extensively determined and used for taxonomy study. To date, a number of mitochondrial genomes of fish have been available, and the mitochondrial genomes or mitogenomes of fish have been provided in a database at http://mitofish.aori.u-tokyo.ac.jp/ (Satoh et al. 2016). Furthermore, DNA bar codes of fish are derived from the 5′ end of the cytochrome c oxidase subunit I gene of mitochondrial gene sequences (Kochzius et al. 2010); international participants have been called to submit the bar codes of all fishes worldwide at http://www.fishbol.org/.

Functional genomics describes gene expression, function, and interactions on a genome-wide scale. Functional genomics integrates bioinformation from large-scale and high-throughput analysis to explore dynamics of gene expression in a range of processes including transcription and translation under various experimental or environmental conditions. Functional genomics also investigates the complex relationship between genotype (both protein-coding genes and regulatory noncoding regions) and phenotype during various biological processes such as growth, development, metabolism, immunity, and reproduction (Rossi et al. 2007; Panhuis et al. 2011; Sun et al. 2013). The most common technologies for functional genomics in aquatic animals have been sequencing-based approaches such as expressed sequence tags (ESTs) and high-throughput sequencing of mRNA or RNA sequencing (RNA-Seq) and hybridization-based microarray analysis (Rossi et al. 2007; Panhuis et al. 2011; Liu et al. 2012; Qian et al. 2014; Salem et al. 2015). ESTs are generated from the 5′ or 3′ end of cDNA libraries. ESTs provide information of transcription-active regions or transcriptomics, which are a primary source for gene databases. The EST database has contributed important genomic bioinformation for the identification of gene expression. For instance, EST resources provide sequence databases for gene discovery, identification of single

nucleotide polymorphisms (SNPs) and microsatellites, microarray development, and genome annotation. To date, EST data in public databases have been available for various aquatic animals including at the NCBI (http://www.ncbi.nlm.nih.gov/dbEST/), the Unigene database (http://www.ncbi.nlm.nih.gov.unigene), the Gene Index database (http://compbio.dfci.harvard.edu/tgi/), the Sigenae EST Contig (http://publiccontigbrower.sigenae.org:2020/index.html), and the USDA National Animal Genome Project (http://www.genome.iastate.edu/bioinfo/). RNA sequencing (RNA-Seq) uses high-throughput sequencing techniques to provide transcriptomic information or the complete set of transcription in both quantitative and qualitative manners. RNA-Seq information have been published for various aquatic animals (Sun et al. 2013; Liu et al. 2013; Salem et al. 2015). Microarray has been a useful technique for analyzing gene expression profiles at a transcription level in an organism under various developmental stages, involving the immune system, disease resistance, and response to environmental conditions (Peatman et al. 2007; Drivenes et al. 2012; Matsumoto et al. 2014). Microarray analysis and construction should be compliant with the Minimum Information About a Microarray Experiments guidelines (MIAME guidelines) and meet the standards of the Microarrays Gene Expression Data (MGED) society. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) has been commonly used to evaluate the individual gene expression to confirm EST, RNA-Seq, and microarray analysis results. To date, functional genomics offers databases for the application of SNP analysis and quantitative trait loci (QTL) mapping which provide the valuable bioinformatics for powerful DNA markers.

Other genomic tools have been gaining attention recently in the aquatic sciences (Ardura et al. 2011; Williams et al. 2014; Moghadam et al. 2015). For example, the study of epigenetics focusing on heritable modification in gene expression does not involve changes to the DNA sequence. The epigenetic modifications affect gene expression due to several known mechanisms such as DNA methylation and histone modification (Chatterjee and Eccles 2015). Epigenomics refers to the whole bioinformation of epigenetics which offers an understanding of transcriptional regulation. In addition, metagenomics which has been referred to environmental genomics or ecogenomics provides the bioinformation on the genetic material of microbial ecology. Since there are a number of uncultivated microorganisms in nature, which cannot be determined by cultivation-based methods, PCR-directed sequencing (shotgun) offers a useful methodology to explore the entire microorganism community in nature (Xing et al. 2013).

5 Concluding Remarks

This chapter offers a brief overview of the potential of proteomics-based technologies in aquaculture management strategies describing its use in factors like welfare, nutrition, safety, or diseases, which pose some of the main constrains in this industry nowadays. Limitations to the use of this technology are mostly related

to the lack of gene annotation for most fish-farmed species. Here, we are sure to see a major change with the development of high-throughput sequencing facilities and sequencing cost reductions. It is also most likely that proteomic technologies will move away from 2-DE and rely more directly on gel-free approaches.

The need of integration with other OMICS technologies like genomics or metabolomics together with the more broad use of bottom-up proteomic techniques, the development of protein arrays, the increased capacity of centralized databases, networks, data repositories, and contingency plans, and, in particular, antibody microarrays might hold potential for a boost of application of proteomics in aquaculture.

Ethical issues also need to be considered as a possible hindrance. New practices such as genetic modification (transgenic and gene editing) will potentially lead to welfare issues, and the functional outputs of such changes will be increasingly assessed by proteomics.

Acknowledgments Rodrigues's lab was supported by project ALG-01-0247-FEDER-003520—ALISSA—Healthy and sustainable feeds for farmed fish, supported by Portugal and the European Union through FEDER, COMPETE 2020, and CRESC Algarve 2020, in the framework of Portugal 2020. Denise Schrama acknowledges scholarship on project ALG-01-0247-FEDER-003520—ALISSA—Healthy and sustainable feeds for farmed fish, supported by Portugal and the European Union through FEDER, COMPETE 2020, and CRESC Algarve 2020, in the framework of Portugal 2020. Márcio Moreira was financially supported by BONAQUA (0433_BONAQUA_5_E) and DIVERSIAQUA projects. Marcelo Sousa (MarceloArt84) for the artwork.

References

Adams A, Thompson KD (2006) Biotechnology offers revolution to fish health management. Trends Biotechnol 24(5):201–205. https://doi.org/10.1016/j.tibtech.2006.03.004

Addis MF, Cappuccinelli R, Tedde V, Pagnozzi D, Porcu MC, Bonaglini E, Roggio T, Uzzau S (2010) Proteomic analysis of muscle tissue from gilthead sea bream (Sparus aurata, L.) farmed in offshore floating cages. Aquaculture 309(1–4):245–252. https://doi.org/10.1016/j.aquaculture.2010.08.022

Ahsan N, Rao RSP, Gruppuso PA, Ramratnam B, Salomon AR (2016) Targeted proteomics: current status and future perspectives for quantification of food allergens. J Proteome 143:15–23. https://doi.org/10.1016/j.jprot.2016.04.018

Almeida AM, Bassols A, Bendixen E, Bhide M, Ceciliani F, Cristobal S, Eckersall PD, Hollung K, Lisacek F, Mazzucchelli G, McLaughlin M, Miller I, Nally JE, Plowman J, Renaut J, Rodrigues P, Roncada P, Staric J, Turk R (2015) Animal board invited review: advances in proteomics for animal and food sciences. Animal 9(1):1–17. https://doi.org/10.1017/S1751731114002602

Alves RN, Cordeiro O, Silva TS, Richard N, de Vareilles M, Marino G, Di Marco P, Rodrigues PM, Conceição LEC (2010) Metabolic molecular indicators of chronic stress in gilthead seabream (Sparus aurata) using comparative proteomics. Aquaculture 299(1–4):57–66. https://doi.org/10.1016/j.aquaculture.2009.11.014

Ardura A, Planes S, Garcia-Vazquez E (2011) Beyond biodiversity: fish metagenomes. PLoS One 6(8):e22592. https://doi.org/10.1371/journal.pone.0022592

Ashley PJ (2007) Fish welfare: current issues in aquaculture. Appl Anim Behav Sci 104 (3–4):199–235. https://doi.org/10.1016/j.applanim.2006.09.001

Ballmer-Weber BK, Fernandez-Rivas M, Beyer K, Defernez M, Sperrin M, Mackie AR, Salt LJ, Hourihane JOB, Asero R, Belohlavkova S, Kowalski M, de Blay F, Papadopoulos NG, Clausen M, Knulst AC, Roberts G, Popov T, Sprikkelman AB, Dubakiene R, Vieths S, van Ree R, Crevel R, Mills ENC (2015) How much is too much? Threshold dose distributions for 5 food allergens. J Allergy Clin Immunol 135(4):964–971. https://doi.org/10.1016/j.jaci.2014.10.047

Barbosa EB, Vidotto A, Polachini GM, Henrique T, de Marqui ABT, Tajara EH (2012) Proteomics: methodologies and applications to the study of human diseases. Rev Assoc Med Bras 58 (3):366–375

Baumgarner BL, Bharadwaj AS, Inerowicz D, Goodman AS, Brown PB (2013) Proteomic analysis of rainbow trout (Oncorhynchus mykiss) intestinal epithelia: physiological acclimation to short-term starvation. Comp Biochem Physiol Part D Genomics Proteomics 8(1):58–64. https://doi.org/10.1016/j.cbd.2012.11.001

Berrill IK, Cooper T, MacIntyre CM, Ellis T, Knowles TG, Jones EKM, Turnbull JF (2012) Achieving consensus on current and future priorities for farmed fish welfare: a case study from the UK. Fish Physiol Biochem 38(1):219–229. https://doi.org/10.1007/s10695-010-9399-2

Berrini A, Tepedino V, Borromeo V, Secchi C (2006) Identification of freshwater fish commercially labelled "perch" by isoelectric focusing and two-dimensional electrophoresis. Food Chem 96(1):163–168. https://doi.org/10.1016/j.foodchem.2005.04.007

Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, Noël B, Bento P, Da Silva C, Labadie K, Alberti A, Aury J-M, Louis A, Dehais P, Bardou P, Montfort J, Klopp C, Cabau C, Gaspin C, Thorgaard GH, Boussaha M, Quillet E, Guyomard R, Galiana D, Bobe J, Volff J-N, Genêt C, Wincker P, Jaillon O, Crollius HR, Guiguen Y (2014) The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. Nat Commun 5:3657. https://doi.org/10.1038/ncomms4657

Boltaña S, Rey S, Roher N, Vargas R, Huerta M, Huntingford FA, Goetz FW, Moore J, Garcia-Valtanen P, Estepa A, MacKenzie S (2013) Behavioural fever is a synergic signal amplifying the innate immune response. Proc R Soc B Biol Sci 280(1766). https://doi.org/10.1098/rspb.2013.1381

Booth NJ, Bilodeau-Bourgeois AL (2009) Proteomic analysis of head kidney tissue from high and low susceptibility families of channel catfish following challenge with Edwardsiella ictaluri. Fish Shellfish Immunol 26(1):193–196. https://doi.org/10.1016/j.fsi.2008.03.003

Braceland M, Bickerdike R, Tinsley J, Cockerill D, McLoughlin MF, Graham DA, Burchmore RJ, Weir W, Wallace C, Eckersall PD (2013) The serum proteome of Atlantic salmon, Salmo salar, during pancreas disease (PD) following infection with salmonid alphavirus subtype 3 (SAV3). J Proteome 94:423–436. https://doi.org/10.1016/j.jprot.2013.10.016

Braceland M, McLoughlin MF, Tinsley J, Wallace C, Cockerill D, McLaughlin M, Eckersall PD (2015) Serum enolase: a non-destructive biomarker of white skeletal myopathy during pancreas disease (PD) in Atlantic salmon Salmo salar L. J Fish Dis 38(9):821–831. https://doi.org/10.1111/jfd.12296

Brownridge P, de Mello LV, Peters M, McLean L, Claydon A, Cossins AR, Whitfield PD, Young IS (2009) Regional variation in parvalbumin isoform expression correlates with muscle performance in common carp (Cyprinus carpio). J Exp Biol 212(2):184–193. https://doi.org/10.1242/Jeb.021857

Buján N, Hernández-Haro C, Monteoliva L, Gil C, Magariños B (2015) Comparative proteomic study of Edwardsiella tarda strains with different degrees of virulence. J Proteome 127(Part B):310–320. https://doi.org/10.1016/j.jprot.2015.05.008

Burge CA, Friedman CS, Getchell R, House M, Lafferty KD, Mydlarz LD, Prager KC, Sutherland KP, Renault T, Kiryu I, Vega-Thurber R (2016) Complementary approaches to diagnosing marine diseases: a union of the modern and the classic. Philos Trans R Soc B 371(1689). https://doi.org/10.1098/rstb.2015.0207

Carrera M, Canas B, Pineiro C, Vazquez J, Gallardo JM (2006) Identification of commercial hake and grenadier species by proteomic analysis of the parvalbumin fraction. Proteomics 6(19):5278–5287. https://doi.org/10.1002/pmic.200500899

Carrera M, Canas B, Lopez-Ferrer D, Pineiro C, Vazquez J, Gallardo JM (2011) Fast monitoring of species-specific peptide biomarkers using high-intensity-focused-ultrasound-assisted tryptic digestion and selected MS/MS ion monitoring. Anal Chem 83(14):5688–5695. https://doi.org/10.1021/ac200890w

Carrera M, Canas B, Gallardo JM (2012) Rapid direct detection of the major fish allergen, parvalbumin, by selected MS/MS ion monitoring mass spectrometry. J Proteome 75(11):3211–3220. https://doi.org/10.1016/j.jprot.2012.03.030

Carrera M, Cañas B, Gallardo JM (2013) Proteomics for the assessment of quality and safety of fishery products. Food Res Int 54(1):972–979. https://doi.org/10.1016/j.foodres.2012.10.027

Castillo J, Teles M, Mackenzie S, Tort L (2009) Stress-related hormones modulate cytokine expression in the head kidney of gilthead seabream (Sparus aurata). Fish Shellfish Immunol 27(3):493–499. https://doi.org/10.1016/j.fsi.2009.06.021

Chatterjee A, Eccles MR (2015) DNA methylation and epigenomics: new technologies and emerging concepts. Genome Biol 16(1):103. https://doi.org/10.1186/s13059-015-0674-5

Chen Z, Peng B, Wang S, Peng X (2004) Rapid screening of highly efficient vaccine candidates by immunoproteomics. Proteomics 4(10):3203–3213. https://doi.org/10.1002/pmic.200300844

Chicano-Gálvez E, Asensio E, Cañavate JP, Alhama J, López-Barea J (2015) Proteomic analysis through larval development of Solea senegalensis flatfish. Proteomics 15(23–24):4105–4119. https://doi.org/10.1002/pmic.201500176

Coates CJ, Decker H (2016) Immunological properties of oxygen-transport proteins: hemoglobin, hemocyanin and hemerythrin. Cell Mol Life Sci 1–25. https://doi.org/10.1007/s00018-016-2326-7

Cordeiro O, Silva T, Alves R, Costas B, Wulff T, Richard N, de Vareilles M, Conceição LC, Rodrigues P (2012) Changes in liver proteome expression of Senegalese Sole (Solea senegalensis) in response to repeated handling stress. Mar Biotechnol 14(6):714–729. https://doi.org/10.1007/s10126-012-9437-4

de Vareilles M, Conceição LEC, Gómez-Requeni P, Kousoulaki K, Richard N, Rodrigues PM, Fladmark KE, Rønnestad I (2012) Dietary lysine imbalance affects muscle proteome in zebrafish (Danio rerio): a comparative 2D-DIGE study. Mar Biotechnol 14(5):643–654. https://doi.org/10.1007/s10126-012-9462-3

Di Girolamo F, Muraca M, Mazzina O, Lante I, Dahdah L (2015) Proteomic applications in food allergy: food allergenomics. Curr Opin Allergy Clin Immunol 15(3):259–266. https://doi.org/10.1097/aci.0000000000000160

Drivenes Ø, Taranger GL, Edvardsen RB (2012) Gene expression profiling of Atlantic cod (Gadus morhua) embryogenesis using microarray. Mar Biotechnol 14(2):167–176. https://doi.org/10.1007/s10126-011-9399-y

Dumpala PR, Gülsoy N, Lawrence ML, Karsi A (2010) Proteomic analysis of the fish pathogen Flavobacterium columnare. Proteome Sci 8(1):26. https://doi.org/10.1186/1477-5956-8-26

Ellis T, Yildiz HY, López-Olmeda J, Spedicato MT, Tort L, Øverli Ø, Martins CIM (2012) Cortisol and finfish welfare. Fish Physiol Biochem 38(1):163–188. https://doi.org/10.1007/s10695-011-9568-y

Encinas P, Rodriguez-Milla MA, Novoa B, Estepa A, Figueras A, Coll J (2010) Zebrafish fin immune responses during high mortality infections with viral haemorrhagic septicemia rhabdovirus. A proteomic and transcriptomic approach. BMC Genomics 11(1):518. https://doi.org/10.1186/1471-2164-11-518

Enyu YL, Shu-Chien AC (2011) Proteomics analysis of mitochondrial extract from liver of female zebrafish undergoing starvation and refeeding: proteomics analysis of mitochondrial extract of female zebrafish. Aquac Nutr 17(2):e413–e423. https://doi.org/10.1111/j.1365-2095.2010.00776.x

European P (2005) Commission regulation (EC) No 2073/2005 of 15 November 2005 on microbiological criteria for foodstuffs. Off J Eur Union L338:1–26
European P (2013) Council regulation (EU) No 1379/2013 of 11 December 2013 on the common organisation of the markets in fishery and aquaculture products. Off J Eur Union L354:1–21
FAO (2012) Code of practice for fish and fishery products. Quality and safety of fish and fish products. Fisheries and Aquaculture department. http://www.fao.org/fishery/quality_safety/en#container. Accessed 25 Oct 2016
Fernández-Álvarez C, Gijón D, Álvarez M, Santos Y (2016) First isolation of Aeromonas salmonicida subspecies salmonicida from diseased sea bass, Dicentrarchus labrax (L.), cultured in Spain. Aquaculture Rep 4:36–41. https://doi.org/10.1016/j.aqrep.2016.05.006
Forné I, Abián J, Cerdà J (2010) Fish proteome analysis: model organisms and non-sequenced species. Proteomics 10(4):858–872. https://doi.org/10.1002/pmic.200900609
Galland C, Dupuy C, Loizeau V, Danion M, Auffret M, Quiniou L, Laroche J, Pichereau V (2015) Proteomic analysis of the European flounder Platichthys flesus response to experimental PAH–PCB contamination. Mar Pollut Bull 95(2):646–657. https://doi.org/10.1016/j.marpolbul.2015.04.038
Gebriel M, Prabhudesai S, Uleberg K-E, Larssen E, Piston D, Bjørnstad AH, Møller SG (2014) Zebrafish brain proteomics reveals central proteins involved in neurodegeneration. J Neurosci Res 92(1):104–115. https://doi.org/10.1002/jnr.23297
Ghaedi G, Keyvanshokooh S, Azarm HM, Akhlaghi M (2016) Proteomic analysis of muscle tissue from rainbow trout (Oncorhynchus mykiss) fed dietary β-glucan. Iran J Vet Res 17(3):184
Ghisaura S, Anedda R, Pagnozzi D, Biosa G, Spada S, Bonaglini E, Cappuccinelli R, Roggio T, Uzzau S, Addis MF (2014) Impact of three commercial feed formulations on farmed gilthead sea bream (Sparus aurata, L.) metabolism as inferred from liver and blood serum proteomics. Proteome Sci 12(1):1. https://doi.org/10.1186/s12953-014-0044-3
Gómez-Requeni P, de Vareilles M, Kousoulaki K, Jordal A-EO, Conceição LEC, Rønnestad I (2011) Whole body proteome response to a dietary lysine imbalance in zebrafish Danio rerio. Comp Biochem Physiol Part D Genomics Proteomics 6(2):178–186. https://doi.org/10.1016/j.cbd.2011.02.002
Hamza N, Silvestre F, Mhetli M, Khemis IB, Dieu M, Raes M, Cahu C, Kestemont P (2010) Differential protein expression profile in the liver of pikeperch (Sander lucioperca) larvae fed with increasing levels of phospholipids. Comp Biochem Physiol Part D Genomics Proteomics 5(2):130–137. https://doi.org/10.1016/j.cbd.2010.03.005
Hill BJ (2005) The need for effective disease control in international aquaculture. Dev Biol (Basel) 121:3–12
Hogstrand C, Balesaria S, Glover CN (2002) Application of genomics and proteomics for study of the integrated response to zinc exposure in a non-model fish species, the rainbow trout. Comp Biochem Physiol B Biochem Mol Biol 133(4):523–535. https://doi.org/10.1016/S1096-4959(02)00125-2
Hosseini M, Kolangi Miandare H, Hoseinifar SH, Yarahmadi P (2016) Dietary Lactobacillus acidophilus modulated skin mucus protein profile, immune and appetite genes expression in gold fish (Carassius auratus gibelio). Fish Shellfish Immunol 59:149–154. https://doi.org/10.1016/j.fsi.2016.10.026
Huntingford F, Kadri S (2014) Defining, assessing and promoting the welfare of farmed fish. Rev Sci Tech (Int Off Epizootics) 33(1):233–244
Ibarz A, Costa R, Harrison AP, Power DM (2010a) Dietary keto-acid feed-back on pituitary activity in gilthead sea bream: effects of oral doses of AKG. A proteomic approach. Gen Comp Endocrinol 169(3):284–292. https://doi.org/10.1016/j.ygcen.2010.09.010
Ibarz A, Martin-Perez M, Blasco J, Bellido D, de Oliveira E, Fernandez-Borras J (2010b) Gilthead sea bream liver proteome altered at low temperatures by oxidative stress. Proteomics 10(5):963–975. https://doi.org/10.1002/pmic.200900528
Jensen LB (2015) Nutritional and environmental impacts on skin and mucus condition in Atlantic salmon (*Salmo salar* L.). Doctoral Thesis. University of Bergen, Norway

Jensen LB, Provan F, Larssen E, Bron JE, Obach A (2015) Reducing sea lice (*Lepeophtheirus salmonis*) infestation of farmed Atlantic salmon (Salmo salar L.) through functional feeds. Aquac Nutr 21(6):983–993. https://doi.org/10.1111/anu.12222

Jessen F, Wulff T, Mikkelsen JB, Hyldig G, Nielsen H (2012) Vegetable based fish feed changes protein expression in muscle of rainbow trout (Oncorhynchus mykiss). In: Rodrigues P, Eckersall D, de Almeida A (eds) Farm animal proteomics. Wageningen Academic, Wageningen, pp 134–137. https://doi.org/10.3920/978-90-8686-751-6_31

Johnson SL, Villarroel M, Rosengrave P, Carne A, Kleffmann T, Lokman PM, Gemmell NJ (2014) Proteomic analysis of chinook salmon (Oncorhynchus tshawytscha) ovarian fluid. PLoS One 9(8):e104155

Johnston I, Dunn J (1987) Temperature acclimation and metabolism in ectotherms with particular reference to teleost fish. In: Symposia of the society for experimental biology. Cambridge University Press, Cambridge, pp 67–93

Jury DR (2005) Proteomic analysis of the effects of diet in zebrafish liver. University of Akron, Akron

Jury DR, Kaveti S, Duan Z-H, Willard B, Kinter M, Londraville R (2008) Effects of calorie restriction on the zebrafish liver proteome. Comp Biochem Physiol Part D Genomics Proteomics 3(4):275–282. https://doi.org/10.1016/j.cbd.2008.07.003

Kao DY, Cheng YC, Kuo TY, Lin SB, Lin CC, Chow LP, Chen WJ (2009) Salt-responsive outer membrane proteins of Vibrio anguillarum serotype O1 as revealed by comparative proteome analysis. J Appl Microbiol 106(6):2079–2085. https://doi.org/10.1111/j.1365-2672.2009.04178.x

Karim M, Puiseux-Dao S, Edery M (2011) Toxins and stress in fish: proteomic analyses and response network. Toxicon 57(7–8):959–969. https://doi.org/10.1016/j.toxicon.2011.03.018

Keyvanshokooh S, Tahmasebi-Kohyani A (2012) Proteome modifications of fingerling rainbow trout (Oncorhynchus mykiss) muscle as an effect of dietary nucleotides. Aquaculture 324–325:79–84. https://doi.org/10.1016/j.aquaculture.2011.10.013

Kobayashi A, Kobayashi Y, Shiomi K (2016a) Fish allergy in patients with parvalbumin-specific immunoglobulin E depends on parvalbumin content rather than molecular differences in the protein among fish species. Biosci Biotechnol Biochem 80(10):2018–2021. https://doi.org/10.1080/09168451.2016.1189318

Kobayashi Y, Yang T, Yu C-T, Ume C, Kubota H, Shimakura K, Shiomi K, Hamada-Sato N (2016b) Quantification of major allergen parvalbumin in 22 species of fish by SDS–PAGE. Food Chem 194:345–353. https://doi.org/10.1016/j.foodchem.2015.08.037

Kochzius M, Seidel C, Antoniou A, Botla SK, Campo D, Cariani A, Vazquez EG, Hauschild J, Hervet C, Hjorleifsdottir S, Hreggvidsson G, Kappel K, Landi M, Magoulas A, Marteinsson V, Nolte M, Planes S, Tinti F, Turan C, Venugopal MN, Weber H, Blohm D (2010) Identifying fishes through DNA barcodes and microarrays. PLoS One 5(9):e12620. https://doi.org/10.1371/journal.pone.0012620

Kolder ICRM, van der Plas-Duivesteijn SJ, Tan G, Wiegertjes GF, Forlenza M, Guler AT, Travin DY, Nakao M, Moritomo T, Irnazarow I, den Dunnen JT, Anvar SY, Jansen HJ, Dirks RP, Palmblad M, Lenhard B, Henkel CV, Spaink HP (2016) A full-body transcriptome and proteome resource for the European common carp. BMC Genomics 17(1):701. https://doi.org/10.1186/s12864-016-3038-y

Kolditz C, Lefèvre F, Borthaire M, Médale F (2007) Transcriptome and proteome analysis of changes induced in trout liver by suppression of dietary fish oil. FASEB J 21(6):A1402–A1403

Kolditz CI, Paboeuf G, Borthaire M, Esquerre D, SanCristobal M, Lefevre F, Medale F (2008) Changes induced by dietary energy intake and divergent selection for muscle fat content in rainbow trout (Oncorhynchus mykiss), assessed by transcriptome and proteome analysis of the liver. BMC Genomics 9(1):506. https://doi.org/10.1186/1471-2164-9-506

Krogdahl Å, Gajardo K, Kortner TM, Penn M, Gu M, Berge GM, Bakke AM (2015) Soya saponins induce enteritis in Atlantic salmon (Salmo salar L.) J Agric Food Chem 63(15):3887–3902. https://doi.org/10.1021/jf506242t

Król E, Douglas A, Tocher DR, Crampton VO, Speakman JR, Secombes CJ, Martin SAM (2016) Differential responses of the gut transcriptome to plant protein diets in farmed Atlantic salmon. BMC Genomics 17(1):156. https://doi.org/10.1186/s12864-016-2473-0

Kuehn A, Swoboda I, Arumugam K, Hilger C, Hentges F (2014) Fish allergens at a glance: variable allergenicity of parvalbumins, the major fish allergens. Front Immunol 5:179. https://doi.org/10.3389/fimmu.2014.00179

Kumar G, Hummel K, Ahrens M, Menanteau-Ledouble S, Welch TJ, Eisenacher M, Razzazi-Fazeli E, El-Matbouli M (2016) Shotgun proteomic analysis of Yersinia ruckeri strains under normal and iron-limited conditions. Vet Res 47(1):100. https://doi.org/10.1186/s13567-016-0384-3

Kwasek KA (2012) The nutritional and genetic effects on body growth, reproduction and molecular mechanisms responsible for muscle growth in yellow perch Perca flavescens. The Ohio State University, Columbus

Lee KH (2001) Proteomics: a technology-driven and technology-limited discovery science. Trends Biotechnol 19(6):217–222. https://doi.org/10.1016/S0167-7799(01)01639-0

Lerebours A, Bignell JP, Stentiford GD, Feist SW, Lyons BP, Rotchell JM (2013) Advanced diagnostics applied to fish liver tumours: relating pathology to underlying molecular aetiology. Mar Pollut Bull 72(1):94–98. https://doi.org/10.1016/j.marpolbul.2013.04.016

Lien S, Koop BF, Sandve SR, Miller JR, Kent MP, Nome T, Hvidsten TR, Leong JS, Minkley DR, Zimin A, Grammes F, Grove H, Gjuvsland A, Walenz B, Hermansen RA, von Schalburg K, Rondeau EB, Di Genova A, Samy JKA, Olav Vik J, Vigeland MD, Caler L, Grimholt U, Jentoft S, Inge Våge D, de Jong P, Moen T, Baranski M, Palti Y, Smith DR, Yorke JA, Nederbragt AJ, Tooming-Klunderud A, Jakobsen KS, Jiang X, Fan D, Hu Y, Liberles DA, Vidal R, Iturra P, Jones SJM, Jonassen I, Maass A, Omholt SW, Davidson WS (2016) The Atlantic salmon genome provides insights into rediploidization. Nature 533(7602):200–205. https://doi.org/10.1038/nature17164

Liu GY, Nie P, Zhang J, Li N (2008) Proteomic analysis of the sarcosine-insoluble outer membrane fraction of Flavobacterium columnare. J Fish Dis 31(4):269–276. https://doi.org/10.1111/j.1365-2761.2007.00898.x

Liu S, Zhang Y, Zhou Z, Waldbieser G, Sun F, Lu J, Zhang J, Jiang Y, Zhang H, Wang X, Rajendran K, Khoo L, Kucuktas H, Peatman E, Liu Z (2012) Efficient assembly and annotation of the transcriptome of catfish by RNA-Seq analysis of a doubled haploid homozygote. BMC Genomics 13(1):595. https://doi.org/10.1186/1471-2164-13-595

Liu L, Li Q, Lin L, Wang M, Lu Y, Wang W, Yuan J, Li L, Liu X (2013) Proteomic analysis of epithelioma papulosum cyprini cells infected with spring viremia of carp virus. Fish Shellfish Immunol 35(1):26–35. https://doi.org/10.1016/j.fsi.2013.03.367

Lorgen M, Casadei E, Król E, Douglas A, Birnie Mike J, Ebbesson Lars OE, Nilsen Tom O, Jordan William C, Jørgensen EH, Dardente H, Hazlerigg David G, Martin Samuel AM (2015) Functional divergence of type 2 deiodinase paralogs in the Atlantic salmon. Curr Biol 25 (7):936–941. https://doi.org/10.1016/j.cub.2015.01.074

Madeira D, Araújo JE, Vitorino R, Capelo JL, Vinagre C, Diniz MS (2016) Ocean warming alters cellular metabolism and induces mortality in fish early life stages: a proteomic approach. Environ Res 148:164–176. https://doi.org/10.1016/j.envres.2016.03.030

Mahanty A, Purohit GK, Banerjee S, Karunakaran D, Mohanty S, Mohanty BP (2016) Proteomic changes in the liver of Channa striatus in response to high temperature stress. Electrophoresis 37(12):1704–1717. https://doi.org/10.1002/elps.201500393

Mäkinen H, Papakostas S, Vøllestad LA, Leder EH, Primmer CR (2015) Plastic and evolutionary gene expression responses are correlated in European grayling (Thymallus thymallus) subpopulations adapted to different thermal environments. J Hered. https://doi.org/10.1093/jhered/esv069

Marco-Ramell A, de Almeida AM, Cristobal S, Rodrigues P, Roncada P, Bassols A (2016) Proteomics and the search for welfare and stress biomarkers in animal production in the one-health context. Mol BioSyst 12:2024–2035. https://doi.org/10.1039/c5mb00788g

Martin SAM, Cash P, Blaney S, Houlihan DF (2001) Proteome analysis of rainbow trout (Oncorhynchus mykiss) liver proteins during short term starvation. Fish Physiol Biochem 24(3):259–270. https://doi.org/10.1023/A:1014015530045

Martin SAM, Vilhelmsson O, Medale F, Watt P, Kaushik S, Houlihan DF (2003) Proteomic sensitivity to dietary manipulations in rainbow trout. Biochim Biophys Acta 1651(1–2):17–29. https://doi.org/10.1016/S1570-9639(03)00231-0

Martinez I, Jakobsen Friis T (2004) Application of proteome analysis to seafood authentication. Proteomics 4(2):347–354. https://doi.org/10.1002/pmic.200300569

Martinez I, Slizyte R, Dauksas E (2007) High resolution two-dimensional electrophoresis as a tool to differentiate wild from farmed cod (Gadus morhua) and to assess the protein composition of klipfish. Food Chem 102(2):504–510. https://doi.org/10.1016/j.foodchem.2006.03.037

Matos E, Silva TS, Wulff T, Valente LMP, Sousa V, Sampaio E, Goncalves A, Silva JMG, de Pinho PG, Dinis MT, Rodrigues PM, Dias J (2013) Influence of supplemental maslinic acid (olive-derived triterpene) on the post-mortem muscle properties and quality traits of gilthead seabream. Aquaculture 396:146–155. https://doi.org/10.1016/j.aquaculture.2013.02.044

Matsumoto T, Feroudj H, Kikuchi R, Kawana Y, Kondo H, Hirono I, Mochizuki T, Nagashima Y, Kaneko G, Ushio H, Kodama M, Watabe S (2014) DNA microarray analysis on the genes differentially expressed in the liver of the pufferfish, Takifugu rubripes, following an intramuscular administration of Tetrodotoxin. Microarrays (Basel) 3(4):226–244. https://doi.org/10.3390/microarrays3040226

Mazzeo MF, Siciliano RA (2016) Proteomics for the authentication of fish species. J Proteome 147:119–124. https://doi.org/10.1016/j.jprot.2016.03.007

Mazzeo MF, De Giulio B, Guerriero G, Ciarcia G, Malorni A, Russo GL, Siciliano RA (2008) Fish authentication by MALDI-TOF mass spectrometry. J Agric Food Chem 56(23):11071–11076. https://doi.org/10.1021/jf8021783

Mente E, Pierce GJ, Antonopoulou E, Stead D, Martin SAM (2017) Postprandial hepatic protein expression in trout Oncorhynchus mykiss a proteomics examination. Biochem Biophys Rep 9:79–85. https://doi.org/10.1016/j.bbrep.2016.10.012

Micallef G, Cash P, Fernandes JM, Rajan B, Tinsley JW, Bickerdike R, Martin SA, Bowman AS (2017) Dietary yeast cell wall extract alters the proteome of the skin mucous barrier in Atlantic salmon (Salmo salar): increased abundance and expression of a calreticulin-like protein. PLoS One 12(1):e0169075. https://doi.org/10.1371/journal.pone.0169075

Moghadam H, Morkore T, Robinson N (2015) Epigenetics—potential for programming fish for aquaculture. J Mar Sci Eng 3:175–192. https://doi.org/10.3390/jmse3020175

Nakamura R, Satoh R, Nakajima Y, Kawasaki N, Yamaguchi T, Sawada J, Nagoya H, Teshima R (2009) Comparative study of GH-transgenic and non-transgenic amago salmon (Oncorhynchus masou ishikawae) allergenicity and proteomic analysis of amago salmon allergens. Regul Toxicol Pharmacol 55(3):300–308. https://doi.org/10.1016/j.yrtph.2009.08.002

Nuez-Ortín WG, Carter CG, Wilson R, Cooke I, Nichols PD (2016) Preliminary validation of a high Docosahexaenoic acid (DHA) and -Linolenic acid (ALA) dietary oil blend: tissue fatty acid composition and liver proteome response in Atlantic salmon (Salmo salar) Smolts. PLoS One 11(8):e0161513. https://doi.org/10.1371/journal.pone.0161513

Oskoueian E, Eckersall PD, Bencurova E, Dandekar T (2016) Application of proteomic biomarkers in livestock disease management. In: Salekdeh GH (ed) Agricultural proteomics, Environmental stresses, vol 2. Springer, Cham, pp 299–310. https://doi.org/10.1007/978-3-319-43278-6_14

Panhuis TM, Broitman-Maduro G, Uhrig J, Maduro M, Reznick DN (2011) Analysis of expressed sequence tags from the placenta of the live-bearing fish Poeciliopsis (Poeciliidae). J Hered. https://doi.org/10.1093/jhered/esr002

Park SB, Aoki T, Jung TS (2012) Pathogenesis of and strategies for preventing Edwardsiella tarda infection in fish. Vet Res 43(1):67. https://doi.org/10.1186/1297-9716-43-67

Parrington J, Coward K (2002) Use of emerging genomic and proteomic technologies in fish physiology. Aquat Living Resour 15(3):193–196

Peatman E, Baoprasertkul P, Terhune J, Xu P, Nandi S, Kucuktas H, Li P, Wang S, Somridhivej B, Dunham R, Liu Z (2007) Expression analysis of the acute phase response in channel catfish (Ictalurus punctatus) after infection with a Gram-negative bacterium. Dev Comp Immunol 31(11):1183–1196. https://doi.org/10.1016/j.dci.2007.03.003

Peng X-X (2013) Proteomics and its applications to aquaculture in China: infection, immunity, and interaction of aquaculture hosts with pathogens. Dev Comp Immunol 39(1–2):63–71. https://doi.org/10.1016/j.dci.2012.03.017

Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R (2009) Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. Cell 138(4):795–806. https://doi.org/10.1016/j.cell.2009.05.051

Pineiro C, Vazquez J, Marina AI, Barros-Velazquez J, Gallardo JM (2001) Characterization and partial sequencing of species-specific sarcoplasmic polypeptides from commercial hake species by mass spectrometry following two-dimensional electrophoresis. Electrophoresis 22 (8):1545–1552. https://doi.org/10.1002/1522-2683(200105)22:8<1545::Aid-Elps1545>3.0.Co;2-5

Piovesana S, Capriotti AL, Caruso G, Cavaliere C, La Barbera G, Zenezini Chiozzi R, Lagana A (2016) Labeling and label free shotgun proteomics approaches to characterize muscle tissue from farmed and wild gilthead sea bream (Sparus aurata). J Chromatogr A 1428:193–201. https://doi.org/10.1016/j.chroma.2015.07.049

Ponnerassery SS, Benjamin RL, Douglas RC, William FS, Scott EL, Gregory DW, Kenneth DC (2007) Identification of potential vaccine target antigens by immunoproteomic analysis of a virulent and a non-virulent strain of the fish pathogen Flavobacterium psychrophilum. Dis Aquat Org 74(1):37–47

Provan F, Jensen LB, Uleberg KE, Larssen E, Rajalahti T, Mullins J, Obach A (2013) Proteomic analysis of epidermal mucus from sea lice–infected Atlantic salmon, Salmo salar L. J Fish Dis 36(3):311–321. https://doi.org/10.1111/jfd.12064

Prunet P, Øverli Ø, Douxfils J, Bernardini G, Kestemont P, Baron D (2012) Fish welfare and genomics. Fish Physiol Biochem 38(1):43–60. https://doi.org/10.1007/s10695-011-9522-z

Purushothaman S, Saxena S, Meghah V, Meena Lakshmi MG, Singh SK, Brahmendra Swamy CV, Idris MM (2015) Proteomic and gene expression analysis of zebrafish brain undergoing continuous light/dark stress. J Sleep Res 24(4):458–465. https://doi.org/10.1111/jsr.12287

Qian X, Ba Y, Zhuang Q, Zhong G (2014) RNA-Seq technology and its application in fish transcriptomics. Omics: J Integr Biol 18(2):98–110. https://doi.org/10.1089/omi.2013.0110

Quinn NL, Gutierrez AP, Koop BF, Davidson WS (2012) Genomics and genome sequencing: benefits for finfish aquaculture. INTECH Open Access Publisher

Rasmussen RS, Morrissey MT (2008) DNA-Based methods for the identification of commercial fish and seafood species. Compr Rev Food Sci Food Saf 7(3):280–295. https://doi.org/10.1111/j.1541-4337.2008.00046.x

Richard N, Fernández I, Wulff T, Hamre K, Cancela L, Conceição LEC, Gavaia PJ (2014) Dietary supplementation with vitamin K affects transcriptome and proteome of Senegalese Sole, improving larval performance and quality. Mar Biotechnol 16(5):522–537. https://doi.org/10.1007/s10126-014-9571-2

Richard N, Silva TS, Wulff T, Schrama D, Dias JP, Rodrigues PML, Conceição LEC (2016) Nutritional mitigation of winter thermal stress in gilthead seabream: associated metabolic pathways and potential indicators of nutritional state. J Proteome 142:1–14. https://doi.org/10.1016/j.jprot.2016.04.037

Rodrigues PM, Silva TS, Dias J, Jessen F (2012) Proteomics in aquaculture: applications and trends. J Proteome 75(14):4325–4345. https://doi.org/10.1016/j.jprot.2012.03.042

Rodrigues PM, Schrama D, Campos A, Osório H, Freitas M (2016) Applications of proteomics in aquaculture. In: Salekdeh GH (ed) Agricultural proteomics, Crops, horticulture, farm animals, food, insect and microorganisms, vol 1. Springer, Cham, pp 165–199. https://doi.org/10.1007/978-3-319-43275-5_10

Rossi F, Chini V, Cattaneo AG, Bernardini G, Terova G, Saroglia M, Gornati R (2007) EST-based identification of genes expressed in perch (Perca fluviatilis, L.) Gene Expr 14(2):117–127. https://doi.org/10.3727/105221607783417600

Rufino-Palomares E, Reyes-Zurita FJ, Fuentes-Almagro CA, de la Higuera M, Lupianez JA, Peragon J (2011) Proteomics in the liver of gilthead sea bream (Sparus aurata) to elucidate the cellular response induced by the intake of maslinic acid. Proteomics 11(16):3312–3325. https://doi.org/10.1002/pmic.201000271

Russell S, Hayes MA, Simko E, Lumsden JS (2006) Plasma proteomic analysis of the acute phase response of rainbow trout (Oncorhynchus mykiss) to intraperitoneal inflammation and LPS injection. Dev Comp Immunol 30(4):393–406. https://doi.org/10.1016/j.dci.2005.06.002

Salas-Leiton E, Cánovas-Conesa B, Zerolo R, López-Barea J, Cañavate JP, Alhama J (2009) Proteomics of Juvenile Senegal Sole (Solea senegalensis) affected by gas bubble disease in hyperoxygenated ponds. Mar Biotechnol 11(4):473–487. https://doi.org/10.1007/s10126-008-9168-8

Salem M, Paneru B, Al-Tobasei R, Abdouni F, Thorgaard GH, Rexroad CE, Yao J (2015) Transcriptome assembly, gene annotation and tissue gene expression atlas of the rainbow trout. PLoS One 10(3):e0121778. https://doi.org/10.1371/journal.pone.0121778

Santos GA, Schrama JW, Mamauag REP, Rombout JHWM, Verreth JAJ (2010) Chronic stress impairs performance, energy metabolism and welfare indicators in European seabass (Dicentrarchus labrax): the combined effects of fish crowding and water quality deterioration. Aquaculture 299(1–4):73–80. https://doi.org/10.1016/j.aquaculture.2009.11.018

Saptarshi SR, Sharp MF, Kamath SD, Lopata AL (2014) Antibody reactivity to the major fish allergen parvalbumin is determined by isoforms and impact of thermal processing. Food Chem 148:321–328. https://doi.org/10.1016/j.foodchem.2013.10.035

Satoh TP, Miya M, Mabuchi K, Nishida M (2016) Structure and variation of the mitochondrial genome of fishes. BMC Genomics 17(1):719. https://doi.org/10.1186/s12864-016-3054-y

Schrama D, Richard N, Silva TS, Figueiredo FA, Conceição LEC, Burchmore R, Eckersall D, Rodrigues PML (2016) Enhanced dietary formulation to mitigate winter thermal stress in gilthead sea bream (Sparus aurata): a 2D-DIGE plasma proteome study. Fish Physiol Biochem. https://doi.org/10.1007/s10695-016-0315-2

Shinn AP, Pratoomyot J, Bron JE, Paladini G, Brooker EE, Brooker AJ (2015) Economic costs of protistan and metazoan parasites to global mariculture. Parasitology 142(Special Issue 01):196–270. https://doi.org/10.1017/S0031182014001437

Siciliano RA, d'Esposito D, Mazzeo MF (2016) Food authentication by MALDI MS: MALDI-TOF MS analysis of fish species. In: Cramer R (ed) Advances in MALDI and laser-induced soft ionization mass spectrometry. Springer, Cham, pp 263–277. https://doi.org/10.1007/978-3-319-04819-2_14

Silva TS, Cordeiro O, Richard N, Conceicao LE, Rodrigues PM (2011) Changes in the soluble bone proteome of reared white seabream (Diplodus sargus) with skeletal deformities. Comp Biochem Physiol Part D Genomics Proteomics 6(1):82–91. https://doi.org/10.1016/j.cbd.2010.03.008

Silva TS, Matos E, Cordeiro OD, Colen R, Wulff T, Sampaio E, Sousa V, Valente LMP, Gonçalves A, Silva JMG, Bandarra N, Nunes ML, Dinis MT, Dias J, Jessen F, Rodrigues PM (2012) Dietary tools to modulate glycogen storage in gilthead seabream muscle: glycerol supplementation. J Agric Food Chem 60(42):10613–10624. https://doi.org/10.1021/jf3023244

Silva TS, da Costa AM, Conceicao LE, Dias JP, Rodrigues PM, Richard N (2014a) Metabolic fingerprinting of gilthead seabream (Sparus aurata) liver to track interactions between dietary factors and seasonal temperature variations. PeerJ 2:e527. https://doi.org/10.7717/peerj.527

Silva TS, Richard N, Dias JP, Rodrigues PM (2014b) Data visualization and feature selection methods in gel-based proteomics. Curr Protein Pept Sci 15(1):4–22. https://doi.org/10.2174/1389203715666140221112334

Silvestre F, Linares-Casenave J, Doroshov SI, Kültz D (2010) A proteomic analysis of green and white sturgeon larvae exposed to heat stress and selenium. Sci Total Environ 408 (16):3176–3188. https://doi.org/10.1016/j.scitotenv.2010.04.005

Sissener NH (2009) Genetically modified plants as fish feed ingredients. Roundup Ready® soy, MON810 maize, Atlantic salmon, zebrafish

Sissener NH, Martin SAM, Cash P, Hevrøy EM, Sanden M, Hemre G-I (2010) Proteomic profiling of liver from Atlantic salmon (Salmo salar) fed genetically modified soy compared to the near-isogenic non-GM line. Mar Biotechnol 12(3):273–281. https://doi.org/10.1007/s10126-009-9214-1

Sletten G, Van Do T, Lindvik H, Egaas E, Florvaag E (2010) Effects of industrial processing on the immunogenicity of commonly ingested fish species. Int Arch Allergy Immunol 151(3):223–236

Smith RW, Cash P, Ellefsen S, Nilsson GE (2009) Proteomic changes in the crucian carp brain during exposure to anoxia. Proteomics 9(8):2217–2229. https://doi.org/10.1002/pmic.200800662

Smith RW, Wang J, Mothersill CE, Lee LEJ, Seymour CB (2015) Proteomic responses in the gills of fathead minnows (Pimephales promelas, Rafinesque, 1820) after 6 months and 2 years of continuous exposure to environmentally relevant dietary 226Ra. Int J Radiat Biol 91(3):248–256. https://doi.org/10.3109/09553002.2014.988894

Spaink HP, Jansen HJ, Dirks RP (2014) Advances in genomics of bony fish. Brief Funct Genomics 13(2):144–156. https://doi.org/10.1093/bfgp/elt046

Srinivasa Rao PS, Yamada Y, Tan YP, Leung KY (2004) Use of proteomics to identify novel virulence determinants that are required for Edwardsiella tarda pathogenesis. Mol Microbiol 53(2):573–586. https://doi.org/10.1111/j.1365-2958.2004.04123.x

Stentiford GD, Viant MR, Ward DG, Johnson PJ, Martin A, Wenbin W, Cooper HJ, Lyons BP, Feist SW (2005) Liver tumors in wild Flatfish: a histopathological, proteomic, and metabolomic study. OMICS 9(3):281–299. https://doi.org/10.1089/omi.2005.9.281

Sun F, Liu S, Gao X, Jiang Y, Perera D, Wang X, Li C, Sun L, Zhang J, Kaltenboeck L, Dunham R, Liu Z (2013) Male-biased genes in catfish as revealed by RNA-Seq analysis of the testis transcriptome. PLoS One 8(7):e68452. https://doi.org/10.1371/journal.pone.0068452

Sveinsdóttir H, Steinarsson A, Gudmundsdóttir Á (2009) Differential protein expression in early Atlantic cod larvae (Gadus morhua) in response to treatment with probiotic bacteria. Comp Biochem Physiol Part D Genomics Proteomics 4(4):249–254. https://doi.org/10.1016/j.cbd.2009.06.001

Swoboda I, Bugajska-Schretter A, Verdino P, Keller W, Sperr WR, Valent P, Valenta R, Spitzauer S (2002) Recombinant carp parvalbumin, the major cross-reactive fish allergen: a tool for diagnosis and therapy of fish allergy. J Immunol 168(9):4576–4584

Tedesco S, Mullen W, Cristobal S (2014) High-throughput proteomics: a new tool for quality and safety in fishery products. Curr Protein Pept Sci 15(2):118–133

Teklemariam AD, Tessema F, Abayneh T (2015) Review on evaluation of safety of fish and fish products. Int J Fish Aquat Stud 3(2):111–117

Tomm JM, van Do T, Jende C, Simon JC, Treudler R, von Bergen M, Averbeck M (2013) Identification of new potential allergens from nile perch (Lates niloticus) and cod (Gadus morhua). J Investig Allergol Clin Immunol 23(3):159–167

Varó I, Rigos G, Navarro JC, del Ramo J, Calduch-Giner J, Hernández A, Pertusa J, Torreblanca A (2010) Effect of ivermectin on the liver of gilthead sea bream Sparus aurata: a proteomic approach. Chemosphere 80(5):570–577. https://doi.org/10.1016/j.chemosphere.2010.04.030

Vasanth G, Kiron V, Kulkarni A, Dahle D, Lokesh J, Kitani Y (2015) A microbial feed additive abates intestinal inflammation in Atlantic salmon. Front Immunol 6:409. https://doi.org/10.3389/fimmu.2015.00409

Veiseth-Kent E, Pedersen ME, Hollung K, Ytteborg E, Bæverfjord G, Takle H, Åsgård TE, Ørnsrud R, Lock E-J, Albrektsen S (2013) Changes in protein abundance in the vertebral column of Atlantic salmon (Salmo salar) fed variable dietary P levels. In: de Almeida A, Eckersall D, Bencurova E et al (eds) Farm animal proteomics 2013. Wageningen Academic, Wageningen, pp 188–191. https://doi.org/10.3920/978-90-8686-776-9_48

Vilhelmsson OT, Martin SAM, Medale F, Kaushik SJ, Houlihan DF (2004) Dietary plant-protein substitution affects hepatic metabolism in rainbow trout (Oncorhynchus mykiss). Br J Nutr 92 (1):71–80. https://doi.org/10.1079/Bjn20041176

Wenne R, Boudry P, Hemmer-Hansen J, Lubieniecki KP, Was A, Kause A (2007) What role for genomics in fisheries management and aquaculture? Aquat Living Resour 20(3):241–255. https://doi.org/10.1051/alr:2007037

Williams TD, Mirbahai L, Chipman JK (2014) The toxicological application of transcriptomics and epigenomics in zebrafish and other teleosts. Brief Funct Genomics 13(2):157–171. https://doi.org/10.1093/bfgp/elt053

Wu Y, Wang S, Peng X (2004) Serum acute phase response (APR)-related proteome of loach to trauma. Fish Shellfish Immunol 16(3):381–389. https://doi.org/10.1016/j.fsi.2003.06.003

Wulff T, Jokumsen A, Hojrup P, Jessen F (2012) Time-dependent changes in protein expression in rainbow trout muscle following hypoxia. J Proteome 75(8):2342–2351. https://doi.org/10.1016/j.jprot.2012.02.010

Xing M, Hou Z, Yuan J, Liu Y, Qu Y, Liu B (2013) Taxonomic and functional metagenomic profiling of gastrointestinal tract microbiome of the farmed adult turbot (Scophthalmus maximus). FEMS Microbiol Ecol 86(3):432–443. https://doi.org/10.1111/1574-6941.12174

Xiong X-P, Dong C-F, Xu X, Weng S-P, Liu Z-Y, He J-G (2011) Proteomic analysis of zebrafish (Danio rerio) infected with infectious spleen and kidney necrosis virus. Dev Comp Immunol 35 (4):431–440. https://doi.org/10.1016/j.dci.2010.11.006

Yadetie F, Bjørneklett S, Garberg HK, Oveland E, Berven F, Goksøyr A, Karlsen OA (2016) Quantitative analyses of the hepatic proteome of methylmercury-exposed Atlantic cod (Gadus morhua) suggest oxidative stress-mediated effects on cellular energy metabolism. BMC Genomics 17(1):554. https://doi.org/10.1186/s12864-016-2864-2

Ye C-X, Wan F, Sun Z-Z, Cheng C-H, Ling R-Z, Fan L-F, Wang A-L (2016) Effect of phosphorus supplementation on cell viability, anti-oxidative capacity and comparative proteomic profiles of puffer fish (Takifugu obscurus) under low temperature stress. Aquaculture 452:200–208. https://doi.org/10.1016/j.aquaculture.2015.10.039

Zhou S, Wan Q, Huang Y, Huang X, Cao J, Ye L, Lim T-K, Lin Q, Qin Q (2011) Proteomic analysis of Singapore grouper iridovirus envelope proteins and characterization of a novel envelope protein VP088. Proteomics 11(11):2236–2248. https://doi.org/10.1002/pmic.200900820

Zhou X, Ding Y, Wang Y (2012) Proteomics: present and future in fish, shellfish and seafood. Rev Aquac 4(1):11–20. https://doi.org/10.1111/j.1753-5131.2012.01058.x

The Use of Proteomics to Study Biomarkers of Stress and Welfare in Farm Animals

Anna Marco-Ramell, Ana M. Gutiérrez, Antonio Velarde, Jose J. Cerón, and Anna Bassols

Abstract Animal welfare and stress are important issues mainly because of public perception, marketing, product acceptance and production efficiency, quality and quantity. They are complex conditions that include physical and psychological stress, as well as the beneficial or deleterious effects that the environment may have on the welfare of the individual. Although a lot has been done in the establishment of protocols to ensure an adequate environment for livestock throughout their lives and their way to the slaughterhouse, there is still a significant lack of information about biological markers that can be easily and objectively measured in the laboratory and can provide information about the biological consequences of suboptimal living conditions in the individual. These biomarkers have to be measured in biological samples that have to fulfil several criteria: they should be easy to obtain, even in a repetitive manner from each individual if necessary, and should mirror the biological processes occurring inside the cells and organs as a consequence of challenging environmental conditions. Proteomics, a technology that allows protein identification in complex samples from a holistic, non-hypothesis-driven approach, is a very suitable method for the search of biomarkers in animal science, including those related to stress and welfare.

A. Marco-Ramell
Departament de Bioquímica i Biologia Molecular, Facultat de Veterinària, Universitat Autònoma de Barcelona, Bellaterra, Spain

Departament de Nutrició, Ciències de l'Alimentació i Gastronomia, Facultat de Farmàcia i Ciències de l'Alimentació, Universitat de Barcelona, Barcelona, Spain

A.M. Gutiérrez • J.J. Cerón
Department of Animal Medicine and Surgery, Regional Campus of International Excellence "Campus Mare Nostrum", University of Murcia, Murcia, Spain

A. Velarde
Animal Welfare Program, IRTA, Monells, Spain

A. Bassols (✉)
Departament de Bioquímica i Biologia Molecular, Facultat de Veterinària, Universitat Autònoma de Barcelona, Bellaterra, Spain
e-mail: anna.bassols@uab.cat

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_17

Keywords Animal welfare • Stress • Biomarkers • Proteomics • Farm animals

List of Abbreviations

1D	One-dimensional
2D	Two-dimensional
APP	Acute-phase protein
BALF	Bronchoalveolar fluid
BRD	Bovine respiratory disease
DIGE	Difference gel electrophoresis
ELF	Epithelial lung fluid
EU	European Union
GPx	Glutathione peroxidase
HPA	Hypothalamic-pituitary-adrenal axis
Hsp	Heat-shock protein
iTRAQ	Isobaric tags for relative and absolute quantitation
LC	Liquid chromatography
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NEB	Negative energy balance
PBMC	Peripheral blood mononuclear cell
PSE	Pale, soft and exudative
SWL	Seasonal weight loss
SOD	Superoxide dismutase

1 Animal Welfare Definition

1.1 Current Practices in Animal Welfare Assessment

Animal welfare can be defined in a number of different ways, but there is a growing consensus that whatever be the definition it has to include three main elements (Duncan and Fraser 1997). The first element is based on its ability to express certain species-specific normal behaviours (Hughes and Duncan 1988). The second one, the emotional state of the animal, refers to the absence of negative emotions such as pain or chronic fear and the presence of positive ones. The third element is the animal's adequate biological functioning, when coping with its environment (Broom 1986).

Welfare comprises physical and mental health (Dawkins 2004; Webster et al. 2004) and includes several aspects such as the absence of thirst, hunger, discomfort, disease, pain, injuries and stress and the expression of normal behaviour, referred to as the Five Freedoms of the Farm Animal Welfare Council (FAWC 1992). When

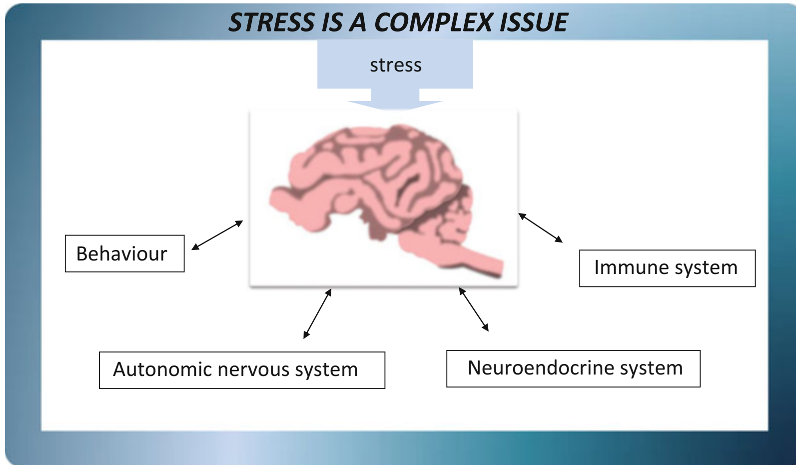

Fig. 1 Physiological mechanisms of response to stress: when the situation moves away from the ideal state, a wide range of reactions take place in an attempt to cope with stress

the situation moves away from the ideal state, the animal will use a wide range of physiological mechanisms and behaviours to cope with its environment (Fig. 1). The impossibility to cope with an aversive situation may lead to the appearance of injuries or diseases and consequently to pain and suffering (Broom and Johnson 1993).

More recently, the integrated project cofinanced by the European Commission, Welfare Quality® (www.welfarequalitynetwork.net), built on and extended the Five Freedoms to four main welfare principles: good feeding, good housing, good health and appropriate behaviour (Blokhuis et al. 2008). Each of these four principles comprises two to four criteria, with an overall total of 12 criteria: good feeding includes the absence of prolonged hunger and absence of prolonged thirst; good housing includes comfort around resting, thermal comfort and ease of movement; good health includes absence of injuries, disease and pain; and appropriate behaviour includes the expression of social and other behaviours, good human-animal relationship and positive emotional state. These 12 animal welfare criteria provide a very useful framework for assessing animal welfare. The acquired data provide feedback to animal unit managers about the welfare status of their animals and can be translated into accessible and understandable information on the welfare status of food-producing animals for consumers and others (Velarde and Dalmau 2012).

The emphasis of Welfare Quality® protocols is placed on animal-based measures (also called outcome or performance measures), in an attempt to estimate the actual welfare state of the animals in terms of, for instance, their behaviour, fearfulness, health or physical condition. Animal-based measures include a range of different parameters (Fig. 2):

– *Physiological measurements*. These measurements are generally related to the stress response, including plasma levels of glucocorticoids and heart rate (Broom and Johnson 1993). Secretion of glucocorticoids by the adrenal cortex is the final

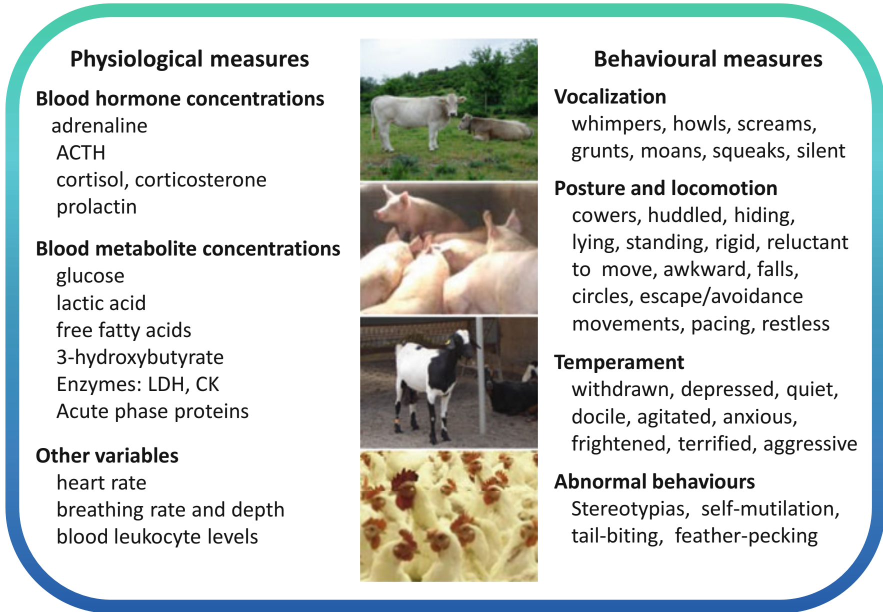

Fig. 2 Physiological and behavioural measures used as indicators of stress in farm animals

step in a cascade of events beginning in the brain (Sapolsky 1992). In aversive situations, plasma concentration of glucocorticoids increases within a few minutes. Uncertainty and anxiety are among the most potent stimuli for the hypothalamic-pituitary-adrenal axis (HPA) activation, whereas reward and sense of control result in suppression of HPA activity (De Kloet et al. 1991). Other hormones increased by stress are adrenaline, ACTH and prolactin. Plasma acute-phase proteins (APPs) and other parameters such as glucose, lactic acid, lipid metabolites or skeletal muscle enzymes are also used to monitor certain aspects of stress.

– *Behavioural measures*. Appropriate behaviour includes the expression of social and other behaviours. The amount of fighting that animals show is of particular interest, as fighting may cause both injuries and stress. Fighting occurs as a consequence of social mixing, and the amount of fighting depends on the species, sex, age and handling conditions (Velarde 2007). Other abnormal behaviours such as stereotypes, self-mutilation, tail-biting in pigs and feather-pecking in hens are often an indicator of poor welfare because they tend to develop in apparently aversive and stressful environments (Tuyttens 2007). Reluctance to move and turning back have been validated as behavioural measures to assess fear in domestic pigs (Dalmau et al. 2009).

1.2 European Requirements

Over the years, consumers' concerns about the ethical treatments of animals have become more and more important. Animal welfare is indeed seen as attribute of an

'overall food quality concept', which is strictly correlated to other issues such as animal health, food safety and food quality, safeguarding environmental protection and sustainability. Over the past 40 years, the European Union (EU) animal welfare legislation has evolved on the basis of sound scientific knowledge, improving the quality of animals' lives in accordance with citizens' expectations and market demands.

In the EU, legislation on animal welfare has a long-standing tradition. It has been the commonest way of protecting farm animal welfare and ensures that animals do not endure avoidable pain, distress or suffering while obliging the owner or any other persons dealing with animals to respect minimum welfare requirements. EU legislation is primarily founded on the principle that animals are 'sentient beings', as recognized by the EU Lisbon Treaty (EU 2009). Recently, the European Commission has adopted a new Strategy for the Protection and Welfare of Animals 2012–2015 to improve welfare conditions for animals kept on farms, living in zoos and used for experiments (https://ec.europa.eu/food/sites/food/files/animals/docs/aw_brochure_strategy_en.pdf). Currently, there is a range of EU Directives and Regulations specifying general and species-specific requirements, conditions and practices to ensure good animal welfare for different species. These cover areas such as animal housing and management, transport and slaughter.

2 The Search for Novel Stress Markers

The laboratory parameters currently addressed to evaluate the stress of farm animals involve most notably the stress hormones. The most important are adrenaline, the fast-acting hormone released by the adrenal gland under the control of the sympathetic nervous system, and the glucocorticoids cortisol and corticosterone, secreted by the adrenal gland after activation of the HPA axis (Broom and Johnson 1993; Bassols et al. 2014). However, the use of these hormones has some limitations: the short half-life of adrenaline, the high intra- and interindividual variability of cortisol and the influence of circadian rhythms on its secretion (Mormede et al. 2007). Although both hormones are good markers for acute stress conditions, they are not adequate as indicators for chronic stress. Other hormones and proteins have been proposed as markers for stress (Fig. 2). For instance, APPs such as Pig-MAP and haptoglobin have been proposed as indicators of stress in cattle and pigs (Murata 2007). Nevertheless, their levels are also modified during inflammation and thus they lack specificity. Other conditions, such as nutritional stress, may be monitored by using metabolic plasma markers, as nonesterified fatty acids and 3-hydroxybutyrate, which are increased in plasma as a result of lipid mobilization from the adipose tissue and synthesis of ketone bodies by the liver. Stress-released adrenaline activates glycolysis, giving rise to blood glucose and lactic acid. The skeletal muscle is rich in the enzymes lactate dehydrogenase and creatine kinase, and as a consequence, any kind of physical stress resulting in muscle damage will provoke an increase in the blood concentration of

these enzymes, but again these parameters are not specific for stress. As a consequence, specific and objective laboratorial criteria to evaluate animal stress and welfare are still lacking, and proteomics may be a useful approach to achieve this goal. The global assessment of animal welfare would greatly benefit from objective laboratory markers to complement the information provided by the behavioural tests.

Proteomics is a global non-hypothesis-driven approach widely used for the search of new biomarkers. This search is being fuelled by the advances in mass spectrometry (MS)-based technologies, which are widely reviewed in other chapters of this book (Fig. 3), as well as by the completion of protein/peptide sequence databases. The genomes of cows (Elsik et al. 2009), pigs (Groenen et al. 2012) and chicken (Hillier et al. 2004) have been already published, and this achievement is being translated into the improvement of databases, essential for the translation of proteomic results to functional conclusions. The most widely used publicly accessible protein databases are the UniProtKB and NCBI databases that offer protein sequence information and contain direct links to numerous other protein information sites and tools. UniProtKB present both manually annotated and reviewed entries (Swiss-Prot) and unreviewed automatically annotated data (TrEMBL), whereas RefSeq databases at NCBI present only one single canonical entry for every gene and protein species. The main provider of farm animal proteome data is PeptideAtlas, which contains sequences for cow (Bislev et al. 2012), pig (Hesselager et al. 2016) and horse (Bundgaard et al. 2014).

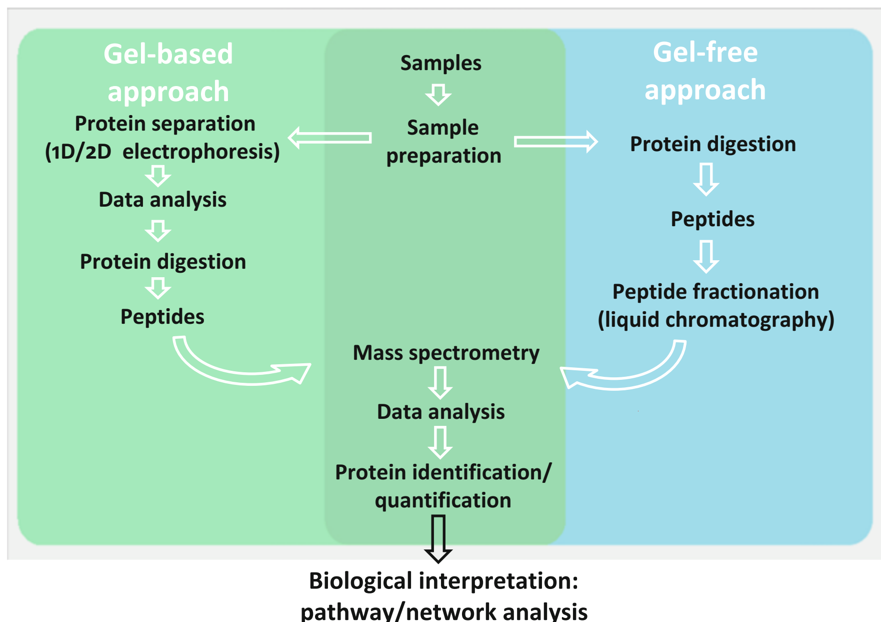

Fig. 3 Main approaches in the proteomic investigation of biological samples

2.1 Biomarkers in Serum, Plasma and Blood Cells

Plasma is obtained after collecting blood in the presence of an anticoagulant, and hence, it contains fibrinogen as well as other coagulation factors. Serum is obtained by allowing the blood to coagulate, and it lacks most of the coagulation factors, including fibrinogen, which is the fourth most abundant protein in plasma (Anderson and Anderson 2002). Serum, EDTA plasma, heparin plasma and citrate plasma have been used for proteomic approaches, but all exhibit shortcomings. Depending on analytical objectives and/or target peptides or proteins, use of either plasma or serum may impact both method and results. Serum may have components, especially proteolytic fragments, coming from the coagulation process. EDTA and citrate are negatively charged polycarboxylic acids and form soluble complexes with metal ions and prevent them from further reaction. They may not be useful if the endpoint measurement of interest involves assays where divalent cations are necessary. Heparin can result in interference in some affinity processes. In any case, platelet-depleted plasma should be prepared to avoid the presence of platelet-related proteins. A very useful set of recommendations for sample collection and handling have been already raised for human studies (Rai et al. 2005). Although in many cases it is very difficult to follow these guidelines in animal studies, attention should be paid to keep them in mind as much as possible. Indeed, biological variability is an important issue when dealing with stress and pathological conditions. Keeping this variability as low as possible by imposing restrictions in sample collection will always increase the chances of positive results in biomarker searches.

The protein content of serum is dominated by a handful of high-abundance proteins including albumin, immunoglobulins and transferrin, with their estimated concentration exceeding the low-abundance proteins highly by ten orders of magnitude (Anderson and Anderson 2002). To detect proteins that are present in low levels, it is advisable to process the samples prior to the proteomic analysis. Many strategies have been developed for the selective removal of high-abundance proteins or for the reduction of the orders of magnitude of protein abundance (Fuchs et al. 2005; Luque-Garcia and Neubert 2007; de Roos et al. 2008; Apweiler et al. 2009; Marco-Ramell and Bassols 2010; Lista et al. 2013; Lygirou et al. 2015; Mesmin et al. 2016; Gianazza et al. 2016; Chutipongtanate et al. 2017).

Blood cells (white blood cells, erythrocytes and platelets) are key components of the immune response and the coagulation cascade, and they have also other functions such as iron and oxygen transportation and protection of blood vessels. The analysis of the proteome of these blood cells can provide new insights on the molecular mechanisms involved in stress-related conditions (Thadikkaran et al. 2005). Blood cells are widely used in proteomic research, but special care has to be taken during collection, handling and storage of this type of samples as they are fragile, and thus the extent of cell haemolysis should be tightly controlled and minimized, i.e. by separation at low centrifugation speeds. Moreover, they are very sensitive to the activation or inactivation, e.g. blood-clotting cascade, thus necessitating stringent sample-preparation procedures (time to processing, temperature of processing) to avoid artefacts (Vuckovic 2012).

2.2 Biomarkers in Saliva

Saliva is an important biological fluid present in the oral cavity. It is composed of more than 99% water but also a variety of electrolytes, proteins, enzymes, mucins and nitrogenous products (Humphrey and Williamson 2001; Nieuw Amerongen et al. 2007; Schipper et al. 2007). The term saliva is usually used to refer to either gland-specific saliva or whole saliva. Gland-specific saliva is directly collected from individual salivary glands. Whole saliva is a mixture of oral fluids and includes secretions from major and minor salivary glands and several constituents of non-salivary origin, such as gingival crevicular fluid; expectorated secretions; serum and blood derivatives; bacteria, fungi and viruses; desquamated epithelial cells; and food debris (Kaufman and Lamster 2002). Whole saliva, usually just called saliva, is most frequently studied for the evaluation of systemic disorders.

Standardization of the salivary collection procedure is of utmost importance in order to obtain reproducible results, particularly when the concentration of intrinsic salivary constituents is used as diagnostic criterion (Nieuw Amerongen et al. 2007). Saliva can be obtained either unstimulated or stimulated. Unstimulated saliva is very rich in mucins with high viscosity that could cause difficulties in handling and analysis (Mau et al. 2009). Stimulation for saliva collection could be performed by chemical or mechanical procedures. Chemical stimulation, by using citric acid, reduces the total amount of saliva proteins in comparison to other collection methods (Topkas et al. 2012), and it has not been widely employed in animals. On the other hand, mechanically stimulated saliva is commonly collected with devices such as Salivette® (Sarstedt AG & Co., Nümbrecht, Germany) that consists of a small cotton roll chewed by the animal until it is moistened. Some technical considerations should be taken into account. For example, saliva is collected in pigs with cotton buds (Parrott et al. 1989) or sponges (Gutierrez et al. 2009), which could be clipped to a flexible metal rod to minimize human disturbances. Moreover, cotton buds could be also enclosed in nylon netting with twine extensions (Bushong et al. 2000). However, sponges have been considered as the optimal material for porcine saliva collection since they are less absorbent and release more saliva following centrifugation than cotton buds (Gutierrez et al. 2009). The use of these devices reduces salivary viscosity, allowing an easier processing of the samples (Amado et al. 2013). Mechanical stimulation has the advantage that saliva is easily obtained in high volumes, thus not diluting the total salivary protein concentration in comparison to passive drooling (Topkas et al. 2012).

Advantages of saliva over serum samples as an analytical tool have been postulated. First, saliva testing has facilitated the safe, efficient and low-cost collection of large numbers of diagnostic samples (Ramirez et al. 2012). Moreover, saliva samples may be obtained in remote sites and by unskilled personnel (Prickett and Zimmerman 2010).

So far the most extensive salivary studies have been carried out on primates, especially humans, and rodents, and much less attention has been paid to farm animals. The composition of bovine saliva is different from that of most other

species studied. It has high levels of carbonic anhydrase but lacks amylase, probably reflecting its role in ruminant metabolism (Mau et al. 2009; Ang et al. 2011). Porcine saliva composition has physiological similarities to human saliva (Verma et al. 2011), but some differences, i.e. amylase and lipocalin levels, have been also observed (Gutiérrez et al. 2013).

Saliva renders a sample type by which animal systemic health could be monitored since salivary components include local and systemic proteins that could act as biomarkers (Lawrence 2002; Lamy and Mau 2012). Similar to serum, the protein profile of saliva is dominated by a subset of abundant proteins. In order to measure low-abundance analytes, several depletion strategies have been developed for saliva samples (Krief et al. 2011).

2.3 Other Biological Samples

Although blood and saliva have been the preferred biological samples to identify stress biomarkers, other types of samples have been also used to search for biomarkers. In the following, we briefly describe them.

Bronchoalveolar fluid (BALF) contains the secretory products and the remnants of the epithelial fluid of the lungs (ELF) and represents the first line of defence against inhaled opportunistic pathogens (Mitchell et al. 2008). BALF is obtained through bronchoscopy, which is a safe but invasive technique. The use of BALF in proteome research is in an exponential growth phase, and protocols on BALF sample collection and preparation have been optimized (Wattiez and Falmagne 2005).

The liver is the largest organ in the body and has a central role in the metabolism of biomolecules, as well as in food digestion, detoxification and immunity. The analysis of its proteome is highly valuable in human and animal research. Indeed the liver proteome has been used to define biomarkers of hepatitis, cirrhosis, hepatotoxicity and fatty liver disease, among others. The Human Liver Proteome Project (HLPP) is an international cooperative effort to integrate academic, industrial and government activities into a comprehensive study of liver proteins (He 2005).

There is also a wide range of proteomic applications in neuroscience. For instance, proteomics has been used to identify biomarkers of neurological diseases such as Alzheimer, Parkinson or schizophrenia. Protocols for brain sample preparation and analysis have been described (Lubec et al. 2003; Fernandez-Gomez et al. 2014).

In addition, other biological matrices such as milk and mammary glands (Danielsen et al. 2010; van Altena et al. 2016), wool (Almeida et al. 2014) and meat (Paredi et al. 2012; Hollung et al. 2014; Piras et al. 2016) have been used for biomarker search through proteomics. In addition, faeces and hair have been also employed as biological matrices for the quantification of the stress hormones cortisol and corticosterone, but no proteomic attempts with them have been made so far.

3 Search for Stress Markers in Pigs by Using Proteomics

Pigs are social animals and prone to experience stress. Welfare problems and stress reactions might result from a number of environmental stimuli in commercial porcine farms, including feeding and housing conditions, social environment and management practices (Merlot et al. 2011). A high degree of stress is related to an increased susceptibility to suffer diseases and detrimental effects in meat quality.

3.1 Management Stress

3.1.1 Biomarkers from Serum Samples

High stocking density is a common housing condition that growing and finishing pigs generally suffer in commercial farms. Although minimum space requirements have been stipulated by the European Council (Directive 2001/88/EC), the space allowance during the latter part of the growing and the finishing periods often fails to meet these dimensions (Leek et al. 2004). The differential expression of serum proteins between high and low stocking density was analysed by two-dimensional difference gel electrophoresis (2D-DIGE). The cytoplasmic β-actin, a component of the cytoskeleton, was identified as a differentially expressed protein by matrix-assisted laser desorption/ionization (MALDI-MS), pointing out to cellular or tissue damage associated to high stocking density (Marco-Ramell et al. 2011).

Individual confinement of pregnant sows is a recently banned housing system by the European Union (CD 66 2001/88/EC) but was widely used under field conditions in the past. The serum proteome of gilts housed in group and subjected to individual confinement was compared by 2D-DIGE and isobaric tags for relative and absolute quantitation (iTRAQ) by Marco-Ramell et al. to identify serum stress biomarkers before the actual legislation came into force. Differentially expressed proteins comprised APPs, antioxidant enzymes, apolipoproteins, heat-shock proteins, metabolic enzymes and structural cytoplasmic proteins, indicating that individual confinement of gilts led to a complex physiological response that is characterized by the activation of the immune system, lipid mobilization, redox imbalance and cellular or tissue damage (Marco-Ramell et al. 2016).

Peripheral blood mononuclear cells (PBMCs), namely, lymphocytes and monocytes, were used to search for stress biomarkers in pigs housed under different management conditions. For instance, one group of pigs was kept under commercial farm conditions and the other one to a close human-animal relationship, with daily personal care of the animals by the same, well-trained person for 2 months. PBMC samples were subjected to 2D-DIGE followed by liquid chromatography tandem mass spectrometry (LC-MS/MS) identification, and several proteins were identified, many of them being potential targets of cortisol (Valent et al. 2017).

3.1.2 Biomarkers from Saliva Samples

To overcome the stress induced by blood sampling itself, alternative methods such as the measurement of stress biomarkers in saliva samples have been developed (Mormede et al. 2007). There is only one report in which potential new salivary biomarkers of stress have been investigated using proteomic approaches. Three experimental models of acute stress were studied: snaring restraint followed by simulated sampling of vena cava blood, brief transport by road and restriction of movement in a digestibility cage. Two proteins were postulated as possible markers of stress, odorant-binding protein and fragments of albumin (Fuentes-Rubio et al. 2014).

Porcine salivary protein patterns were characterized using one-dimensional (1D) and 2D electrophoresis in healthy individuals and in animals with subclinical infections (Gutierrez et al. 2011). Medium- to high-abundance proteins were identified as differential among both groups of animals including adenosine deaminase, calcium-binding proteins S-100 A8 and A9, submandibular gland-like double-headed protease inhibitor, haptoglobin, lipocalin 1, pancreatic α-amylase, protein S-100 A2, salivary lipocalin and serum albumin (Gutiérrez et al. 2013).

In addition, glycopattern alterations in pigs suffering from acute rectal prolapse have been observed in saliva, but not in serum samples (Gutiérrez et al. 2016). These results highlight the potential of saliva samples to provide additional information on disease condition and require intensive exploration in further studies.

Criado et al. studied the protein composition in saliva in two groups of pigs subjected to short and long transport, as conditions of low and high stress level, through 1D and 2D electrophoresis. Four proteins were found to be differentially regulated, such as serum albumin, haptoglobin, immunoglobulin L chain and lipocalin (Criado et al. 2014).

3.2 Heat Stress

Heat stress annually causes important losses to the porcine industry, including animal performance decrease, metabolic disorders and health problems, but the physiological mechanisms underlying this condition still remain elusive.

To characterize the skeletal muscle cellular response to acute heat stress in pigs, the sarcoplasmic proteome of the semitendinosus muscle was analysed by 2D-DIGE followed by LC-MS and MS/MS. Several proteins involved in glycolysis, glycogenesis and glycogenolysis were increased or modified, indicating enhanced glycolytic capacity in response to heat stress. At the same time, abundance of tubulins and other proteins decreased, indicating a loss of microtubule structure. Changes in redox metabolism were also observed with increased manganese superoxide dismutase (SOD) and decreased peroxiredoxin 2 abundances, a sign for an antioxidant response to heat stress (Cruzen et al. 2015).

Furthermore, the effects of chronic heat stress on the small intestine were analysed by using proteomics in finishing pigs maintained at 30 °C for 3 weeks. Heat stress damaged the intestinal mucosa, as observed by changes in morphology. In addition, the proteome analysis of the jejunum mucosa by 2D electrophoresis followed by MS confirmed changes in proteins related to cell structure and showed an impaired energy metabolism due to a downregulation of proteins involved in tricarboxylic acid cycle, electron transport chain and oxidative phosphorylation (Cui and Gu 2015). Similar responses in the ileum were shown by a 2D-DIGE approach. Stress response enzymes peroxiredoxin-1 and peptidyl-prolyl cis-trans isomerase A were altered by heat stress as well, suggesting a reduced intestinal integrity (Pearce et al. 2015).

4 Search for Stress Markers in Cattle by Using Proteomics

4.1 Management Conditions

4.1.1 Biomarkers from Blood Samples

Marco-Ramell et al. compared the serum proteome of a group of cows that was maintained on a diet of cultivated pastures during the winter and a group of cows that lived in semi-feral conditions through 2D-DIGE followed by MALDI-MS. Several serum proteins were identified as mechanisms of adaptation to harsh housing conditions including the antioxidant enzymes paraoxonase-1 and glutathione peroxidase (GPx) 3, proteins related to the complement cascade and the APPs α2-HS-glycoprotein and AMBP precursor. Some of these proteins were validated as stress biomarkers in cows in a third group of cows that were living in very harsh conditions during the winter (Marco-Ramell et al. 2012).

In addition to stress conditions, proteomics has been also used to differentiate stress from a viral infection in cows. Bovine respiratory disease (BRD) is a leading cause of morbidity and mortality in feedlot cattle, and it results from the combined infection of bacterial and viral agents. Animals subjected to stress such as weaning, transportation and processing are prone to suffer this disease. Aich et al. employed an experimental infection of BRD together with several stressors (a combination of weaning, transportation and social reorganization in young calves). Different analytical approaches (proteomics, metabolomic and elemental profiles) on serum samples allowed differentiating a stress response from a primary viral respiratory infection. Differential regulation of serum amyloid A, apolipoproteins A-I and C-III and haptoglobin, among others, and changes in the metabolome were found (Aich et al. 2007).

4.1.2 Biomarkers from Bronchoalveolar Fluid

Several proteomic studies have employed BALF to identify stress and disease (BRD) biomarkers. Mitchell et al. hypothesized that stress would change ELF composition, which would increase the susceptibility of the animal to suffer respiratory diseases such as BRD. The study revealed increases in antioxidant enzymes as SOD and GPx and changes in several APPs including α2-HS-glycoprotein, fibrinogen and annexin, among others (Mitchell et al. 2008).

4.2 *Nutritional Stress*

4.2.1 Biomarkers from Blood

Proteomics has also helped to understand the physiological basis of the many problems arising during the transition period of dairy cows, normally characterized by a status of negative energy balance (NEB) because the ingestion of nutrients is insufficient to meet the energy demands of milk production and secretion (Kuhla et al. 2007). Increased energy requirements lead to hypoglycaemia, increased lipid mobilization to the liver and further hepatic synthesis of ketone bodies. Ketosis is a main problem in dairy cow production (Turk et al. 2008), and persistent ketosis is accompanied by acidosis, redox imbalance and impaired immune function, which lead to a higher susceptibility to diseases and metabolic disorders (Loor et al. 2013).

The proteome of plasma from ketotic cows was analysed by 2D electrophoresis and MALDI-MS, and an increase of the APPs haptoglobin, α1-acid glycoprotein and α-chymotrypsin levels and a decrease in other APPs including apolipoprotein A-IV and α2-HS-glycoprotein were detected (Cairoli et al. 2006; Yang et al. 2012). These were validated by iTRAQ and the enhanced sensitivity of this technique also allowed the identification of the antimicrobial peptides cathelicidins, the antioxidant enzyme GPx and several components of the complement system and coagulation cascade (factor V, α2-antiplasmin and prothrombin), demonstrating the role of these proteins in inflammatory and immune processes (Yang et al. 2012).

4.2.2 Biomarkers from Liver and Brain Samples

In the liver, important changes in carbohydrate, fatty acid, amino acid and nucleotide metabolism were identified by proteomics. For example, regulatory metabolic enzymes involved in the Krebs cycle and gluconeogenesis were deregulated during early lactation indicating the adaptation of the liver to cope with the metabolic imbalance (Xu and Wang 2008; Moyes et al. 2013). Due to the lipid mobilization to the liver, an exaggerated deposit of lipids in this organ could occur, giving rise to ‘fatty liver’. In addition to its intrinsic importance in dairy cow production, research

on hepatic steatosis is of major importance since it is closely related to metabolic disorders in humans such as obesity, metabolic syndrome and diabetes.

Kuhla et al. analysed the proteome of cows subjected to nutrient deprivation through 2D electrophoresis followed by MALDI-MS and MS/MS. An increase in liver total fat content and a decrease in the concentration of glucose and glycogen, similar to the hepatic proteome observed in dairy cows during the periparturient period, were found in feed-restricted cows. The analysis also showed that the enzymes involved in the oxidation of fatty acids were downregulated after feed restriction, which could cause an excessive fat deposition in the liver (Kuhla et al. 2009). Concomitantly, there is a decrease in proteins involved in glycogen synthesis and the TCA cycle and an increase in proteins related to the glycolysis, fatty acid degradation and lactate production, thus providing the Cori cycle with substrates for the hepatic gluconeogenesis. In addition the organism counteracts the energy deprivation by obtaining amino acids from muscle proteins that can be used for glucose production via gluconeogenesis (Kuhla et al. 2011). Additional insights of nutritional stress in cows were obtained by analysing the proteome of brain tissue (Kuhla et al. 2007, 2010).

4.3 Heat Stress

Heat stress has an enormous economic impact on the dairy industry. In order to identify the mechanisms underlying heat stress, plasma from cows subjected to neutral thermal conditions and to heat stress were compared through iTRAQ proteomic approach. Database searches combined with Gene Ontology and KEGG pathway enrichment analyses revealed that many components of the complement and coagulation cascades were altered in heat-stressed cows. For instance, several complement components were downregulated, whereas components of the coagulation system such as coagulation factors, vitamin K-dependent proteins and fibrinogens were upregulated by heat stress (Min et al. 2016).

5 Search for Stress Markers in Small Ruminants by Using Proteomics: Seasonal Weight Loss

Seasonal weight loss (SWL) is the most important limitation in small ruminant production in the Tropical and Mediterranean regions, thereby conditioning producer's incomes and the nutritional status of rural communities. SWL significantly affects the productivity of milk, meat and wool farms.

The mitochondrial proteome of the mammary gland was analysed in two goat breeds with different levels of adaptation to nutritional stress (tolerant and susceptible to nutritional stress). A gel-based strategy (Blue Native-PAGE) followed by

MALDI-MS/MS enabled the identification of several proteins related to mitochondrial oxidative metabolism associated to weight loss, including NADH-ubiquinone oxidoreductase, lamin B1, guanine nucleotide-binding protein subunit beta-2, cytochrome b-c1 complex subunit 1 and chain D of ATP synthase (Cugno et al. 2016). Hernandez-Castellano et al. performed a label-free LC-MS/MS analysis of the whole proteome of the mammary gland of the same goat breeds that found similar results, including a number of proteins related to protein, carbohydrate and fat metabolism. Furthermore, the tolerant breed showed higher concentrations of immune system-related proteins, whereas the susceptible breed had higher abundance of apoptosis-related proteins (Hernández-Castellano et al. 2016).

The proteomic investigation of the consequences of SWL on skeletal muscle revealed putative biomarkers of tolerance to weight loss: desmin, troponin T, phosphoglucomutase and the histidine triad nucleotide-binding protein 1 (Almeida et al. 2016). SWL also affects the structure and protein composition of wool fibres from merino sheep. A significant reduction in the fibre diameter and an increase in the expression of the high sulphur protein KAP13.1 and proteins from the high glycine-tyrosine protein KAP6 family were found in the wool of animals on a restricted diet. Such findings have strong implications for the wool industry that favours finer wool (Almeida et al. 2014).

6 Searching for Stress Markers in Poultry by Using Proteomics

6.1 Management Conditions

Poultry production can be extremely intensive and subjected to stressful situations. For instance, animals are kept in cages in groups and transported to the abattoir by road. Hazard et al. studied the effects of restraining transport on thigh muscle in chickens subjected to a 2-h road transport through transcriptomics, proteomics and metabolomics. Animals presented increased levels of corticosterone, the stress hormone in birds, changes in the myofibrillar protein profile and impaired glycogenolysis and glycolysis metabolic processes (Hazard et al. 2011).

6.2 Heat Stress

Heat stress also has detrimental effects in the poultry industry. For example, it compromises feed intake and conversion rates and increases the susceptibility to diseases and infections, resulting in higher mortality rates. The effects of heat stress on the liver proteome of two breeds of foie-producing ducks were studied by Zeng et al. through 2D electrophoresis followed by MALDI-MS. Heat-shock proteins (Hsp) Hsp70 and Hsp10 were clearly increased in both species in response to heat stress, whereas α-enolase and S-adenosylmethionine synthetase showed a different regulation

pattern, revealing that heat stress affected the glycolytic pathway in both duck species. In addition, processes such as antigen processing/presentation and apoptosis seem to essentially affect the heat-stress-resistant duck breed (Zeng et al. 2013).

The effects of acute stress on the testis proteome have been investigated in roosters by Wang et al. Two-dimensional-DIGE followed by MALDI-MS revealed that heat stress has an impact on protein folding and degradation linked to spermatogenesis, contributing to the abnormal, apoptotic, spermatogenic cells observed in this condition. Proteins with antioxidant properties were also differentially regulated, showing a compensatory mechanism that could attenuate heat-stress consequences and could be used as putative markers of tolerance to heat stress in chickens (Wang et al. 2014).

7 Stress and Meat Quality

Plenty of attention has been devoted to changes in muscle proteins after slaughter. It is well known that any cause of preslaughter stress affects meat quality giving rise to pale, soft and exudative meat (PSE or PSE-like meat). The phenomenon occurs essentially in pigs and poultry. It was first described in homozygous (nn) Pietrain pigs carrying the halothane mutation affecting the Ryr1 gene; however, PSE-like meat can also occur in heterozygous (Nn) carriers for this mutation and noncarriers (NN) if animals are slaughtered following an intense effort or stress. Two proteomic approaches based on 2D gels analysed the metabolic changes associated to PSE-like meat in pigs. Changes in Hsp proteins (Hsp72 and Hsp27) and metabolic enzymes indicated that the muscle of nn pigs possesses less anti-oxidative and repair capacities than NN and Nn genotypes. Furthermore, glyceraldehyde 3-phosphate dehydrogenase, an enzyme involved in multiple routes as amino acid, lipid and carbohydrate metabolism and antioxidant defence, has been proposed as stress biomarker in carriers of the nn mutation (Kwasiborski et al. 2008; Picard et al. 2010).

Desai et al. profiled through 2D electrophoresis the muscle proteome in chicken and identified differentially abundant proteins in PSE-like breast samples, including mainly structural proteins. Mitochondrial metabolic enzymes such as phosphoglycerate kinase, creatine kinase M type, beta-enolase, carbonic anhydrase 2, pyruvate kinase and malate dehydrogenase were more abundant in PSE-like meat, whereas glycolytic enzymes including phosphoglycerate mutase-1, alpha-enolase, ATP-dependent 6-phosphofructokinase and fructose 1,6-bisphosphatase were more abundant in non-PSE-like meat (Desai et al. 2016).

8 Final Conclusion

The complementarity of animal behaviour, biochemistry, physiology and immunology, as well as the use of modern, high-throughput technologies such as proteomics, will help to create a more comprehensive scientific basis for animal care and

management. When appropriate biomarkers will be available to assess animal welfare, stress tolerance and their effects on the final product, the quality and effectiveness of animal production will improve, leading to a higher acceptance by the public.

Acknowledgements The authors would like to thank all the colleagues and the funding agencies that have contributed and supported research on farm animal proteomics. Especially, the authors are most grateful to COST Action FA1002—Farm Animal Proteomics. We thank Helena Ariño for her contribution to the figure illustrations.

References

Aich P, Jalal S, Czuba C et al (2007) Comparative approaches to the investigation of responses to stress and viral infection in cattle. OMICS 11:413–434. https://doi.org/10.1089/omi.2007.0023

Almeida AM, Plowman JE, Harland DP et al (2014) Influence of feed restriction on the wool proteome: a combined iTRAQ and fiber structural study. J Proteomics 103:170–177. https://doi.org/10.1016/j.jprot.2014.03.036

Almeida AM, Palhinhas RG, Kilminster T et al (2016) The effect of weight loss on the muscle proteome in the damara, dorper and Australian Merino ovine breeds. PLoS One 11:e0146367. https://doi.org/10.1371/journal.pone.0146367

Amado FML, Ferreira RP, Vitorino R (2013) One decade of salivary proteomics: current approaches and outstanding challenges. Clin Biochem 46:506–517. https://doi.org/10.1016/j.clinbiochem.2012.10.024

Anderson NL, Anderson NG (2002) The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics 1:845–867

Ang C-S, Binos S, Knight MI et al (2011) Global survey of the Bovine salivary proteome: integrating multidimensional prefractionation, targeted, and glycocapture strategies. J Proteome Res 10:5059–5069. https://doi.org/10.1021/pr200516d

Apweiler R, Aslanidis C, Deufel T et al (2009) Approaching clinical proteomics: current state and future fields of application in fluid proteomics. Clin Chem Lab Med 47:724–744. https://doi.org/10.1515/CCLM.2009.167

Bassols A, Turk R, Roncada P (2014) A proteomics perspective: from animal welfare to food safety. Curr Protein Pept Sci 15:156–168

Bislev SL, Deutsch EW, Sun Z et al (2012) A Bovine PeptideAtlas of milk and mammary gland proteomes. Proteomics 12:2895–2899. https://doi.org/10.1002/pmic.201200057

Blokhuis HJ, Keeling LJ, Gavinelli A, Serratosa J (2008) Animal welfare's impact on the food chain. Trends Food Sci Technol. https://doi.org/10.1016/j.tifs.2008.09.007

Broom DM (1986) Indicators of poor welfare. Br Vet J 142:524–526. https://doi.org/10.1016/0007-1935(86)90109-0

Broom DM, Johnson KG (1993) Stress and animal welfare. Anim Welf 2:195–218

Bundgaard L, Jacobsen S, Sørensen MA et al (2014) The Equine PeptideAtlas: a resource for developing proteomics-based veterinary research. Proteomics 14:763–773. https://doi.org/10.1002/pmic.201300398

Bushong DM, Friend TH, Knabe DA (2000) Salivary and plasma cortisol response to adrenocorticotropin administration in pigs. Lab Anim 34:171–181. https://doi.org/10.1258/002367700780457482

Cairoli F, Battocchio M, Veronesi MC et al (2006) Serum protein pattern during cow pregnancy: acute-phase proteins increase in the peripartum period. Electrophoresis 27:1617–1625. https://doi.org/10.1002/elps.200500742

Chutipongtanate S, Chatchen S, Svasti J (2017) Plasma prefractionation methods for proteomic analysis and perspectives in clinical applications. Proteomics Clin Appl 11(7–8):1600135. https://doi.org/10.1002/prca.201600135

Criado L, Arroyo L, Valent D et al (2014) Proteomic characterization of molecular changes in saliva proteins in stressed pigs. In: de Almeida A, Ceciliani F, Eckersall P et al (eds) Farm animal proteomics. Wageningen Academic Publishers, Wageningen, pp 37–40

Cruzen SM, Pearce SC, Baumgard LH et al (2015) Proteomic changes to the sarcoplasmic fraction of predominantly red or white muscle following acute heat stress. J Proteomics 128:141–153. https://doi.org/10.1016/j.jprot.2015.07.032

Cugno G, Parreira JR, Ferlizza E et al (2016) The goat (Capra hircus) mammary gland mitochondrial proteome: a study on the effect of weight loss using blue-native PAGE and two-dimensional gel electrophoresis. PLoS One 11:e0151599. https://doi.org/10.1371/journal.pone.0151599

Cui Y, Gu X (2015) Proteomic changes of the porcine small intestine in response to chronic heat stress. J Mol Endocrinol 55:JME-15-0161. https://doi.org/10.1530/JME-15-0161

Dalmau A, Fabrega E, Velarde A (2009) Fear assessment in pigs exposed to a novel object test. Appl Anim Behav Sci 117:173–180

Danielsen M, Codrea MC, Ingvartsen KL et al (2010) Quantitative milk proteomics—host responses to lipopolysaccharide-mediated inflammation of bovine mammary gland. Proteomics 10:2240–2249. https://doi.org/10.1002/pmic.200900771

Dawkins M (2004) Using behaviour to assess animal welfare. Anim Welf 13:S3–S7

De Kloet ER, Joels M, Oitzl M (1991) Implication of brain corticosteroid receptor diversity for the adaptation syndrome concept. In: Jasmin G, Cantin M (eds) Stress revisited II. Systemic effects of stress. Karger, New York

de Roos B, Duthie SJ, Polley AC et al (2008) Proteomic methodological recommendations for studies involving human plasma, platelets, and peripheral blood mononuclear cells. J Proteome Res 7:2280–2290. https://doi.org/10.1021/pr700714x

Desai MA, Jackson V, Zhai W et al (2016) Proteome basis of pale, soft, and exudative-like (PSE-like) broiler breast (*Pectoralis major*) meat. Poult Sci 95:2696–2706. https://doi.org/10.3382/ps/pew213

Duncan I, Fraser D (1997) Understanding animal welfare. In: Appleby M, Hughes BO (eds) Animal Welfare. CAB International, Wallingford

Elsik CG, Tellam RL, Worley KC et al (2009) The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science 324:522–528. https://doi.org/10.1126/science.1169588

EU (2009) Treaty of Lisbon amending the treaty on European Union and the treaty establishing the European Community, p 51. http://eur-lex.europa.eu/legal-content/en/ALL/?uri=OJ:C:2007:306:TOC

FAWC (1992) Farm Animal Welfare Council updates the five freedoms. Vet Rec 17:357

Fernandez-Gomez F-J, Jumeau F, Derisbourg M et al (2014) Consensus brain-derived protein, extraction protocol for the study of human and murine brain proteome using both 2D-DIGE and mini 2DE immunoblotting. J Vis Exp. https://doi.org/10.3791/51339

Fuchs D, Winkelmann I, Johnson IT et al (2005) Proteomics in nutrition research: principles, technologies and applications. Br J Nutr 94:302–314

Fuentes-Rubio M, Cerón JJ, de Torre C et al (2014) Porcine salivary analysis by 2-dimensional gel electrophoresis in 3 models of acute stress: a pilot study. Can J Vet Res 78:127–132

Gianazza E, Miller I, Palazzolo L et al (2016) With or without you — Proteomics with or without major plasma/serum proteins. J Proteomics 140:62–80. https://doi.org/10.1016/j.jprot.2016.04.002

Groenen MA, Archibald AL, Uenishi H et al (2012) Analyses of pig genomes provide insight into porcine demography and evolution. Nature 491:393–398. https://doi.org/10.1038/nature11622

Gutierrez AM, Martinez-Subiela S, Soler L et al (2009) Use of saliva for haptoglobin and C-reactive protein quantifications in porcine respiratory and reproductive syndrome affected pigs in field conditions. Vet Immunol Immunopathol 132:218–223. https://doi.org/10.1016/j.vetimm.2009.06.013

Gutierrez AM, Miller I, Hummel K et al (2011) Proteomic analysis of porcine saliva. Vet J 187:356–362. https://doi.org/10.1016/j.tvjl.2009.12.020

Gutiérrez AM, Nöbauer K, Soler L et al (2013) Detection of potential markers for systemic disease in saliva of pigs by proteomics: a pilot study. Vet Immunol Immunopathol 151:73–82. https://doi.org/10.1016/j.vetimm.2012.10.007

Gutiérrez AM, Miller I, Kolarich D et al (2016) Detection and first characterization of an uncommon haptoglobin in porcine saliva of pigs with rectal prolapse by using boronic acid sample enrichment. Animal:1–9. https://doi.org/10.1017/S1751731116002159

Hazard D, Fernandez X, Pinguet J et al (2011) Functional genomics of the muscle response to restraint and transport in chickens. J Anim Sci 89:2717–2730. https://doi.org/10.2527/jas.2010-3288

He F (2005) Human liver proteome project: plan, progress, and perspectives. Mol Cell Proteomics 4:1841–1848. https://doi.org/10.1074/mcp.R500013-MCP200

Hernández-Castellano LE, Ferreira AM, Nanni P et al (2016) The goat (Capra hircus) mammary gland secretory tissue proteome as influenced by weight loss: a study using label free proteomics. J Proteomics. https://doi.org/10.1016/j.jprot.2016.03.030

Hesselager MO, Codrea MC, Sun Z et al (2016) The pig PeptideAtlas: a resource for systems biology in animal production and biomedicine. Proteomics 16:634–644. https://doi.org/10.1002/pmic.201500195

Hillier LW, Miller W, Birney E et al (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695–716. https://doi.org/10.1038/nature03154

Hollung K, Timperio AM, Olivan M et al (2014) Systems biology: a new tool for farm animal science. Curr Protein Pept Sci 15:100–117

Hughes B, Duncan I (1988) The notion of ethological “need”, models of motivation and animal welfare. Anim Behav 36:1696–1707

Humphrey SP, Williamson RT (2001) A review of saliva: normal composition, flow, and function. J Prosthet Dent 85:162–169. https://doi.org/10.1067/mpr.2001.113778

Kaufman E, Lamster IB (2002) The diagnostic applications of saliva—a review. Crit Rev Oral Biol Med 13:197–212

Krief G, Deutsch O, Gariba S et al (2011) Improved visualization of low abundance oral fluid proteins after triple depletion of alpha amylase, albumin and IgG. Oral Dis 17:45–52. https://doi.org/10.1111/j.1601-0825.2010.01700.x

Kuhla B, Kuhla S, Rudolph PE et al (2007) Proteomics analysis of hypothalamic response to energy restriction in dairy cows. Proteomics 7:3602–3617. https://doi.org/10.1002/pmic.200700248

Kuhla B, Albrecht D, Kuhla S, Metges CC (2009) Proteome analysis of fatty liver in feed-deprived dairy cows reveals interaction of fuel sensing, calcium, fatty acid, and glycogen metabolism. Physiol Genomics 37:88–98. https://doi.org/10.1152/physiolgenomics.90381.2008

Kuhla B, Albrecht D, Bruckmaier R et al (2010) Proteome and radioimmunoassay analyses of pituitary hormones and proteins in response to feed restriction of dairy cows. Proteomics 10:4491–4500. https://doi.org/10.1002/pmic.201000383

Kuhla B, Nürnberg G, Albrecht D et al (2011) Involvement of skeletal muscle protein, glycogen, and fat metabolism in the adaptation on early lactation of dairy cows. J Proteome Res 10:4252–4262. https://doi.org/10.1021/pr200425h

Kwasiborski A, Sayd T, Chambon C et al (2008) Pig Longissimus lumborum proteome: part II: relationships between protein content and meat quality. Meat Sci 80:982–996. https://doi.org/10.1016/j.meatsci.2008.04.032

Lamy E, Mau M (2012) Saliva proteomics as an emerging, non-invasive tool to study livestock physiology, nutrition and diseases. J Proteomics 75:4251–4258. https://doi.org/10.1016/j.jprot.2012.05.007
Lawrence HP (2002) Salivary markers of systemic disease: noninvasive diagnosis of disease and monitoring of general health. J Can Dent Assoc 68:170–174
Leek ABG, Sweeney BT, Duffy P et al (2004) Effect of stocking density and social regrouping stressors on growth performance, carcass characteristics, nutrient digestibility and physiological stress responses in pigs. Anim Sci 79:109–119
Lista S, Faltraco F, Hampel H (2013) Biological and methodical challenges of blood-based proteomics in the field of neurological research. Prog Neurobiol 101–102:18–34. https://doi.org/10.1016/j.pneurobio.2012.06.006
Loor JJ, Bionaz M, Drackley JK (2013) Systems physiology in dairy cattle: nutritional genomics and beyond. Annu Rev Anim Biosci 1:365–392. https://doi.org/10.1146/annurev-animal-031412-103728
Lubec G, Krapfenbauer K, Fountoulakis M (2003) Proteomics in brain research: potentials and limitations. Prog Neurobiol 69:193–211
Luque-Garcia JL, Neubert TA (2007) Sample preparation for serum/plasma profiling and biomarker identification by mass spectrometry. J Chromatogr A 1153:259–276. https://doi.org/10.1016/j.chroma.2006.11.054
Lygirou V, Makridakis M, Vlahou A (2015) Biological sample collection for clinical proteomics: existing SOPs. In: Methods in molecular biology (Clifton, N.J.). Humana Press, New York, pp 3–27
Marco-Ramell A, Bassols A (2010) Enrichment of low-abundance proteins from bovine and porcine serum samples for proteomic studies. Res Vet Sci 89:340–343. https://doi.org/10.1016/j.rvsc.2010.03.019
Marco-Ramell A, Pato R, Pena R et al (2011) Identification of serum stress biomarkers in pigs housed at different stocking densities. Vet J 190:e66–e71. https://doi.org/10.1016/j.tvjl.2011.01.003
Marco-Ramell A, Arroyo L, Saco Y et al (2012) Proteomic analysis reveals oxidative stress response as the main adaptative physiological mechanism in cows under different production systems. J Proteomics 75:4399–4411. https://doi.org/10.1016/j.jprot.2012.04.002
Marco-Ramell A, Arroyo L, Peña R et al (2016) Biochemical and proteomic analyses of the physiological response induced by individual housing in gilts provide new potential stress markers. BMC Vet Res 12:265. https://doi.org/10.1186/s12917-016-0887-1
Mau M, Kaiser TM, Südekum K-H (2009) Evidence for the presence of carbonic anhydrase 29-kDa isoenzyme in salivary secretions of three ruminating species and the gelada baboon. Arch Oral Biol 54:354–360. https://doi.org/10.1016/j.archoralbio.2008.12.004
Merlot E, Mounier AM, Prunier A (2011) Endocrine response of gilts to various common stressors: a comparison of indicators and methods of analysis. Physiol Behav 102:259–265. https://doi.org/10.1016/j.physbeh.2010.11.009
Mesmin C, van Oostrum J, Domon B (2016) Complexity reduction of clinical samples for routine mass spectrometric analysis. Proteomics Clin Appl 10:315–322. https://doi.org/10.1002/prca.201500135
Min L, Cheng J, Zhao S et al (2016) Plasma-based proteomics reveals immune response, complement and coagulation cascades pathway shifts in heat-stressed lactating dairy cows. J Proteomics 146:99–108. https://doi.org/10.1016/j.jprot.2016.06.008
Mitchell GB, Clark ME, Siwicky M, Caswell JL (2008) Stress alters the cellular and proteomic compartments of bovine bronchoalveolar lavage fluid. Vet Immunol Immunopathol 125:111–125. https://doi.org/10.1016/j.vetimm.2008.05.005
Mormede P, Andanson S, Auperin B et al (2007) Exploration of the hypothalamic-pituitary-adrenal function as a tool to evaluate animal welfare. Physiol Behav 92:317–339. https://doi.org/10.1016/j.physbeh.2006.12.003

Moyes KM, Bendixen E, Codrea MC, Ingvartsen KL (2013) Identification of hepatic biomarkers for physiological imbalance of dairy cows in early and mid lactation using proteomic technology. J Dairy Sci 96:3599–3610. https://doi.org/10.3168/jds.2012-5900
Murata H (2007) Stress and acute phase protein response: an inconspicuous but essential linkage. Vet J 173:473–474. https://doi.org/10.1016/j.tvjl.2006.05.008
Nieuw Amerongen AV, Ligtenberg AJM, Veerman ECI (2007) Implications for diagnostics in the biochemistry and physiology of saliva. Ann NY Acad Sci 1098:1–6. https://doi.org/10.1196/annals.1384.033
Paredi G, Raboni S, Bendixen E et al (2012) "Muscle to meat" molecular events and technological transformations: the proteomics insight. J Proteomics 75:4275–4289. https://doi.org/10.1016/j.jprot.2012.04.011
Parrott RF, Misson BH, Baldwin BA (1989) Salivary cortisol in pigs following adrenocorticotrophic hormone stimulation: comparison with plasma levels. Br Vet J 145:362–366. https://doi.org/10.1016/0007-1935(89)90034-1
Pearce SC, Lonergan SM, Huff-Lonergan E et al (2015) Acute heat stress and reduced nutrient intake alter intestinal proteomic profile and gene expression in pigs. PLoS One 10:e0143099. https://doi.org/10.1371/journal.pone.0143099
Picard B, Berri C, Lefaucheur L et al (2010) Skeletal muscle proteomics in livestock production. Brief Funct Genomics 9:259–278. https://doi.org/10.1093/bfgp/elq005
Piras C, Roncada P, Rodrigues PM et al (2016) Proteomics in food: quality, safety, microbes, and allergens. Proteomics 16:799–815. https://doi.org/10.1002/pmic.201500369
Prickett JR, Zimmerman JJ (2010) The development of oral fluid-based diagnostics and applications in veterinary medicine. Anim Heal Res Rev 11:207–216. https://doi.org/10.1017/S1466252310000010
Rai AJ, Gelfand CA, Haywood BC et al (2005) HUPO Plasma Proteome Project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. Proteomics 5:3262–3277. https://doi.org/10.1002/pmic.200401245
Ramirez A, Wang C, Prickett JR et al (2012) Efficient surveillance of pig populations using oral fluids. Prev Vet Med 104:292–300. https://doi.org/10.1016/j.prevetmed.2011.11.008
Sapolsky RM (1992) Neuroendocrinology of the stress response. In: Becker J, Breedlove S, Crews D (eds) Behavioral endocrinology. Massachusetts Institute of Technology, Cambridge, MA
Schipper RG, Silletti E, Vingerhoeds MH (2007) Saliva as research material: biochemical, physicochemical and practical aspects. Arch Oral Biol 52:1114–1135. https://doi.org/10.1016/j.archoralbio.2007.06.009
Thadikkaran L, Siegenthaler MA, Crettaz D et al (2005) Recent advances in blood-related proteomics. Proteomics 5:3019–3034. https://doi.org/10.1002/pmic.200402053
Topkas E, Keith P, Dimeski G et al (2012) Evaluation of saliva collection devices for the analysis of proteins. Clin Chim Acta 413:1066–1070. https://doi.org/10.1016/j.cca.2012.02.020
Turk R, Juretic D, Geres D et al (2008) Influence of oxidative stress and metabolic adaptation on PON1 activity and MDA level in transition dairy cows. Anim Reprod Sci 108:98–106. https://doi.org/10.1016/j.anireprosci.2007.07.012
Tuyttens F (2007) Stereotypies. In: Velarde A, Geers R (eds) On farm monitoring of pig welfare. Wageningen Academic Publishers, Wageningen, pp 41–46
Valent D, Arroyo L, Peña R et al (2017) Effects on pig immunophysiology, PBMC proteome and brain neurotransmitters caused by group mixing stress and human-animal relationship. PLoS One 12:e0176928. https://doi.org/10.1371/journal.pone.0176928
van Altena SEC, de Klerk B, Hettinga KA et al (2016) A proteomics-based identification of putative biomarkers for disease in bovine milk. Vet Immunol Immunopathol 174:11–18. https://doi.org/10.1016/j.vetimm.2016.04.005
Velarde A (2007) Agonistic behaviour. In: Velarde A, Geers R (eds) On farm monitoring of pig welfare. Wageningen Academic Publishers, Wageningen, pp 53–56
Velarde A, Dalmau A (2012) Animal welfare assessment at slaughter in Europe: moving from inputs to outputs. Meat Sci 92:244–251

Verma N, Rettenmeier AW, Schmitz-Spanke S (2011) Recent advances in the use of Sus scrofa (pig) as a model system for proteomic studies. Proteomics 11:776–793. https://doi.org/10.1002/pmic.201000320

Vuckovic D (2012) Current trends and challenges in sample preparation for global metabolomics using liquid chromatography–mass spectrometry. Anal Bioanal Chem 403:1523–1548. https://doi.org/10.1007/s00216-012-6039-y

Wang S-H, Cheng C-Y, Chen C-J et al (2014) Changes in protein expression in testes of L2 strain Taiwan country chickens in response to acute heat stress. Theriogenology 82:80–94. https://doi.org/10.1016/j.theriogenology.2014.03.010

Wattiez R, Falmagne P (2005) Proteomics of bronchoalveolar lavage fluid. J Chromatogr B 815:169–178. https://doi.org/10.1016/j.jchromb.2004.10.029

Webster A, Main D, Whay H (2004) Welfare assessment: indices from clinical observation. Anim Welf 13:93–98

Xu C, Wang Z (2008) Comparative proteomic analysis of livers from ketotic cows. Vet Res Commun 32:263–273. https://doi.org/10.1007/s11259-007-9028-4

Yang YX, Wang JQ, Bu DP et al (2012) Comparative proteomics analysis of plasma proteins during the transition period in dairy cows with or without subclinical mastitis after calving. Czech J Anim Sci 57:481–489

Zeng T, Jiang X, Li J et al (2013) Comparative proteomic analysis of the hepatic response to heat stress in Muscovy and Pekin ducks: insight into thermal tolerance related to energy metabolism. PLoS One 8:e76917. https://doi.org/10.1371/journal.pone.0076917

Bioinformatics Support for Farm Animal Proteomics

Aivett Bilbao and Frédérique Lisacek

Abstract In this chapter, we attempt to compile information published in the most recent reviews and regular publications highlighting the use of bioinformatics in the field of veterinary proteomics. We present a summary of the data resources and popular end user-oriented computational tools that do not require advanced informatics skills.

1 Introduction

The application of proteomics in veterinary science is lagging behind in comparison to studies that have explored the potential of advanced proteomic technologies in human research. The situation is particularly acute in clinical medicine (Ceciliani et al. 2014). This slow start may turn out an advantage, as the recent boost in veterinary proteomics is contemporary with technological development (e.g. targeted proteomics or data-independent acquisition) and improvement of method accuracy and coverage. This progress is in turn challenging the design of automation procedures necessary to cope with ever-increasing amounts of data, thereby justifying a dedicated chapter on bioinformatics in this book.

Several generic reviews summarise the advent of shotgun proteomics (Aebersold 2003; Nesvizhskii 2010) that provides the original context for software development. The corresponding methodology mainly focused on protein identification based on sequence database search engines is now mature with its recognised shortcomings such as a strong bias towards abundant proteins, and the present chapter is centred on more recent efforts applicable to veterinary science.

A. Bilbao
Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA
e-mail: bilbao.aivett@gmail.com

F. Lisacek (✉)
Proteome Informatics Group, SIB Swiss Institute of Bioinformatics, Geneva, Switzerland
e-mail: frederique.lisacek@sib.swiss

© Springer International Publishing AG 2018
A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_18

One of the remarkable paradigm shifts from shotgun proteomics with data-dependent acquisition (DDA) is data-independent acquisition (DIA). It introduces the systematic fragmentation of precursor ions as opposed to a limited selection based on high-intensity peaks. Within successive and overlapping sliding windows along the *m*/*z* dimension, all precursor ions are fragmented resulting in a very populated tandem MS collection of redundant data (Chapman et al. 2014; Bilbao et al. 2015a). The sheer amount of mass spectra requires smart and accurate software for data analysis (Bilbao et al. 2015a). DIA is now well established for its ability to monitor detectable peptides with high sensitivity and reproducibility across multiple samples. Specifically, since DIA fragmentation of all detectable ions within a wide *m*/*z* range is systematically carried out regardless of intensity, then extracted ion chromatograms (XICs) can be generated at the fragment ion level. It is therefore particularly suitable to perform more consistent and accurate quantification. A direct consequence of using DIA methods is the expanding use of spectral libraries in particular as an alternative approach to sequence database search. Matching experimental to reference spectra is considerably faster and less error-prone than checking all possible theoretical spectra of the tryptic digest of proteins. Corresponding software tools for spectral library searching are reviewed in Griss (2016). However, the lack of a standardised file format and the possible incomplete coverage of spectral libraries are still limiting expansion. At present, library search is recognised as a helpful complementing approach to database search.

The second major development in recent proteomics is quantification and in particular targeted proteomics (Picotti and Aebersold 2012). Selecting the appropriate peptides, optimising fragment prediction and integrating these steps in a pipeline are the main bioinformatics challenges as summarised in Reker and Malmström (2012). These are definitely usable across the board, irrespective of the application.

Finally, a crucial point remains that of data and processing quality. Indeed, data should be appropriately processed with robust software ensuring reproducible and accurate results. Even with robust software and optimised settings, low-quality data yield poor and questionable results. Reproducible and high confidence results strongly rely on software usability and the ability to choose the most appropriate parameters, since different parameters could lead to different results and wrong interpretations or conclusions. Within all the steps in the proteomics workflow, from sample collection to data processing, "Mount Bioinformatics" remains the last and the highest peak to climb (Aebersold 2009). The challenges of bioinformatics in software for quantitative proteomics have been previously described (Cappadona et al. 2012). Important issues that directly impact the effectiveness of proteomic quantitation and common tasks in computational solutions to correct for the occurrence of these factors are well depicted. This chapter surveys the different resources and software tools that are currently in use for data reference and analysis in the field of proteomics and tackles veterinary proteomics issues in this context either referring to published work or to prospective solutions.

2 Data Resources

As presented by Perez-Riverol et al. (2015), the information generated in proteomics experiments is organised in three levels:

(1) Raw MS data
(2) Processed experimental data
(3) Interpreted biological results

Level (1) corresponds to MS data collection through the ProteomeXchange protocol (Vizcaıno et al. 2014; Deutsch et al. 2017). In the past few years, this international initiative has allowed channelling all reported/published experimental data into three main repositories through standardised submission and dissemination pipelines.

Level (2) encompasses a significant number of databases storing peptide/protein identification and quantification whenever available.

Level (3) is associated with the concept of "knowledgebase" in which protein information is curated and recorded and can be searched, compared or mined.

2.1 Generic Proteomics Databases in a Nutshell

Several recent reviews comprehensively cover the topic of proteomics databases (Martens 2010; Perez-Riverol et al. 2015). Suffice to say that the collection of identified proteins is steadily growing and this broadens the extent of comparative or integrative approaches of data analysis. This is made possible through the generalised use of shared data formats and standards acknowledged by the Proteomics Standard Initiative (PSI; http://psidev.info). Table 1 summarises the range of formats in use in the current proteomics databases.

Martens and Vizcaıno (2017) very recently praised the "golden age" of proteomic data sharing precisely based on the availability and broad usage of standards.

2.1.1 Mass Spectrometry Data

File formats commonly used in MS-based proteomics are reviewed by Deutsch (2012). Mass spectra are stored in multiple repositories: PeptideAtlas (Farrah et al. 2013), GPMDB (Craig and Beavis 2004), Massive (https://massive.ucsd.edu) and PRIDE (Vizcaıno et al. 2015) to cite the most renowned. The vast majority of these publicly available datasets are generated for human and the main model organisms (*S. cerevisiae*, *D. melanogaster*, *C. elegans*, etc.). However, some dedicated resources have been created for farm animal proteomics as detailed in Sect. 2.2.1. One of the obvious advantages of such raw data availability is the potential for reuse and reanalysis. Reanalysis can be done individually, but note that PeptideAtlas and

Table 1 Formats in use in the current proteomics databases

Encoding purpose	Standard name	References
Mass spectra	mzXML	Pedrioli et al. (2004)
	mzML	Martens (2010)
Peptide/protein	pepXML/protXML	Keller et al. (2005)
identifications	mzIdentML	Jones et al. (2012)
Quantitative analysis	mzQuantML	Walzer et al. (2013)
SRM transitions	TraML	Deutsch et al. (2012)

GPMDB, for instance, routinely reprocess many datasets with in-house bioinformatics pipelines.

The second remarkable feature of these repositories is the option of defining a characteristic spectral library. Despite the existence of institutions such as the National Institute of Standards and Technology (NIST) where reference spectral libraries are collected, it may be of interest to constitute a specialised library. As mentioned in the introduction, the current shift from DDA to DIA approaches emphasises the need for quality spectral libraries.

It is important to stress that minimal quality checks are undertaken so that the resulting data resources contain uneven quality spectra and subsequent more or less reliable identifications of peptides and proteins.

2.1.2 Integrated Data

UniProt (www.uniprot.org) and protein data collected at NCBI (https://www.ncbi.nlm.nih.gov/protein) are the main sequence sources used in database search engines. UniProt however has a greater level of data integration and contains a wealth of information beyond sequence features including protein expression and structure. It is, as such, the most popular resource used for characterising proteins as finely as possible.

Recently proteogenomics has become a popular approach for merging information originating from genomics, transcriptomics and proteomics studies. Indeed, DNA sequence and RNA expression data accumulation over the past three decades provides rich sources to be combined with ever-increasing proteomics data. Proteogenomics currently significantly contributes to the identification of sequence variants, especially in humans, and the tools developed in this context could easily benefit farm animal proteomics studies. At the time of writing this chapter, no farm animal proteogenomics study has been published.

2.2 Farm Animal Proteomics Dedicated Databases

Farm animal dedicated databases are usually focused on genetic mapping information. Though apparently discontinued, ArkDB (http://www.thearkdb.org/arkdb) still hosts genome mapping data from farmed and other animal species spanning genetic linkage and QTL (quantitative trait locus). A related website called ResSpecies (http://www.resspecies.org/resspecies/) gathers tools for displaying or exporting genotyping and phenotype data as well as population coverage and markers. Along the same lines, QTLdb [http://www.animalgenome.org/QTLdb, (Hu et al. 2015)] collects publicly available trait mapping data for a smaller range of animal and is associated with CorrDB, the animal trait correlation database (http://www.animalgenome.org/cgi-bin/CorrDB). In order to complete this abundant genetic information, some effort was invested into developing proteomic-centred resources especially in capturing mass spectrometry data.

2.2.1 Farm Animal Mass Spectrometry Data

As mentioned in Sect. 2.1.1, mass spectrometry data repositories have recently imposed a paradigm shift in considering published data. PeptideAtlas pioneered in collecting datasets dedicated to farm animals though admittedly, these are not as often updated or enriched with new data as the human collection is, as shown in the homepage where all newly included sets are listed.

The frequent use of mass spectrometry in studying milk and its constituents naturally led to the first initiative devoted to collecting MS data in bovine milk and mammary gland (Bislev et al. 2012). Then the Equine PeptideAtlas (Bundgaard et al. 2014) and the Pig PeptideAtlas were introduced (Hesselager et al. 2016), and finally the PeptideAtlas for the domestic chicken (McCord et al. 2017) is the most recent addition to the collection. This data can be queried from the generic interface (“Queries” tab), and corresponding data can be downloaded.

2.2.2 Farm Animal Integrated Data

The success of proteomic-based investigations largely depends on the availability of complete and annotated databases containing the gene and protein sequence information for the animal species of interest (Soares et al. 2012). Nonetheless, once proteins are identified by matching mass spectrometry data with sequence data, according to various strategies detailed in Sect. 3, stored and in the best of cases integrated knowledge of protein properties is very useful to potentially rationalise the content of the studied sample. There are very few such resources for FAP, and the following three cover the range:

1. ProteINSIDE specifically supports the annotation of farm animal proteomics experiments (http://www.proteinside.org; Kaspric et al. 2015). It stores

information on bovine, sheep and goat. An in-built workflow processes lists of proteins input by a user to extract information on protein function, subcellular location (secreted/cytoplasmic) and interacting partners.

2. AgBase is an alternative resource for functional annotation though not proteomics-oriented (http://www.agbase.msstate.edu; McCarthy et al. 2010). It covers more farm animal organisms.
3. paxDB stores proteome-wide protein abundance information across organisms and tissues (http://pax-db.org; Wang et al. 2015b). Pig, bovine, horse and chicken are included.

3 Data Analysis Software

The use of two-dimensional gel electrophoresis (2DE) approaches combined with MS already allowed the characterisation of several distinct proteomes in different fields of animal science (Soares et al. 2012). However, the complexity of most proteomic samples challenges the separation power of such traditional techniques. More importantly, even with the current improvements in 2DE, it is still a manual and time-demanding process. In proteomics studies, liquid chromatography (LC) has been increasingly used as a replacement technique for gel electrophoresis, since it can be employed to analyse large numbers of samples in a faster, automated and more repeatable fashion. This trend can be observed in Fig. 1.

As one of the core technologies routinely used in advanced proteomics research, we focus this section on software and computational methods for analysing LC-MS data. We particularly describe the data and associated algorithms from a

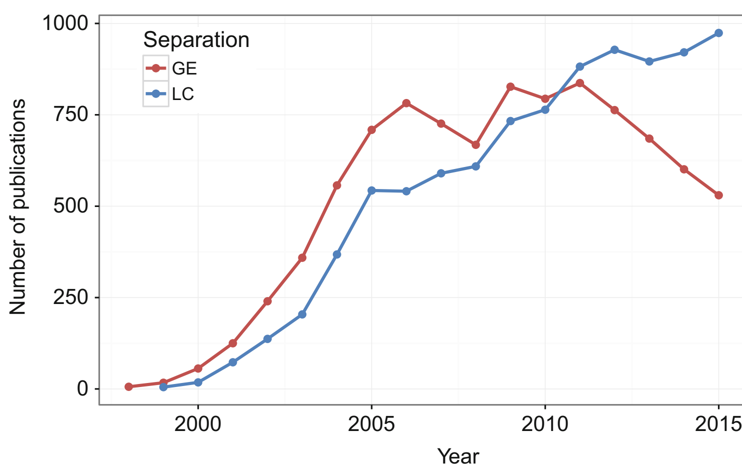

Fig. 1 Estimated number of MS-based proteomics publications per year: gel electrophoresis (GE, *red*) vs. liquid chromatography (LC, *blue*). PubMed queries: “proteomics gel electrophoresis mass spectrometry” and “proteomics liquid chromatography mass spectrometry”

perspective of the acquisition methods by which it is generated, because understanding the process of data generation may turn out to be critical for successful data analysis and result interpretation.

Prior to LC separation, complex protein samples are digested into peptides. Since the chemistry of peptides is more uniform than the chemistry of proteins, tandem MS methods to sequence peptides are nowadays robust and mature (Yates III 2015). The most popular enzyme used for protein digestion is trypsin, which leads to peptides with C-terminally protonated amino acids, providing an advantage in subsequent MS-based peptide sequencing (Aebersold 2003). As peptides elute from the LC column, they are ionised by electrospray ionisation (ESI), and resulting ions are analysed by the mass spectrometer.

The MS analysis can be performed with different instruments and different operation modes. Broadly speaking, they can be classified as targeted and shotgun methods. The characteristics of the spectra and informatics approaches for these methods are discussed in the following subsections.

As illustrated in Fig. 2, we emphasise examples of chromatogram-based MS quantification workflows using the Skyline software tool (MacLean et al. 2010), developed and maintained by the MacCoss lab. Skyline is a Windows client, versatile and robust platform that can be used to analyse the different types of MS data here described. A key feature of Skyline is the extraction of MS data directly from many instrument vendor formats, that is, a conversion to open file formats is usually not required. Skyline is freely available and open source, with an interactive and rather intuitive graphical user interface (GUI) for visualisation, and several tutorials describing the usage are available. Results can be exported as text reports that researchers can further process using other tools.

Moreover, several tools are available as plug-in modules within Skyline. A framework called “external tools” allows researchers to integrate their tools into Skyline without modifying the Skyline codebase (Broudy et al. 2014). With a uniform interface for installation into Skyline, the external tools can be easily accessed by all users for downstream statistical analyses.

3.1 Targeted MS Methods

Selected reaction monitoring (SRM)—also referred to as multiple reaction monitoring (MRM)—is a targeted MS technique whereby a predefined series of transitions (precursor/fragment ion pairs) are selected by the two mass filters of a triple quadrupole instrument and monitored over chromatographic elution for precise quantification (Lange et al. 2008; Picotti and Aebersold 2012).

Since SRM strictly targets a predetermined set of peptides, it is particularly useful when only a handful of proteins from a complex background, such as those constituting a cellular network or a set of candidate biomarkers, need to be measured across multiple samples in a consistent, reproducible and quantitatively precise manner.

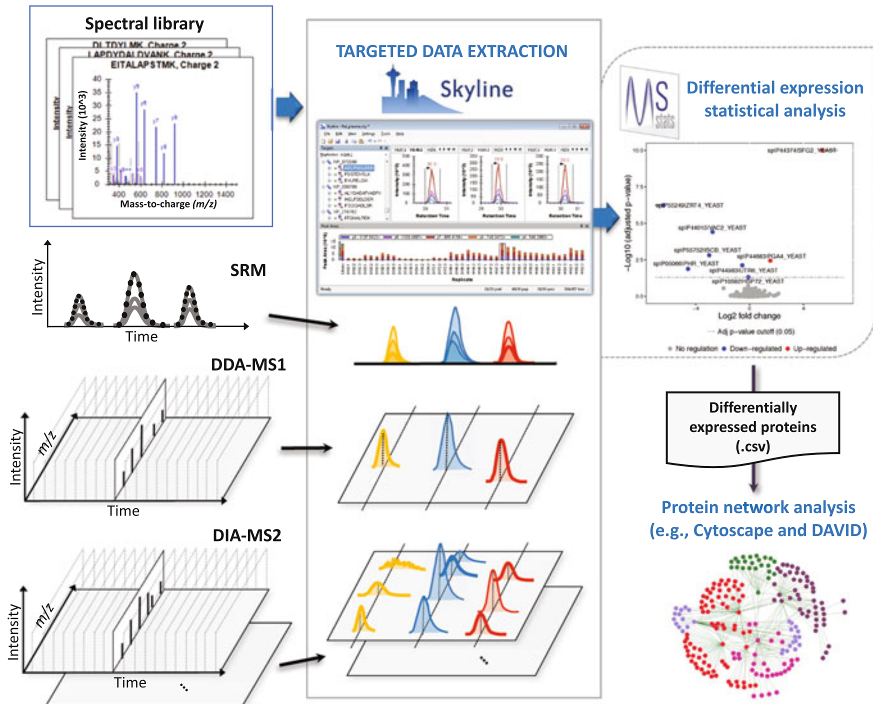

Fig. 2 Chromatogram-based MS quantification workflows in proteomics. Raw MS data generated by LC-MS analysis have different complexities depending on the purpose of the study. Targeted MS methods such as SRM require more effort prior to the actual LC-MS analysis, but the data contains single chromatogram traces for each targeted peptide transition and therefore requires less complex data processing. Shotgun MS methods acquire a full mass spectrum within each cycle. For DDA, MS1 data is used for quantification and MS2 data for identification. For DIA, the MS1 data is optional, and quantification is usually performed with the MS2 data, since the systematic fragmentation acquires continuous data for all fragments across the complete elution profile. The *m*/*z* range is typically divided in several precursor isolation windows, and thus DIA spectra contain several fragment ion maps. The spectral library provides information such as *m*/*z* and expected retention time used by Skyline as seeds for chromatogram extraction. The spectral library is not required for processing SRM data, but it might be used for designing the acquisition method, e.g., for selecting the best peptides and fragment ions. For DDA, the spectral library is used to extract the MS1 chromatograms for each identified peptide. For DIA, the spectral library is used to extract the MS2 chromatograms for the fragment ions of each identified peptide, from the spectra corresponding to the precursor isolation window. Statistical analysis to find differentially expressed proteins can be performed using MSstat, either as one of the external tools available within the Skyline GUI or independently using the exported results for the input of the R MSstat package. Further downstream analysis can be performed with tools such as Cytoscape and DAVID

A useful discussion of the best practices for targeted MS measurements in biology and medicine, also applicable in animal science, can be found in Carr et al. (2014). The authors discussed the analytical goals and the experimental evidence needed to properly describe developed assays according to the required

levels of performance, as well as the computational and statistical tools useful for the analysis of targeted MS.

3.1.1 Design of Targeted Acquisition Methods

In SRM-based proteomics, a significant amount of time is required for the design of the acquisition method or assay, and several informatics tools are available to assist this process. For instance, software to select proteotypic peptides, transitions and best acquisition settings include SRMCollider (Röst et al. 2012), MRMOptimizer (Alghanem et al. 2017) and PREGO (Searle et al. 2015).

We highlight the PNNL Biodiversity Plugin as one of the external tools available in Skyline (Degan et al. 2016). The tool summarises available mass spectrometry data in a pathway-centric view and facilitates querying it from a biological perspective to design quantitative experiments. Selected proteins and their underlying mass spectra are imported to Skyline for further assay design (transition selection). The PNNL Biodiversity Library catalogues MS/MS spectra from over 3 million peptides and 230,000 proteins from 118 distinct organisms across the tree of life all cross-referenced to KEGG pathways for intuitive biological interpretation.

To maximise the number of peptides that can be monitored in a single LC-MS analysis, scheduled SRM methods can be designed (Stahl-Zeng et al. 2007). Using information of the expected peptide elution time in the target list, computer programs automatically generate SRM acquisition methods where the transitions of a specific peptide are only targeted during a time window around its elution time. In this way, the number of peptides measured in a LC-MS run is increased without compromising the limit of detection or the quantitative accuracy.

3.1.2 Data Processing for SRM Quantification

Unlike in other MS-based proteomic techniques, no full mass spectra are recorded in SRM analysis. SRM data consist of a set of chromatographic traces with the retention time and signal intensity for each of the monitored transition. This non-scanning nature translates into an increased sensitivity by one or two orders of magnitude compared with conventional "full-scan" techniques. The two levels of quadrupole mass selection with narrow mass windows result in a high selectivity, as co-eluting background ions are filtered out very effectively.

Integration of the chromatographic peaks for each transition supports the relative or, if suitable heavy isotope-labelled reference standards are used, absolute quantification of the targeted peptides, which are used as a surrogate measure of the proteins of interest. Isotope labelling increases the complexity and costs of an experiment with the benefit of more precise quantification.

Other targeted methods, in which high-resolution full MS/MS spectra are acquired for each target peptide, such as parallel reaction monitoring (PRM) (Peterson et al. 2012), generate data for all detectable fragment ions. The third

quadrupole of a triple quadrupole is substituted with a high-resolution and accurate mass analyser.

3.2 Shotgun MS Methods

In contrast to targeted MS methods, the so-called shotgun or bottom-up approach does not require predefined information about the proteins of interest or analytes in the sample. The MS instrument is operated to record the spectra of many peptides as possible, as they elute from the LC column and are ionised by electrospray.

3.2.1 Data-Dependent Acquisition and MS1 Quantification

Within each DDA cycle, ion signals are recorded in a MS1 or survey scan (precursor ion signals), and the top-*N* most abundant ions are then selected and serially isolated for fragmentation (MS/MS, MS2, or tandem MS) to generate structural information. Since fragmentation models are well characterised for amino acid sequences, theoretical spectra can be generated according to factors including peptide sequence and type of fragmentation, typically collision-induced dissociation (CID). Typically, most of the MS/MS data is highly pure (each spectrum contains fragments from mainly one peptide) and can be annotated with peptide sequences using search tools such as Mascot (Perkins et al. 1999), SEQUEST (Eng et al. 1994), Andromeda (Cox et al. 2011), X!Tandem (Craig and Beavis 2004) and MS-GF+(Kim and Pevzner 2014). A search tool in silico digests the protein sequences into peptides and generates theoretical spectra to score the observed tandem mass spectra against the predicted fragmentation (Nesvizhskii 2010).

While search tools produce a match for almost every input MS/MS spectrum, only a fraction of those peptide to spectrum matches (PSMs) are true. The most commonly used and accepted statistical confidence measure is the false discovery rate (FDR) (Benjamini and Hochberg 1995) adapted in proteomics (Elias and Gygi 2007)—also known as “target-decoy approach” (TDA)—as a summary statistics for the entire collection of PSMs. A decoy database is generated by reversing or shuffling the amino acids in the sequences of the reference database and thus included in the search to estimate the FDR as the expected proportion of incorrect PSMs among all accepted PSM (Nesvizhskii 2010). FDR analysis is typically included within the identification software; however, it can also be performed and refined using other tools such as MAYU (Reiter et al. 2009), Percolator (Käll et al. 2007; The M et al. 2016), iProphet (Shteynberg et al. 2011), PeptideProphet (Keller et al. 2002; Choi and Nesvizhskii 2007) and ProteinProphet (Nesvizhskii et al. 2003).

Properly estimating and controlling the FDR are essential steps in the computational pipeline for preventing subtle but profound errors in high-throughput science.

It is necessary to place less emphasis on the number of identifications achieved and instead to value the work as a whole (Serang and Käll 2015).

After FRD analysis, confidently identified peptides can be used to build a library, which can be imported into Skyline to perform label-free relative quantification. Skyline extracts the precursor ion signals of each peptide from the MS1 raw data and computes the area under the peak or ion chromatogram from each peptide elution profile (Schilling et al. 2012).

Another popular method that has been traditionally used for label-free quantification in shotgun proteomics is the spectral count, which is based on the number of MS/MS spectra identified for each peptide sequence (Liu et al. 2004). Despite providing a rapid and semi-quantitative measure of abundance, spectral count-based quantification is affected by sample complexity; it has been found to often give irreproducible results and being unsuitable for quantifying low-abundance proteins (Cappadona et al. 2012; Ahrné et al. 2013). In contrast, chromatogram-based MS quantification based on integration of the peptide elution profile provides a level of accuracy comparable to labelling approaches.

3.2.2 Data-Independent Acquisition and MS2 Quantification

With recent developments in MS instrumentation, application of alternative MS operation modes such as DIA has become feasible (Chapman et al. 2014; Bilbao et al. 2015a). In contrast to DDA and by means of systematically parallelising the fragmentation, DIA avoids the selection of individual peptide ions during LC-MS analysis, therefore providing several advantages for characterising complex protein digests.

In a single injection or LC-MS analysis, DIA generates a comprehensive and permanent digital record of the sample (Liu et al. 2013). Because of the systematic sampling process, there is no need to reinject the sample for LC-MS analysis, as opposed to DDA and SRM. Acquired once and mineable forever, DIA spectra can be used to test for new hypothesis or reprocessed when a better-quality library is available (e.g. new genome available) or an updated or new processing software tool is released.

At the same time, in DIA, convoluted or multiplexed MS/MS spectra are generated without explicit association between each single precursor and its corresponding fragments. As a result, DDA search engines are not appropriate for processing DIA spectra, and several informatics tools have been developed recently to effectively process these complex datasets for identification: DIA-Umpire (Tsou et al. 2015, 2016), Group-DIA (Li et al. 2015) and MSPLIT-DIA (Wang et al. 2015a).

Here we focus on the targeted data extraction strategy related to the SWATH methodology (Gillet et al. 2012), where a list of peptide transitions (also called assay) built from previous DDA/SRM libraries is required. These libraries can be either collected from public repositories such as PeptideAtlas (www.peptideatlas.org), SRMAtlas (www.srmatlas.org) and SWATHAtlas (www.swathatlas.org) or

generated by analysing the studied sample also in DDA mode to generate reference libraries (for a detailed protocol to generate high-quality reference libraries, see Schubert et al. (2015a)). A multiplexed variant of the SWATH methodology, termed MSX (Egertson et al. 2013, 2015), is also implemented within the Skyline software, supporting both acquisition method design and data processing for quantification.

As for MS1 quantification, Skyline extracts the ion signals of each peptide, but in this case fragment ion signals are extracted from the MS2 raw DIA files. Quantification is therefore performed using the elution profile of the peptide fragments, like for SRM, with the area under the peak or ion chromatogram as the abundance measure. Fragment ion abundances are subsequently aggregated into the corresponding peptides and proteins. A statistical measure of detection confidence is also computed for each peptide, using a similar version of the mProphet algorithm (Reiter et al. 2011) implemented within Skyline.

Other software tools for targeted data extraction are also available: Spectronaut (Bernhardt et al. 2012) (Biognosys proprietary software with free license for academics) and OpenSWATH (Röst et al. 2014) (open-source standalone tool or integrated module into the proteomics software OpenMS) (Sturm et al. 2008; Röst et al. 2016), including tools without GUI such as DIANA (Teleman et al. 2014) and SWATHProphet (Keller et al. 2015).

Recently, the performance of several of these tools was compared (Navarro et al. 2016). The authors observed similar reliable performances after software and parameter optimisation and concluded that targeted data extraction is a valid alternative to isotope-labelling-based methods.

Another consideration related to DIA data processing is the fact that the concurrent fragmentation of peptides has the drawback of increasing the likelihood of interference due to the overlap of fragment ions from different precursors. Several computational strategies can tackle this issue to further expand the benefits of DIA (Zhang et al. 2015; Bilbao et al. 2015b, 2016).

3.3 Statistical Analysis of Quantitative Results

Based on the quantification data, the next step is to determine candidate proteins showing significant differences across several sample types or conditions. The R statistical package MSstats (Choi et al. 2014) can be used as a standalone or as one of the external tools available within Skyline. MSstats can be used to interpret SRM, DDA and DIA quantification results.

The external set of QuaSAR tools (Abbatiello et al. 2010; Mani et al. 2012) automate and assist quantification of stable isotope dilution experiments. QuaSAR produces tabulated results for every peptide for essential statistics such as coefficient of variation, regression slope and intercept (with confidence intervals) and limits of detection and quantification as well as figures summarising their distribution and variation.

3.4 Automated PTM Detection

It is now well established that posttranslational modifications (PTMs) act in isolation or in combination with proteins for modulation and regulation purposes. In recent years, this field of investigation has led to intense bioinformatics development.

3.4.1 Main PTM Bioinformatics Resources

Two main databases are considered as references for storing PTM-related information. UniMod (http://www.unimod.org) is a comprehensive list of protein modifications for mass spectrometry applications. dbPTM (http://dbPTM.mbc.nctu.edu.tw) describes substrate specificity of PTM sites and provides functional annotation of PTM-related substrates and known interacting proteins. Neither specifically distinguishes between species as both aim at increasing numbers of their respective statistical tables. However, with this concern for broad coverage, dbPTM maintains a comprehensive list of databases and prediction software dedicated to individual or groups of PTMs. To complement the summary information associated with each resource, useful and more detailed comments can be found, for instance, in Kamath et al. (2011).

Over the past decades, the accumulation of sequence data led to implementation of PTM site prediction software based on amino acid patterns in aligned sequences. Each method was usually designed to identify individual PTMs. Many of these methods were developed with web interfaces and are hosted on the ExPASy (www.expasy.org/proteomics/post-translational_modification) and the CBS (www.cbs.dtu.dk/databases/PTMpredictions) servers whose creators pioneered in this field. More recently, the accumulation of mass spectrometry data to support PTM detection contributed to refining the reliability of prediction based on more comprehensive experimental data. This is, for instance, the case of phosphorylation sites through the use of resources such as PhosphoSite (http://www.phosphosite.org) that collects published mass spectrometry data for site annotation. In fact, understanding PTM occurrence goes along with studying the corresponding modifying enzyme(s). In many cases, these enzymes are not known, or their target is not precisely defined. For phosphorylation, KinBase (http://kinase.com/kinbase/) is the kinome reference, and the combined use of MS-validated sites and kinase specificity helps in refining site prediction as further explained in Sect. 3.4.2.

Although not fully considered as a PTM, protein cleavage should nonetheless be part of the modification landscape, and proteases have long been collected and classified in the MEROPS database (http://merops.sanger.ac.uk). The connection between proteolysis and PTMs is brought out in TOPFind, the N-/C-terminal modification database (http://clipserve.clip.ubc.ca/topfind). With a strong focus on human and mouse data, TOPFind also attempts to merge protein cleavage

information with protein-protein interactions with the PathFinder tool (Fortelny et al. 2014) through mapping and modelling a protease interaction network.

Nonetheless, at present, besides published articles, UniProt remains the main source compiling information on alternative effectors of PTMs. In fact, examples of protein-protein interaction networks integrating PTM knowledge (occurrence + specific effector) are rare. They tend to be devoted to mapping data collected in eukaryotes such the yeast methylome (Erce et al. 2012) or the phospho-tyrosine interaction network in human (Grossmann et al. 2015).

3.4.2 Predicting PTM and Their Associated Enzymes

The association between a phosphorylated site and the kinase that actually performed the attachment on a serine, threonine or tyrosine residue is far from being obvious to predict, despite the clear need for getting a fuller picture of phosphorylation. So far, the most known tool that combines several sources of data to suggest site-enzyme associations for phosphorylation is networKIN (http://networkin.info). The method first uses a predictor to label a given phosphosite sequence with a kinase or kinase family. This predictor is trained with experimental data to ascertain the relationship between a site and a kinase. For instance, Scansite (http://scansite3.mit.edu) mostly relies on peptide library screening, phage display and mass spectrometry experiments to get enough examples of labelled sites and identify characteristic sequence patterns for a site in definite association with a known kinase. Nonetheless a high level of ambiguity persists and as a second step in order to narrow down the options, networKIN includes contextual information by extracting knowledge of protein-protein interactions centred on the kinases of interest from a database of interactions. By calculating the proximity of the substrate to all kinases in a network of functional relationships, networKIN infers the most likely candidate kinase for each site.

The knowledge of phosphorylation is by far more advanced than that of other common PTMs such as glycosylation. Despite the possible mapping between a glycan structure and the set of enzymes that are required for its synthesis, the characterisation of intact glycopeptides remains a definite challenge. As it is, most glycan structures have been solved after being cleaved off their natural support, while protein glycosylation sites are identified after removing the attached glycans. In the end, key information on the glycoconjugate is lost. The correlation between glycan structures and glycoproteins can be restored manually through literature searches that are both labour- and time-consuming. This is, for example, the purpose of UniCarbKB (www.unicarbkb.org). In this context, the design of prediction tools linking a glycosite with the appropriate glycosyltransferases may happen in a not too distant future. Information on these enzymes has accumulated in the CAZy database (www.cazy.org) over the past two decades.

3.4.3 PTM Discovery

Mass spectrometry is the method of choice for detecting PTMs. Since the early days of software development for analysing mass spectra, the concern for identifying possible mass shifts corresponding to the addition or removal of chemical groups of known masses has been shared by bioinformaticians. Then it appeared that, conversely, the occurrence of regular and identical mass shifts of unknown origin in MS2 data could be a source of new knowledge, and a range of tools was then developed to perform the so-called open modification search, that is, PTM search with no a priori (Ahrné et al. 2010). This approach was scaled to process high-throughput proteome data with the prospect of discovering unexpected modifications (Na et al. 2011; Horlacher et al. 2015). However, data interpretation remains a challenge, and findings require experimental validation. Nonetheless, large-scale processing supports scientists in investigating and discovering new leads.

It is worth noting that the top-down proteomic approach is very promising for generating mass information on PTMs (Smith et al. 2013). The ProSight software that is commonly run to analyse this particular type of data is adapted to the identification of a broad range of PTMs (Fellers et al. 2015).

Recently, the potential of label-free by PRM was demonstrated for targeted phosphoproteome analysis (Lawrence et al. 2016). The authors also created a web-based assay development application that queries the database for optimal peptide selection and retention time scheduling (phosphopedia.gs.washington.edu).

3.4.4 PTM Combination

The next challenge is to identify the constraints that rule PTM cooperative and/or antagonist effects as pointed in Venne et al. (2014). Indeed, PTMs can be mutually exclusive such as the phosphorylation and the O-GlcNAcylation of serine and threonine of signalling proteins presented very early on as the "yin-yang hypothesis" (Hart et al. 1995), but they can also be cooperative as in the well-studied case of histones (Schwammle et al. 2014). Data has accumulated on a few other proteins such as tubulins (Verhey and Gaertig 2007), the FoxO regulator (Calnan and Brunet 2008) or chaperones (Cloutier and Coulombe 2013). However, a critical mass of experimental data is still missing to design appropriate bioinformatics tools supporting the discovery of PTM co-occurrence rules that could potentially explain corresponding effects on protein function. In the meantime, simple co-occurrence is collected, for instance, in the PTMCode database (Minguez et al. 2014) that shows potential trends. This field is likely to blossom in the years to come.

3.5 Interactomics and Data Interpretation

A list of identified proteins is not sufficient for characterising a sample. The challenge is to understand why some proteins are co- or differentially expressed and what are the underlying processes explaining cooperative or concurrent activity.

Detailed information about how MS-based proteomics has been applied to network biology, to detect and quantify perturbation-induced network changes and to correlate network dynamics with cellular phenotypes can be found in Bensimon et al. (2012).

The most common approach is to extract protein information from the gene ontology (GO) that assigns relevant terms from controlled vocabularies to specify the protein subcellular location, its function and its contribution to one or more processes (see www.geneontology.org). Then, when proteins share terms characterising a location or a function or a process, this type of similarity helps in shaping a hypothesis. Alternatively, the knowledge of protein-protein interactions stored in databases such as IntAct (http://www.ebi.ac.uk/intact/) or BioGrid (http://www.biogrid.org) can be used to build an interaction network. Note that readily available interaction networks can be visualised and queried in STRING (http://string-db.org). In all cases, integrated tools in the cited data resources provide support for protein data interpretation. One of the most popular open-source software for integration, visualisation and analysis of biological networks is Cytoscape (Shannon et al. 2003). This popularity has been driven by the ability of extending Cytoscape functionality through plugins (Saito et al. 2012), yielding a powerful and heterogeneous set of tools and enabling a broad community of scientists to contribute. This part is illustrated further in the following section with concrete examples.

4 Applications

There are roughly three types of applications where high-throughput proteomics methods are used and therefore bioinformatics is necessary. In most instances detailed below, resources tailored for the study of model organisms spanning mainly humans, mouse, drosophila and yeast show some limitation in supporting the interpretation of experimental results.

4.1 Animal Health

In much the same way as in medical studies, one of the goals of using proteomics is to identify reliable disease biomarkers for diagnosis or prognosis. For example,

serum protein profiles have been used to detect infectious disease in pigs (Koene et al. 2012). However, the prevalence of genomics approaches remains as illustrated in a recent comprehensive review on bioinformatics tools available to study parasites of veterinary significance (Cantacessi et al. 2012). Omics data integration is still a prospect at this point in time.

As highlighted in Sect. 3.5, interactomics provides the most attractive ground for the interpretation of proteomics data. The identification of protein complexes in a sample is a first step in rationalising and understanding protein co-expression. For example, cellular proteins of the host-forming complexes with specific proteins of the porcine reproductive and respiratory syndrome virus (PRRSV) essential in virus replication were identified by pull-down experiments in Dong et al. (2016). Binding partners of viral proteins were then systematically mapped on known pathways with a piece of software (https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/) to shortlist those that involve the mostly expressed partners. This study along with further experimental work led to identify and validate the role of the HSP70 chaperone as key to transcription and replication of PRRSV. More generally, the investigation of host-pathogen interactions is obviously suited to proteomics-based approaches, and the wealth of information stored in bioinformatics databases provides useful support. Proteomics is increasingly introduced in veterinary medicine (Ceciliani et al. 2014), and data accumulation will soon popularise the use of bioinformatics to allow for information-rich comparative studies.

Furthermore, Bundgaard et al. (2016) applied SRM to investigated levels of eight inflammatory acute phase proteins in interstitial fluid from wounds in horses. Selection of protein-specific peptides was performed using the equine PeptideAtlas website.

Packialakshmi et al. investigated proteomics differences in the plasma of healthy and femoral head necrosis-affected chickens using shotgun MS methods (Packialakshmi et al. 2016). MS/MS data was converted from proprietary format to mzXML files using the instrument vendor software and submitted to global proteome machine (GPM; http://www.thegpm.org) for identification with X! Tandem.

Proteins with at least one unique peptide and 5% FDR were considered true for protein identifications. The results were downloaded as *.xml files for Skyline software. After MS1 signal extraction, label-free quantitation was performed using MSstats as one of the external tools directly available in Skyline. Group comparison function was used for the label-free quantitation and to generate the volcano plot that shows the differentially expressed proteins.

The list of proteins was mapped to the corresponding ensemble gene IDs using Biomart and analysed for relative enrichment, clustering and GO annotations using DAVID (Huang et al. 2009a, b).

4.2 Adipose Tissue and Muscle Studies

The lack of farm animal dedicated resources in proteomics motivated the development of the ProteINSIDE database cited in Sect. 2.2.2. It was used in a study of bovine adipogenesis and myogenesis as well as the balance between these two processes (Kaspric et al. 2015). Previously analysed mass spectrometry data of adipose (Taga et al. 2012) and muscle (Chaze et al. 2009) foetal bovine tissues led to identify proteins, which were poorly annotated. Data and tool integration of ProteINSIDE supported the interpretation of protein lists. In particular protein-protein interactions, as collected from various sources as cited in Sect. 3.5, were used as the main piece of information for identifying clusters of functionally interconnected proteins. In the muscle they were associated with four processes, muscle development, cell proliferation, energetic complex and respiratory chain, and in the adipose tissue with seven, cell proliferation, proteasome complex, complexes I and III of the respiratory chain, redox activity and differentiation and metabolism of adipose tissue. The overlap between the two tissue types led to suggest possible crosstalk mechanisms.

Other authors have used bioinformatics sequence analysis tools to identify proteotypic peptides in an attempt to define biomarkers of meat authenticity following a targeted proteomics approach (Orduna et al. 2015). In the same vein, Stella and co-workers used SRM to quantify 12 potential protein markers of skeletal muscle and detect anabolic treatments with dexamethasone (Stella et al. 2016). The listed proteins were markers identified in a previous study applying a two-dimensional difference gel electrophoresis proteomics approach. A scheduled SRM method was developed using Skyline software to monitor 24 signature peptides from the 12 considered protein markers (two peptides per protein). For each peptide, 3 precursor-to-product ion transitions were targeted. Peptide quantification was achieved using a spike-in dedicated internal standard for each target. To this end, 13C/15N isotopically labelled peptides sharing the same sequence but with a defined mass shift were used. Peptide quantification was achieved using Skyline software integrating the area of the chromatographic peak of each peptide and the corresponding labelled internal standard.

Using R, protein abundances were graphically described using box plots, and potential differences of protein concentration values among different animal groups were explored performing one-way analysis of variance (ANOVA) on the two animal sets.

4.3 Milk Proteome and Glycoproteome

As mentioned in Sect. 2.2.1, the analysis of protein content in farm animal milk has long been the focus of veterinary and biological science. It is not surprising that several groups already undertook global studies such as a more detailed pathway

mapping of 106 human milk proteins (D'Alessandro et al. 2010). In this in silico study, pathway analysis software based on knowledge of metabolism is used to identify cell proliferation and differentiation pathways on top of the usual nutrition and immune functions known to characterise milk proteins. This provides evidence of tissue growth and organ development capacities of milk proteins based on collecting public data and using pathway analysis tools.

To account for quantitative aspects, mass spectrometry and bioinformatics tools were used in two recent studies (Tacoma et al. 2016; Zhang et al. 2016). While the former reference conventionally relies on gene ontology to track functional features of differentially expressed proteins in two dairy cow breeds, the latter brings the comparison of the human and bovine milk proteomes over lactation further. In this study, MS data was first differentially quantified using the MaxQuant software, and interactions between the most co-expressed proteins were then derived from STRING (see Sect. 3.5). Results reveal the interconnected roles of milk proteins in nutrition and protection to the neonate.

Finally, the importance of glycosylation in milk needs to be acknowledged, and quantitative studies are also undertaken in a systematic way. For example, Huang et al. (2016) used SRM to quantitate seven human milk proteins and their glycoforms.

5 Conclusion

The application of proteomics in veterinary studies has been moving from the initial qualitative description towards the quantification stage, where experiments to identify and quantify protein changes in different samples of particular tissues or fluids become more common (Ceciliani et al. 2014). Application of untargeted label-free quantification methods is rapidly increasing in the proteomics field in general, fuelled by the advantages of DIA methods and development of software tools and innovative algorithms, which already have been shown implementations to estimate absolute cellular protein concentrations (Schubert et al. 2015b). In this context, further improvements and developments of new computational strategies for quantification are expected.

Numerous studies have been published to date in domestic and farm animal proteomics; among several challenges and limitations, we highlight the lack of detailed information of MS-based proteome informatics tools in the context of farm animal reviews/resources. Only general notions are described but not specific tools and very few references.

The previously described difficulty to use MS proteomics software (Cappadona et al. 2012), associated with the lack of appropriate documentation or with a poor graphical user interface, is still an issue. Fortunately, this is changing with tools like Skyline. Skyline is an active project and continues expanding, for instance, it currently supports chromatogram extraction from MS1 and MS2 spectra with the additional ion mobility separation recently available in commercial instruments

(Baker et al. 2015). An increasing number of available tools within open-source and collaborative projects can be expected, and we encourage development of more robust MS tools that can be used by researchers nonspecialised in mass spectrometry or informatics.

References

Abbatiello SE, Mani D, Keshishian H, Carr SA (2010) Automated detection of inaccurate and imprecise transitions in peptide quantification by multiple reaction monitoring mass spectrometry. Clin Chem 56:291–305

Aebersold R (2003) Mass spectrometry-based proteomics. Nature 422:198–207

Aebersold R (2009) A stress test for mass spectrometry-based proteomics. Nat Methods 6:411–412

Ahrné E, Müller M, Lisacek F (2010) Unrestricted identification of modified proteins using MS/MS. Proteomics 10:671–686

Ahrné E, Molzahn L, Glatter T, Schmidt A (2013) Critical assessment of proteome-wide label-free absolute abundance estimation strategies. Proteomics 13:2567–2578

Alghanem B, Nikitin F, Stricker T, Duchoslav E, Luban J, Strambio-De-Castillia C, Muller M, Lisacek F, Varesio E, Hopfgartner G (2017) Optimization by infusion of multiple reaction monitoring transitions for sensitive peptides LC-MS quantification. Rapid Commun Mass Spectrom 31(9):753–761

Baker ES, Burnum-Johnson KE, Ibrahim YM, Orton DJ, Monroe ME, Kelly RT, Moore RJ, Zhang X, Théberge R, Costello CE et al (2015) Enhancing bottom-up and top-down proteomic measurements with ion mobility separations. Proteomics 15:2766–2776

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 57:289–300

Bensimon A, Heck AJ, Aebersold R (2012) Mass spectrometry-based proteomics and network biology. Annu Rev Biochem 81:379–405

Bernhardt OM, Selevsek N, Gillet LC, Rinner O, Picotti P, Aebersold R, Reiter L (2012) Spectronaut: a fast and efficient algorithm for MRM-like processing of data independent acquisition (SWATH-MS) data. In: 60th ASMS Conference on mass spectrometry and allied topics. Vancouver, Canada, pp 20–24

Bilbao A, Varesio E, Luban J, Strambio-De-Castillia C, Hopfgartner G, Müller M, Lisacek F (2015a) Processing strategies and software solutions for data-independent acquisition in mass spectrometry. Proteomics 15:964–980

Bilbao A, Zhang Y, Varesio E, Luban J, Strambio-De-Castilla C, Lisacek F, Hopfgartner G (2015b) Ranking fragment ions based on outlier detection for improved label-free quantification in data-independent acquisition LC-MS/MS. J Proteome Res 14:4581–4593

Bilbao A, Lisacek F, Hopfgartner G (2016) Dedicated software enhancing data-independent acquisition methods in mass spectrometry. CHIMIA Int J Chem 70:293–293

Bislev SL, Deutsch EW, Sun Z, Farrah T, Aebersold R, Moritz RL, Bendixen E, Codrea MC (2012) A Bovine PeptideAtlas of milk and mammary gland proteomes. Proteomics 12:2895–2899

Broudy D, Killeen T, Choi M, Shulman N, Mani DR, Abbatiello SE, Mani D, Ahmad R, Sahu AK, Schilling B et al (2014) A framework for installable external tools in Skyline. Bioinformatics 30:2521–2523

Bundgaard L, Jacobsen S, Sørensen MA, Sun Z, Deutsch EW, Moritz RL, Bendixen E (2014) The Equine PeptideAtlas: a resource for developing proteomics-based veterinary research. Proteomics 14:763–773

Bundgaard L, Bendixen E, Sørensen MA, Harman VM, Beynon RJ, Petersen LJ, Jacobsen S (2016) A selected reaction monitoring-based analysis of acute phase proteins in interstitial

fluids from experimental equine wounds healing by secondary intention. Wound Repair Regen 24:525–532

Calnan DR, Brunet A (2008) The FoxO code. Oncogene 27:2276–2288

Cantacessi C, Campbell BE, Jex AR, Young ND, Hall RS, Ranganathan S, Gasser RB (2012) Bioinformatics meets parasitology. Parasite Immunol 34:265–275

Cappadona S, Baker PR, Cutillas PR, Heck AJ, van Breukelen B (2012) Current challenges in software solutions for mass spectrometry-based quantitative proteomics. Amino Acids 43:1087–1108

Carr SA, Abbatiello SE, Ackermann BL, Borchers C, Domon B, Deutsch EW, Grant RP, Hoofnagle AN, Hüttenhain R, Koomen JM et al (2014) Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based assay development using a fit-for-purpose approach. Mol Cell Proteomics 13:907–917

Ceciliani F, Eckersall D, Burchmore R, Lecchi C (2014) Proteomics in veterinary medicine. Vet Pathol 51:351–362

Chapman JD, Goodlett DR, Masselon CD (2014) Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. Mass Spectrom Rev 33:452–470

Chaze T, Meunier B, Chambon C, Jurie C, Picard B (2009) Proteome dynamics during contractile and metabolic differentiation of bovine foetal muscle. Animal 3:980–1000

Choi H, Nesvizhskii AI (2007) Semisupervised model-based validation of peptide identifications in mass spectrometry-based proteomics. J Proteome Res 7:254–265

Choi M, Chang C-Y, Clough T, Broudy D, Killeen T, MacLean B, Vitek O (2014) MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. Bioinformatics 30:2524–2526

Cloutier P, Coulombe B (2013) Regulation of molecular chaperones through post-translational modifications: decrypting the chaperone code. Biochim Biophys Acta Gene Regul Mech 1829:443–454

Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res 10:1794–1805

Craig R, Beavis RC (2004) TANDEM: matching proteins with tandem mass spectra. Bioinformatics 20:1466–1467

D'Alessandro A, Scaloni A, Zolla L (2010) Human milk proteins: an interactomics and updated functional overview. J Proteome Res 9:3339–3373

Degan MG, Ryadinskiy L, Fujimoto GM, Wilkins CS, Lichti CF, Payne SH (2016) A skyline plugin for pathway-centric data browsing. J Am Soc Mass Spectrom 27:1752–1757

Deutsch EW (2012) File formats commonly used in mass spectrometry proteomics. Mol Cell Proteomics 11:1612–1621

Deutsch EW, Chambers M, Neumann S, Levander F, Binz P-A, Shofstahl J, Campbell DS, Mendoza L, Ovelleiro D, Helsens K et al (2012) TraML—a standard format for exchange of selected reaction monitoring transition lists. Mol Cell Proteomics 11:111–15040

Deutsch EW, Csordas A, Sun Z, Jarnuczak A, Perez-Riverol Y, Ternent T, Campbell DS, Bernal-Llinares M, Okuda S, Kawano S et al (2017) The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. Nucleic Acids Res 45: D1100–D1106

Dong S, Liu L, Wu W, Armstrong SD, Xia D, Nan H, Hiscox JA, Chen H (2016) Determination of the interactome of non-structural protein12 from highly pathogenic porcine reproductive and respiratory syndrome virus with host cellular proteins using high throughput proteomics and identification of HSP70 as a cellular factor for virus replication. J Proteomics 146:58–69

Egertson JD, Kuehn A, Merrihew GE, Bateman NW, MacLean BX, Ting YS, Canterbury JD, Marsh DM, Kellmann M, Zabrouskov V et al (2013) Multiplexed MS/MS for improved data-independent acquisition. Nat Methods 10:744–746

Egertson JD, MacLean B, Johnson R, Xuan Y, MacCoss MJ (2015) Multiplexed peptide analysis using data-independent acquisition and Skyline. Nat Protoc 10:887–903

Elias JE, Gygi SP (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods 4:207–214

Eng JK, McCormack AL, Yates JR (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 5:976–989

Erce MA, Pang CNI, Hart-Smith G, Wilkins MR (2012) The methylproteome and the intracellular methylation network. Proteomics 12:564–586

Farrah T, Deutsch EW, Hoopmann MR, Hallows JL, Sun Z, Huang C-Y, Moritz RL (2013) The state of the human proteome in 2012 as viewed through PeptideAtlas. J Proteome Res 12:162–171

Fellers RT, Greer JB, Early BP, Yu X, LeDuc RD, Kelleher NL, Thomas PM (2015) ProSight Lite: graphical software to analyze top-down mass spectrometry data. Proteomics 15:1235–1238

Fortelny N, Cox JH, Kappelhoff R, Starr AE, Lange PF, Pavlidis P, Overall CM (2014) Network analyses reveal pervasive functional regulation between proteases in the human protease web. PLoS Biol 12:e1001869

Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, Bonner R, Aebersold R (2012) Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics. https://doi.org/10.1074/mcp.O111.016717

Griss J (2016) Spectral library searching in proteomics. Proteomics 16:729–740

Grossmann A, Benlasfer N, Birth P, Hegele A, Wachsmuth F, Apelt L, Stelzl U (2015) Phospho-tyrosine dependent protein-protein interaction network. Mol Syst Biol 11:794–794

Hart GW, Greis KD, Dong L-YD, Blomberg MA, Chou T-Y, Jiang M-S, Roquemore EP, Snow DM, Kreppel LK, Cole RN, Comer FI, Arnold CS, Hayes BK (1995) O-linked N-Acetylglucosamine: the "yin-yang" of ser/Thr Phosphorylation? In: Alavi A, Axford JS (eds) Glycoimmunology, Advances in experimental medicine and biology. Springer Nature, Boston, MA, pp 115–123

Hesselager MO, Codrea MC, Sun Z, Deutsch EW, Bennike TB, Stensballe A, Bundgaard L, Moritz RL, Bendixen E (2016) The Pig PeptideAtlas: a resource for systems biology in animal production and biomedicine. Proteomics 16:634–644

Horlacher O, Lisacek F, Müller M (2015) Mining large scale tandem mass spectrometry data for protein modifications using spectral libraries. J Proteome Res 15:721–731

Hu Z-L, Park CA, Reecy JM (2015) Developmental progress and current status of the animal QTLdb. Nucleic Acids Res 44:D827–D833

Huang DW, Sherman BT, Lempicki RA (2009a) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57

Huang DW, Sherman BT, Lempicki RA (2009b) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13

Huang J, Kailemia MJ, Goonatilleke E, Parker EA, Hong Q, Sabia R, Smilowitz JT, German JB, Lebrilla CB (2016) Quantitation of human milk proteins and their glycoforms using multiple reaction monitoring (MRM). Anal Bioanal Chem 409:589–606

Jones AR, Eisenacher M, Mayer G, Kohlbacher O, Siepen J, Hubbard SJ, Selley JN, Searle BC, Shofstahl J, Seymour SL et al (2012) The mzIdentML data standard for mass spectrometry-based proteomics results. Mol Cell Proteomics 11:111–14381

Käll L, Canterbury JD, Weston J, Noble WS, MacCoss MJ (2007) Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nat Methods 4:923–925

Kamath KS, Vasavada MS, Srivastava S (2011) Proteomic databases and tools to decipher post-translational modifications. J Proteomics 75:127–144

Kaspric N, Picard B, Reichstadt M, Tournayre J, Bonnet M (2015) ProteINSIDE to easily investigate proteomics data from ruminants: application to mine proteome of adipose and muscle tissues in bovine foetuses. PLoS One 10:e0128086

Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 74:5383–5392

Keller A, Eng J, Zhang N, Li X, Aebersold R (2005) A uniform proteomics MS/MS analysis platform utilizing open XML file formats. Mol Syst Biol 1(2005):0017

Keller A, Bader SL, Shteynberg D, Hood L, Moritz RL (2015) Automated validation of results and removal of fragment ion interferences in targeted analysis of data independent acquisition MS using SWATHProphet. Mol Cell Proteomics. https://doi.org/10.1074/mcp.O114.044917

Kim S, Pevzner PA (2014) MS-GF + makes progress towards a universal database search tool for proteomics. Nat Commun. https://doi.org/10.1038/ncomms6277

Koene MG, Mulder HA, Stockhofe-Zurwieden N, Kruijt L, Smits MA (2012) Serum protein profiles as potential biomarkers for infectious disease status in pigs. BMC Vet Res 8:32

Lange V, Picotti P, Domon B, Aebersold R (2008) Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol 4:1–14

Lawrence RT, Searle BC, Llovet A, Villén J (2016) Plug-and-play analysis of the human phosphoproteome by targeted high-resolution mass spectrometry. Nat Methods 13(5):431–434

Li Y, Zhong C-Q, Xu X, Cai S, Wu X, Zhang Y, Chen J, Shi J, Lin S, Han J (2015) Group-DIA: analyzing multiple data-independent acquisition mass spectrometry data files. Nat Methods. https://doi.org/10.1038/NMETH.3593

Liu H, Sadygov RG, Yates JR III (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76:4193–4201

Liu Y, Hüttenhain R, Collins B, Aebersold R (2013) Mass spectrometric protein maps for biomarker discovery and clinical research. Expert Rev Mol Diagn 13:811–825

MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 26:966–968

Mani D, Abbatiello SE, Carr SA (2012) Statistical characterization of multiple-reaction monitoring mass spectrometry (MRM-MS) assays for quantitative proteomics. BMC Bioinformatics 13:S9

Martens L (2010) Proteomics databases and repositories. In: Methods in molecular biology. Springer Nature, Boston, MA, pp 213–227

Martens L, Vizcaıno JA (2017) A golden age for working with public proteomics data. Trends Biochem Sci 42(5):333–341

McCarthy FM, Gresham CR, Buza TJ, Chouvarine P, Pillai LR, Kumar R, Ozkan S, Wang H, Manda P, Arick T, Bridges SM, Burgess SC (2010) AgBase: supporting functional modeling in agricultural organisms. Nucleic Acids Res 39:D497–D506

McCord J, Sun Z, Deutsch EW, Moritz RL, Muddiman DC (2017) The PeptideAtlas of the domestic laying Hen. J Proteome Res 16:1352–1363

Minguez P, Letunic I, Parca L, Garcia-Alonso L, Dopazo J, Huerta-Cepas J, Bork P (2014) PTMcode v2: a resource for functional associations of post-translational modifications within and between proteins. Nucleic Acids Res 43:D494–D502

Na S, Bandeira N, Paek E (2011) Fast multi-blind modification search through tandem mass spectrometry. Mol Cell Proteomics 11:M111.010199

Navarro P, Kuharev J, Gillet LC, Bernhardt OM, MacLean B, Röst HL, Tate SA, Tsou C-C, Reiter L, Distler U et al (2016) A multicenter study benchmarks software tools for label-free proteome quantification. Nat Biotechnol 34(11):1130–1136

Nesvizhskii AI (2010) A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. J Proteomics 73:2092–2123

Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 75:4646–4658

Orduna AR, Husby E, Yang CT, Ghosh D, Beaudry F (2015) Assessment of meat authenticity using bioinformatics, targeted peptide biomarkers and high-resolution mass spectrometry. Food Addit Contam Part A 32:1709–1717

Packialakshmi B, Liyanage R, Jackson O, Lay J, Okimoto R, Rath NC (2016) Proteomic changes in the plasma of broiler chickens with femoral head necrosis. Biomark Insights 11:55–62

Pedrioli PG, Eng JK, Hubley R, Vogelzang M, Deutsch EW, Raught B, Pratt B, Nilsson E, Angeletti RH, Apweiler R et al (2004) A common open representation of mass spectrometry data and its application to proteomics research. Nat Biotechnol 22:1459–1466

Perez-Riverol Y, Alpi E, Wang R, Hermjakob H, Vizcauno JA (2015) Making proteomics data accessible and reusable: current state of proteomics databases and repositories. Proteomics 15:930–950

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567

Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ (2012) Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. Mol Cell Proteomics 11:1475–1478

Picotti P, Aebersold R (2012) Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. Nat Methods 9:555–566

Reiter L, Claassen M, Schrimpf SP, Jovanovic M, Schmidt A, Buhmann JM, Hengartner MO, Aebersold R (2009) Protein identification false discovery rates for very large proteomics data sets generated by tandem mass spectrometry. Mol Cell Proteomics 8:2405–2417

Reiter L, Rinner O, Picotti P, Hüttenhain R, Beck M, Brusniak M-Y, Hengartner MO, Aebersold R (2011) mProphet: automated data processing and statistical validation for large-scale SRM experiments. Nat Methods 8:430–435

Reker D, Malmström L (2012) Bioinformatic challenges in targeted proteomics. J Proteome Res 11:4393–4402

Röst H, Malmström L, Aebersold R (2012) A computational tool to detect and avoid redundancy in selected reaction monitoring. Mol Cell Proteomics 11:540–549

Röst HL, Rosenberger G, Navarro P, Gillet L, Miladinovi SM, Schubert OT, Wolski W, Collins BC, Malmström J, Malmström L et al (2014) OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. Nat Biotechnol 32:219–223

Röst HL, Sachsenberg T, Aiche S, Bielow C, Weisser H, Aicheler F, Andreotti S, Ehrlich H-C, Gutenbrunner P, Kenar E et al (2016) OpenMS: a flexible open-source software platform for mass spectrometry data analysis. Nat Methods 13:741–748

Saito R, Smoot ME, Ono K, Ruscheinski J, Wang P-L, Lotia S, Pico AR, Bader GD, Ideker T (2012) A travel guide to Cytoscape plugins. Nat methods 9:1069–1076

Schilling B, Rardin MJ, MacLean BX, Zawadzka AM, Frewen BE, Cusack MP, Sorensen DJ, Bereman MS, Jing E, Wu CC et al (2012) Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in Skyline application to protein acetylation and phosphorylation. Mol Cell Proteomics 11:202–214

Schubert OT, Gillet LC, Collins BC, Navarro P, Rosenberger G, Wolski WE, Lam H, Amodei D, Mallick P, MacLean B et al (2015a) Building high-quality assay libraries for targeted analysis of SWATH MS data. Nat Protoc 10:426–441

Schubert OT, Ludwig C, Kogadeeva M, Zimmermann M, Rosenberger G, Gengenbacher M, Gillet LC, Collins BC, Röst HL, Kaufmann SHE, Sauer U, Aebersold R (2015b) Absolute proteome composition and dynamics during dormancy and resuscitation of mycobacterium tuberculosis. Cell Host Microbe 18:96–108

Schwammle V, Aspalter C-M, Sidoli S, Jensen ON (2014) Large scale analysis of co-existing post-translational modifications in histone tails reveals global fine structure of cross-talk. Mol Cell Proteomics 13:1855–1865

Searle BC, Egertson JD, Bollinger JG, Stergachis AB, MacCoss MJ (2015) Using data independent acquisition (DIA) to model high-responding peptides for targeted proteomics experiments. Mol Cell Proteomics 14:2331–2340

Serang O, Käll L (2015) Solution to statistical challenges in proteomics is more statistics, not less. J Proteome Res. https://doi.org/10.1021/acs.jproteome.5b00568

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504

Shteynberg D, Deutsch EW, Lam H, Eng JK, Sun Z, Tasman N, Mendoza L, Moritz RL, Aebersold R, Nesvizhskii AI (2011) iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. Mol Cell Proteomics 10:111–7690

Smith LM, Kelleher NL et al (2013) Proteoform: a single term describing protein complexity. Nat Methods 10:186–187

Soares R, Franco C, Pires E, Ventosa M, Palhinhas R, Koci K, de Almeida AM, Coelho AV (2012) Mass spectrometry and animal science: protein identification strategies and particularities of farm animal species. J Proteomics 75:4190–4206

Stahl-Zeng J, Lange V, Ossola R, Eckhardt K, Krek W, Aebersold R, Domon B (2007) High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. Mol Cell Proteomics 6:1809–1817

Stella R, Barrucci F, Angeletti R, James P, Montesissa C, Biancotto G (2016) Targeted proteomics for the indirect detection of dexamethasone treatment in bovines. Anal Bioanal Chem 408:8343–8353

Sturm M, Bertsch A, Gröpl C, Hildebrandt A, Hussong R, Lange E, Pfeifer N, Schulz-Trieglaff O, Zerck A, Reinert K et al (2008) OpenMS–an open-source software framework for mass spectrometry. BMC Bioinform 9:163

Tacoma R, Fields J, Ebenstein DB, Lam Y-W, Greenwood SL (2016) Characterization of the bovine milk proteome in early-lactation Holstein and Jersey breeds of dairy cows. J Proteomics 130:200–210

Taga H, Chilliard Y, Meunier B, Chambon C, Picard B, Zingaretti MC, Cinti S, Bonnet M (2012) Cellular and molecular large-scale features of fetal adipose tissue: Is bovine perirenal adipose tissue Brown1685. J Cell Physiol 227:1688–1700

Teleman J, Röst H, Rosenberger G, Schmitt U, Malmström L, Malmström J, Levander F (2014) DIANA-algorithmic improvements for analysis of data-independent acquisition MS data. Bioinformatics. https://doi.org/10.1093/bioinformatics/btu686

The M, MacCoss MJ, Noble WS, Käll L et al (2016) Fast and accurate protein false discovery rates on large-scale proteomics data sets with percolator 3.0. J Am Soc Mass Spectrom 27:1719–1727

Tsou C-C, Avtonomov D, Larsen B, Tucholska M, Choi H, Gingras A-C, Nesvizhskii AI (2015) DIA-Umpire: comprehensive computational framework for data independent acquisition proteomics. Nat Methods 12:258–264

Tsou C-C, Tsai C-F, Teo G, Chen Y-J, Nesvizhskii AI (2016) Untargeted, spectral library-free analysis of data independent acquisition proteomics data generated using Orbitrap mass spectrometers.16(15–16):2257–2271

Venne AS, Kollipara L, Zahedi RP (2014) The next level of complexity: crosstalk of posttranslational modifications. Proteomics 14:513–524

Verhey KJ, Gaertig J (2007) The Tubulin Code. Cell Cycle 6:2152–2160

Vizcaıno JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rıos D, Dianes JA, Sun Z, Farrah T, Bandeira N, Binz P-A, Xenarios I, Eisenacher M, Mayer G, Gatto L, Campos A, Chalkley RJ, Kraus H-J, Albar JP, Martinez-Bartolomé S, Apweiler R, Omenn GS, Martens L, Jones AR, Hermjakob H (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat Biotechnol 32:223–226

Vizcaıno JA, Csordas A, del Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu Q-W, Wang R, Hermjakob H (2015) 2016 update of the PRIDE database and its related tools. Nucleic Acids Res 44:D447–D456

Walzer M, Qi D, Mayer G, Uszkoreit J, Eisenacher M, Sachsenberg T, Gonzalez-Galarza FF, Fan J, Bessant C, Deutsch EW et al (2013) The mzQuantML data standard for mass spectrometry–based quantitative studies in proteomics. Mol Cell Proteomics 12:2332–2340

Wang M, Herrmann CJ, Simonovic M, Szklarczyk D, von Mering C (2015a) Version 4.0 of PaxDb: protein abundance data, integrated across model organisms, tissues, and cell-lines. Proteomics 15:3163–3168

Wang J, Tucholska M, Knight JD, Lambert J-P, Tate S, Larsen B, Gingras A-C, Bandeira N (2015b) MSPLIT-DIA: sensitive peptide identification for data-independent acquisition. Nat Methods. https://doi.org/10.1038/nmeth.3655

Yates JR III (2015) Pivotal role of computers and software in mass spectrometry–SEQUEST and 20 years of tandem MS database searching. J Am Soc Mass Spectrom 26:1804–1813

Zhang Y, Bilbao A, Bruderer T, Luban J, Strambio-De-Castillia C, Lisacek F, Hopfgartner G, Varesio E (2015) The use of variable Q1 isolation Windows improves selectivity in LC–SWATH–MS acquisition. J Proteome Res 14:4359–4371

Zhang L, van Dijk ADJ, Hettinga K (2016) An interactomics overview of the human and bovine milk proteome over lactation. Proteome Sci. https://doi.org/10.1186/s12953-016-0110-0

Peptidomics on Farm Animal Research

Silvina Fadda, Enrique Sentandreu, and Miguel Angel Sentandreu

Abstract Although peptidomics is a discipline complementary to proteomics, since nowadays both mainly rely on analytical strategies based on mass spectrometry, there are fundamental differences. In this chapter, we discuss these differences along with the application of these technologies for the study of the different stages of meat production, from storage to processing to unravel mechanisms that will allow reaching high-quality and safer meat products. The use of peptidomics and the related high-throughput technologies, now relying on mass spectrometry but once also on N-terminal sequencing, is discussed. Clear examples are provided dealing with relevant studies on meat proteolysis and peptide generation occurring during ageing, as well as those produced during ripening of meat products by endogenous and microbial enzymes. Also the involvement of this phenomenon in the development of taste-active compounds is addressed. Finally, the application of novel omics technologies on bacterial identification in food for diagnosis and safety purposes is presented, putting emphasis on their potential advantages and future perspectives.

Keywords Peptidomics • Peptides • Sequence identification • Meat biochemistry • Meat microbiology • Mass spectrometry • Systems biology • Lactic acid bacteria • Meat contamination

S. Fadda
Centro de Referencia para Lactobacilos (CERELA), CONICET, San Miguel de Tucumán, Argentina
e-mail: sfadda@cerela.org.ar

E. Sentandreu • M.A. Sentandreu (✉)
Instituto de Agroquímica y Tecnología de Alimentos (CSIC), Paterna, Valencia, Spain
e-mail: infusoriosentandreu@gmail.com; ciesen@iata.csic.es

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_19

1 General Concepts

1.1 Peptidomics Versus Proteomics

The concept peptidomics is complementary to the more familiar term proteomics, which prompts us to the need to describe the two concepts. Proteomics aims the study of the expression of products of a genome (proteins) in a biological sample with the aim to have a global and integrated vision of the cellular processes at a particular time period. On the other hand, peptidomics will cover the comprehensive study of peptides that are present in a biological sample. Qualitative and quantitative features would be addressed in both cases. The term peptidomics already appeared in some scientific literature published in 2001 (Bergquist and Ekman 2001; Clynen et al. 2001; Minamino 2001; Schulz-Knappe et al. 2001; Verhaert et al. 2001). It is also important to better differentiate between these two terms to establish the difference between a peptide and a protein. It is assumed that peptides are short molecules formed from the linking of various amino acids being in the range from 2 to 50. However, this convention is rather arbitrary and flexible, not being possible in many cases to establish a clear difference, for example, between big polypeptides and small proteins. For that reason, it has also been accepted that a peptide has a less complex structure, being able to contain helices, sheets, other functional subunits or modifications of some particular residues but not reaching the more complex tertiary and quaternary structures as in the case of proteins. Peptides resulting from proteolytic cleavage of proteins will be influenced by these structural elements, which will drive the action of peptidases in the release of peptide products.

1.2 Technological Features in Peptidomics as Compared to Proteomics

Nowadays, proteomics and peptidomics are related disciplines since both rely on analytical strategies based on mass spectrometry (MS), and thus the separation between the two concepts is sometimes blurred. In proteomics, with the aim to facilitate the analysis by mass spectrometry, proteins are usually digested into peptides, which will be further analysed by single MS to obtain their peptide mass fingerprint (PMF) or by tandem mass spectrometry (MS/MS) to get the amino acid sequence of the generated peptides. This is known as the *bottom-up* approach, since the information obtained from the analysis of peptides will serve to reach up to the identity of the protein of origin. Even if the enzymatic digestion will contribute to increase sample complexity, the ease to analyse relatively short peptide sequences as compared to large polypeptide chains largely compensates for this. As a drawback of the digestion process, information about tertiary and quaternary structure of the protein will be lost. In contrast to this, peptidomics

would point at characterizing as much as possible the structure of the target peptides, getting information about the proteolytic processes responsible for their generation and the presence of post-translational modifications. This will be the case of a *top-down* approach, in which it will be necessary to manually evaluate the MS/MS spectra generated and verify that the assignment of peptides is accurate (Schrader et al. 2014). In food peptidomics, this can reveal valuable information about the proteolytic processes and final quality of food and food-derived products.

In *bottom-up* approaches, proteins are cleaved using proteolytic enzymes that have a defined substrate specificity. The most representative example is trypsin, which cleaves proteins and polypeptides having Lys or Arg at the left side (P1 position) of the peptide bond. Due to the frequency in the occurrence of Lys and Arg residues in proteins, this restricted specificity will allow for the generation of medium-size peptides, falling within the range from 7 to 25 amino acids (700–3000 Da), which are ideal for their precise mass determination using current proteomic technology based on the ionization of intact peptides by the so-called "soft" ionization techniques such as matrix-assisted laser desorption/ionization (MALDI) (Karas and Hillenkamp 1988) and electrospray ionization (ESI) (Fenn et al. 1989). Furthermore, fragmentation of peptides based on techniques such as CID (collision-induced dissociation), electron-transfer dissociation (ETD) or some others (Jones and Cooper 2011) will allow for the elucidation of the amino acid sequence in tandem mass spectrometry experiments. Protein digestion using enzymes with defined cleavage sites greatly simplifies the study of the fragmentation mechanisms, making it possible to predict the peptide fragment ions according to the type of MS instrument. This has allowed the development of several bioinformatic algorithms that, together with the expansion of protein databases by massive genome sequencing, makes it possible to analyse these peptide ion product spectra automatically by spectral matching, comparing the obtained MS/MS spectra to theoretical peptide spectra obtained by *in silico* digestion of proteins contained in previously defined protein databases. A probabilistic score derived from these comparisons will allow the identification of the most probable sequence matching our MS/MS data (Panchaud et al. 2012). Several robust and high-throughput MS/MS algorithms have been developed based on the principle of spectral matching (Hernandez et al. 2006).

Apart from making easier the study and prediction of fragment ions, tryptic peptides hold an amino group on their N-terminus, but also most of them have an additional amino group at the side chain of the C-terminal Arg or Lys. If in addition histidine residues occur in the peptide sequence, then charge states higher than 2 can also be observed. This will favour strong peptide signal intensities easily detected by the MS instruments and MS/MS fragmentation spectra of high quality, necessary for an accurate elucidation of the peptide sequence with no ambiguities (Schrader et al. 2014).

Peptides generated by protein digestion are not exclusive of *bottom-up* proteomic studies. As it will be explained below, some *top-down* peptidomic approaches are based in the identification and quantification of some of these sequences as individual entities, constituting the core of the analysis beyond the protein of origin.

Some examples are the use of peptide biomarkers in meat authentication studies (Marbaix et al. 2016; Sentandreu et al. 2010) and nutritional studies (Paolella et al. 2015; Wen et al. 2015) or the generation of bioactive peptides by the action of proteolytic enzymes on animal proteins (Fu et al. 2016; Lafarga et al. 2016; Mirdhayati et al. 2016). Other peptidomic approaches, however, deal with the characterization of the endogenous peptides naturally present in foodstuffs. These peptides, in many cases, will not follow the ideal model of those generated by digestion with known enzymes, having more difficulties for MS analyses. Special attention needs to be paid to the design of extraction protocols in order to minimize the peptide proteolysis after sample collection (Dallas et al. 2015). Because the enzymatic processes that are responsible for the formation of endogenous peptides are complex and most often imply the action of unknown peptidases and/or unspecific cleavages, size can change considerably, having sequences too large or too small for a proper identification and quantification using standard MS applications that have been developed for medium-size peptides in *bottom-up* proteomics. In the case of large peptides, elucidation of the amino acid sequence from MS/MS spectra of peptides having charge states +5 or more becomes more complicated than for peptides having charge states +2 or +3. Moreover, the non-specificity for cleavage sites makes identifications based on spectral matching considerably more complicated and time consuming in database searches. In the case of small peptides (lower than six amino acids), ambiguity in the elucidation of the sequence considerably increases, and database searches usually bring several identifications with similar scores and a confusing variety of parental proteins. For these cases, especially in very small peptides (two–four amino acids), other, less efficient bioinformatic tools not based on database search are required, together with manual verification of the MS/MS sequencing results. Another challenging aspect of small peptides is the difficulty to be retained in the reverse phase C_{18} columns that are massively used for tryptic peptides. Unless the small peptides contain some hydrophobic amino acids, they will not be retained by the reverse phase columns, and thus other separation alternatives such as ion-exchange, size-exclusion or hydrophilic interaction (HILIC) chromatography will become necessary.

1.3 Types of Peptides

According to their synthesis, three classes of peptides can be distinguished in biological samples (Zürbig and Mischak 2008):

(a) *Ribosomal peptides*, which will be those that are synthesized by translation of mRNA. These include peptides that function as hormones and signalling transmitters in superior organisms. In the case of lower organisms, these peptides can exert antibiotic action (Mahlapuu et al. 2016). In general, they are linear, though other nonlinear forms also exist (Craik et al. 2016; McIntosh et al. 2009).

(b) *Nonribosomal peptides*, which are those synthesized using a modular enzyme complex. These peptides are mainly ascribed to unicellular organisms, plants and fungi. They appear as a family of related compounds sharing a common core structure. In general they are cyclic, although linear structures also exist (Horst and Niedermeyer 2016).

(c) *Processed peptides*, which are the result of non-specific proteolysis in digestive/proteolytic processes: There are many events both inside and outside of cells by which proteins or longer peptides are cut into smaller fragments. Most of works dealing with peptidomics in food science and domestic animals revised in the present chapter are part of this group. Indeed, food scientists have mainly worked on the characterization of peptides generated from muscle tissue either by the action of digestive enzymes or by the proteolytic action occurring during the processing of meat and meat-derived products due to the action of endogenous muscle peptidases or peptidases coming from starter cultures. In the literature, apart from muscle tissue, there are also many important contributions dealing with peptides derived from milk proteins in both milk and milk-related products. Since quite recent and interesting reviews have been published, we will not focus on this part but encourage readers to also look at some of these contributions (Picariello et al. 2012; Roncada et al. 2012; Sanchez-Rivera et al. 2014). Other issues dealing with peptide research in domestic animals that have been also recently reviewed or reported and, thus out of the scope of this chapter, are researched into food authentication studies based on identification of species-specific peptide biomarkers (Ortea et al. 2016; Sentandreu and Sentandreu 2011, 2014) and characterization of bioactive peptides derived from farm animal sources (Fu et al. 2017; Lafarga and Hayes 2014; Stadnik and Keska 2015; Udenigwe and Howard 2013).

2 Peptidomics as Related to Postmortem Muscle Proteolysis

The interest of meat scientists in the study of postmortem proteolysis occurring in skeletal muscle immediately after animal slaughter has been a major issue for several decades. The reason is that this phenomenon is strongly linked to the development of the main quality attributes of meat and derived meat products. Among them, tenderness has been traditionally considered as the most important meat quality attribute by consumers (Ouali et al. 2013). Tenderness can be defined as the development of adequate toughness of meat due to the ageing process occurring after slaughter. This is because during this time, generally up to 2–3 weeks in the case of beef, there is an intense breakdown of muscle structure due to the action of several endogenous proteolytic systems, giving rise to a decrease in the mechanical resistance of meat and the generation of large polypeptides. With respect to dry-cured meat products, processing times are generally longer than

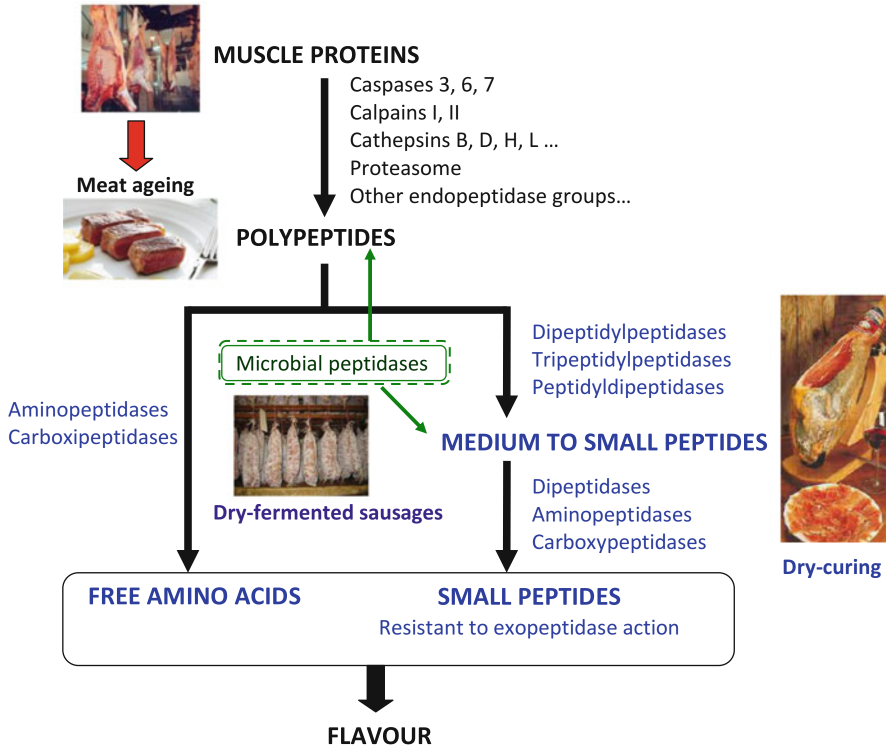

Fig. 1 Overview of the proteolytic systems implicated in the degradation of muscle proteins during both postmortem meat ageing and dry-curing

meat ageing. In the case of dry-cured ham, for example, processing times of 12 months and even more are usual. As a consequence, degradation of muscle proteins will be much more intense, and large polypeptides will be further degraded to generate medium-to-small peptides, together with free amino acids, that will contribute to reach their appreciated final flavour and texture characteristics (Fig. 1). Postmortem proteolysis is not a single process but a complex set of enzymatic actions driven by a wide variety of endo- and exopeptidase groups endogenous of muscle tissue that will be active far beyond animal slaughter. In the case of meat products in which starter cultures have been used during processing, proteolytic reactions will be even more complex due to the additional implication of bacterial proteolytic systems. This complexity of the proteolytic events occurring in postmortem muscle is summarized in Fig. 1. With respect to the proteolytic events occurring early postmortem, triggering of apoptosis is most probably the first step responsible for the degradation of protein cell constituents, with the activation of both initiator and executor caspases (Ouali et al. 2013). After this, other endopeptidase groups such as calpains, cathepsins, proteasome, etc. will continue the degradation of the myofibrillar structure. In that respect, additional

research is needed to clearly understand the role and importance of each peptidase group in the initial steps of muscle degradation (Sentandreu et al. 2002).

During the processing of meat products, the proteolytic action extends far beyond the initial days after slaughter; as a result, large polypeptides generated during the initial steps will be further degraded to smaller peptides and free amino acids due to the action of exopeptidase groups such as di- and tri-peptidilpeptidases, peptidyl dipeptidases, dipeptidases, carboxypeptidases and aminopeptidases (Fig. 1). Recent research works based on mass spectrometry applied to peptidomics have notably contributed to gain knowledge about the main protein substrates and peptides generated during postmortem muscle proteolysis, but it is also true that much work is still needed to explain the whole proteolytic processes as a way to predict and standardize the final quality of meat and processed meat products.

2.1 *Meat Storage and Development of Its Final Quality*

There have been notable efforts during the last two decades in the development of state-of-the-art peptidomic approaches focused on the study of peptides generated during the storage of meat in relation to product quality. Table 1 summarizes the main contributions in this field, highlighting the topics of research and the different strategies that have been employed to achieve the identification of peptides derived from muscle proteins. With respect to the identification of peptides generated during meat storage, one of the pioneer contributions is the work carried out by Nishimura and colleagues. In 1995, this group carried out the identification of a peptide produced during beef meat storage up to 10 days postmortem. To achieve this, trichloroacetic acid (TCA) muscle-soluble extracts were fractionated by HPLC, allowing isolation and collection of the target peptide. Determination of the peptide sequence was achieved by N-terminal sequencing using automated Edman degradation, identifying a 15-amino acid peptide derived from bovine muscle troponin T (Nakai et al. 1995). A similar sequence, found in peptides generated during pig meat ageing, was shown to be able to suppress the sour taste (Okumura et al. 2004). All these results contributed to establish troponin T as one of the key proteins involved in meat tenderization. Following a similar approach, Stoeva et al. (2000) were able to identify small peptide fragments derived from glyceraldehyde-3-phosphate dehydrogenase, troponin T and creatine kinase during bovine meat storage up to 15 days postmortem. In addition to N-terminal sequencing, these authors made use of MALDI-TOF MS for determining the exact molecular mass of the isolated peptides. They proposed the identified peptides as potential good indicators of beef meat quality, suggesting that rapid analytical procedures such as HPLC, capillary electrophoresis or immunoassays could be developed for routine analysis of these quality markers (Voelter et al. 2000).

A different strategy was carried out by Muroya et al. (2004) at the time to reveal the identity of large peptide fragments (25–32 kDa) generated during postmortem ageing of bovine muscle. Western blot analysis revealed that these polypeptides

Table 1 Peptidomics as related to meat storage

Topic	Focus	Peptide separation	Peptide identification	Peptides derived from	Animal model (References)
(1) Peptides generated during meat storage/ageing	(1.A) Small peptides	HPLC	N-terminal sequencing	Troponin T	Beef (Nakai et al. 1995)
				Troponin T, GAPDH	Pork (Okumura et al. 2003)
			MALDI-TOF/N-terminal sequencing	GAPDH, troponin T, CK	Beef (Stoeva et al. 2000)
			ESI-tripleQ/N-terminal sequencing	Troponin T	Pork (Okumura et al. 2004)
		1D-PAGE/HPLC/capillary electrophoresis	ELISA/N-terminal sequencing	Troponin T, GAPDH	Beef (Voelter et al. 2000)
	(1.B) Polypeptides	1D-PAGE	Western blot/N-terminal sequencing	Troponin T	Beef (Muroya et al. 2004)
			MALDI-TOF/LC-ESI-QTOF	MHC	Beef (Sawdy et al. 2004)
				GAPDH, troponin I, CK, AK	Pork (Park et al. 2007)
			MALDI-TOF/TOF	GAPDH, troponin T, Troponin I, MHC, PK	Beef (Sierra et al. 2012)
		2-D-PAGE	2-D-PAGE patterns comparison	Muscle proteins	Pork (Lametsch and Bendixen 2001)
			MALDI-TOF	Actin, myosin, troponin T, GP, CK, PPH, myokinase, PK, DHLST	Pork (Lametsch et al. 2002, 2003)
				Actin, myokinase, CK, α-crystallin	Pork (Morzel et al. 2004)
				Actin, GPDH, CK, PGK, HSC 71, GAPDH, β-enolase	Beef (Laville et al. 2009)
				Actin, MyBPH, HSP27, α-crystallin	Beef (Morzel et al. 2008)
			MALDI-TOF/Western blot	Troponin T	Beef (Muroya et al. 2007)
			LC-ESI-ion trap MS/MS	Actin, CK, GAPDH, MLC, MHC, titin, troponin I, troponin T, UCE	Beef (Polati et al. 2012)

(2) Specificity of enzymes implicated in muscle proteolysis	(2.A) Calpains	1-D/2-D-PAGE	MALDI-TOF	Myofibrillar proteins	Pork (Lametsch et al. 2004)
		1-D-PAGE/HPLC	MALDI-TOF/N-terminal sequencing	Troponin T	Rabbit (Hughes et al. 2001)
			Western blot/N-terminal sequencing	Troponin T	Pork (Kitamura et al. 2005)
		1-D-PAGE	Western blot/N-terminal sequencing	Desmin	Pork (Baron et al. 2004)
	(2.B) Cathepsins	1-D-PAGE/HPLC	[a]PDMS-TOF/N-terminal sequencing	F-actin	Beef (Hughes et al. 1999, 2000)
		1-D-PAGE	Western blot/N-terminal sequencing	Desmin	Pork (Baron et al. 2004)
	(2.C) Caspases	1-D-PAGE	MALDI-TOF/Western blot	Desmin, troponin I, MLC, actin, troponin T	Pork (Kemp and Parr 2008)
			LC-ESI-ion trap MS/MS	Actin, MHC	Pork (Rodriguez-Frometa et al. 2013)

GAPDH glyceraldehyde-3-P dehydrogenase, *GPDH* glycerol-3-P dehydrogenase, *CK* creatine kinase, *AK* adenylate kinase, *GP* glycogen phosphorylase, *PPH* phosphopyruvate hydratase, *PK* pyruvate kinase, *PGK* phosphoglycerate kinase, *DHLST* dihydrolipoamide succinyltransferase, *HSC 71* heat-shock cognate 71 kDa, *MLC* myosin light chain, *MHC* myosin heavy chain, *UCE* ubiquitin-conjugating enzyme E2N-like

[a]*PDMS-TOF* plasma desorption time-of-flight mass spectrometry

were fragments of troponin T. N-terminal sequencing corresponding to the first amino acid residues of these fragments confirmed troponin T as the protein of origin and allowed to define the N-terminal cleavage sites responsible for the generation of these fragments.

With the aim to study changes in muscle proteome during storage by pointing at a larger set of protein coverage, Bendixen and co-workers introduced two-dimensional gel electrophoresis (2-DE) for characterizing changes occurring in pig muscle during postmortem storage. This technique has proved to be particularly suitable for the study of large polypeptides; however, gel electrophoresis approaches are limited in the fractionation of peptides lower than 3–5 kDa since they will be eluted from the gel unless special protocols are applied. Protein changes were initially reported by comparing 2-DE patterns of muscle samples taken at different postmortem times (Lametsch and Bendixen 2001), whereas identification of differently abundant spots was accomplished by trypsin digestion followed by MALDI-TOF MS (Lametsch et al. 2002). This way, they were able to identify 18 polypeptides generated from postmortem storage that could serve as possible markers of pork quality. Furthermore, 2-DE proteome analysis established interesting correlations between tenderness development and postmortem degradation of actin and myosin (Lametsch et al. 2003). As seen in Table 1, the use of 2-DE as a high-throughput resolving technique coupled with different MS approaches for spot identification has been widely used in subsequent works to characterize the generation of considerable amounts of polypeptides during meat storage in relation to tenderness. There are also some works reporting the use of one-dimensional electrophoresis (1-DE) as the fractionation step to separate and further identify, by MS approaches, some proteolytic fragments that were found to correlate with ultimate beef tenderness (Sawdy et al. 2004; Sierra et al. 2012). Compared to automated Edman degradation, current high-throughput peptide sequencing based on tandem mass spectrometry (MS/MS) has made affordable the generation of massive sequencing data and the analysis of meat peptidome at great scale. As a drawback, the standard identification of polypeptides based on *bottom-up* MS approaches requires a prior trypsin digestion step that most often does not give information about the identity of neither N- nor C-terminal ends, thus being difficult to elucidate the cleavage sites responsible for the generation of these polypeptides.

Peptidomic research has been decisive in gaining knowledge about the role that the different endopeptidase groups can have in the development of meat tenderness through the cleavage of key proteins. Works carried out by Hughes and collaborators showed the ability of cathepsins B and D to hydrolyse bovine F-actin. Cathepsin B was able to generate both large (29–35 kDa) and small peptides of around 1 kDa (Hughes et al. 1999), whereas cathepsin D caused an extensive degradation of actin, generating small peptides (0.8–2.2 kDa) that could be mainly detected by HPLC (Hughes et al. 2000). The combined use of N-terminal sequencing and peptide mass determination using mass spectrometry allowed authors characterizing, for the small peptides, both the N- and C-terminal cleavage sites of the cathepsin endopeptidase action. On the contrary, incubation of desmin with

cathepsin B caused hydrolysis of this protein but through its peptidyl dipeptidase action, liberating successive dipeptides from the C-terminal end (Baron et al. 2004).

The action of calpains on myofibrillar proteins has been also a major topic of research since this enzyme group is considered to play also an important role in postmortem myofibril disruption (Sentandreu et al. 2002). Using the aforementioned strategy of combining N-terminal sequencing and mass spectrometry for peptide mass determination, Hughes et al. (2001) were able to characterize the proteolytic action of μ-calpain on rabbit troponin T. These authors were able to identify some peptides, as well as to characterize cleavage sites of the calpain action at both N- and C-terminal sides, for both large polypeptides (18–22 kDa) and medium-to-small peptides (2–5 kDa). A similar work was carried out by Kitamura et al. (2005); however, the molecular mass of generated troponin T peptides was not determined; thus, only the nature of N-terminal cleavage site was elucidated. In a different approach, using 2-DE and spot identification by in-gel digestion and MALDI-TOF MS, Lametsch et al. (2004) confirmed the hydrolysis of troponin T by μ-calpain together with eight additional myofibrillar proteins. They identified polypeptide fragments coming from actin, tropomyosin α4, myosin heavy chain and myosin light chain 1. They also reported that the degree of hydrolysis of both actin and myosin by the action of μ-calpain was rather low compared to other protein substrates. Considering the hypothesis that postmortem muscle proteolysis is a multienzymatic process, involving different peptidase groups in addition to cathepsins and calpains (Sentandreu et al. 2002), further studies were driven to elucidate the ability of other peptidases to hydrolyse actin and myosin more efficiently.

The way to look at the conversion of muscle into meat experienced a step forward from the conceptual point of view when the group of Ouali and co-workers introduced in 2006 the hypothesis of programmed cell death or apoptosis as one of the first events occurring in postmortem muscle after animal bleeding, providing answers to observations that remained unclear by that time. As depicted in Fig. 2, the onset of apoptosis in the muscle would probably imply the activation of caspases as the first peptidase group initiating the cell dismantling process and facilitating the action of the rest of peptidase groups, with important consequences on the development of the final quality of meat (Ouali et al. 2006). Following this idea, Kemp and Parr (2008) investigated the ability of caspase 3 to disrupt myofibrillar proteins by means of SDS-PAGE and MALDI-TOF MS. They found that caspase 3 was able to efficiently degrade desmin and troponin I, observing also the appearance of degradation products at around 32, 28 and 18 kDa, which were identified as fragments of actin, troponin T and myosin light chain 3, respectively. These findings confirmed the hypothesis of caspase action on postmortem muscle proteolysis and development of meat tenderness. This was also supported by the results obtained by Rodriguez-Frometa et al. (2013), who studied the action of caspases 3 and 7 on both bovine and porcine myofibrillar proteins based on SDS-PAGE followed by identification of fragments using LC-ESI-MS/MS. According to their results, caspase 7, but not caspase 3, efficiently hydrolysed myosin heavy chain of both animal species, generating a polypeptide of around

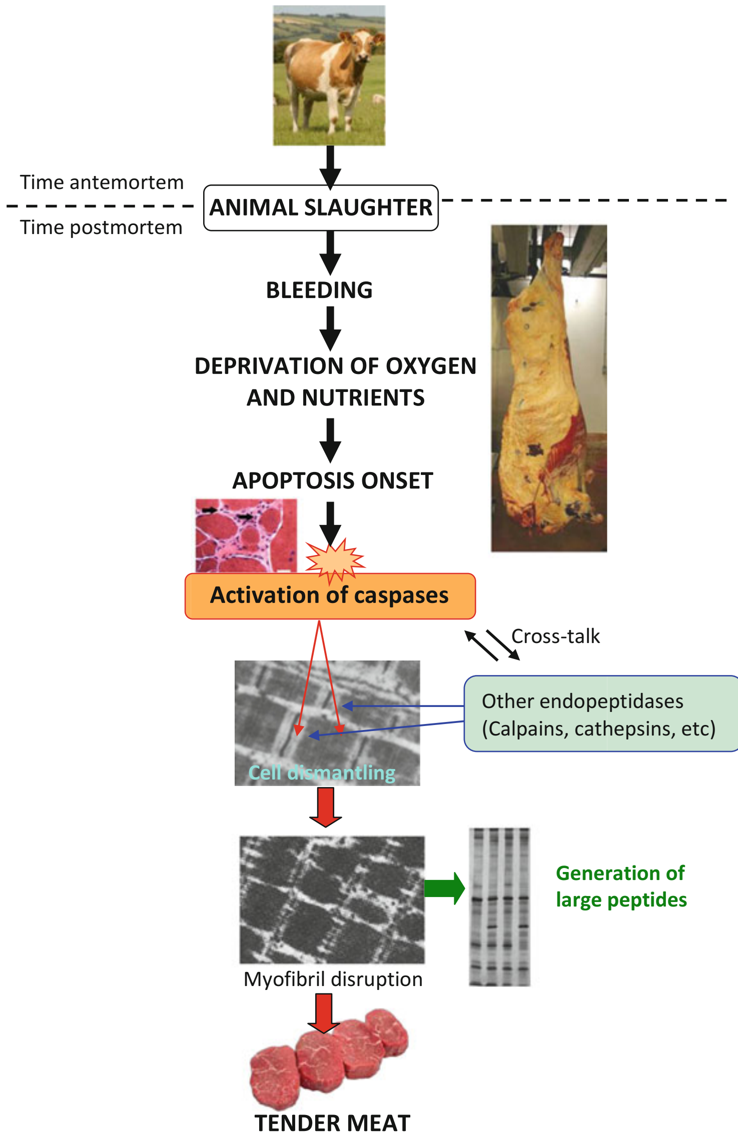

Fig. 2 Schematic view illustrating the conversion of muscle into meat after animal slaughtering. Triggering of apoptosis by caspase activation would be the first step in the dismantling of muscle myofibrillar structure, following the concerted and complementary action of other endogenous endopeptidase groups

200 kDa. On the other hand, porcine and bovine myosin light chain 1 was extensively hydrolysed by caspase 3 but not by caspase 7, suggesting a complementary and coordinated action between these two effector caspases during cell dismantling. But probably the main hallmark of caspase-mediated proteolysis on both porcine

and bovine myofibrillar proteins was the generation of a polypeptide of around 30 kDa of remarkable intensity, indistinctive of either caspase 3 or caspase 7 action. MS/MS data of tryptic peptides generated from these polypeptides revealed that it was a fragment of actin in all cases, supporting the idea that caspases would be probably the main peptidase group, through triggering of apoptosis, implicated in the early postmortem degradation of the actomyosin complex during the development of meat tenderness (Ouali et al. 2013). Despite the commented advances on peptidomic research in relation to the conversion of muscle into meat, there is still a long way to do if we want to fully characterize the main proteolytic products of muscle peptidases to really assess the real contribution of each group to the final tenderness and overall meat quality.

2.2 Proteolysis During the Processing of Meat Products

Studies on Non-Fermented Meats The main contributions of peptidomic research dealing with the quality of processed meat product are summarized in Table 2. First works were focused on the characterization of taste-related peptides present in different types of dry-cured ham. Sforza et al. (2001) attempted to characterize small peptides present in Parma ham in relation to the occurrence of a bitter taste. Since peptide analysis was carried out by HPLC coupled with single MS detection, the exact peptide sequence was difficult to achieve for each full scan spectrum; however, they were able to determine the presence of lipophilic amino acids in peptides, establishing a correlation to bitter taste. Following a similar approach but using N-terminal sequencing, these authors were able to elucidate the sequence of a 25-amino acid peptide having a correlation to ageing time but not to bitter taste, supporting the assumption that the main contribution to bitter taste may be due to free amino acids and small peptides (Sforza et al. 2003). This was in agreement to the results obtained by Sentandreu et al. (2003), who determined the sequence of peptides contained in the size-exclusion savoury fractions of Serrano dry-cured ham, most of them being dipeptides. More recently, taste-active peptides associated to umami taste from Jinhua and Parma hams were elucidated by Dang et al. (2015) using MALDI-TOF/TOF MS.

Research has also been applied to characterize the great variety of peptides generated during the ripening of dry-cured meat products, as can be seen in Table 2. Usual approaches include fractionation of deproteinized ham extracts by classical size-exclusion chromatography and HPLC followed by peptide sequence identification using tandem mass spectrometry. In the work carried out by Sentandreu et al. (2007), they identified four oligopeptides derived from the degradation of actin in Serrano dry-cured ham using both MALDI-TOF MS and ESI-MS/MS. Some of the cleavage sites associated to the generation of these fragments were shown to be produced by the action of cathepsin D on bovine F-actin (Hughes et al. 2000), thus supporting a relevant action of this enzyme during the processing of dry-cured ham. Peptides identified in Serrano dry-cured ham also include many others derived from

Table 2 Peptidomics in processed meat products

Topic	Peptide fractionation/ separation	Peptide identification	Product	References
(1) Study of peptide generation during the processing period	SDS-PAGE + HPLC	ESI-MS/MS + Western blot	San Daniele, Parma and Toscano ham	Fabbro et al. (2016)
		MALDI-TOF/TOF	Argentinean fermented sausages	López et al. (2015a)
	2-DE	MALDI-TOF	Parma and San Daniele ham	Di Luccia et al. (2005), Picariello et al. (2006)
	2-DE + HPLC	MALDI-TOF+ ESI-MS/MS	Bayonne ham	Theron et al. (2011)
	Size-exclusion + HPLC	MALDI-TOF/TOF	Serrano ham	Mora et al. (2010)
		MALDI-TOF + ESI-MS/MS	Serrano ham	Sentandreu et al. (2007), Mora et al. (2009a, b), Gallego et al. (2015b)
		ESI-MS/MS	Serrano ham	Mora et al. (2011a, b, 2015b), Gallego et al. (2014, 2016)
			Meat-fermented models	López et al. (2015b, c)
			Serrano ham + dry-fermented sausages	Mora et al. (2015c)
	Solid phase extraction + HPLC	ESI-MS/MS	Serrano ham	Gallego et al. (2015a)
	Deproteinized extract	MALDI-TOF/TOF	Dry-fermented sausages	Mora et al. (2015a)
(2) Characterization of taste peptides	Size-exclusion + HPLC	N-terminal sequencing	Serrano ham	Sentandreu et al. (2003)
		MALDI-TOF/TOF	Jinhua and Parma ham	Dang et al. (2015)
		ESI-MS	Parma ham	Sforza et al. (2001)
		N-terminal sequencing + ESI-MS	Parma ham	Sforza et al. (2003)
(3) Characterization of raw and/or final product according to animal characteristics	2-DE + HPLC	MALDI-TOF+ ESI-MS/MS	Kraski prsut ham	Skrlep et al. (2011)
	HPLC	ESI-MS/MS	Serrano ham	Mora et al. (2016)

myofibrillar proteins such as troponin T (Mora et al. 2010), titin (Gallego et al. 2015b) or myosin light chains (Mora et al. 2009b, 2011a), for example, which will be directly related to the final texture and taste characteristics of this product. In the same line, an intense proteolysis has also been reported for glycolytic enzymes at the end of dry-curing with the identification of small and medium peptides (Mora et al. 2009a, 2011b, 2015b). Under the same approach, Gallego et al. (2016) studied the evolution of peptide formation along the dry-curing process with the aim to identify potential biomarkers of the curing time and final product quality.

In the case of Parma and San Daniele hams, the use of 2-DE, trypsin hydrolysis and MALDI-TOF MS allowed reporting the presence of large polypeptides of myosin heavy chain still at 14 months of curing, whereas no such polypeptides were found for actin and myosin light chain, indicating an extensive proteolysis, thus as proof of a ripening time of at least 12 months (Di Luccia et al. 2005). The same authors reported a separation approach, alternative to conventional 2-DE, which combined acetic acid, urea, and Triton X-100 polyacrylamide gel electrophoresis, with standard SDS-PAGE. This would allow achieving a higher resolution in the separation of complex mixtures of basic proteins having also similar molecular weights, which is the case of water-soluble extracts of meat and meat products. Concomitant to the disappearance of several sarcoplasmic proteins, they reported the generation of several polypeptides ranging from 14 to 38 kDa in several Italian dry-cured meat products (Picariello et al. 2006). 2-DE coupled with MS identification of spots of interest has also been used in the study of proteome changes and polypeptide formation in other dry-cured ham types. In the case of Kraski prsut ham, pig genotype influenced the quantity of an actin polypeptide, whereas the salt level had a marked effect on the quantity of several protein fragments, indicating its important effect on proteolysis during dry-curing (Skrlep et al. 2011). This is in accordance to the results obtained in Bayonne dry-cured ham, where a more intense formation of protein fragments was reported in the *biceps femoris* compared to *semimembranosus* muscle as consequence of a higher proteolysis due to the fact that the enzyme inhibitory effect of salt occurs later in this muscle, which is protected by rind and fat (Theron et al. 2011).

Studies on Fermented Meats In the case of meat products in which bacterial fermentation occurred and led mainly by lactic acid bacteria (LAB), meat protein degradation will be a quite more complex process. As mentioned before, the proteolytic system of the microorganisms growing spontaneously or added as starter cultures will complement meat endogenous proteolytic machinery to achieve protein hydrolysis during ripening. Prediction and/or characterization of the final quality of meat-fermented products is a difficult task since a high number of factors such as the type of raw materials, technologies and starter cultures are involved in the process. The identification of the protein fragments naturally generated during sausage fermentation and ripening would be beneficial in order to better understand proteolysis and flavour development mechanisms that occur during the processing of fermented products. Recently, López et al. (2015a, b, c) have focused their studies on fermented meat proteolysis applying both proteomic and peptidomic

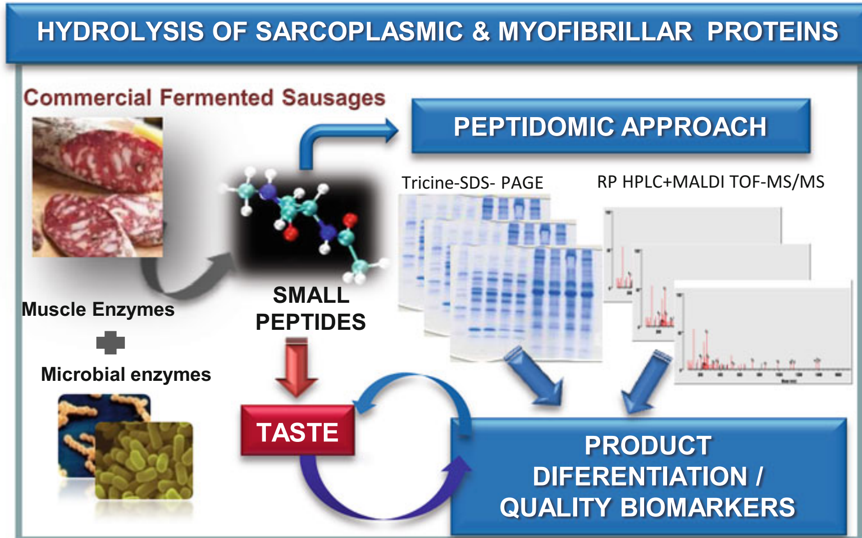

Fig. 3 Schematic summary of proteolysis that occurred in Argentinean fermented sausages carried out by López et al. (2015a) from a peptidomic approach

approaches. In the first work, low molecular weight (LMW) peptides (< 3 kDa) and protein profiles from commercial Argentinean fermented sausages obtained by tricine-SDS-PAGE and RP-HPLC-MS, respectively, allowed to distinguish two different types of fermented sausages, although no specific biomarkers relating to commercial brands or quality were recognized (Fig. 3). Thirty-six LMW peptides arising from sarcoplasmic (28) and myofibrillar (8) proteins were identified. These peptides had been originated from α-actin, myoglobin and creatine kinase M-type but also from the hydrolysis of other proteins not previously reported. Results showed that although muscle enzymes exerted a major role on peptidogenesis, microbial contribution cannot be excluded. This work represents the first peptidomic approach for fermented sausages, thereby providing a baseline to define key peptides acting as potential biomarkers (López et al. 2015a). Subsequently, the study of the degradation of proteins was approached in a beaker model system inoculated with a selected autochthonous starter culture to evaluate the role of the microorganisms in the proteolytic process. Sausage models inoculated with *Lactobacillus curvatus* CRL705 and *Staphylococcus vitulinus* GV318 were incubated 10 days at 22 °C. Low molecular weight peptides (<3 kDa) derived from sarcoplasmic and myofibrillar proteins were analysed by 2-DE and LC-MS/MS and complemented with amino acid profiles, in order to provide a whole map of proteolysis results (Fig. 4). A diverse number of protein fragments were identified. Results indicated that peptides mainly arose from myoglobin, creatine kinase, glyceraldehyde-3-phosphate dehydrogenase and fructose-biphosphate aldolase (ALDOA). Also the hydrolysis of actin, myosin light chain 1/3 (MLC 1/3), myosin

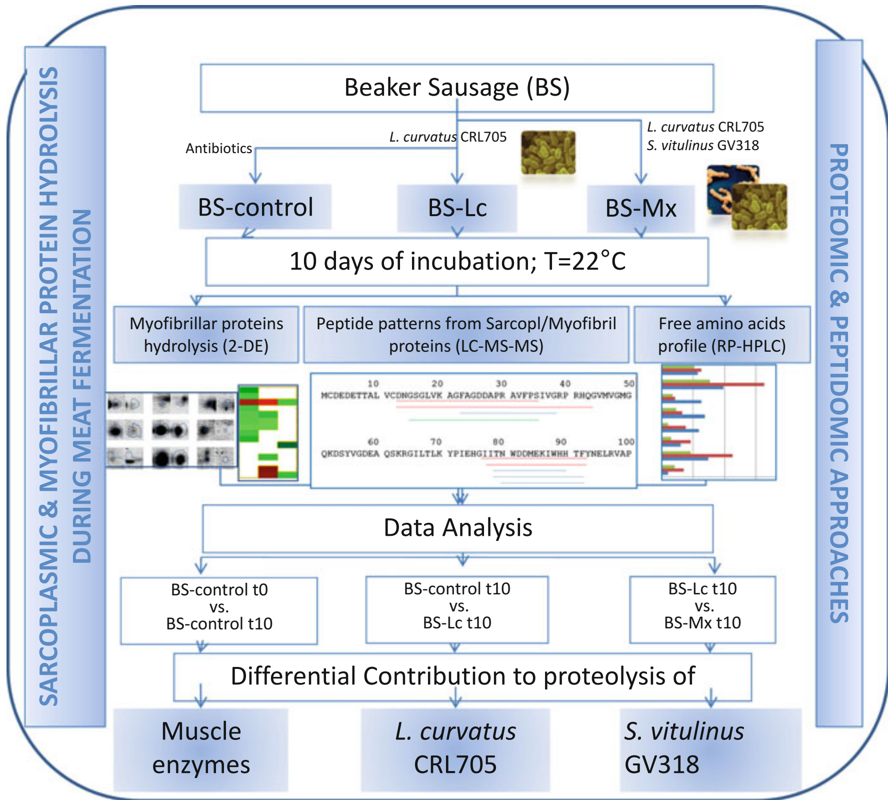

Fig. 4 Graphical summary of López et al.'s (2015b, c) research in which proteolysis of sarcoplasmic and myofibrillar protein hydrolysis by the action of an autochthonous starter culture was studied in a beaker sausage model using a proteomic and peptidomic approach

regulatory light chain 2 (MRLC-2) and myosin heavy chain (MHC) were evidenced by 2-DE. Thirty-three peptides arisen from troponin T, MRLC-2 and particularly from actin were identified. These results showed that, in addition to the endogenous enzymes, the starter culture enhanced the hydrolysis of the above-mentioned proteins. *L. curvatus* CRL705 highly enriched both peptide pattern and amino acid concentrations. The primary structure of actin was highly susceptible to degradation by the starter culture especially in three different regions. Additionally, the essential role of exopeptidases—from meat and bacteria—was evidenced by the diversity of actin-derived peptides during fermentation. These studies improved the knowledge of the proteolysis of sarcoplasmic and myofibrillar proteins, as well as the role of the studied autochthonous starter culture. In fact, the use of a specific autochthonous starter culture guarantees hygiene and tipicity of fermented sausages. The identification of new peptides as well as new target proteins by means of peptidomics represents a significant step towards the elucidation of the role of microorganisms in meat proteolysis. Moreover, these peptides may be further used

as biomarkers capable to certify the use of the applied autochthonous starter culture described here (López et al. 2015b, c).

3 Meat Microbiology: Introduction to Meat Ecosystems

In animals under good condition, tissues that eventually develop into meat are usually considered to be germ-free. However, if not properly handled, processed and preserved, meat will support the growth of microorganisms, thereby creating a significant health risk. Usually contamination arises during the slaughter, dressing, cutting and/or packaging process (Koutsoumanis and Sofos 2004; Sharma and Chattopadhyay 2015). Sources of contamination are diverse: faeces, ingestion, hides, lymph nodes and/or intestines of the animals. Although some differences exist between animal species regarding the type and frequency of microorganisms isolated from carcasses, the same bacterial species from beef, pork, sheep and even chicken carcasses are generally isolated (Zhao et al. 2001).

3.1 Meat Contamination: Pathogenic Microorganisms

In recent years, the major short-term health risks associated with the consumption of meat have been from infections with enteric pathogens. *Campylobacter*, *Salmonella* and pathogenic *E. coli* all colonize the gastrointestinal tracts of a wide range of wild and domestic animals, especially animals raised for human consumption. *Salmonella* is the main causative agent of food-borne outbreaks reported in Europe in 2014 (EFSA European Food Safety Authority 2015). High levels of *Salmonella* in the meat may arise from animal production practices at the rearing stage as well as cross-contamination after slaughter (McEvoy et al. 2006). This pathogen is most frequently detected in poultry meat followed by turkey and pork meat (Carraturo et al. 2016). Food-poisoning staphylococci are also widely distributed; meat contamination is often connected to poor hygienic practices during slaughtering, transportation, chopping, storage and points of sale by the personnel implicated in the manufacture process. Moreover contaminated meat can move staphylococci to different processing surfaces and packaging materials (Karmi 2013). The etiologic agent is *Staphylococcus aureus*, and its related heat-stable enterotoxins are a major cause of food-poisoning cases and outbreaks worldwide (Liu et al. 2006). For instance, 50 isolates of methicillin-resistant *Staphylococcus* spp. were detected recently in raw meat samples in Nigeria (52% from pork, 28% from beef and 20% from chicken samples) (Igbinosa et al. 2016).

On the other hand, Shiga toxin-producing *Escherichia coli* (STEC) emerged as a food-borne pathogen more significant than other well-known ones because of the severe consequences of infection, its low infection dose, its acid tolerance and its association with ruminants used for food (Hussein and Bollinger 2005). STEC is of

major concern for the sustainability of the meat industry and a serious threat for public health. As such, STEC is still a very serious menace that can jeopardize the sustainability of beef- and pork-meat chain. In fact pigs are also reservoirs of STEC. The entrance of these strains into the food chain implies a risk for consumers because of the severity of haemolytic uremic syndrome. Colello et al. (2016) have recently reported the prevalence and characterization of STEC during pork production.

Listeria monocytogenes has continued to raise food safety concerns, especially with respect to ready-to-eat (RTE) products. Listeriosis, caused by this pathogen, is a significant public health concern as a result of its clinical severity and high mortality rates (Vaillant et al. 2005).

Species within the genus *Campylobacter* and *Yersinia* have also emerged as pathogens of human public health concerns (Buncic et al. 2014). Campylobacteriosis was the most commonly reported zoonosis in the European Union (EU) since 2008. In food, the incidence of *Campylobacter* remained high in broiler meat. Positive findings for *Yersinia* were mainly reported in pork meat and products thereof (EFSA 2015).

As mentioned before, food contamination with these pathogens occurs generally along the food chain (production, processing, distribution, retail marketing and handling). In consequence, microbial food safety is an increasing public health worry worldwide. Many epidemiological studies have implicated foods of animal origin as the major vehicles associated with illnesses caused by food-borne pathogens. Contaminated raw or undercooked poultry and red meats are particularly important in transmitting these pathogens (Zhao et al. 2001). To prevent food-borne outbreaks related to harmful microorganisms, the official control of meat is critical in order to verify and ensure the safety of products and, therefore, guarantee public health (Carraturo et al. 2016). As proposed by many authors, effective control of meat-borne pathogens requires an integrated approach involving all the actors of meat chain, the use of Good Manufacturing/Good Hygienic Practice (GMP/GHP) and Hazard Analysis and Critical Control Point (HACCP) principles (Buncic 2006; Nørrung et al. 2009).

3.2 Meat Contamination: Spoilage Microorganisms

Some of the microorganisms that grow in food can cause unacceptable sensory alterations, through the production of metabolites, such as off-flavours or changes in texture or appearance (Ellis and Goodacre 2001). These are the microorganisms defined as specific spoilage organisms for food, as other microorganisms may also grow in food but without causing any sensory changes.

Cold storage of meat will decrease bacterial growth, only 10% of the bacteria initially present being able to grow at refrigeration temperatures. As meat is a selective agent for aerobic microbiota, a consortium of bacteria commonly dominated by *Pseudomonas* spp. is responsible for the spoilage of meat stored

aerobically at temperatures between −1 and 25 °C. Cold-tolerant *Enterobacteriaceae*, *Brochothrix thermosphacta* and LAB also occur in chilled meat stored aerobically, but in terms of numbers, they do not contribute to dominate microbial associations (Nychas et al. 2008). Oxygen restriction by the use of vacuum or modified atmospheres will drastically reduce the presence of *Pseudomonas*, and bacterial microbiota will be gradually selected towards CO_2-tolerant organisms. Under these conditions, the dominating microorganisms involve *Brochothrix thermosphacta*, *Enterobacteriaceae* and LAB.

Systems for retail meat distribution and commercialization are mainly based on vacuum packaging (VP) or modified atmosphere packaging (MAP) of meat cuts using low gas permeability films and refrigeration. These methods using gas mixtures containing variable O_2 and CO_2 concentrations and low temperatures proved to be very effective in extending the shelf life of perishable foods such as muscle foods preventing the growth of pathogens such as pseudomonads and *Enterobacteriaceae* so that spoilage typically occurs associated with the growth of psychrotolerant LAB (Jones 1999; McMillin 2008). Both LAB and *B. thermosphacta*, the most important cause of spoilage commonly associated with meat packaged under VP and MAP conditions, result from the competition between facultative anaerobic Gram (+) biota (Russo et al. 2006). The LAB most often isolated from meat belong to the genera *Carnobacterium*, *Lactobacillus*, *Leuconostoc* and *Weissella*. They show varying potential to cause spoilage, e.g. by souring, discoloration, swelling and slime or off-odour compound production (Nieminen et al. 2011). The metabolites produced by LAB do not tend to cause spoilage of packaged meat until LAB species have reached maximum numbers and their metabolic by-products accumulated to sensorial detectable levels (Borch et al. 1996).

3.3 Bacteria of Technological Interest in Meat: The Focus on Lactic Acid Bacteria

Lactic acid bacteria (LAB) are cocci, coccobacilli or rods Gram (+), non-spore forming with a low G+C content (less than 53%). They involve a heterogeneous group of microorganisms that produce lactic acid from the fermentation of sugars as the main metabolic product. They are generally nonrespiratory organisms and lack catalase, fermenting glucose primarily to lactic acid, CO_2 and ethanol. LAB grow anaerobically, but they are aerotolerant and possess superoxide dismutase and peroxidase enzymes to detoxify peroxide radicals (Carr et al. 2002). The genera included in the LAB group are *Lactobacillus*, *Lactococcus*, *Streptococcus*, *Pediococcus*, *Weissella*, *Carnobacterium*, *Leuconostoc*, *Enterococcus*, *Tetragenococcus*, *Oenococcus* and *Vagococcus* (Holzapfel et al. 2001). Although most of LAB members are non-pathogenic being generally recognized as safe (GRAS) organisms, only a few of them are pathogenic, such as some members of

the genus *Streptococcus*. In addition, LAB are heterotrophic and generally have complex nutritional requirements because they lack many biosynthetic capabilities; most species have multiple requirements for amino acids and vitamins. Because of these nutritional exigencies, they can be found in environments where these requirements can be provided such as meat and meat products, milk and dairy products and cereals. They are often associated with animal and human oral cavities and intestines (e.g. *Streptococcus mutans*, *Enterococcus faecalis*), plants (*Lactobacillus*, *Leuconostoc*) and decaying plant materials.

LAB are the microorganisms of major technological importance in the food industry. Due to their characteristics and metabolic properties, they are essential to carry out the fermentation in fermented food, such as yogurt, cheese and cured sausages. Specifically in fermented meat products, LAB contribute to the development of the texture and flavour of the final product and inhibit food spoilage bacteria by producing growth-inhibiting substances and large amounts of lactic acid (Fadda et al. 2010). LAB are either present as contaminants or deliberately added as starter cultures to guarantee the optimal fermentation process. They are also considered the most important microorganisms responsible for the health-promoting effects of fermented foods, such as immunomodulation, intestinal integrity and pathogen resistance (Venema and do Carmo 2015). Indeed, strains of some species have traditionally been used as probiotics and added as functional bacteria in various food commodities (Florou-Paneri et al. 2013). Due to the huge economic significance of industrial application of LAB as starters, biopreservatives and probiotics, a research emphasis on their metabolism, genetics and applications has been placed in the last 30 years (Das and Goyal 2012).

On the other hand, in response to consumer's demands related to food free from pathogens, with minimal processing and less preservatives and additives, but keeping its sensorial quality, present trends in the meat technology include the research of alternative inhibitors of pathogen and spoilage microorganisms. Especially, biopreservation has gained increasing attention as a means of naturally controlling the shelf life and safety of foods. Certainly the use of bioprotective cultures to ensure the hygienic quality of food is a promising tool, and LAB stand as the best candidates to be used because of their prevalence during storage, their GRAS status and its inhibitory potential towards certain pathogen and spoilage microorganisms. In fact, the inhibitory properties of LAB are ascribed mainly to the production of organic acids, especially lactic and acetic, and production of hydrogen peroxide or bacteriocins. Bacteriocins are peptides ribosomally synthesized by bacteria that have antibacterial activity towards closely related strains, including food pathogens (*L. monocytogenes* and *Clostridium*) and other meat spoilage bacteria such as *B. thermosphacta*, which play an important role in food preservation (Pérez et al. 2014). New insights into LAB metabolism present perspectives for the application of a novel generation of functional cultures which contribute to the safety or offering technological, nutritional and sensorial advantages. However, the application of biopreservation technology in meat and meat products, by means of LAB, constitutes only an additional hurdle that complements good manufacturing and processing practices (Vignolo et al. 2015).

3.4 *Omics for Bacterial Quantification and Identification*

Microbial contamination of food from animal origin such as milk and meat has been a constant nuisance and an unavoidable problem throughout history. The role of the analytical tools for food products, involving the detection of spoilage and pathogenic organisms through quantitative and qualitative processes, is of utmost importance for food product appraisal and in the promotion of public health. In an attempt to preserve food quality manufacturing and production levels while safeguarding the public health, HACCP system has been implemented worldwide.

In an effort to achieve the necessities of food testing, research is being directed towards rapid methods of detection that could be implemented in highly automated processing facilities. Numerous different techniques have been explored and utilized over the years for the detection and quantification of microbiological contamination in food products, such as microscopy, polymerase chain reaction, ATP bioluminescence, nucleic acid probing, immunoassay methods and electronic nose techniques with positive and negative results into their application (Jay 2000; Commas-Riu and Rius 2009; Flint et al. 2006; Luo et al. 2009). Major drawbacks of these techniques include the high demand on operator skills, time consuming and relatively slow sample turnaround times, which limit them for daily use in the food industry.

The emergence of MALDI-TOF MS as a simple, rapid and economical technique for characterization of bacteria, largely at the species level and occasionally at the strain level, has greatly revolutionized microbial diagnostics (Nomura 2015; Cheng et al. 2016). Already in 1996, two reports indicated that MALDI-TOF mass spectral "fingerprints" could be simply and rapidly obtained from whole bacterial cells without any pretreatment before the MS analysis (Holland et al. 1996; Claydon et al. 1996). The use of small sample quantities and simple analysis techniques and the applicability on heterogeneous food samples are some of the potential advantages of the technique. Although other MS techniques, such as LC-MS, have also been used for microbial identification (Everley et al. 2008), MALDI-TOF MS has been adopted for routine diagnostics primarily because of the simple sample preparation during analysis that enables greater automation and high throughput.

MALDI-TOF MS-Based Identification: How Does It Work? A small number of cells from a colony cultured onto the agar plate are directly smeared onto the MALDI target plate and overlaid with the matrix solution. A solution of alpha-cyano-4-hydroxycinnamic acid in a mixture of organic solvents and water is used as the matrix for routine identification. The spotted mixture is air-dried and then inserted into the mass spectrometer for automated measurement. The final mass spectral signature is composed of peaks ranging from 1,000 to 30,000 m/z. MALDI usually produces singly charged ions; thus, the m/z of an analyte corresponds to its mass. The spectra obtained are compared with a library of known spectra, and a result is generated (Fig. 5). The characteristic spectrum pattern of this proteomic fingerprint is used to reliably and accurately identify a particular microorganism by

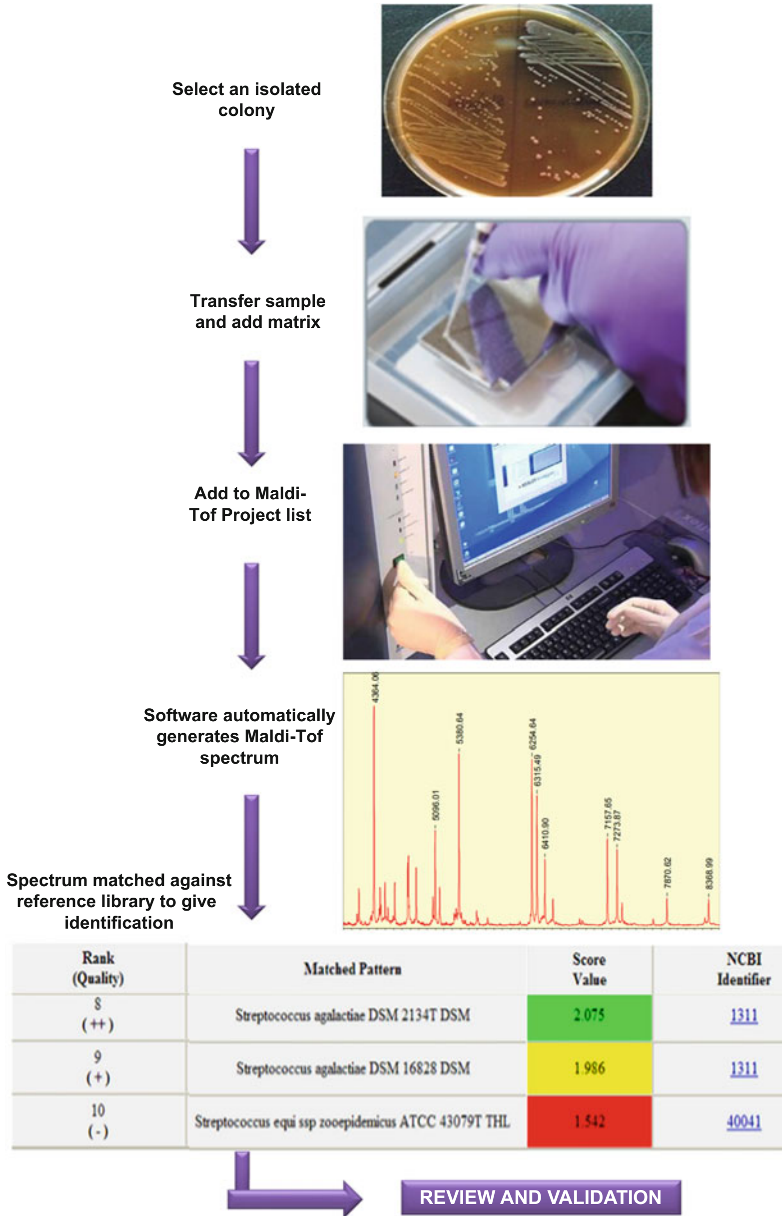

Fig. 5 MALDI-TOF mass spectrometry workflow in the microbiology laboratory with the Bruker MALDI Biotyper® system

matching thousands of reference spectra from microorganisms. The characterization of bacteria by MALDI-TOF MS involves the analysis of the whole bacterial proteomes. The profile obtained is known to include mainly ribosomal proteins that are expected to be minimally affected by changes in culture conditions (Wieser et al. 2012; Teramoto et al. 2007). In addition to ribosomal proteins, nucleic acid-binding proteins and cold shock proteins were assigned (Ryzhov and Fenselau 2001). It is noteworthy, however, that the identity of each peak is not necessarily relevant for bacterial identification by MALDI-TOF MS which is based on the following facts: (1) spectral fingerprints vary between microorganisms; (2) among the compounds detected in the spectra, some peaks (molecular masses) are specific to the genus, species and sometimes subspecies levels; (3) spectra are reproducible as long as the bacteria are grown under the same conditions. Currently, two systems for applications of MALDI-TOF MS-based bacterial identification are widely used (including their specific databases): the Bruker Biotyper® (Bruker Daltonics; https://www.bruker.com/products/mass-spectrometry-and-separations/maldi-biotyper-systems.html) and the VITEK® MS (bioMerieux; http://www.biomerieux-diagnostics.com/sites/clinic/files/9300819-002-gb-a_vitek-ms.pdf). The analytical principles of the two systems are similar, although there are differences in database construction and also in the algorithms used to identify microorganisms.

MALDI-TOF MS-Based Identification: Current Applications and Future Perspectives Numerous studies already have explored the subtyping ability of MALDI-TOF MS for several pathogens such as *Streptococcus pneumoniae* (Williamson et al. 2008), *L. monocytogenes* (Hsueh et al. 2014), *Streptococcus agalactiae* (Lartigue et al. 2009), *Staphylococcus aureus* (Wang et al. 2013) and *Enterococcus* species (Quintela-Qunitela-Baluja et al. 2013). Most of the above-mentioned studies have mainly focused on subtypification of pathogens during epidemiological or clinical research. In contrast, Siegrist et al. (2007) successfully used MALDI-TOF MS to classify environmental isolates of *E. coli* according to their origin.

This technology has been used extensively in clinical diagnostics; however, its use for food pathogen detection is comparatively less explored (Jadhav et al. 2014; Pavlovic et al. 2013). A few reports show MALDI-TOF MS for the identification of isolates from food and beverage sources. These reports describe different aspects of the usefulness of MALDI-TOF to food microbiology and spanned from the classification of lactic acid bacteria in fermented meat and the inspection of probiotics to strain identification and characterization of biogenic amine-producing bacteria (Nguyen et al. 2012; Angelakis et al. 2011; Ruiz-Moyano et al. 2012; Fernández-No et al. 2010). The analysis of microbes in food is a challenging task because food-borne pathogens may be dispersed in low concentrations in food and high concentrations of harmless background microbiota. Then, a pre-concentration step is necessary, which separates the target bacteria from other bacteria with the complex background of food products. In this sense, Ochoa and Harrington (2005) successfully evaluated a rapid method for the isolation and identification of the enterohaemorrhagic *E. coli* O157:H7 in ground beef by using microscopic

magnetic beads coated with specific O157 antibodies covalently bonded to the surface. The immunomagnetic separation increased the sensitivity of the method and permitted the detection and identification of bacteria in meat by using MALDI-TOF MS. On the other hand, Jadhav et al. (2014) reported a proof-of-concept study using a proteomics-based MALDI-TOF MS approach that proposes a detection scheme that is more rapid and simple compared to conventional methods of *Listeria* detection. Very low levels of the pathogen could be identified from different food samples (milk, cheese or chicken paté) post-enrichment in selective broths. Also another standardized MALDI-TOF MS scheme can provide an inexpensive, rapid and simple solution for detecting and source tracking *L. monocytogenes* isolates obtained from different food processing environments (Jadhav et al. 2015). Biomarker peaks specific to isolates obtained from different sources were identified. However, standardization of culture conditions may play an important role in assessing or testing the robustness of the technique.

These rapid and accurate merits of MALDI-TOF MS (only 4 min per sample) in relation to bacterial detection and enumeration (up to 2 days) make this approach a very good candidate for potential use in the dairy and meat industries. Such methods are urgently needed within HACCP, and such a general spoilage detection technique will be important for improving consumer safety and product quality. Particularly in the case of pathogen detection in foods, MALDI-TOF MS analysis represents an attractive tool for detection and subtyping of multiple pathogens, considering its rapid turnaround time according to the timely information required by industry to avoid product recalls or delays in product release. The bottleneck for successful identification is related to there being no or insufficient information on the organism in the database. This is indicating that an up-to-date and comprehensive database is a critical requirement for accurate identification of isolates using MALDI-TOF MS. As mentioned before, also culture conditions such as culture media, time and temperature can affect MALDI-TOF MS efficiency. Following this, there are reports showing difficulties of MALDI-TOF MS for differentiating closely related species (van Veen et al. 2010). This topic leads to some controversial results as some studies report excellent differentiation up to the subspecies level and some studies report that the MALDI-TOF MS was unable to identify particular isolates at all. Finally, the discussion thus far has centred on the identification of bacteria from cultured colonies. Direct analysis of clinical or food samples without the need for prior culturing might shorten the time required for identifying microorganisms and would therefore further increase the usefulness of the MALDI-TOF MS-based approach, this point remaining as a pending issue in the next future.

4 Conclusion and Future Prospects

Despite the terms "peptidomics" and "proteomics" referring to different concepts, it is also true that the difference between them is somewhat blurred. Among the reason for that, one is the fact that, nowadays, both are mainly based on mass

spectrometry analyses of peptides of different nature. Many proteomic workflows usually follow a *bottom-up* approach, which implies the prior digestion of proteins with enzymes of known specificity into peptides, to be further analysed by mass spectrometry as a way to get information about the protein of origin. Peptidomics, on the contrary, would aim at analysing peptides from its intact structure in a *top-down approach* with the objective to get the maximum information about the proteolytic processes responsible for their generation in the matrix of study. High-throughput analysis of peptides has become possible with the advent of the so-called soft ionization techniques (MALDI and ESI) together with massive genome sequencing linked to the expansion of protein databases and the development of powerful bioinformatic algorithms, capable to automatically elucidate the sequence of thousands of peptides from MS/MS spectra generated by modern MS equipment.

Peptidomics has become a relevant subject in meat science since it is directly related to postmortem muscle proteolysis, contributing to better understand the role of the different endopeptidase groups in myofibrillar protein disruption and development of meat tenderness. In the case of processed meat products, peptidomics has notably contributed to identify those protein fragments present at the different stages of the ripening period, reporting the intense proteolysis occurring during the whole period by the action of both endo- and exopeptidases and also, in the case of fermented meats, by the activity of microbial peptidases. Peptide characterization is important since they contribute, directly or indirectly, to the typical flavour characteristics of these products. On the other hand, MALDI-TOF MS has emerged as a promising technique to achieve a rapid, simple and economical characterization of the bacterial population that can be present in meat and meat-derived products. This would allow the detection of both pathogenic and non-pathogenic microorganisms in short times, something of outmost importance to assure safety of food and product appraisal. Future prospects in this approach will have to face the current limitations of the technique such as the need for standardized culture conditions prior to species identification and the difficulty to differentiate between closely related species. In addition to this, peptidomics needs to face the challenge to develop new and powerful bioinformatic tools enabling high-throughput and confident analysis of peptides that do not follow the ideal chemical structure of peptides obtained by cleavage with enzymes of known specificity, as in the case of trypsin in *bottom-up* proteomic approaches. This is even more remarkable in the case of peptides originated from proteins whose sequence is hosted in protein databases, for which peptide identification algorithms based on spectral matching and comparison with theoretical spectra cannot be applied. For those cases, improvements should address the development of tailored high-throughput peptide identification tools based on *de novo* sequencing from MS/MS spectral information independent of protein databases.

References

Angelakis E, Million M, Henry M, Raoult D (2011) Rapid and accurate bacterial identification in probiotics and yoghurts by MALDI-TOF mass spectrometry. J Food Sci 76:M568–M571

Baron CP, Jacobsen S, Purslow PP (2004) Cleavage of desmin by cysteine proteases: Calpains and cathepsin B. Meat Sci 68:447–456

Bergquist J, Ekman R (2001) Future aspects of psychoneuroimmunology - lymphocyte peptides reflecting psychiatric disorders studied by mass spectrometry. Arch Physiol Biochem 109:369–371

Buncic S (2006) Integrated food safety and veterinary public health. CABI International Publishing, Wallingford

Borch E, Kant-Muemansb ML, Blixt Y (1996) Bacterial spoilage of meat products and cured meat. Int J Food Microbiol 33:103–120

Buncic S, Nychas GI, Lee MRF, Koutsoumanis K, Hébraud M, Desvaux M, Chorianopoulos N, Bolton D, Blagojevic B, Antic D (2014) Microbial pathogen control in the beef chain: recent research advances. Meat Sci 97:288–297

Carr FJ, Hill D, Maida N (2002) The lactic acid bacteria: a literature survey. Crit Rev Microbiol 28:281–370

Carraturo F, Gargiulo G, Giorgio A, Aliberti F, Guida M (2016) Prevalence, distribution, and diversity of *Salmonella* spp. in meat samples collected from Italian slaughterhouses. J Food Sci 81(10):M2545–M2551. https://doi.org/10.1111/1750-3841.13430.

Cheng K, Chui H, Domish L, Hernandez D, Wang G (2016) Recent development of mass spectrometry and proteomics applications in identification and typing of bacteria. Proteomics Clin Appl 10:346–357. https://doi.org/10.1002/prca.201500086

Claydon MA, Davey SN, Edwards-Jones VD, Gordon B (1996) The rapid identification of intact microorganisms using mass spectrometry. Nat Biotechnol 14:1584–1586

Clynen E, Baggerman G, Veelaert D, Cerstiaens A, Van der Horst D, Harthoorn L, Derua R, Waelkens E, De Loof A, Schoofs L (2001) Peptidomics of the pars intercerebralis-corpus cardiacum complex of the migratory locust, Locusta migratoria. Eur J Biochem 268:1929–1939

Colello R, Cáceres ME, Ruiz MJ, Sanz M, Etcheverría A, Padola NL (2016) From farm to table: follow-up of Shiga toxin-producing *Escherichia coli* throughout the pork production chain in Argentina. Front Microbiol 7:93. https://doi.org/10.3389/fmicb.2016.00093

Commas-Riu J, Rius NJ (2009) Flow cytometry applications in the food industry. J Ind Microbiol Biotechnol 36:999–1011

Craik DJ, Young Shim Y, Goransson U, Moss GP, Tan N, Jadhav PD, Shen J, Reaney MJ (2016) Nomenclature of homodetic cyclic peptides produced from ribosomal precursors: an IUPAC task group interim report. Biopolymers 106:917–924

Dallas DC, Guerrero A, Parker EA, Robinson RC, Gan JN, German JB, Barile D, Lebrilla CB (2015) Current peptidomics: applications, purification, identification, quantification, and functional analysis. Proteomics 15:1026–1038

Dang YL, Gao XC, Ma FM, Wu XQ (2015) Comparison of umami taste peptides in water-soluble extractions of Jinhua and Parma hams. Lwt-Food Sci Technol 60:1179–1186

Das D, Goyal A (2012) Lactic Acid Bacteria in Food Industry. In: Satyanarayana T, Johri B, Prakash A (eds) Microorganisms in sustainable agriculture and biotechnology. Springer, Dordrecht

Di Luccia A, Picariello G, Cacace G, Scaloni A, Faccia M, Liuzzi V, Alviti G, Musso SS (2005) Proteomic analysis of water soluble and myofibrillar protein changes occurring in dry-cured hams. Meat Sci 69:479–491

EFSA (European Food Safety Authority) and ECDC (European Centre for Disease Prevention and Control) (2015) The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2014. EFSA J 2015 13:191. https://doi.org/10.2903/j.efsa.2015.4329

Ellis DI, Goodacre R (2001) Rapid and quantitative detection of the microbial spoilage of muscle foods: current status and future trends. Trends Food Sci Technol 12:414–424

Everley RA, Mott TM, Wyatt SA, Toney DM, Croley TR (2008) Liquid chromatography mass spectrometry characterization of *Escherichia coli* and *Shigella* species. J Am Soc Mass Spectrom 19:1621–1628

Fabbro A, Bencivenni M, Piasentier E, Sforza S, Stecchini ML, Lippe G (2016) Proteolytic resistance of actin but not of myosin heavy chain during processing of Italian PDO (protected designation of origin) dry-cured hams. Eur Food Res Technol 242:881–889

Fadda S, López MC, Vignolo G (2010) Role of lactic acid bacteria during meat conditioning and fermentation: peptides generated as sensorial and hygienic biomarkers. Meat Sci 86:66–79

Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Electrospray ionization for mass spectrometry of large biomolecules. Science 246:64–71

Fernández-No IC, Böhme K, Gallardo JM, Barros-Velázquez J, Cañas B, Calo-Mata P (2010) Differential characterization of biogenic amine-producing bacteria involved in food poisoning using MALDI-TOF mass fingerprinting. Electroph 31:1116–1127

Flint S, Drocourt J-L, Walker K, Stevenson B, Dwyer M, Clarke I, McGill D (2006) A rapid, two-hour method for the enumeration of total viable bacteria in samples from commercial milk powder and whey protein concentrate powder manufacturing plants. Int Dairy J 16:379–384

Florou-Paneri P, Christaki E, Bonos E (2013) Chapter 25: Lactic acid bacteria as source of functional ingredients. In: Marcelino Kongo (ed) Lactic acid bacteria - R & D for food, health and livestock purposes. Biochemistry, genetics and molecular biology. https://doi.org/10.5772/47766

Fu Y, Young JF, Lokke MM, Lametsch R, Aluko RE, Therkildsen M (2016) Revalorisation of bovine collagen as a potential precursor of angiotensin 1-converting enzyme (ACE) inhibitory peptides based on in silico and in vitro protein digestions. J Funct Foods 24:196–206

Fu Y, Young JF, Therkildsen M (2017) Bioactive peptides in beef: endogenous generation through postmortem aging. Meat Sci 123:134–142

Gallego M, Mora L, Fraser PD, Aristoy MC, Toldra F (2014) Degradation of LIM domain-binding protein three during processing of Spanish dry-cured ham. Food Chem 149:121–128

Gallego M, Mora L, Aristoy MC, Toldra F (2015a) Evidence of peptide oxidation from major myofibrillar proteins in dry-cured ham. Food Chem 187:230–235

Gallego M, Mora L, Aristoy MC, Toldra F (2015b) Titin-derived peptides as processing time markers in dry-cured ham. Food Chem 167:326–339

Gallego M, Mora L, Toldra F (2016) Peptidomics as a tool for quality control in dry-cured ham processing. J Proteomics 147:98–107

Hernandez P, Muller M, Appel RD (2006) Automated protein identification by tandem mass spectrometry: issues and strategies. Mass Spectrom Rev 25:235–254

Holland RD, Wilkes JG, Rafii F, Sutherland JB, Persons CC, Voorhees KJ, Lay JO Jr (1996) Rapid identification of intact whole bacteria on spectral patterns using matrix assisted laser desorption/ionization with time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 10:1227–1232

Holzapfel WH, Haberer P, Geisen R, Björkroth J, Schillinger U (2001) Taxonomy and important features of probiotic microorganisms in food nutrition. Am J Clin Nutr 73:365S–373S

Horst T, Niedermeyer J (2016) Annotating and interpreting linear and cyclic peptide tandem mass spectra. In: Evans BS (ed) Nonribosomal peptide and polyketide biosynthesis: methods and protocols, methods in molecular biology. Springer, New York, pp 199–207

Hsueh P-R, Lee T-F, Du S-H, Teng SH, Liao CH, Sheng WH, Teng LJ (2014) Bruker biotypermatrix-assisted laser desorption ionization-time of flightmass spectrometry system for identification of *Nocardia*, *Rhodococcus*, *Kocuria*, *Gordonia*, *Tsukamurella*, and *Listeria* Species. J Clin Microbiol 52:2371–2379

Hughes MC, O'Neill EE, McSweeney PLH, Healy A (1999) Proteolysis of bovine F-actin by cathepsin B. Food Chem 64:525–530

Hughes MC, Healy A, McSweeney PLH, O'Neill EE (2000) Proteolytic specificity of cathepsin D on bovine F-actin. Meat Sci 56:165–172

Hughes MC, Geary S, Dransfield E, McSweeney PLH, O'Neill EE (2001) Characterization of peptides released from rabbit skeletal muscle troponin-T by mu-calpain under conditions of low temperature and high ionic strength. Meat Sci 59:61–69

Hussein HS, Bollinger LM (2005) Prevalence of Shiga toxin-producing Escherichia coli in beef. Meat Sci 71:676–689. https://doi.org/10.1016/j.meatsci.2005.05.012

Igbinosa EO, Beshiru A, Akporehe LU, Oviasogie FE, Igbinosa OO (2016) Prevalence of methicillin-resistant *Staphylococcus aureus* and other *Staphylococcus* species in raw meat samples intended for human consumption in Benin city, Nigeria: implications for public health. Int J Environ Res Public Health 13(10):949. https://doi.org/10.3390/ijerph13100949

Jadhav S, Sevior D, Bhave M, Palombo E (2014) Detection of *Listeria monocytogenes* from selective enrichment broth using MALDI-TOF mass spectrometry. J Proteomics 97 (31):100–106

Jadhav S, Gulati V, Fox E, Karpe A, Bealec D, Sevior D, Bhave M, Palombo E (2015) Rapid identification and source-tracking of *Listeria monocytogenes* using MALDI-TOF mass spectrometry. Int J Food Microbiol 202:1–9

Jay JM (2000) Modern food microbiology. Chapman and Hall, London

Jones RJ (1999) Immunisation against lactic-acid bacteria as a technique to extend the chilled storage life of vacuum-packed lamb. Food Agric Immunol 11:75–81

Jones AW, Cooper HJ (2011) Dissociation techniques in mass spectrometry-based proteomics. Analyst 136:3419–3429

Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 60:2299–2301

Karmi M (2013) Prevalence of methicillin-resistant *Staphylococcus aureus* in poultry meat in Qena, Egypt. Vet World 6:711–715

Kemp CM, Parr T (2008) The effect of recombinant caspase 3 on myofibrillar proteins in porcine skeletal muscle. Animal 2:1254–1264

Kitamura SI, Muroya S, Tanabe S, Okumura T, Chikuni K, Nishimura T (2005) Mechanism of production of troponin T fragments during postmortem aging of porcine muscle. J Agric Food Chem 53:4178–4181

Koutsoumanis K, Sofos J (2004) Microbial contamination of carcasses and cuts. In: Jensens W (ed) Encyclopedia of meat sciences. Elsevier Academic Press, Amsterdam, pp 727–737

Lafarga T, Hayes M (2014) Bioactive peptides from meat muscle and by-products: generation, functionality and application as functional ingredients. Meat Sci 98:227–239

Lafarga T, Wilm M, Wynne K, Hayes M (2016) Bioactive hydrolysates from bovine blood globulins: generation, characterisation, and in silico prediction of toxicity and allergenicity. J Funct Foods 24:142–155

Lametsch R, Bendixen E (2001) Proteome analysis applied to meat science: characterizing post mortem changes in porcine muscle. J Agric Food Chem 49:4531–4537

Lametsch R, Roepstorff P, Bendixen E (2002) Identification of protein degradation during post-mortem storage of pig meat. J Agric Food Chem 50:5508–5512

Lametsch R, Karlsson A, Rosenvold K, Andersen HJ, Roepstorff P, Bendixen E (2003) Postmortem proteome changes of porcine muscle related to tenderness. J Agric Food Chem 51:6992–6997

Lametsch R, Roepstorff P, Moller HS, Bendixen E (2004) Identification of myofibrillar substrates for mu-calpain. Meat Sci 68:515–521

Lartigue M-F, Héry-Arnaud G, Haguenoer E, Domelier A-S, Schmit P-O, van der Mee-Marquet N, Lanotte P, Mereghetti L, Kostrzewa M, Quentin R (2009) Identification of *Streptococcus agalactiae* isolates from various phylogenetic lineages by matrix-assisted laser desorption ionization-time of flight mass spectrometry. J Clin Microbiol 47:2284–2287

Laville E, Sayd T, Morzel M, Blinet S, Chambon C, Lepetit J, Renand G, Hocquette JF (2009) Proteome changes during meat aging in tough and tender beef suggest the importance of

apoptosis and protein solubility for beef aging and tenderization. J Agric Food Chem 57:10755–10764

Liu X, Chen Y, Fan Y, Wang M (2006) Foodborne diseases occurred in 2003 - report of the national foodborne diseases surveillance system, China. Chin J Hyg Res 35:201–204

López CM, Sentandreu MA, Vignolo GM, Fadda SG (2015a) Proteomic and peptidomic insights on myofibrillar protein hydrolysis in a sausage model during fermentation with autochthonous starter cultures. Food Res Int 78:41–49

López CM, Bru E, Vignolo GM, Fadda SG (2015b) Identification of small peptides arising from hydrolysis of meat proteins in dry fermented sausages. Meat Sci 104:20–29

López CM, Sentandreu MA, Vignolo GM, Fadda SG (2015c) Low molecular weight peptides derived from sarcoplasmic proteins produced by an autochthonous starter culture in a beaker sausage model. EuPA Open Proteomics 7:54–63

Luo J, Liu X, Tian Q, Yue W, Zeng J, Chen G, Cai X (2009) Disposable bioluminescence-based biosensor for detection of bacterial count in food. Anal Biochem 394:1–6. https://doi.org/10.1016/j.ab.2009.05.021.

Mahlapuu M, Hakansson J, Ringstad L, Bjorn C (2016) Antimicrobial peptides: an emerging category of therapeutic agents. Front Cell Infect Microbiol 6:194

Marbaix H, Budinger D, Dieu M, Fumiere O, Gillard N, Delahaut P, Mauro S, Raes M (2016) Identification of proteins and peptide biomarkers for detecting banned Processed Animal Proteins (PAPs) in meat and bone meal by mass spectrometry. J Agric Food Chem 64:2405–2414

McEvoy, J. M., A. M. Doherty, M. Finnerty, J. J. Sheridan, L., McGuire, I. S. Blair, D. A. McDowell, and D. Harrington. (2000) The relationship between hide cleanliness and bacterial numbers on beef carcasses at a commercial abattoir. Lett Appl Microbiol 30, 390–395.

McIntosh JA, Donia MS, Schmidt EW (2009) Ribosomal peptide natural products: bridging the ribosomal and nonribosomal worlds. Nat Prod Rep 26:537–559

McMillin K (2008) Where is MAP going? A review and future potential of modified atmosphere packaging for meat. Meat Sci 80:43–65

Minamino N (2001) Peptidome: the fact-database for endogenous peptides. Tanpakushitsu kakusan koso. Protein Nucleic Acid Enzyme 46:1510–1517

Mirdhayati I, Hermanianto J, Wijaya CH, Sajuthi D, Arihara K (2016) Angiotensin converting enzyme (ACE) inhibitory and antihypertensive activities of protein hydrolysate from meat of Kacang goat (Capra aegagrus hircus). J Sci Food Agric 96:3536–3542

Mora L, Sentandreu MA, Fraser PD, Toldra F, Bramley PM (2009a) Oligopeptides arising from the degradation of creatine kinase in Spanish dry-cured ham. J Agric Food Chem 57:8982–8988

Mora L, Sentandreu MA, Koistinen KM, Fraser PD, Toldra F, Bramley PM (2009b) Naturally generated small peptides derived from myofibrillar proteins in serrano dry-cured ham. J Agric Food Chem 57:3228–3234

Mora L, Sentandreu MA, Toldra F (2010) Identification of small troponin T peptides generated in dry-cured ham. Food Chem 123:691–697

Mora L, Sentandreu MA, Toldra F (2011a) Intense degradation of myosin light chain isoforms in Spanish dry-cured ham. J Agric Food Chem 59:3884–3892

Mora L, Valero ML, del Pino MMS, Sentandreu MA, Toldra F (2011b) Small peptides released from muscle glycolytic enzymes during dry-cured ham processing. J Proteomics 74:442–450

Mora L, Escudero E, Toldra F, Aristoy M-C (2015a) A peptidomic approach to study the contribution of added casein proteins to the peptide profile in Spanish dry-fermented sausages. Int J Food Microbiol 212:41–48

Mora L, Gallego M, Aristoy MC, Fraser PD, Toldra F (2015b) Peptides naturally generated from ubiquitin-60S ribosomal protein as potential biomarkers of dry-cured ham processing time. Food Control 48:102–107

Mora L, Gallego M, Escudero E, Reig M, Toldra F, Aristoy M-C (2015c) Small peptides hydrolysis in dry-cured meats. Int J Food Microbiol 212:9–15

Mora L, Calvo L, Escudero E, Toldra F (2016) Differences in pig genotypes influence the generation of peptides in dry-cured ham processing. Food Res Int 86:74–82

Morzel M, Chambon C, Hamelin M, Sante-Lhoutellier V, Sayd T, Monin G (2004) Proteome changes during pork meat ageing following use of two different pre-slaughter handling procedures. Meat Sci 67:689–696

Morzel M, Terlouw C, Chambon C, Micol D, Picard B (2008) Muscle proteome and meat eating qualities of Longissimus thoracis of "Blonde d'Aquitaine" young bulls: a central role of HSP27 isoforms. Meat Sci 78:297–304

Muroya S, Kitamura S, Tanabe S, Nishimura T, Nakajima I, Chikuni K (2004) N-terminal amino acid sequences of troponin T fragments, including 30 kDa one, produced during postmortem aging of bovine longissimus muscle. Meat Sci 67:19–24

Muroya S, Ohnishi-Kameyama M, Oe M, Nakajima I, Chikuni K (2007) Postmortem changes in bovine troponin T isoforms on two-dimensional electrophoretic gel analyzed using mass spectrometry and western blotting: the limited fragmentation into basic polypeptides. Meat Sci 75:506–514

Nakai Y, Nishimura T, Shimizu M, Arai S (1995) Effects of freezing on the proteolysis of beef during storage at 4 degrees C. Biosci Biotech Bioch 59:2255–2258

Nguyen DT, Van Hoorde K, Cnockaert M, De Brandt E, Aerts M, Le Thanh B, Vandamme P (2012) Validation of MALDI-TOF MS for rapid classification and identification of lactic acid bacteria, with a focus on isolates from traditional fermented foods in Northern Vietnam. Lett Appl Microbiol 55:265–273

Nieminen TT, Vihavainen E, Paloranta A, Lehto J, Paulin L, Auvinen P, Solismaa M, Björkroth KJ (2011) Characterization of psychrotrophic bacterial communities in modified atmosphere-packed meat with terminal restriction fragment length polymorphism. Int J Food Microbiol 144:360–366

Nomura F (2015) Proteome-based bacterial identification using matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS): a revolutionary shift in clinical diagnostic microbiology. Biochim Biophys Acta 1854:528–537

Nørrung B, Andersen JK, Buncic S (2009) Main concerns of pathogenic microorganisms in meat. In: Todra F (ed) Safety of meat and processed meat. Springer, New York, pp 3–29

Nychas G-J, Skandamis P, Tassou C, Koutsoumanis K (2008) Meat spoilage during distribution. Meat Sci 78:77–89

Ochoa ML, Harrington PB (2005) Immunomagnetic isolation of enterohemorrhagic *Escherichia coli* O157:H7 from ground beef and identification by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and database searches. Anal Chem 77:5258–5267

Okumura T, Yamada R, Nishimura T (2003) Survey of conditioning indicators for pork loins: changes in myofibrils, proteins and peptides during postmortem conditioning of vacuum-packed pork loins for 30 days. Meat Sci 64:467–473

Okumura T, Yamada R, Nishimura T (2004) Sourness-suppressing peptides in cooked pork loins. Biosci Biotech Bioch 68:1657–1662

Ortea I, O'Connor G, Maquet A (2016) Review on proteomics for food authentication. J Proteomics 147:212–225

Ouali A, Herrera-Mendez CH, Coulis G, Becila S, Boudjellal A, Aubry L, Sentandreu MA (2006) Revisiting the conversion of muscle into meat and the underlying mechanisms. Meat Sci 74:44–58

Ouali A, Gagaoua M, Boudida Y, Becila S, Boudjellal A, Herrera-Mendez CH, Sentandreu MA (2013) Biomarkers of meat tenderness: present knowledge and perspectives in regards to our current understanding of the mechanisms involved. Meat Sci 95:854–870

Panchaud A, Affolter M, Kussmann M (2012) Mass spectrometry for nutritional peptidomics: how to analyze food bioactives and their health effects. J Proteomics 75:3546–3559

Paolella S, Falavigna C, Faccini A, Virgili R, Sforza S, Dall'Asta C, Dossena A, Galaverna G (2015) Effect of dry-cured ham maturation time on simulated gastrointestinal digestion: characterization of the released peptide fraction. Food Res Int 67:136–144

Park BY, Kim NK, Lee CS, Hwang IH (2007) Effect of fiber type on postmortem proteolysis in longissimus muscle of landrace and Korean native black pigs. Meat Sci 77:482–491

Pavlovic M, Huber I, Konrad R, Busch U (2013) Application of MALDI-TOF MS for the identification of food borne bacteria. Open Microbiol J 7:135–141

Pérez RH, Zendo T, Sonomoto K (2014) Novel bacteriocins from lactic acid bacteria (LAB): various structures and applications. Microb Cell Factories 13(Suppl 1):S3. http://www.microbialcellfactories.com/content/13/S1/S3

Picariello G, De Martino A, Mamone G, Ferranti P, Addeo F, Faccia M, SpagnaMusso S, Di Luccia A (2006) Proteomic study of muscle sarcoplasmic proteins using AUT-PAGE/SDS-PAGE as two-dimensional gel electrophoresis. J Chromatogr B 833:101–108

Picariello G, Mamone G, Addeo F, Ferranti P (2012) Novel mass spectrometry-based applications of the 'omic' sciences in food technology and biotechnology. Food Technol Biotech 50:286–305

Polati R, Menini M, Robotti E, Millioni R, Marengo E, Novelli E, Balzan S, Cecconi D (2012) Proteomic changes involved in tenderization of bovine Longissimus dorsi muscle during prolonged ageing. Food Chem 135:2052–2069

Quintela-Qunitela-Baluja M, Böhme K, Fernández-No IC, Morandi S, Alnakip ME, Caamaño-Antelo S, Velázquez JB, Calo-Mata P (2013) Characterization of different food-isolated *Enterococcus* strains by MALDI-TOFmass fingerprinting. Electrophoresis 34:2240–2250

Rodriguez-Frometa RA, Rodriguez-Diaz J, Sentandreu E, Ouali A, Sentandreu MA (2013) Proteolytic action of caspases 3 and 7 on the hydrolysis of bovine and porcine muscle myofibrillar proteins. In: de Almeida A, Eckersall D, Bencurova E, Dolinska S, Mlynarcik P, Vincova M, Bhide M (eds) Farm animal proteomics 2013. Wageningen Academic Publishers, Wageningen, pp 278–281

Roncada P, Piras C, Soggiu A, Turk R, Urbani A, Bonizzi L (2012) Farm animal milk proteomics. J Proteomics 75:4259–4274

Ruiz-Moyano S, Nannan T, Underwood MA, Millls DA (2012) Rapid discrimination of *Bifidobacterium animalis* subspecies by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Food Microbiol 30:432–437

Russo F, Ercolini D, Mauriello G, Villani F (2006) Behaviour of Brochothrix thermosphacta in presence of other meat spoilage microbial groups. Food Microbiol 23:797–802

Ryzhov V, Fenselau C (2001) Characterization of the protein subset desorbed by MALDI from whole bacterial cells. Anal Chem 73:746–750

Sanchez-Rivera L, Martinez-Maqueda D, Cruz-Huerta E, Miralles B, Recio I (2014) Peptidomics for discovery, bioavailability and monitoring of dairy bioactive peptides. Food Res Int 63:170–181

Sawdy JC, Kaiser SA, St-Pierre NR, Wick MP (2004) Myofibrillar 1-D fingerprints and myosin heavy chain MS analyses of beef loin at 36 h postmortem correlate with tenderness at 7 days. Meat Sci 67:421–426

Schrader M, Schulz-Knappe P, Fricker LD (2014) Historical perspective of peptidomics. EuPA Open Proteomics 3:171–182

Schulz-Knappe P, Zucht HD, Heine C, Jurgens M, Hess R, Schrader M (2001) Peptidomics: the comprehensive analysis of peptides in complex biological mixtures. Comb Chem High Throughput Screen 4:207–217

Sentandreu MA, Sentandreu E (2011) Peptide biomarkers as a way to determine meat authenticity. Meat Sci 89:280–285

Sentandreu MA, Sentandreu E (2014) Authenticity of meat products: tools against fraud. Food Res Int 60:19–29

Sentandreu MA, Coulis G, Ouali A (2002) Role of muscle endopeptidases and their inhibitors in meat tenderness. Trends Food Sci Tech 13:400–421

Sentandreu MA, Stoeva S, Aristoy MC, Laib K, Voelter W, Toldra F (2003) Identification of small peptides generated in Spanish dry-cured ham. J Food Sci 68:64–69

Sentandreu MA, Armenteros M, Calvete JJ, Ouali A, Aristoy MC, Toldra F (2007) Proteomic identification of actin-derived oligopeptides in dry-cured ham. J Agric Food Chem 55:3613–3619

Sentandreu MA, Fraser PD, Halket J, Patel R, Bramley PM (2010) A proteomic-based approach for detection of chicken in meat mixes. J Proteome Res 9:3374–3383

Sforza S, Pigazzani A, Motti M, Porta C, Virgili R, Galaverna G, Dossena A, Marchelli R (2001) Oligopeptides and free amino acids in Parma hams of known cathepsin B activity. Food Chem 75:267–273

Sforza S, Boni M, Ruozi R, Virgili R, Marchelli R (2003) Identification and significance of the N-terminal part of swine pyruvate kinase in aged Parma hams. Meat Sci 63:57–61

Sharma KP, Chattopadhyay UK (2015) Assessment of microbial load of raw meat samples sold in the open markets of city of Kolkata. J Agric Vet Sci 8:24–27

Siegrist TJ, Anderson PD, Huen WH, Kleinheinz GT, McDermott CM, Sandrin TR (2007) Discrimination and characterization of environmental strains of *Escherichia coli* by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). J Microbiol Methods 68:554–562

Sierra V, Fernandez-Suarez V, Castro P, Osoro K, Vega-Naredo I, Garcia-Macia M, Rodriguez-Colunga P, Coto-Montes A, Olivan M (2012) Identification of biomarkers of meat tenderisation and its use for early classification of Asturian beef into fast and late tenderising meat. J Sci Food Agric 92:2727–2740

Skrlep M, Candek-Potokar M, Mandelc S, Javornik B, Gou P, Chambon C, Sante-Lhoutellier V (2011) Proteomic profile of dry-cured ham relative to PRKAG3 or CAST genotype, level of salt and pastiness. Meat Sci 88:657–667

Stadnik J, Keska P (2015) Meat and fermented meat products as a source of bioactive peptides. Acta Sci Pol Technol Aliment 14:181–190

Stoeva S, Byrne CE, Mullen AM, Troy DJ, Voelter W (2000) Isolation and identification of proteolytic fragments from TCA soluble extracts of bovine M-longissimus dorsi. Food Chem 69:365–370

Teramoto K, Sato H, Sun L, Torimura M, Tao H, Yoshikawa H, Hotta Y, Hosoda A, Tamura H (2007) Phylogenetic classification of *Pseudomonas putida* strains by MALDI-MS using ribosomal subunit proteins as biomarkers. Anal Chem 79:8712–8719

Theron L, Sayd T, Pinguet J, Chambon C, Robert N, Sante-Lhoutellier V (2011) Proteomic analysis of semimembranosus and biceps femoris muscles from Bayonne dry-cured ham. Meat Sci 88:82–90

Udenigwe CC, Howard A (2013) Meat proteome as source of functional biopeptides. Food Res Int 54:1021–1032

Vaillant V, De Valk H, Baron E, Ancelle T, Colin P, Delmas MC, Dufour B, Pouillot R, Le Strat Y, Weinbreck P, Jougla E, Desenclos JC (2005) Foodborne infections in France. Foodborne Pathog Dis 2:221–232

van Veen SQ, Claas EC, Kuijper EJ (2010) High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. J Clin Microbiol 48:900–907

Venema K, do Carmo AP (eds) (2015) Probiotic and prebiotics: current status and future trends. Caister Academic Press, Norfolk, pp 1–3

Verhaert P, Uttenweiler-Joseph S, de Vries M, Loboda A, Ens W, Standing KG (2001) Matrix-assisted laser desorption/ionization quadrupole time-of-flight mass spectrometry: an elegant tool for peptidomics. Proteomics 1:118–131

Vignolo, G., Castellano, P., and Fadda S. (2015). Part III, Chapter 15: Starter cultures: bioprotective cultures. In: Toldrá F, Talon R (eds) Handbook of fermented meat and poultry, 2nd edn. Blackwell, Malden, MA, pp 129–137. isbn: 978-1-118-52269-1

Voelter W, Stoeva S, Echner H, Beck A, Schutz J, Lehmann R, Haring HU, Schleicher E, Mullen AM, Casserly U et al (2000) Analytical tools for rapid, sensitive, quantitative identification of potential meat quality markers. J Prakt Chem Chem Ztg 342:179–191

Wang YR, Chen Q, Cui SH, Li FQ (2013) Characterization of *Staphylococcus aureus* isolated from clinical specimens by matrix assisted laser desorption/ionization time-of-flight mass spectrometry. Biomed Environ Sci 26:430–436

Wen SY, Zhou GH, Song SX, Xu XL, Voglmeir J, Liu L, Zhao F, Li MJ, Li L, Yu XB et al (2015) Discrimination of in vitro and in vivo digestion products of meat proteins from pork, beef, chicken, and fish. Proteomics 15:3688–3698

Wieser A, Schneider L, Jung J, Schubert S (2012) MALDI-TOF MS in microbiological diagnostics-identification of microorganisms and beyond (mini review). Appl Microbiol Biotechnol 93:965–974. https://doi.org/10.1007/s00253-011-3783-4

Williamson YM, Moura H, Woolfitt AR, Pirkle JL, Barr JR, Carvalho MDG, Ades ED, Carlone GM, Sampson JS (2008) Differentiation of *Streptococcus pneumonia* conjunctivitis outbreak isolates by matrix-assisted laser desorption ionization-time of flight mass spectrometry. Appl Environ Microbiol 74:5891–5897

Zhao C, Ge B, De Villena J, Sudler R, Yeh E, Zhao S, White DG, Wagner D, Meng J (2001) Prevalence of *Campylobacter* spp., *Escherichia coli*, and *Salmonella* serovars in retail chicken, turkey, pork, and beef from the greater Washington, D.C. area. Appl Environ Microbiol 67:5431–5436

Zürbig P, Mischak H (2008) Peptidomics approach to proteomics. In: Soloviev M, Andrén P, Shaw C (eds) Peptidomics methods and applications. Wiley, Hoboken, NJ, pp 155–175

Studying the Animal Transcriptome: State of the Art and Challenges in the Context of Animal and Veterinary Sciences

José Ricardo Parreira and Susana de Sousa Araújo

Abstract Knowledge in animal physiology has significantly advanced due to transcriptomics, a complementary tool to proteomics, and this is particularly relevant for farm animals. Transcriptomics aims to study the transcriptome, the complete set of coding (mRNAs) and noncoding (e.g., small RNAs) transcripts encoded by the genome in a specific spatiotemporal context. Various technologies, including hybridization and sequencing-based approaches, have been developed to infer and quantify the transcriptome changes.

Transcriptomics has become an option for many research studies in the main farmed animals, herein limited to bovines, pig, chicken, sheep, goat, and salmon. So far, the transcriptomic studies performed in these species have disclosed potential candidate genes associated with muscle growth, meat quality, lactation, reproduction efficiency, or response to diseases. Coupled to this, transcriptional and post-transcriptional regulatory mechanisms controlling the expression of these genes have been uncovered.

This chapter focuses on the recent contributions that transcriptomics has brought to improve our knowledge in farmed animal physiology. The current limitations associated with the application of this methodology, as well as the possible implications of using transcriptome data to develop new strategies to improve animal health, welfare, and production, are also discussed.

Keywords Farmed animals • Gene expression • mRNA • miRNA • Transcriptomics

---

J.R. Parreira
Plant Cell Biotechnology Laboratory, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal

S. de Sousa Araújo (✉)
Plant Cell Biotechnology Laboratory, Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal

Plant Biotechnology Laboratory, Department of Biology and Biotechnology 'L. Spallanzani', Università degli Studi di Pavia, Pavia, Italy
e-mail: saraujo@itqb.unl.pt

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_20

1 Introduction

Transcriptomics is a broad term that refers to the study of the transcriptome, the complete set of coding (mRNAs) and noncoding (e.g., small RNAs) transcripts encoded by the genome, in a specific spatiotemporal context. Broadly, the main objective of a transcriptomic study is to identify the differentially expressed transcripts between different/contrasting experimental conditions, by means of bioinformatics methodology. The main goals of a transcriptome analysis are to catalogue all species of transcripts; to determine the transcriptional structure of genes, in terms of their start sites, 5′ and 3′ ends, splicing patterns, and other posttranscriptional modifications; and to quantify the expression levels (abundance) of each transcript during development or under distinct conditions (Wang et al. 2009).

Transcriptome-wide technologies allowed significant advances on the knowledge of molecular mechanisms that underlie animal responses toward production and health constraints. Several reviews have been published about the major outcomes obtained from transcriptomic studies for the most important farmed animals (Shahzad and Loor 2012; Loor et al. 2013, 2015; Hekman et al. 2015; Schroyen and Tuggle 2015; Sullivan et al. 2015).

Besides providing a snapshot of the different metabolic pathways activated in target tissues, post-genomic studies such as transcriptomics have contributed to the identification of the so-called biomarkers. Biomarkers are indicators of biological processes and pathological states that can reveal a variety of health and disease traits (Biomarkers Definitions Working Group 2001). Biomarkers are particularly relevant in animal and veterinary research where they have an important role in the characterization of animal health and diseases (Moore et al. 2007) or the study of stress and welfare in animal production (Marco-Ramell et al. 2016).

As said previously, whole transcriptome characterization or differential gene expression studies using RNA-Seq or microarrays have become an option for many experimental approaches in animal and veterinary sciences. By browsing PubMed in the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/pubmed), we retrieved 644 research articles (15th December 2016), in which transcriptomic studies were conducted in the main farmed animals, herein limited to bovines, pig, chicken, sheep, goat, and salmon. The results of this non-exhaustive search are depicted in Fig. 1.

More than 38% of the publications (249 out of 644) bring advances in our knowledge of bovine physiology. This highlights the relevance of this species in an animal production context and their impact as source of meat and dairy products for humans. Chicken and pig research are ranked second and third, respectively, which also reflects their overall relevance in farming and meat-producing systems. The cumulative amount of transcriptomic studies published in sheep, goat, and salmon transcriptomics does not exceed 132 research manuscripts.

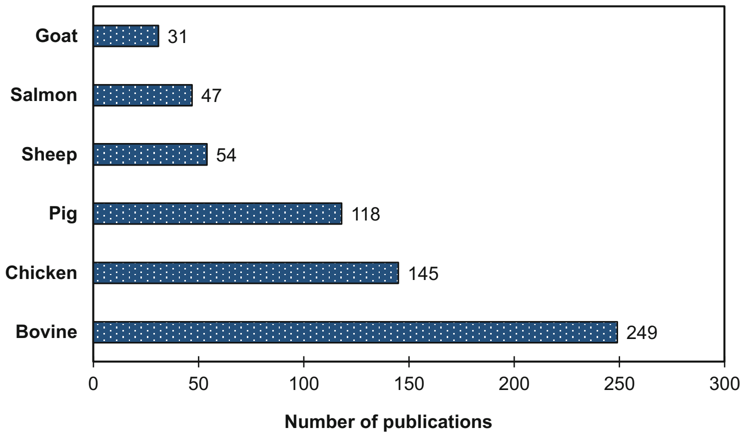

Fig. 1 Query of research scientific papers on farmed animals found at NCBI between 2006 and 2016. The search was performed individually for each animal (bovine, chicken, pig, goat, sheep, and salmon) using the following setting for an advanced search: [*transcriptome* (Title/Abstract) AND *animal* (Title/Abstract)] NOT *review* [Publication Type]. Review manuscripts were not included. The search was conducted on the 15th of December 2016

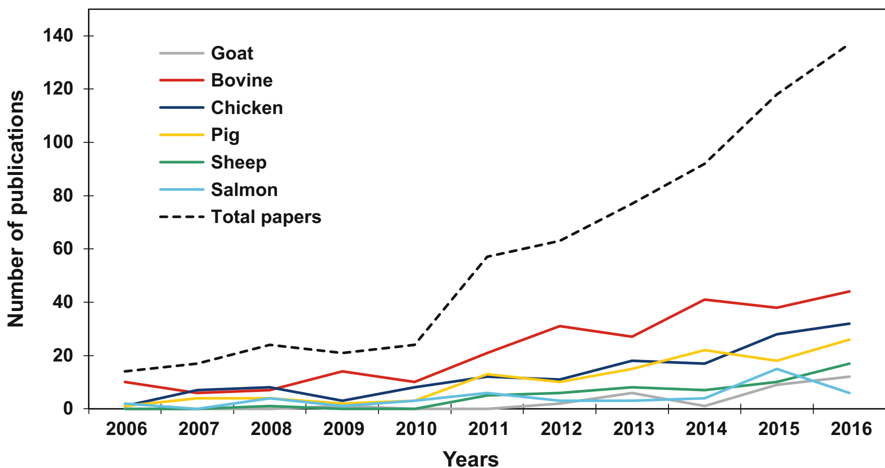

Fig. 2 Total number of research scientific papers published per year on farmed animal transcriptomics found in NCBI between 2006 and 2016. The NCBI query was conducted as described in Fig. 1

Another interesting aspect is the evolution in the number of transcriptome publications seen in the last decade which is depicted in Fig. 2, which also reflects the results described in Fig. 1.

One of the most interesting features is that the number of publications on farmed animal transcriptomics increased almost six fold since 2010 (Fig. 2). This was

probably boosted by the development of next-generation sequencing technologies, the general decrease in the costs to perform high-throughput sequencing, and the concomitant release of several animal genomes. Interestingly, in 2015 the transcriptomic research outputs publicly available for salmon (15 manuscripts) approached the number of pig research outputs (18 manuscripts). Such results highlight the growing value of farming fish systems, also known as aquaculture, to supply the growing demand of healthier animal protein for human consumption. Based on the overall analysis of Fig. 2, it is tempting to claim that the number of transcriptome publications is expected to increase in forthcoming years.

This book chapter focuses on the contributions that transcriptomics, which has emerged as complementary tool to proteomics, has brought to improve our knowledge in farm animal physiology. The possible implications of using transcriptome data to develop new strategies to improve animal health, welfare, and production are also discussed.

2 Studying the Animal Transcriptome Reprogramming: Which Tools?

Various technologies, including hybridization and sequencing-based approaches, have been developed to infer and quantify the transcriptome. In the 1990s, hybridization-based microarray technologies allowed, for the first time, the simultaneous measurement of the abundance of thousands of transcripts in different cells or tissues (Guo et al. 1994; Schena et al. 1995; Schena 1996; Duggan et al. 1999). The development of next-generation sequencing technology in the middle 2000s has revolutionized the way eukaryotic transcriptomes are analyzed (Wang et al. 2009; Metzker 2010). With the development of RNA-sequencing (RNA-Seq) technologies, gene expression can be analyzed on a global scale, thereby providing numerous advantages over microarray analyses, such as the ability to quantify all genes present and not only those targets found on the array (Wickramasinghe et al. 2014).

2.1 Microarray Technology

Microarray analysis is based on the affinity of single-stranded DNA sequences (targets) to bind to complementary sequences of oligonucleotides or cDNA molecules (probes) deposited in an array. An array or DNA chip is an orderly arrangement of DNA molecules that have been chemically bonded to physical surfaces, such as a glass slide also known as chips (Schena et al. 1995). Typically, an array contains thousands of probe sequences allowing a genome-wide assessment of gene expression. The labeled cRNAs or cDNAs are then hybridized to the microarray,

the array is washed, and the signal is detected by measuring fluorescence at each spot (Bumgarner 2013). Signal intensity values from hybridization of transcripts with the probes are then used to calculate the expression level of transcripts. A graphical representation of the microarray hybridization workflow can be visualized in Schena (1996). A considerable number of array and microarray platforms have been developed reflecting differences in terms of the number of labeled samples to be used (single color, one sample; two colors, two samples) in a single chip [for a comparison of platforms, please refer to Woo et al. (2004)]. Other differences are, for instance, the type of probe type placed onto the arrays (oligonucleotides or cDNA) or the process by which probes are created and affixed to the slide (Bumgarner 2013).

Despite the development of technologies that allow the analysis of RNA through cDNA sequencing (RNA-Seq), transcriptome profiling by microarrays is still being widely used to understand specific questions in animal sciences. Nowadays, commercial microarrays from Affymetrix© and Agilent© are available to study transcriptome changes in livestock (e.g., bovine, chicken, rabbit, ovine, among others) and farmed fish (salmon). As an example, microarrays have been recently used to provide new insights into cell-specific reprogramming and differentiation processes that transforms somatic follicular cells into luteal cells in bovines (Romereim et al. 2017) or to understand the molecular mechanisms underlying heart and skeletal muscle inflammation disease in farmed Atlantic salmon (Johansen et al. 2016). Other examples, in which microarrays were used to address main scientific questions in farmed animals, will be provided in the following sections.

Although microarrays have been widely useful in these applications, a considerable number of limitations associated with this technique are well described. The main limitation is undoubtedly associated with the technology: it is only possible to detect sequences that the array was designed to detect (Bumgarner 2013). Background levels of hybridization (i.e., hybridization to a probe that occurs irrespective of the corresponding transcript's expression level) limit the accuracy of expression measurements, particularly for transcripts with low abundance (Marioni et al. 2008). There is also a risk for DNA sequences with a high level of homology (e.g., gene families or multiple splice variants) to bind the same probe on the array, especially if probes are not targeting exons (Gardina et al. 2006).

2.2 RNA Sequencing

High-throughput next-generation sequencing (NGS) of DNA, also known as deep sequencing, has been revolutionizing biomedical research in recent years. The introduction of RNA-Seq technologies has provided a more complete characterization and quantification of RNA transcripts at a massive scale (Ozsolak and Milos 2011a). Such technology allows transcriptome analysis at a wider resolution and dynamic range, independently of the transcript size and without requiring a prior

knowledge of the genome sequence of the target organism (Marguerat and Bähler 2010). Indeed, the short-sequenced fragments produced (reads) are countable requiring, afterward, huge computing resources for de novo assembling and quantifying the number of reads. If a reference genome is available, reads are mapped back to reference genome, being the number of mapped reads a measure of expression level for that gene (Malone and Oliver 2011).

Most of the microarray drawbacks were largely overcome by RNA-Seq, namely, the issues of hybridization background or probe cross-hybridization, the restriction on transcript detection imposed by the array design, and the need for a sequenced genome (Wang et al. 2009; Costa et al. 2010). As RNA-Seq is quantitative, it can be used to determine RNA expression levels more accurately than microarrays, allowing the detection of novel and/or low-abundance transcripts and splice variations (Wang et al. 2016) as well as the identification of noncoding microRNAs (miRNAs) (Ioannidis and Donadeu 2016). Additionally, RNA-Seq is also well suited for detecting single nucleotide polymorphisms (SNPs), small insertions and deletions (INDELs) (Chitwood et al. 2013; Xing et al. 2016), and parental allele-specific differences in gene expression (Chen et al. 2016).

Although direct sequencing of RNA molecules is possible (Ozsolak et al. 2009; Ozsolak and Milos 2011b), commonly RNA-Seq experiments are carried out on instruments that sequence complementary DNA (cDNA) molecules; thus, cDNA library preparation from RNA is a required step. The cDNA library preparation method depends on the platform used for sequencing but also on the RNA species under investigation (e.g., mRNAs or small RNAs), which can differ in size, sequence, structural features, and abundance (for details see Hrdlickova et al. 2017).

The major commercially available NGS platforms used for RNA-Seq are the Roche 454 Genome Sequencer FLX+ System (Roche Applied, http://www.454.com), the Illumina/HiSeq© Genome Analyzer (Solexa, http://www.illumina.com), and the ABI SOLID system (Applied Biosystems, http://www.appliedbiosystems.com). In 2011, new NGS platforms such as the Ion Torrent's Personal Genome Machine, the Pacific Biosciences' RS, and the Illumina MiSeq have been released (Quail et al. 2012). Comprehensive reviews describing each platform, their applications advantages, and limitation may be found in Ansorge (2009), Wang et al. (2009), Metzker (2010), Ozsolak and Milos (2011a), Quail et al. (2012), and Hrdlickova et al. (2017).

An enormous number of bibliographic reports is now available, describing the use of such approaches in studies targeting a broad range of farmed species, including cattle (Berton et al. 2016), pig (Cardoso et al. 2017), sheep (Jäger et al. 2011), and salmon (Valenzuela-Muñoz et al. 2017). An exponential amount of sequencing data resulting from multiple studies using NGS technologies is available and can be mined in public sequence data repositories, such as the NCBI Sequence Read Archive (Wheeler et al. 2008).

3 Transcriptome Studies Are Driving Advances in Meat Production, Lactation, Nutrition, Reproduction, and Disease Responses

Transcriptome-based studies have emerged as a complementary tool to proteomics providing another layer of knowledge in farm animal physiology (D'Alessandro and Zolla 2013), namely, in the identification of regulatory pathways driving gene expression. As seen previously, a considerable number of research studies have uncovered new knowledge on molecular processes related with meat and milk production and their quality, as well as reproduction efficiency and responses to diseases. Whenever possible, the implications of using transcriptome data, namely, candidate genes, to develop new strategies to improve animal, health, welfare, and production will be presented. Herein we will provide some of these achievements by reviewing some very recent studies conducted in bovines, poultry, pigs, small ruminants, and farmed fishes.

3.1 Bovine Transcriptomic Studies

World bovine production has increased largely over the last 50 years (FAOSTAT 2017). According to FAOSTAT, the world production of bovine meat commodity reached the 67.5 million tonnes. Similarly, more than 653 million tonnes of whole fresh milk were produced. In this context, it became clear the relevance of cattle production to reach the demand for livestock products to sustain human population growth (Thornton 2010). Several markers for growth-associated traits in beef have been established, particularly in genetics- and transcriptomics-based studies (Almeida et al. 2017). In cattle, transcriptome-based studies have made significant contributions to understand muscle biochemical pathways affected by the production system (Guo et al. 2015; Tizioto et al. 2016) or to identify markers of meat quality (Zhao et al. 2012; Bongiorni et al. 2016). A considerable research effort was also applied to understand the molecular mechanisms controlling sexual reproduction and animal development (Vigneault et al. 2009; Cánovas et al. 2014; Baufeld et al. 2017; Capra et al. 2017; Scolari et al. 2017). Not less studied are the molecular mechanisms associated with the inflammatory changes occurring in the mammary gland infected with an invading microbial agent known as mastitis (Fang et al. 2016; Wang et al. 2016).

Differences between cattle production systems can influence the nutritional and sensory characteristics of beef, particularly its fatty acid composition (Sweeney et al. 2016). Meat products derived from pasture-based systems generally attract consumer's preferences, who are looking for meat with enhanced nutritional quality but are concerned with cattle health and welfare and the sustainability of the livestock production system. Pasture-produced meat has generally a higher commercial value. Consequently, authentication or certification tools that could attest

the type of production system in which the animals were produced are needed, since there are implications to the beef production industry, regulatory bodies, and ultimately the consumer (Almeida et al. 2017). In the scope of this thematic, Sweeney et al. (2016) conducted a microarray study to identify differentially expressed genes in the muscle of two groups of cattle raised in distinct production systems. The first group grazed outdoors and was pasture-fed, with a diet consisting predominantly of *Lolium perenne* L., *Poa* spp., and *Trifolium repens* L. The second group was permanently housed and fed on concentrate and barley straw. The authors also aimed to identify a panel of genes suitable to classify the meat based on the production system. The outcome revealed 26 genes differentially expressed between muscle samples coming from the two production systems, most of them related to fatty acid metabolism. The study also highlights that the expression profiles of delta-aminolevulinate dehydratase (*ALAD*), eukaryotic translation initiation factor 4E-binding protein 1 (*EIF4EBP1*), and nephronectin (*NPNT*) almost completely separate the samples based on production system. The authors propose that these genes could be used as potential transcriptomic biomarkers and/or molecular signatures to develop methods of authentication for meat products coming from contrasting production systems (Sweeney et al. 2016). Nevertheless, they also highlighted that the use of transcriptomic data in authenticating production systems is still preliminary and requires more exploration across a range of contexts and breeds.

Embryo survival is a major factor affecting production and economic efficiency in all systems of ruminant milk and meat production (Diskin and Morris 2008). In cattle, most of the pathways and factors from the oviduct implicated in the cross talk between this organ and gametes/early embryo remain unknown. Mounting evidences show that the oviduct environment plays a major role in sperm storage and capacitation, fertilization, and early embryo development in cattle (Maillo et al. 2015). Therefore, studying the oviduct environment is crucial to improve our understanding of the regulatory mechanisms controlling fertilization and embryo development (Maillo et al. 2016a). Gonella-Diaza et al. (2015) conducted a differential gene expression study using RNA-Seq to investigate the transcriptome changes occurring in ampulla and isthmus samples of *Bos taurus indicus* animals with large or small growth of the preovulatory follicle, as result of hormonal manipulations. Among the candidate genes identified, the results showed that chemokine (*C-X-C motif*) receptor 4 (*CXCR4*), heparanase (*HSPE*), derived growth factor D (*PDGFD*), and regulator of G protein signaling 20 (*RGS20*) were more expressed in animals with large preovulatory follicle growth and in the ampulla. MET proto-oncogene receptor tyrosine kinase (*C-MET*), platelet transforming growth factor beta 3 (*TGFB3*), and vinculin (*VCL*) were more expressed with large preovulatory follicle growth but in the isthmus. Endothelin 1 (*EDN1*) was more expressed in the small preovulatory follicle growth group and in the isthmus. Another study by Maillo et al. (2016b) characterized the transcriptome changes occurring in the isthmus in pregnant and cyclic heifers from a Charolais and Limousin cross. Additionally, focusing in pregnant heifers, the authors also investigated the tissue-specific responses occurring on the isthmus and ampulla regions

of the oviduct at a critical time-point (3 days after estrus/insemination) in which the embryo is exposed to these environments. Microarray analysis using the Bovine Gene ST 1.0 microarray chip (Affymetrix, USA) allowed the identification of 2287 genes significantly differentially expressed between the ampulla and isthmus of the oviduct ipsilateral to the *corpus luteum* in pregnant animals. Gene ontology analysis revealed that the main overrepresented biological processes in the isthmus were synthesis of nitrogen, lipids, nucleotides, steroids, and cholesterol as well as vesicle-mediated transport, cell cycle, apoptosis, endocytosis, and exocytosis (Maillo et al. 2016b). Contrastingly, cell motion, motility and migration, DNA repair, calcium ion homeostasis, carbohydrate biosynthesis, and regulation of cilium movement and beat frequency were overrepresented in the ampulla samples.

The abovementioned studies allowed the identification of candidate genes and metabolic pathways activated in the oviduct environment that may control fertilization and embryo development. Differences in the periovulatory sex steroid milieu are reflected on the different gene expression profiles in the oviduct. A tissue-specific molecular signature also described two distinct regions of the oviduct, which also reflects the morphological and functional differences. Such information is relevant to develop new strategies to improve fertilization, embryo survival, and development in animals throughout the manipulation of the environment in the oviduct.

In this context, it became clear that an enormous research effort has been devoted to identifying molecular signatures and metabolic pathways. They could be used to develop new strategies to improve animal production at multiple levels but also to identify markers associated with meat nutritional quality and value.

3.2 Poultry Transcriptomic Studies

Poultry is increasingly preferred in many regions as an affordable source of animal protein, which unlike pork or beef is accepted for consumption by most of the major religions in the world (Van Horne and Achterbosch 2008). Poultry meat production, as a commodity, presents approximately a 1.45-fold change in the world production during the 2003–2013 (FAOSTAT 2017). According to FAOSTAT, the world production of poultry meat commodity, which includes meat from chicken and other domestic birds, reached in 2013 the 108.5 million tonnes. Similarly, more than 73.8 million tonnes of eggs were also produced in 2013. Consequently, high production efficiency in poultry is essential to meet the demands of the growing human population.

Chicken production accomplishes two major end uses. While some birds are grown for meat production, also known as broilers, other breeds are used to produce eggs. Broodiness, a common habit of most domestic fowls, reduces the egg production and is a major restricting element to industrial-scale production. Although genetic, phenotypic, and physiological factors influencing broodiness in poultry have been extensively reviewed (Nestor et al. 1996; Romanov 2001;

Shimmura et al. 2015; Johnson et al. 2015), the information available about the molecular regulatory mechanisms is still scarce. Goose laying performance is closely related to the development of follicles and the establishment of the follicular hierarchy (Yu et al. 2016a). Recently, Yu et al. (2016b) explored the molecular mechanisms of broodiness occurring in geese. The authors conducted a transcriptome profiling of large yellow follicles and small yellow and large white follicles during egg-laying and broody stages in geese. Several hormone-related genes and autophagy- and redox-related transcription factors were retrieved from follicular transcriptomes, providing new insights in the metabolic pathways activated during egg-laying and broody stages in geese that could be extrapolated to other fowls.

Poultry breeding has been focused in the identification and production of breeds with improved meat production efficiency through rapid growth and high feed efficiency (Fairfull and Chambers 1984). Kong et al. (2017) investigated the molecular mechanisms regulating rapid muscle growth and high feed efficiency in a modern broiler chicken breed. A global gene expression analysis (RNA-Seq) on breast muscle tissue was conducted to compare the transcriptome profiles of modern pedigree male (PeM) broilers (rapid growth and muscle development) compared with the foundational BPR chickens (slow growth and lower efficiency). Among differently expressed genes detected, leiomodin 2 (*LMOD2*), carbonic anhydrase III (*CA3*), troponin T type 2 (*TNNT2*), and myosin-binding protein C1 (*MYBPC1*) were upregulated in PeM and described as mainly expressed in skeletal or cardiac muscles particularly in slow-twitch fibers (type 1). This suggests that breast muscle in PeM chickens can have higher slow fiber composition compared to BPR muscle. On the other side, the expression of myoglobin (*MB*) and Myozenin 2 (*MYOZ2*) were also highly upregulated (10.85 and 5.90-fold, respectively) in PeM compared to BPR. The expression of *MB*, *MYOZ2*, and *MyBPC1* has been previously associated to subclinical myopathy observed in the commercial broiler breast tissue that can lead to muscle hardening, named as "wooden breast" myopathy (Mutryn et al. 2015). This myopathy affects the pectoralis major muscle of the commercial broiler; thus, it has an increasing negative impact in poultry meat production (Sihvo et al. 2014), leading to remarkable economic losses, as well as a potential welfare problem of the birds. Importantly, this study also shows that a strong breeding effort for specific traits may trigger undesirable genetic side effects that need to be considered in future breeding programs.

Colibacillosis is an economically important syndromic disease affecting the immune system in poultry caused by extraintestinal avian pathogenic *Escherichia coli* (APEC) (Kemmett et al. 2013). Beside the associated production losses, poultry products contaminated with APEC are a potential source of foodborne extraintestinal pathogenic *Escherichia coli* (ExPEC) infection to humans, posing a threat to human health (Johnson et al. 2007). Sandford et al. (2011) conducted a study in chickens to identify host traits that are associated with colibacillosis resistance. The spleen tissue was studied, since it has been frequently used in avian studies, to infer immune system strength in birds (Smith and Hunt 2004). To better define effective mechanism of host resistance, a global and multivariate

gene expression profiling was conducted using the Agilent 4 × 44 Chicken microarray. In the experimental scheme, birds were divided into vaccinated (V) and non-vaccinated (NV), challenged (C) and non-challenged (NC), and sampled on 1 (D1) or 5 (D5) days post challenge, and the non-vaccinated challenged birds on both days were subdivided into mild (M) and severe (S) pathology. The results revealed a large difference in splenic transcriptome profiles between birds with mild and severe lesions in response to APEC infection (Sandford et al. 2011). For example, 1101 genes were significantly differentially expressed between severely infected and non-infected groups on D1 and 1723 on D5. Between birds exhibiting mild and severe pathology, there were 2 significantly differentially expressed genes on D1 and 799 on D5. Gene ontology (GO) analysis also revealed a time frame of metabolic events occurring during the progression of the infection. For example, the NV-C severe group on day 1 had many GO terms related to response to bacteria, inflammation, and circulatory processes, along with a few receptor signaling terms. Among the differentially expressed genes, avian beta-defensins and Toll-like receptors (*TLR*) were identified and linked to inflammation and circulatory processes potentially responsible for observed lesion phenotype. Later at day 5 and for the same experimental group, the GO terms found are related to regulation of white blood cells, localization, and transport. Importantly, the expression changes detected in the Janus kinase and two signal transducer and activator of transcription (*Jak*-*STAT*) pathway and cytokine-cytokine receptor signaling reinforce the importance of proper signaling cascades to fight infection (Sandford et al. 2011).

The transcriptomic studies herein described provide insights into the genetic alterations accumulated in modern poultry breeds upon selective breeding conducted over several decades. Transcriptome data obtained also contributed to a more comprehensive understanding of poultry responses to diseases. In due time, this knowledge could potentially be used in the development of breeds with enhanced traits such as host genetic resistance and be applied to derive new diagnosis tools or vaccines and ultimately ensure food safety for human population.

3.3 Pig Transcriptomic Studies

Despite fat being considered an unhealthy constituent of meat for consumers, fat depots and intramuscular fat content have a role in the taste, tenderness, and juiciness of cooked pig meat (Wood et al. 2008; Wang et al. 2015). Consequently, traits related with fat deposition during muscle growth are economically important in pig breeding because these can influence meat quality and carcass composition. Undoubtedly, muscle growth and fat deposition in pigs is breed dependent, but the response to environmental factors such as the composition of feed plays a major role in modulating these physiological processes.

Pig breeds with distinct patterns of muscle growth were used by Wang et al. (2015) as models for identifying the functional genes responsible for the molecular mechanisms that control the aforementioned economical traits. In this study,

Chinese breeds (Diannan Small-ear pig and Tibetan pig) with lower growth rate, more fat deposition, and better meat quality were compared to the introduced breeds (Landrace and Yorkshire) which are lean-type pig breeds, characterized by a fast growth rate and high lean meat content. To achieve this goal, differential transcriptome profiles of *longissimus dorsi* muscle tissues were established, to study the genome-wide expression and gene-miRNA interaction in the four breeds. The combination of the coding (mRNA) and noncoding (miRNAs) datasets allowed the identification of a network of 46 genes and 18 miRNAs regulating muscle growth. For lipid deposition, a network of 15 genes and 16 miRNAs was described. For example, adenosine receptor A1 (*ADORA1*) showed a downregulated expression in the Chinese breeds by comparison with introduced breeds. *ADORA1* is known to play important roles in lipid catabolism (Heseltine et al. 1995), which suggests that its increased expression in the introduced breeds is conducive to lipolysis. The combination of the miRNA-Seq and the RNA-Seq data provided experimental evidences that posttranscriptional gene regulation occurs, by comparing the relationship between differentially expressed miRNAs and the target mRNAs. One interesting example is described for caveolin-2 (*CAV2*), which is the target of miR-29b and miR-122. In Chinese breeds, *CAV2* was upregulated, which is consistent with the approximately seven- and fivefold lower expression of miR-29b and miR-122 seen in Chinese breeds than in introduced breeds. The protein encoded by *CAV2* is involved in essential cellular functions, including signal transduction, lipid metabolism, cellular growth control, and apoptosis thus having a relevant role in muscle growth. With this study, a comprehensive list of candidate genes, miRNAs, and metabolic pathways in pigs were identified (Wang et al. 2015) that could possibly be involved in the regulation of muscle growth and lipid deposition in pigs.

Nutrigenomics is an emerging discipline which aims to increase our knowledge about the mechanisms by which nutrition affects the metabolic pathways underlying homeostatic control (Ouhtit 2014). Recently, much attention has been given to the nutritional and health impacts of polyunsaturated omega-6 and omega-3 fatty acids intake, which can be obtained from food/feed sources (Salem and Eggersdorfer 2015; Bernardi et al. 2016). Being a monogastric species, the pig is amenable to changes in the fatty acid composition of adipose tissue and muscle using diets containing different oils (Wood et al. 2008). Different diet patterns elicit a variety of changes in gene expression, protein abundance, and metabolite production which impact animal physiology. Such diet-induced patterns and their effects are described by nutrigenomics as "signature dietary patterns" (Ouhtit 2014). Ogłuszka et al. (2017) investigated the changes in the pig *gluteus medius* muscle transcriptome and regulated biological functions as a result of an increased omega-3 and omega-6 fatty acid concentration. In this study, pigs were subjected to either a control diet or a diet supplemented with linseed and rapeseed oil to increase polyunsaturated fatty acid content. Differential expression analysis identified 749 genes significantly differing at least in twofold of change between two groups of animals fed with divergent level of omega-3 and omega-6 fatty acids. A "signature dietary pattern" was described in fatty acid-supplemented animals by

comparison with non-supplemented ones. In fatty acid-supplemented animals, a downregulation of the expression of genes encoding for chemokines C-C motif [e.g., chemokine ligand 2 (*CCL2*)], complement components [e.g., complement component 4-binding protein alpha (*C4BPA*)], as well as apolipoproteins [e.g., apolipoprotein A1 (*APOA1*)] was noticed. Overall, the results support usefulness of omega-3 and omega-6 fatty acids pig supplementation in the management and modulation of inflammatory, signaling, autoimmune responses, as well as lipid metabolism.

3.4 Small Ruminant Transcriptomic Studies: Goats and Sheep

Small ruminants, like sheep and goats, are particularly important in the tropics and the Mediterranean regions and are a major source of income and food in small-scale subsistence farming systems (Palma et al. 2016a). Goats have long been used for their milk, meat, hair (including cashmere), and skins throughout much of the world (MacHugh and Bradley 2001). It has been suggested that sheep were initially reared for meat and milk and only later for wool (Chessa et al. 2009). In these animals, dairy production depends upon the knowledge of the morphology and regulation of the mammary gland and lactation. High-throughput technologies, as omics-based ones, are allowing a much broader and detailed knowledge on the biology of the mammary gland (Ferreira et al. 2013).

Goats are one of the best models to study lactation, as well as mammary gland biology and physiology (Lérias et al. 2014; Cugno et al. 2016; Palma et al. 2016b; Hernández-Castellano et al. 2016). Several studies have proposed that miRNAs are important regulators of the gene expression in ruminant mammary glands (Lin et al. 2013; Mobuchon et al. 2015). MiRNAs are endogenous RNA molecules, approximately 22 nucleotides in length that regulate gene expression, mostly at the posttranscriptional level by translational repression or degradation of their targets (Samir et al. 2016). To the present date, the number of annotated and validated miRNA is still limited for goats, and the profiles of miRNAs across stages of lactation in dairy goats are relatively unknown. In this context, Wang et al. (2017) conducted a comprehensive study of the miRNA transcriptome using tissue samples from mammary gland of Saanen dairy goats harvested at early, mid, and peak lactation. This study identified a total of 796 conserved miRNAs, 263 new miRNAs, and 821 pre-miRNAs. After comparative analysis, a total of 37 miRNAs had significant differences in expression over the lactation cycle. Among those, the miR-145 was previously described to have a role in lipid metabolism in adipose tissues (Lin et al. 2014). To investigate the biological relevance of miR-145 during lactation, the same authors determined the overexpression and silencing of miR-145 to evaluate if such an approach resulted in a change in fatty acid metabolism in goat mammary epithelial cells. The results showed miR-145 modulation

resulted in changes in the fatty acid contents. Later, the authors proved that miR-145 acts as a negative regulator of insulin-induced gene 1 (*INSIG1*) expression by binding to their 3′-UTR in goat mammary epithelial cells. Downregulation of *INSIG1* has been associated with triacylglycerol accumulation in parallel with an increase in the biosynthesis of unsaturated lipids (Dong et al. 2012). Importantly, this study revealed the potential of modulating miRNA expression as an approach for improving beneficial milk components in ruminant milk or modulating lactation performance.

In 2014, approximately more than half a billion animals were slaughtered to meet the demand for sheep meat in the world (FAOSTAT 2017). Asia is one of the biggest producers of sheep meat, with a production bigger than 4 millions of tonnes in the same year (FAOSTAT 2017). In this context, much effort has been put in sheep breeding programs to improve not only carcass lean meat yield but also traits related to meat quality. One of these examples is the Qianhua Mutton Merino (QHMM), which is a new sheep bred for both meat and wool, presenting better meat performance when compared with the traditional local variety Small Tail Han (STH) sheep (Sun et al. 2016). Sun et al. (2016) compared the transcriptome profiles (RNA-Seq) of muscle tissues of QHMM and STH sheep to understand the molecular mechanisms behind improved performance of the new breed. Among the 960 differentially expressed genes, 405 were upregulated, while 555 were downregulated. As an example, a differential expression of muscle regulatory factors (MFRs) was observed between lines. Myogenin (*MYOG*) and myogenic differentiation (*MYOD*) were upregulated, while myogenic factor 6 (*MYF6)* was downregulated in QHMM. MFRs have been previously described as candidate genes for meat production traits in pigs (Ropka-Molik et al. 2011). An upregulation of MRFs expression in the postnatal skeletal muscles has been previously associated with an increase in muscle fiber diameter and eventually affected the meat quality of the skeletal muscles in adult chickens (Yin et al. 2015). Interestingly, the authors conclude that their candidate genes need to be functionally validated.

This highlights an important aspect of the candidate gene approach: the functional validation to link a candidate gene to a specific trait needs to be performed. Functional information from gene knockout in model transgenic animals and cellular models could provide distinct clues about the validity of candidate genes responsible for the phenotypes of interest. However, there is little practical information available because of the difficulty of producing gene knockout and transgenic animals in livestock (Zhu and Zhao 2007). This aspect still constitutes a challenge and open research avenue for animal and veterinary sciences.

3.5 Farmed Fish Transcriptomic Studies

Farmed Atlantic salmon (*Salmo salar*) is now a "super-commodity," a uniform product available on demand around the globe, which has revolutionized the economic and political landscape of the fishing industry (Eagle et al. 2004).

An ongoing discussion argues that Atlantic salmon is likely the most efficient domesticated farm animal, as 100 kg dry feed yields 65 kg Atlantic salmon fillets compared to only 20 kg of poultry fillets or 12 kg of pork fillets (Torrissen et al. 2011). Nevertheless, more studies are needed to generate a consensus about aquaculture production efficiency, sustainability, product quality, and profitability in support of the commercial sector and for the benefit of consumers (Abdelrahman et al. 2017).

Since the early 1970s, salmon sea lice, *Caligus rogercresseyi*, has been one of the most significant pathogenic marine parasites affecting farmed salmon and, thus, represents one of the most important threats to salmon farming throughout the world (Hamilton-West et al. 2012). Salmons cultured in Chile, namely, Atlantic and coho (*Oncorhynchus kisutch*) salmons, demonstrate different levels of susceptibility to the sea louse constituting interesting models to study the mechanisms that trigger salmonid immune responses to *C. rogercresseyi* infestation. Valenzuela-Muñoz et al. (2017) investigated the transcriptomic profiles of Atlantic and coho salmon subjected to a 14-day infestation trial with *C. rogercresseyi*, focusing on changes on iron metabolism as a proxy for nutritional immune responses. The results showed differences among the species which reflects their resistance toward the parasite. In the susceptible Atlantic salmon, genes related to heme degradation and iron transport such as hepcidin, transferrin, and haptoglobin were primary upregulated. Contrastingly, in coho salmon, an upregulation of aminolevulinic acid synthase and coproporphyrinogenase genes, with an important role in heme biosynthesis, was seen. In summary, the Atlantic salmon (more susceptible to infestation) presents molecular mechanisms to deplete cellular iron availability, suggesting putative mechanisms of nutritional immunity. In contrast, resistant coho salmons were less affected by sea lice, mainly activating pro-inflammatory mechanisms to cope with infestation.

With the rapid expansion of the salmon farming industry and the limited availability of wild-caught fish, the future of salmon farming will depend to a large extent on sustainable dietary protein alternatives to fish meal (Bostock et al. 2010). Significant progress has been made over the past decade in reducing levels of animal protein (fishmeal) throughout the increase of plant protein in commercial feeds for farmed fish (Hardy 2010). Grain legumes, also called pulses, yield nutritive grains, constituting therefore an important source of plant protein for human and animal consumption. Beans, peas, soybeans, and lupins have been used as protein sources for aquaculture feeds, but the challenge associated with replacing fish meal with plant protein concentrates is associated with the presence of anti-nutritional compounds (Hardy 2010; Molina-Poveda et al. 2013). The anti-nutritional factors present in plant-based feed have been described to promote gut inflammation (enteritis) and compromise fish health (Baeverfjord and Krogdahl 1996). The transcriptome responses of distal gut to varying dietary levels (0–45%) of soy protein concentrate (SPC) and faba bean (*Vicia faba*) protein concentrate (BPC) were investigated in Atlantic salmon parr (Król et al. 2016). After an 8-week feeding trial, fish were sampled for body composition analysis. Distal gut samples were collected for transcriptome analysis using a microarray platform developed

and validated for Atlantic salmon. The results demonstrated that different plant protein materials generated substantially different gene expression profiles in the distal gut of Atlantic salmon relative to guts from fish fed a FM-based diet used as controls. When SPC and BPC were simultaneously included in the diet, a less extensive alteration of gut transcriptome was seen than diets with either SPC or BPC separately, probably due to reduced levels of individual anti-nutritional factors. The mixed plant protein diets were also associated with improved body composition of fish relative to the single plant protein diets, which may provide evidence for a link between the magnitude of changes in gut transcriptome, gut health, and whole-animal performance. Importantly, this study indicates that gut transcriptomic profiling provides a useful tool for testing the suitability of new sources of protein for aquaculture feeds. It can also contribute for designing diets that have a reduced impact on fish health while improving aquaculture systems productivity.

One of the biggest obstacles in intensive aquaculture practice is the highly variable and unpredictable gamete quality, with eggs being the highest concern. Little is known about the determinants of egg quality, and this is also reflected on the scarce existence of reliable methods for assessing egg quality (Migaud et al. 2013). The establishment of transcriptomic profiles or identification of candidate genes associated with high-quality eggs could be a promising tool allowing the improvement of husbandry practices in aquaculture. One relevant study is the one conducted by Żarski et al. (2017), in which they correlated transcriptomic profile of sea bass (*Dicentrarchus labrax*) eggs with an assessment of its quality. Using a microarray approach, the authors identified several genes differentially expressed between high- and low-quality eggs. Some of these genes have already been reported to be of potential application to assess fish egg quality, among them an E3 ubiquitin-protein ligase (*RNF213*) and interferon regulatory factor 7 (*IRF7*). New candidate genes were also found as ubiquitin carboxyl-terminal hydrolase 5 (*USP5*), plectin-like isoform (*PLEC*), and centromere protein F (*CENPF*), but they lack validation on other fish species. One global overview of the metabolic pathways activated was also provided. In sea bass, protein ubiquitination, translation, DNA repair, cell structure, and cell architecture seem likely to be mechanisms contributing to egg developmental competence (Żarski et al. 2017).

4 Integrating Transcriptomic and Proteomic Data: The Quest for Systems Biology Approaches

Protein production for human consumption is the ultimate goal in animal farming. Consequently, research tools used to investigate proteins play a major role in farm animal and meat science (D'Alessandro and Zolla 2013). Similarly to transcriptomics, proteomics has been applied to the various fields of farmed animal

science, from lactation (Ferreira et al. 2013) and meat quality (Hollung et al. 2014) to animal production and welfare (Marco-Ramell et al. 2016).

One of the major challenges of systems biology is the integration of the different large omics datasets, such as transcriptomics and proteomics, originated from the same biological sample to attain a holistic overview of all regulatory processes and reactions (Parreira et al. 2016). The characterization of the observed changes from a gene-to-protein perspective provides a platform to cross-validate candidate genes and proteins associated with a given experimental condition. However, the most innovative knowledge will be the identification of regulatory mechanisms associated with gene expression, as the translational or posttranscriptional ones (Nie et al. 2007). While studies on animal cells are taking full advantage of omics data integration [e.g., see Wilmes et al. (2013) and Villar et al. (2015)], the number of reports available for farmed animals is limited, indicating that this is still a vast challenge for animal and veterinary researchers. Nevertheless, we will herein refer to some successful case studies, in which transcriptomic and proteomic data integration approaches have contributed with significant advances on animal physiology.

Hornshøj et al. (2009) conducted a comparative study of three high-throughput technologies for multi-sample expression profiling in heart and skeletal muscle (*Longissimus dorsi*) of Hampshire gilts. In this study, the same tissue samples were used in expression profiling with 454-sequencing, cDNA microarray, and iTRAQ-based proteomics to ensure an integrative approach. Although the authors compared the transcript profiles coming from different transcriptomic platforms and tissues, the comparison between transcript and protein abundance levels yielded the most interesting results. Transcript and protein ratios across heart and skeletal muscle were computed, and the results showed that 148 genes were detected by all three technologies and positively correlated. The positive correlation between transcript and protein ratios suggests that a relative different transcript level leads to a similar difference in the protein level, irrespective of the tissue studied. For genes without transcript-protein concordance, the authors hypothesized that this may arise from annotation errors or differential regulation of translation, turnover, or alternative splicing (Hornshøj et al. 2009). These results herein presented are among the first showing the importance of studying the regulation of transcript and protein abundances in mammalian tissues.

In another study, Timperio et al. (2009) investigated the transcriptomic and proteomic changes in liver metabolism in Chianina and Holstein Friesian cow breeds. These cattle breeds are representative of extreme selection for milk and meat traits, respectively, with significant changes in metabolism resulting from human selection over the past centuries. Despite a high number of proteins and genes differentially expressed between Chianina and Holstein Friesian breeds, the most interesting outcome came from the integrative analysis performed either for independent or merged datasets from both transcriptomic and proteomic approaches. Pathway analysis (Ingenuity® Systems, www.ingenuity.com) in the merged datasets indicated that in Chianina the top pathways activated were "amino acid metabolism," "small molecule biochemistry," and "posttranslational

modifications." In Holstein Friesian, the same approach revealed that the top pathways were "cellular compromise," "posttranslational modifications," and "protein folding." Network analysis in Chianina liver samples indicated a central core network of "lipid metabolism, amino acid metabolism, and molecular transport" pathway. On the other hand, in Holstein Friesian samples, constant central core of heat shock proteins and proteins involved in stress responses was retrieved. This study identified important metabolic differences that have arisen upon modification of relatively few genes/proteins (or gene/protein networks) elucidating about biological processes which have been molded in Holstein Friesian and Chianina over a thousand years of human breeding selection.

Recently, Yang et al. (2016) investigated the molecular basis underlying the higher meat quality in muscle samples of Shaziling pigs versus the Yorkshire ones using an integrative analysis of transcriptome and proteome. The Shaziling pig is a well-known breed, with origins in the Hunan Province, China. Its meat presents a high content of intramuscular fat, a proxy of high meat quality, when compared with other commercial breeds such as the Yorkshire pig (Yang et al. 2003). Yang et al. (2016) identified the differentially expressed genes and proteins from the *longissimus dorsi* muscles of Shaziling and Yorkshire pig breeds. RNA-Seq analysis identified 488 differentially expressed genes, of which 297 were upregulated in the Shaziling pig and 192 in the Yorkshire pig. Proteomics (2DE coupled to MALDI-TOF-MS/MS) identified 38 differentially abundant proteins, of which 27 protein spots were upregulated in the Shaziling pigs and 11 in the Yorkshire pigs. Additionally, the same authors correlated the mRNA expression and protein accumulation levels and found that proteomic results were consistent with transcriptomic results for only three genes, namely, Enolase 1 (*ENO1)* and alpha-actin (*ACTC1*) (overexpressed in Shaziling pigs) and ATP synthase subunit beta, mitochondrial (*ATP5)* (overexpressed in Yorkshire pigs). In the majority of the cases, transcriptome and proteome data were divergent. Differently expressed transcripts and proteins belong to the lipid mobilization, energy metabolism, the cytoskeleton, and signal transduction functional categories.

As seen in the described examples, the integrative analysis of transcriptome changes coupled to proteomic changes allowed an assessment of whether transcript abundance is a proxy of protein abundance. Although correlation between mRNA expression levels and protein abundance is expected to be strong based on the central dogma of molecular genetics (Nie et al. 2007), the evidences collected from the different studies described here are not always in agreement. Effective integration of high-throughput transcriptomic and proteomic data depends not only on experimental design but also on data quality and the availability of statistical methods and algorithms that can address the complexity of the data, extracting usable information. Nie et al. (2007) provide additional considerations for researchers aiming to develop approaches envisaging data integration. Indeed, working with species with limited genomic resources, such as farmed animals, is an added challenge and another level of complexity.

5 Conclusions

Recent advances in animal physiology owe a lot to transcriptomic studies, and this is particularly relevant for farm animals. The major scientific accomplishments obtained in the above described studies have provided valuable resources for animal breeders to develop new strategies to improve animal, health, welfare, and production. Genomics advances, the development of high-throughput transcriptomic platforms, and the development of user-friendly software for large-scale data analysis have been essential to accomplish such goal. Nevertheless, important drawbacks still need to be overcome such as the cost and access to cutting-edge analytic platforms and advanced data analysis expertise to integrate the complex datasets and extract valuable biological information.

So far, the transcriptomic approaches performed have disclosed potential candidate genes associated with muscle growth, meat quality, lactation, reproduction efficiency, or response to diseases. The candidate genes encode for functional or regulatory proteins, but also regulatory mechanisms mediated by hormones, transcriptional, posttranscriptional, and translational mechanisms were identified. We have also collected studies describing posttranscriptional regulation mechanisms mediated by miRNAs, which have been implicated in the modulation of muscle growth, lactation, and lipid metabolism in several animals. However, functional validation is still required to extend this knowledge.

Accumulating evidences show that epigenetic marks influence gene expression and phenotypic outcome in livestock species (Ibeagha-Awemu and Zhao 2015). Epigenetic mechanisms alter gene expression with impact on the phenotype without changing the DNA sequence. In this context, research on epigenetic regulation still constitutes an open research topic for animal and veterinary sciences due to its evidenced roles in genome integrity, developmental processes, stress responses, natural variation, and even transgenerational memory.

Animal researchers, farmers, and breeders are continuously facing the challenge of producing superior animals and processed livestock products, as a result from the combination of elite breeds with livestock management methodologies. Several genes associated with enhanced muscle development, accumulation of intramuscular fat, or disease resistance have been identified and can be used as biomarkers to screen for improved traits in response to environmental constraints, such as feed limitation or quality. Nevertheless, a validation of these strategies still needs to be performed at the interface of the academic/farm environment. Most of these studies were conducted in a restricted number of breeds or species of cattle, porcine, poultry, small ruminants, and, recently, farmed fishes. Therefore, the aforementioned validation needs to be extended to a much higher number of relevant species and breeds in order to result in effective guidelines for animal breeders. In due time, such knowledge and guidelines can contribute to face the growing demand for animal protein in a sustainable environment and animal welfare concerned society.

Acknowledgments The financial support from Fundação para a Ciência e a Tecnologia (F.C.T., Lisbon, Portugal) is acknowledged through research unit GREEN-it "Bioresources for Sustainability" (UID/Multi/04551/2013), SSA Post-Doctoral grant (SFRH/BPD/108032/2015), and also JRP Plants for Life PhD grant (PD/BD/113474/2015) within the scope of the PhD program Plants for Life (PD/00035/2013).

References

Abdelrahman H, ElHady M, Alcivar-Warren A et al (2017) Aquaculture genomics, genetics and breeding in the United States: current status, challenges, and priorities for future research. BMC Genomics 18:191. https://doi.org/10.1186/s12864-017-3557-1

Almeida AM, Nanni P, Ferreira AM et al (2017) The longissimus thoracis muscle proteome in Alentejana bulls as affected by growth path. J Proteomics 152:206–215. https://doi.org/10.1016/j.jprot.2016.10.020

Ansorge WJ (2009) Next-generation DNA sequencing techniques. N Biotechnol 25:195–203. https://doi.org/10.1016/j.nbt.2008.12.009

Baeverfjord G, Krogdahl A (1996) Development and regression of soybean meal induced enteritis in Atlantic salmon, *Salmo salar* L., distal intestine: a comparison with the intestines of fasted fish. J Fish Dis 19:375–387. https://doi.org/10.1046/j.1365-2761.1996.d01-92.x

Baufeld A, Koczan D, Vanselow J (2017) Induction of altered gene expression profiles in cultured bovine granulosa cells at high cell density. Reprod Biol Endocrinol 15:3. https://doi.org/10.1186/s12958-016-0221-6

Bernardi DM, Bertol TM, Pflanzer SB et al (2016) ω-3 in meat products: benefits and effects on lipid oxidative stability. J Sci Food Agric 96:2620–2634. https://doi.org/10.1002/jsfa.7559

Berton MP, Fonseca LFS, Gimenez DFJ et al (2016) Gene expression profile of intramuscular muscle in Nellore cattle with extreme values of fatty acid. BMC Genomics 17:972. https://doi.org/10.1186/s12864-016-3232-y

Biomarkers Definitions Working Group (2001) Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther 69:89–95. https://doi.org/10.1067/mcp.2001.113989

Bongiorni S, Gruber CEM, Chillemi G et al (2016) Skeletal muscle transcriptional profiles in two Italian beef breeds, Chianina and Maremmana, reveal breed specific variation. Mol Biol Rep 43:253–268. https://doi.org/10.1007/s11033-016-3957-3

Bostock J, McAndrew B, Richards R et al (2010) Aquaculture: global status and trends. Philos Trans R Soc Lond B Biol Sci 365:2897–2912. https://doi.org/10.1098/rstb.2010.0170

Bumgarner R (2013) Overview of DNA microarrays: types, applications, and their future. Curr Protoc Mol Biol 101:22.1.1–22.1.11. https://doi.org/10.1002/0471142727.mb2201s101

Cánovas A, Reverter A, DeAtley KL et al (2014) Multi-tissue omics analyses reveal molecular regulatory networks for puberty in composite beef cattle. PLoS One 9:e102551. https://doi.org/10.1371/journal.pone.0102551

Capra E, Turri F, Lazzari B et al (2017) Small RNA sequencing of cryopreserved semen from single bull revealed altered miRNAs and piRNAs expression between high- and low-motile sperm populations. BMC Genomics 18:14. https://doi.org/10.1186/s12864-016-3394-7

Cardoso TF, Cánovas A, Canela-Xandri O et al (2017) RNA-seq based detection of differentially expressed genes in the skeletal muscle of Duroc pigs with distinct lipid profiles. Sci Rep 7:40005. https://doi.org/10.1038/srep40005

Chen Z, Hagen DE, Wang J et al (2016) Global assessment of imprinted gene expression in the bovine conceptus by next generation sequencing. Epigenetics 11:501–516. https://doi.org/10.1080/15592294.2016.1184805

Chessa B, Pereira F, Arnaud F et al (2009) Revealing the history of sheep domestication using retrovirus integrations. Science 324:532–536. https://doi.org/10.1126/science.1170587

Chitwood JL, Rincon G, Kaiser GG et al (2013) RNA-seq analysis of single bovine blastocysts. BMC Genomics 14:350. https://doi.org/10.1186/1471-2164-14-350

Costa V, Angelini C, De Feis I et al (2010) Uncovering the complexity of transcriptomes with RNA-Seq. J Biomed Biotechnol 2010:1–19. https://doi.org/10.1155/2010/853916

Cugno G, Parreira JR, Ferlizza E et al (2016) The goat (*Capra hircus*) mammary gland mitochondrial proteome: a study on the effect of weight loss using Blue-Native PAGE and two-dimensional gel electrophoresis. PLoS One 11:e0151599. https://doi.org/10.1371/journal.pone.0151599

D'Alessandro A, Zolla L (2013) Meat science: from proteomics to integrated omics towards system biology. J Proteomics 78:558–577. https://doi.org/10.1016/j.jprot.2012.10.023

Diskin M, Morris D (2008) Embryonic and early foetal losses in cattle and other ruminants. Reprod Domest Anim 43:260–267. https://doi.org/10.1111/j.1439-0531.2008.01171.x

Dong X-Y, Tang S-Q, Chen J-D (2012) Dual functions of Insig proteins in cholesterol homeostasis. Lipids Health Dis 11:173. https://doi.org/10.1186/1476-511X-11-173

Duggan DJ, Bittner M, Chen Y et al (1999) Expression profiling using cDNA microarrays. Nat Genet 21:10–14. https://doi.org/10.1038/4434

Eagle J, Naylor R, Smith W (2004) Why farm salmon outcompete fishery salmon. Mar Policy 28:259–270. https://doi.org/10.1016/j.marpol.2003.08.001

Fairfull RW, Chambers JR (1984) Breeding for feed efficiency: poultry. Can J Anim Sci 642:513–527. https://doi.org/10.4141/cjas84-062

Fang L, Hou Y, An J et al (2016) Genome-wide transcriptional and post-transcriptional regulation of innate immune and defense responses of bovine mammary gland to *Staphylococcus aureus*. Front Cell Infect Microbiol 6:193. https://doi.org/10.3389/fcimb.2016.00193

FAOSTAT (2017) Production (livestock primary) and food balance (commodity balances - livestock and fish primary equivalent) datasets. Food and Agriculture Organization of the United Nations Statistics. http://www.fao.org/faostat/en/#data

Ferreira AM, Bislev SL, Bendixen E et al (2013) The mammary gland in domestic ruminants: a systems biology perspective. J Proteomics 94:110–123. https://doi.org/10.1016/j.jprot.2013.09.012

Gardina PJ, Clark TA, Shimada B et al (2006) Alternative splicing and differential gene expression in colon cancer detected by a whole genome exon array. BMC Genomics 7:325. https://doi.org/10.1186/1471-2164-7-325

Gonella-Diaza AM, Andrade SCS, Sponchiado M et al (2015) Size of the ovulatory follicle dictates spatial differences in the oviductal transcriptome in cattle. PLoS One 10:e0145321. https://doi.org/10.1371/journal.pone.0145321

Guo Z, Guilfoyle RA, Thiel AJ et al (1994) Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. Nucleic Acids Res 22:5456–5465

Guo B, Greenwood PL, Cafe LM et al (2015) Transcriptome analysis of cattle muscle identifies potential markers for skeletal muscle growth rate and major cell types. BMC Genomics 16:177. https://doi.org/10.1186/s12864-015-1403-x

Hamilton-West C, Arriagada G, Yatabe T et al (2012) Epidemiological description of the sea lice (*Caligus rogercresseyi*) situation in southern Chile in August 2007. Prev Vet Med 104:341–345. https://doi.org/10.1016/j.prevetmed.2011.12.002

Hardy RW (2010) Utilization of plant proteins in fish diets: effects of global demand and supplies of fishmeal. Aquac Res 41:770–776. https://doi.org/10.1111/j.1365-2109.2009.02349.x

Hekman JP, Johnson JL, Kukekova AV (2015) Transcriptome analysis in domesticated species: challenges and strategies. Bioinf Biol Insights 9:21–31. https://doi.org/10.4137/BBI.S29334

Hernández-Castellano LE, Ferreira AM, Nanni P et al (2016) The goat (*Capra hircus*) mammary gland secretory tissue proteome as influenced by weight loss: a study using label free proteomics. J Proteomics 145:60–69. https://doi.org/10.1016/j.jprot.2016.03.030

Heseltine L, Webster JM, Taylor R (1995) Adenosine effects upon insulin action on lipolysis and glucose transport in human adipocytes. Mol Cell Biochem 144:147–151. https://doi.org/10.1007/BF00944394

Hollung K, Timperio A, Olivan M et al (2014) Systems biology: a new tool for farm animal science. Curr Protein Pept Sci 15:100–117. https://doi.org/10.2174/1389203715666140221120642

Hornshøj H, Bendixen E, Conley LN et al (2009) Transcriptomic and proteomic profiling of two porcine tissues using high-throughput technologies. BMC Genomics 10:30. https://doi.org/10.1186/1471-2164-10-30

Hrdlickova R, Toloue M, Tian B (2017) RNA-Seq methods for transcriptome analysis. Wiley Interdiscip Rev RNA 8:e1364. https://doi.org/10.1002/wrna.1364

Ibeagha-Awemu EM, Zhao X (2015) Epigenetic marks: regulators of livestock phenotypes and conceivable sources of missing variation in livestock improvement programs. Front Genet 6:302. https://doi.org/10.3389/fgene.2015.00302

Ioannidis J, Donadeu FX (2016) Circulating miRNA signatures of early pregnancy in cattle. BMC Genomics 17:184. https://doi.org/10.1186/s12864-016-2529-1

Jäger M, Ott C-E, Grünhagen J et al (2011) Composite transcriptome assembly of RNA-seq data in a sheep model for delayed bone healing. BMC Genomics 12:158. https://doi.org/10.1186/1471-2164-12-158

Johansen L-H, Dahle MK, Wessel Ø et al (2016) Differences in gene expression in Atlantic salmon parr and smolt after challenge with Piscine orthoreovirus (PRV). Mol Immunol 73:138–150. https://doi.org/10.1016/j.molimm.2016.04.007

Johnson TJ, Kariyawasam S, Wannemuehler Y et al (2007) The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. J Bacteriol 189:3228–3236. https://doi.org/10.1128/JB.01726-06

Johnson PA, Stephens CS, Giles JR (2015) The domestic chicken: causes and consequences of an egg a day. Poult Sci 94:816–820. https://doi.org/10.3382/ps/peu083

Kemmett K, Humphrey T, Rushton S et al (2013) A longitudinal study simultaneously exploring the carriage of APEC virulence associated genes and the molecular epidemiology of faecal and systemic *E. coli* in commercial broiler chickens. PLoS One 8:e67749. https://doi.org/10.1371/journal.pone.0067749

Kong B-W, Hudson N, Seo D et al (2017) RNA sequencing for global gene expression associated with muscle growth in a single male modern broiler line compared to a foundational Barred Plymouth Rock chicken line. BMC Genomics 18:82. https://doi.org/10.1186/s12864-016-3471-y

Król E, Douglas A, Tocher DR et al (2016) Differential responses of the gut transcriptome to plant protein diets in farmed Atlantic salmon. BMC Genomics 17:156. https://doi.org/10.1186/s12864-016-2473-0

Lérias JR, Hernández-Castellano LE, Suárez-Trujillo A et al (2014) The mammary gland in small ruminants: major morphological and functional events underlying milk production—a review. J Dairy Res 81:304–318. https://doi.org/10.1017/S0022029914000235

Lin X, Luo J, Zhang L et al (2013) MicroRNAs synergistically regulate milk fat synthesis in mammary gland epithelial cells of dairy goats. Gene Expr 16:1–13. https://doi.org/10.3727/105221613X13776146743262

Lin Y-Y, Chou C-F, Giovarelli M et al (2014) KSRP and microRNA 145 are negative regulators of lipolysis in white adipose tissue. Mol Cell Biol 34:2339–2349. https://doi.org/10.1128/MCB.00042-14

Loor JJ, Bionaz M, Drackley JK (2013) Systems physiology in dairy cattle: nutritional genomics and beyond. Annu Rev Anim Biosci 1:365–392. https://doi.org/10.1146/annurev-animal-031412-103728

Loor JJ, Vailati-Riboni M, McCann JC et al (2015) TRIENNIAL LACTATION SYMPOSIUM: nutrigenomics in livestock: systems biology meets nutrition. J Anim Sci 93:5554. https://doi.org/10.2527/jas.2015-9225

MacHugh DE, Bradley DG (2001) Livestock genetic origins: goats buck the trend. Proc Natl Acad Sci USA 98:5382–5384. https://doi.org/10.1073/pnas.111163198
Maillo V, Gaora PO, Forde N et al (2015) Oviduct-embryo interactions in cattle: two-way traffic or a one-way street? Biol Reprod 92:144–144. https://doi.org/10.1095/biolreprod.115.127969
Maillo V, de Frutos C, O'Gaora P et al (2016a) Spatial differences in gene expression in the bovine oviduct. Reproduction 152:37–46. https://doi.org/10.1530/REP-16-0074
Maillo V, Sánchez-Calabuig MJ, Lopera-Vasquez R et al (2016b) Oviductal response to gametes and early embryos in mammals. Reproduction 152:R127–R141. https://doi.org/10.1530/REP-16-0120
Malone JH, Oliver B (2011) Microarrays, deep sequencing and the true measure of the transcriptome. BMC Biol 9:34. https://doi.org/10.1186/1741-7007-9-34
Marco-Ramell A, de Almeida AM, Cristobal S et al (2016) Proteomics and the search for welfare and stress biomarkers in animal production in the one-health context. Mol BioSyst 12:2024–2035. https://doi.org/10.1039/C5MB00788G
Marguerat S, Bähler J (2010) RNA-seq: from technology to biology. Cell Mol Life Sci 67:569–579. https://doi.org/10.1007/s00018-009-0180-6
Marioni JC, Mason CE, Mane SM et al (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18:1509–1517. https://doi.org/10.1101/gr.079558.108
Metzker ML (2010) Sequencing technologies - the next generation. Nat Rev Genet 11:31–46. https://doi.org/10.1038/nrg2626
Migaud H, Bell G, Cabrita E et al (2013) Gamete quality and broodstock management in temperate fish. Rev Aquac 5:S194–S223. https://doi.org/10.1111/raq.12025
Mobuchon L, Marthey S, Boussaha M et al (2015) Annotation of the goat genome using next generation sequencing of microRNA expressed by the lactating mammary gland: comparison of three approaches. BMC Genomics 16:285. https://doi.org/10.1186/s12864-015-1471-y
Molina-Poveda C, Lucas M, Jover M (2013) Evaluation of the potential of Andean lupin meal (*Lupinus mutabilis* Sweet) as an alternative to fish meal in juvenile *Litopenaeus vannamei* diets. Aquaculture 410:148–156. https://doi.org/10.1016/j.aquaculture.2013.06.007
Moore RE, Kirwan J, Doherty MK et al (2007) Biomarker discovery in animal health and disease: the application of post-genomic technologies. Biomark Insights 2:185–196
Mutryn MF, Brannick EM, Fu W et al (2015) Characterization of a novel chicken muscle disorder through differential gene expression and pathway analysis using RNA-sequencing. BMC Genomics 16:399. https://doi.org/10.1186/s12864-015-1623-0
Nestor KE, Noble DO, Zhu NJ et al (1996) Direct and correlated responses to long-term selection for increased body weight and egg production in turkeys. Poult Sci 75:1180–1191. https://doi.org/10.3382/ps.0751180
Nie L, Wu G, Culley DE et al (2007) Integrative analysis of transcriptomic and proteomic data: challenges, solutions and applications. Crit Rev Biotechnol 27:63–75. https://doi.org/10.1080/07388550701334212
Ogłuszka M, Szostak A, Te Pas MFW et al (2017) A porcine gluteus medius muscle genome-wide transcriptome analysis: dietary effects of omega-6 and omega-3 fatty acids on biological mechanisms. Genes Nutr 12:4. https://doi.org/10.1186/s12263-017-0552-8
Ouhtit A (2014) Nutrigenomics: from promise to practice. Sultan Qaboos Univ Med J 14:e1–e3
Ozsolak F, Milos PM (2011a) RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12:87–98. https://doi.org/10.1038/nrg2934
Ozsolak F, Milos PM (2011b) Single-molecule direct RNA sequencing without cDNA synthesis. Wiley Interdiscip Rev RNA 2:565–570. https://doi.org/10.1002/wrna.84
Ozsolak F, Platt AR, Jones DR et al (2009) Direct RNA sequencing. Nature 461:814–818. https://doi.org/10.1038/nature08390
Palma M, Hernández-Castellano LE, Castro N et al (2016a) NMR-metabolomics profiling of mammary gland secretory tissue and milk serum in two goat breeds with different levels of

tolerance to seasonal weight loss. Mol Biosyst 12:2094–2107. https://doi.org/10.1039/c5mb00851d

Palma M, Scanlon T, Kilminster T et al (2016b) The hepatic and skeletal muscle ovine metabolomes as affected by weight loss: a study in three sheep breeds using NMR-metabolomics. Sci Rep 6:39120. https://doi.org/10.1038/srep39120

Parreira JR, Branco D, Almeida AM et al (2016) Systems biology approaches to improve drought stress tolerance in plants: state of the art and future challenges. In: Hossain AM, Wani HS, Bhattacharjee S et al (eds) Drought stress tolerance in plants, Vol 2: Molecular and genetic perspectives. Springer, Cham, pp 433–471. https://doi.org/10.1007/978-3-319-32423-4_16

Quail M, Smith ME, Coupland P et al (2012) A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. BMC Genomics 13:341. https://doi.org/10.1186/1471-2164-13-341

Romanov MN (2001) Genetics of broodiness in poultry - A review. Asian-Australasian J Anim Sci 14:1647–1654. https://doi.org/10.5713/ajas.2001.1647

Romereim SM, Summers AF, Pohlmeier WE et al (2017) Gene expression profiling of bovine ovarian follicular and luteal cells provides insight into cellular identities and functions. Mol Cell Endocrinol 439:379–394. https://doi.org/10.1016/j.mce.2016.09.029

Ropka-Molik K, Eckert R, Piórkowska K (2011) The expression pattern of myogenic regulatory factors MyoD, Myf6 and Pax7 in postnatal porcine skeletal muscles. Gene Expr Patterns 11:79–83. https://doi.org/10.1016/j.gep.2010.09.005

Salem N, Eggersdorfer M (2015) Is the world supply of omega-3 fatty acids adequate for optimal human nutrition? Curr Opin Clin Nutr Metab Care 18:147–154. https://doi.org/10.1097/MCO.0000000000000145

Samir M, Vaas LAI, Pessler F (2016) MicroRNAs in the host response to viral infections of veterinary importance. Front Vet Sci 3:86. https://doi.org/10.3389/fvets.2016.00086

Sandford EE, Orr M, Balfanz E et al (2011) Spleen transcriptome response to infection with avian pathogenic *Escherichia coli* in broiler chickens. BMC Genomics 12:469. https://doi.org/10.1186/1471-2164-12-469

Schena M (1996) Genome analysis with gene expression microarrays. BioEssays 18:427–431. https://doi.org/10.1002/bies.950180513

Schena M, Shalon D, Davis RW et al (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470. https://doi.org/10.1126/science.270.5235.467

Schroyen M, Tuggle CK (2015) Current transcriptomics in pig immunity research. Mamm Genome 26:1–20. https://doi.org/10.1007/s00335-014-9549-4

Scolari SC, Pugliesi G, Strefezzi R de F et al (2017) Dynamic remodeling of endometrial extracellular matrix regulates embryo receptivity in cattle. Reproduction 153:49–61. https://doi.org/10.1530/REP-16-0237

Shahzad K, Loor JJ (2012) Application of top-down and bottom-up systems approaches in ruminant physiology and metabolism. Curr Genomics 13:379–394. https://doi.org/10.2174/138920212801619269

Shimmura T, Maruyama Y, Fujino S et al (2015) Persistent effect of broody hens on behaviour of chickens. Anim Sci J 86:214–220. https://doi.org/10.1111/asj.12253

Sihvo H-K, Immonen K, Puolanne E (2014) Myodegeneration with fibrosis and regeneration in the pectoralis major muscle of broilers. Vet Pathol 51:619–623. https://doi.org/10.1177/0300985813497488

Smith KG, Hunt JL (2004) On the use of spleen mass as a measure of avian immune system strength. Oecologia 138:28–31. https://doi.org/10.1007/s00442-003-1409-y

Sullivan CV, Chapman RW, Reading BJ et al (2015) Transcriptomics of mRNA and egg quality in farmed fish: some recent developments and future directions. Gen Comp Endocrinol 221:23–30. https://doi.org/10.1016/j.ygcen.2015.02.012

Sun L, Bai M, Xiang L et al (2016) Comparative transcriptome profiling of longissimus muscle tissues from Qianhua Mutton Merino and Small Tail Han sheep. Sci Rep 6:33586. https://doi.org/10.1038/srep33586

Sweeney T, Lejeune A, Moloney AP et al (2016) The application of transcriptomic data in the authentication of beef derived from contrasting production systems. BMC Genomics 17:746. https://doi.org/10.1186/s12864-016-2851-7

Thornton PK (2010) Livestock production: recent trends, future prospects. Philos Trans R Soc B 365:2853–2867. https://doi.org/10.1098/rstb.2010.0134

Timperio AM, D'Alessandro A, Pariset L et al (2009) Comparative proteomics and transcriptomics analyses of livers from two different *Bos taurus* breeds: "Chianina and Holstein Friesian". J Proteomics 73:309–322. https://doi.org/10.1016/j.jprot.2009.09.015

Tizioto PC, Coutinho LL, Oliveira PSN et al (2016) Gene expression differences in Longissimus muscle of Nelore steers genetically divergent for residual feed intake. Sci Rep 6:39493. https://doi.org/10.1038/srep39493

Torrissen O, Olsen RE, Toresen R et al (2011) Atlantic Salmon (*Salmo salar*): the "Super-Chicken" of the sea? Rev Fish Sci 19:257–278. https://doi.org/10.1080/10641262.2011.597890

Valenzuela-Muñoz V, Boltaña S, Gallardo-Escárate C (2017) Uncovering iron regulation with species-specific transcriptome patterns in Atlantic and coho salmon during a *Caligus rogercresseyi* infestation. J Fish Dis 40(9):1169–1184. https://doi.org/10.1111/jfd.12592

Van Horne PLM, Achterbosch TJ (2008) Animal welfare in poultry production systems: impact of EU standards on world trade. Worlds Poult Sci J 64:40–51. https://doi.org/10.1017/S0043933907001705

Vigneault C, Gravel C, Vallee M et al (2009) Unveiling the bovine embryo transcriptome during the maternal-to-embryonic transition. Reproduction 137:245–257. https://doi.org/10.1530/REP-08-0079

Villar M, Ayllon N, Alberdi P et al (2015) Integrated metabolomics, transcriptomics and proteomics identifies metabolic pathways affected by *Anaplasma phagocytophilum* infection in tick cells. Mol Cell Proteomics 14(12):3154–3172. https://doi.org/10.1074/mcp.M115.051938

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63. https://doi.org/10.1038/nrg2484

Wang Z, Li Q, Chamba Y et al (2015) Identification of genes related to growth and lipid deposition from transcriptome profiles of pig muscle tissue. PLoS One 10:e0141138. https://doi.org/10.1371/journal.pone.0141138

Wang XG, Ju ZH, Hou MH et al (2016) Deciphering transcriptome and complex alternative splicing transcripts in mammary gland tissues from cows naturally infected with *Staphylococcus aureus mastitis*. PLoS One 11:e0159719. https://doi.org/10.1371/journal.pone.0159719

Wang H, Shi H, Luo J et al (2017) MiR-145 regulates lipogenesis in goat mammary cells via targeting *INSIG1* and epigenetic regulation of lipid-related genes. J Cell Physiol 232:1030–1040. https://doi.org/10.1002/jcp.25499

Wheeler DL, Barrett T, Benson DA et al (2008) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 36:D13–D21. https://doi.org/10.1093/nar/gkm1000

Wickramasinghe S, Cánovas A, Rincón G et al (2014) RNA-sequencing: a tool to explore new frontiers in animal genetics. Livest Sci 166:206–216. https://doi.org/10.1016/j.livsci.2014.06.015

Wilmes A, Limonciel A, Aschauer L et al (2013) Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. J Proteomics 79:180–194. https://doi.org/10.1016/j.jprot.2012.11.022

Woo Y, Affourtit J, Daigle S et al (2004) A comparison of cDNA, oligonucleotide, and Affymetrix GeneChip gene expression microarray platforms. J Biomol Tech 15:276–284

Wood JD, Enser M, Fisher AV et al (2008) Fat deposition, fatty acid composition and meat quality: a review. Meat Sci 78:343–358. https://doi.org/10.1016/j.meatsci.2007.07.019

Xing K, Zhu F, Zhai L et al (2016) Identification of genes for controlling swine adipose deposition by integrating transcriptome, whole-genome resequencing, and quantitative trait loci data. Sci Rep 6:23219. https://doi.org/10.1038/srep23219

Yang S-L, Wang Z-G, Liu B et al (2003) Genetic variation and relationships of eighteen Chinese indigenous pig breeds. Genet Sel Evol 35(7):657–671. https://doi.org/10.1186/1297-9686-35-7-657.

Yang H, Xu X, Ma H et al (2016) Integrative analysis of transcriptomics and proteomics of skeletal muscles of the Chinese indigenous Shaziling pig compared with the Yorkshire breed. BMC Genetics 17:80. https://doi.org/10.1186/s12863-016-0389-y

Yin H, Li D, Wang Y et al (2015) Myogenic regulatory factor (MRF) expression is affected by exercise in postnatal chicken skeletal muscles. Gene 561:292–299. https://doi.org/10.1016/j.gene.2015.02.044

Yu J, Lou Y, He K et al (2016a) Goose broodiness is involved in granulosa cell autophagy and homeostatic imbalance of follicular hormones. Poult Sci 95:1156–1164. https://doi.org/10.3382/ps/pew006

Yu J, Lou Y, Zhao A (2016b) Transcriptome analysis of follicles reveals the importance of autophagy and hormones in regulating broodiness of Zhedong white goose. Sci Rep 6:36877. https://doi.org/10.1038/srep36877

Żarski D, Nguyen T, Le Cam A et al (2017) Transcriptomic profiling of egg quality in sea bass (*Dicentrarchus labrax*) sheds light on genes involved in ubiquitination and translation. Mar Biotechnol 19:102–115. https://doi.org/10.1007/s10126-017-9732-1

Zhao C, Tian F, Yu Y et al (2012) Muscle transcriptomic analyses in Angus cattle with divergent tenderness. Mol Biol Rep 39:4185–4193. https://doi.org/10.1007/s11033-011-1203-6

Zhu M, Zhao S (2007) Candidate gene identification approach: progress and challenges. Int J Biol Sci 3:420–427. https://doi.org/10.7150/ijbs.3.420

NMR Metabolomics *pari passu* with Proteomics: Two Relevant Tools for Animal Sciences Combined

Mariana Palma, André M. de Almeida, and Manolis Matzapetakis

Abstract Metabolomics is a field of growing importance in farm animal science and agricultural research. It may be based on either mass spectrometry or NMR (nucleic magnetic resonance) spectroscopy and is particularly advantageous for the quantification and identification of metabolites of various sizes that can provide clues on how biological systems cope or react to any studied factor. Metabolomics is particularly interesting when seen from the perspective of other so-called post-genomic tools, particularly transcriptomics and proteomics in a systems biology approach. In this chapter, we focus on the importance of NMR metabolomics, providing brief concepts and the necessary workflow to work with this very powerful analytical tool. Furthermore, we provide two study cases where NMR metabolomics has been used to address research topics in animal science, namely, the response of small ruminants to seasonal weight loss. We finally conclude the chapter with brief considerations on the future of NMR metabolomics, its integration with proteomics research and its context in animal science.

Keywords Animal science • NMR • Metabolomics

M. Palma
Center for Functional Ecology, Department of Live Sciences, Universidade de Coimbra, Coimbra, Portugal

A.M. de Almeida (✉)
Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, Lisbon, Portugal
e-mail: aalmeida@isa.ulisboa.pt

M. Matzapetakis
Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Oeiras, Portugal

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_21

List of Abbreviations

1D	One dimension
DHA	Docosahexaenoic acid
DPA	Docosapentaenoic acid
EPA	Eicosapentaenoic acid
HR-MAS	High-resolution magic angle spinning
K	Kelvin
kDa	Kilodalton
MALDI-TOF/TOF MS	Matrix-assisted laser desorption ionization—time of flight—mass spectrometry
MHz	Megahertz
MS	Mass spectrometry
MW	Molecular weight
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser enhancement spectroscopy
PCA	Principal component analysis
SWL	Seasonal weight loss

1 Introduction

Together with crops, animal production is among the most important areas of human endeavour and was ultimately responsible for the expansion of mankind throughout the world and the development of civilization. Traditionally and similarly to other fields of knowledge, animal sciences were essentially empiric and most progress was made through trial and error. As a result, progress was limited and widespread adoption of new technologies was slow. The situation changed considerably with the onset of industrial revolution that introduced modern farming and animal production methods, as well as the basis for most varieties and breeds that are presently used in world agriculture and animal production.

Despite their deep roots in the transformations occurring in the eighteenth century, most massive developments in modern agriculture and animal production occurred after World War II, moving from a small farm, family-based model to large-scale corporation-based operations with tremendous improvement on particular fields such as mechanization, infrastructures and housing, feeding and genetics that combined with cool-chain storage, and modern distribution changed forever the farming business. These improvements are particularly noticeable in specific sectors like poultry, pig or dairy cattle production. During the last 50 years, agriculture and animal sciences have moved to an era of molecular-based studies where the characterization of genes of plants and animals are related to production traits of interest to humans. This revolution was made possible by the era of genomics, and understanding gene function was the most important challenge. Nevertheless, during the last 25 years, other disciplines have emerged, the so-called post-genomic disciplines. These, based on information created through

a deep characterization of the genome, allowed to understand gene function as a whole. This approach aims not only to characterize the genes and genomes of plants and animals of agricultural interest but mainly to understand how they work from an end-product perspective and also how the system worked in a dynamic, whole context. It allowed a deep and full characterization of living systems, contributing significantly to the improvement of agricultural and animal science productivity and ultimately product quality.

Post-genomic tools span several different fields, and since they were inspired by the genome (or the discipline that studies genes and gene functions—genomics), they are generally called the omics and include, namely, proteomics (the science that studies the proteome or the proteins in a given cell, tissue, organ, organism, population or fluid) and metabolomics (the science that studies the metabolome or the metabolites in a given cell, tissue, organ, organism, population or fluid). Metabolomics has been used in many studies, of very diverse nature. Some of them are in the field of animal science and include, for instance, approaches on sea bass aquaculture (Mannina et al. 2008; Viegas et al. 2011), beef quality (Kodani et al. 2017), swine adiposity (Jegou et al. 2016) or markers of production performance and metabolic health in dairy cows (Huber et al. 2016; Dervishi et al. 2017). These different fields or tools are frequently combined in the so-called systems biology approach. This topic has been extensively reviewed elsewhere in other chapters of this book and previously in other studies (Almeida et al. 2015).

We have seen elsewhere in this book the important interactions existing between proteomics and transcriptomics. In this chapter, we aim to show the complementary nature of metabolomics and proteomics, highlighting the need for a combination of both fields in animal science research. This chapter is divided in five different sections. After an introduction, we will mention some metabolomics concepts, highlighting the differences between the two major metabolomics techniques: mass spectrometry based and nucleic magnetic resonance (NMR) based, focusing on the advantages and disadvantages of each technique. In a third part, we will overview experimental aspects and the workflow used in NMR-based metabolomics. Two case studies on the importance of NMR-based metabolomics in small ruminants will follow, highlighting the main achievements and how they interact with proteomics. Finally, we will provide the main conclusions and highlights of this chapter.

2 Metabolomics: NMR and MS

Metabolomics studies are generally tasked with the identification and quantification of small metabolites (compounds of low-molecular weight, MW $<$ 1.5 kDa) that are the result of numerous metabolic processes. Since the metabolome is highly affected by gene expression, protein activity and/or the metabolic activities (Bernini et al. 2011; Griffin and Shockcor 2004; Nicholson and Wilson 2003), it is highly complementary to other omics-based studies such as proteomics and

transcriptomics, contributing to a more comprehensive overview of the animal physiology and production potential (Kadarmideen 2014).

Numerous metabolomics studies have been developed since the debut of the concept, and since then, the number of metabolomics-related publications has been increasing steadily. Most of these studies are developed using mass spectrometry (MS) or nuclear magnetic resonance techniques (Gowda and Raftery 2015). Each methodology has specific features that could constitute advantages or disadvantages. The selection of the technique that better fits each work is dependent on the objectives to be achieved (Gowda and Raftery 2015). Usually, due to their complementary character, both techniques are used.

NMR spectroscopy offers important advantages that highlight its usefulness in metabolic studies. For instance, it is both quantitative and qualitative, non-destructive, non-equilibrium disturbing (Beckonert et al. 2007), highly reproducible and has the ability to detect compounds within a wide range of physiochemical properties in the micromolar to millimolar range (Beckonert et al. 2007; Bernini et al. 2011). On the other hand, the main advantages of MS for metabolomics studies are its higher sensitivity that allows the identification of secondary metabolites in the picomole to the femtomole range, the possibility to use different ionization techniques that increase the number of potentially detectable compounds. It is also generally better suited for targeted analysis, and it is less expensive to purchase and maintain the equipment (Emwas 2015).

NMR-based metabolomics studies have been used in farm animal sciences for different purposes: analysis of food quality and safety parameters (Beauclercq et al. 2016; Bertram et al. 2002; Bonnefont et al. 2014; Coco et al. 2009; Hu et al. 2007), milk metabolomics (Klein et al. 2010; Sundekilde et al. 2013), sample provenance discrimination (Mannina et al. 2008; Rezzi et al. 2007; Sacco et al. 2009; Standal et al. 2008), diagnosis (Pears et al. 2007; Maher et al. 2012; Sundekilde et al. 2013; Sun et al. 2014; Ceccarelli et al. 2015; Scano et al. 2015), biomarker identification (Mannina et al. 2008; Nyberg et al. 2010; Lu et al. 2013; Palma et al. 2016a, b; Wang et al. 2016; Yang et al. 2016; Xu et al. 2016) and animal welfare (Li et al. 2011).

3 NMR Metabolomics: Concepts and Workflow

Nuclear magnetic resonance or NMR is a technique that probes the interaction of radiofrequency range electromagnetic radiation with atomic nuclei, specifically the spin and nuclear magnetism, under the influence of external magnetic fields. Detailed explanation of the theoretical background of the method is beyond the scope of this chapter and has been extensively documented (Roth 1984; Levitt 2008). We will, however, briefly introduce the main elements of an NMR spectrum that are needed for the identification of metabolite signatures. These are chemical shift, J-couplings, peak area and signal relaxation.

Chemical shift refers to the position of each signal in the spectrum and results from the slight differences in resonance frequency due to the magnetic microenvironment generated by the chemical surroundings of each nucleus as well as that of its neighbours. It carries information about the chemical surroundings of each nucleus (chemical structure). A second level of information can be derived from the splitting of each signal which results from direct magnetic interactions of chemically bonded nuclei manifesting themselves in the form of J-couplings. It contains information on the number and different types of neighbouring nuclei. In the simpler cases, the number of peaks each signal splits into is equal to the number of neighbouring nuclei plus one. Perhaps the element that is more relevant to metabolomics studies is that of the signal magnitude which is proportional to the number of atoms giving rise to it and the signal can therefore be used for quantification. Finally, signal decay affects peak widths, and since it can vary even within the same molecules, quantification is done using peak areas instead of peak heights (Roth 1984; Levitt 2008) (Fig. 1).

Among the types of nuclei that can be studied by NMR, those that are more relevant for metabolomics are ^{1}H and ^{13}C, although ^{15}N and ^{31}P also have uses but are not often explored. ^{1}H is the type that is most used because of its high sensitivity, high natural abundance and chemically rich information that can be extracted from it. It is also present in the vast majority of organic compounds.

In terms of sample considerations, the main concern of NMR is sensitivity as it is not a very sensitive technique. However, many biological fluids such as blood, urine, spinal fluid, etc. have been found to be suitable for the identification and quantification of 40–60 of the most concentrated compounds when modern high-field instruments are used (500–800 MHz). The use of high-sensitivity cryoprobes can boost sensitivity and decrease collection time to only a few minutes per sample. Their use is therefore highly advised although not strictly necessary. Other considerations include the composition of buffers that are used together with the deuterated solvents that are needed for measurement stability purposes. Buffers are needed because variations in the pH of a solution can cause shifts in the

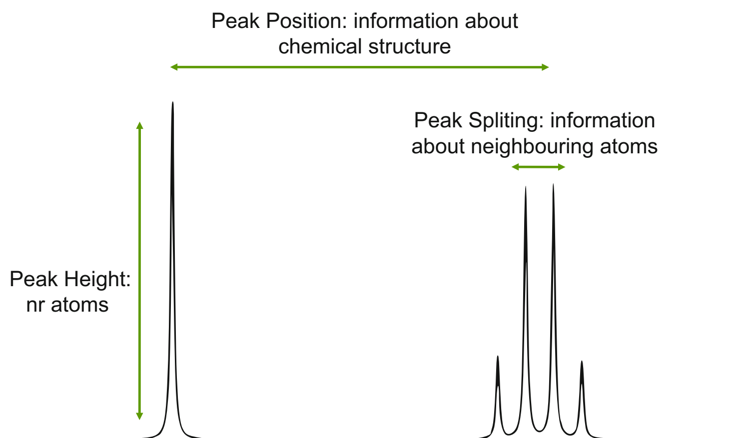

Fig. 1 Example of an NMR spectrum (1D ^{1}H) with the main chemical information available

position of the NMR signals which can make comparison of multiple samples very difficult.

Most of the NMR metabolomics studies are based on single 1D spectra of liquid samples that allow the identification and quantification of metabolites, require low optimization concerning parameter acquisition and are usually fast to acquire. Even though some samples require previous preparation, this is usually an easy and swift process (Bernini et al. 2011).

The most commonly used experiments are those of 1D with solvent (usually water) suppression. The pulse sequence that is mostly used for this purpose is that of the 1D NOESY with pre-saturation. In the case of very high protein content which can hinder the identification and quantification of metabolites, different pulse sequences can be used to suppress the signals of proteins. Alternatively, removal of proteins with a protocol involving extractions can be tried. This step is usually necessary in the case of solid samples such as muscle, various glands and other organs. Alternatively, the high-resolution magic angle spinning (HR-MAS) technique can also be applied and requires only a small intact piece of tissue, without any previous preparation (Gowda and Raftery 2015; Moestue et al. 2011).

After acquisition, data interpretation can be achieved via a complete spectral assignment and quantification—targeted approach (profiling) or via an untargeted approach (chemometric method). Profiling requires the identification and quantification of all detectable metabolites in a mixture (Mercier et al. 2011). Signal assignment is achieved using databases of compound spectral data to allow peak identification while quantification is achieved by comparison of peak volumes with those of standards (Ellinger et al. 2012). Targeted analysis yields information of the composition of each sample; however, it is limited by the number of compounds that can be reliably identified. NMR acquisition followed by profiling has been used in small ruminants, especially for diagnosis and biomarker identification (Li et al. 2011; Mickiewicz et al. 2015; Wang et al. 2016). This approach has the added advantage of providing information about the physiological metabolome profile of the studied animals, which can be useful for future works. On the other hand, chemometric analysis is based on the quantification of signals in discrete regions of the spectra without any signal interpretation prior to the analysis. This is only possible when results are analysed with statistical tools. While chemometrics is fast and convenient, it has the possible drawbacks of masking low concentrated spectral signals that are located nearby high-concentrated ones (Worley and Powers 2012) and being affected by small variations in chemical shifts (Salek et al. 2011). The chemometric approach does not allow sample profiling; however, it could be a useful and faster option when the profile has already been established or when it is only needed to identify one compound. Its use was already described in production animals for diagnosis and quality control (Castejón et al. 2015; Nyberg et al. 2010).

4 NMR Metabolomics and Proteomics: Small Ruminants and Tolerance to Seasonal Weight Loss

In the tropics and the Mediterranean, animal production varies considerably throughout the year. In those regions, the existence of two different seasons (rainy and dry) regulates pasture availability, affecting animal production. In fact, during the dry season, due to the scarcity of pastures, animals may lose up to 40% of their body weight, a condition usually known as seasonal weight loss (SWL) (Lamy et al. 2012; Cardoso and Almeida 2013). SWL is one of the major drawbacks in animal production, as had been demonstrated in South Africa (Almeida et al. 2006, 2007), Western Africa (Almeida and Cardoso 2008a, b), Western Australia (Almeida et al. 2013; Scanlon et al. 2013; Palma et al. 2016b) and the Canary Islands (Lérias et al. 2013, 2015; Palma et al. 2016a). Meat and milk production is strongly affected by SWL. In general, quantity and quality decrease and reduce income. To counter the effects of SWL, farmers use supplementation to balance the nutritional need of the animals. However, supplementation is expensive and difficult to implement in extensive production systems in developing countries or remote locations. An alternative method for addressing the effects of SWL is the use of breeds naturally adapted to this constraint or selected by domestication that are able to thrive and more effectively produce in such difficult environments.

Our research group has been studying tolerance to seasonal weight loss in small ruminants over the last 15 years. During recent years, emphasis has been placed on breed comparison, contrasting breeds that are tolerant to SWL with breeds that are not. This research has been done using two major species, meat-producing sheep (focusing on skeletal muscle and liver tissues) and dairy goats (focusing on the secretory tissue of the mammary gland and on milk). We have extensively used omics, particularly proteomics and NMR-based metabolomics to conduct this research. The two major topics are illustrated on Fig. 2, specifically referring to the metabolomics studies, whereas in Fig. 3, we present examples of the spectra for the different tissues and fluids studied.

4.1 NMR Metabolomics and Tolerance to Seasonal Weight Loss in Meat-Producing Sheep

In this study, we compared three sheep breeds with different levels of tolerance to SWL: Damara (highly tolerant), Australian Merino (susceptible) and Dorper (intermediate). These breeds exist in Western Australia, where an animal trial was conducted. Briefly, it is important to mention that the Damara is a fat-tailed breed that originates in the fringes of the Kalahari Desert in Namibia, is particularly well adapted to semiarid environments and has therefore a very high tolerance to SWL (Almeida 2011). The Merino is a breed of European origin, and the Dorper is a breed that includes in its genesis both fat-tailed African and European breeds.

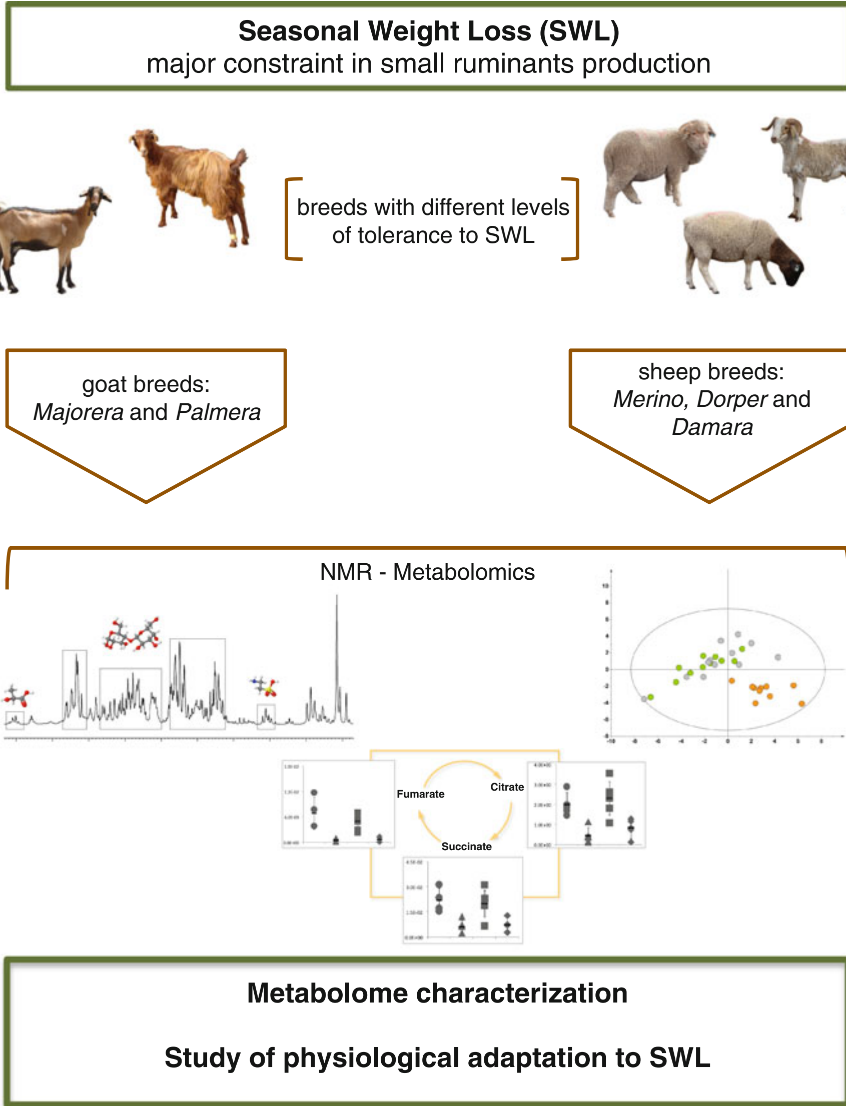

Fig. 2 Major research topics involving seasonal weight loss (SWL) tolerance in small ruminants and specifically the NMR metabolomics approach. On the *left*, we have contrasted two goat breeds with different levels of tolerance to SWL, specifically at the level of the mammary gland secretory tissue, whereas on the right, we have contrasted three meat-producing sheep breeds with different levels of tolerance to SWL. In both approaches, we aimed on one hand to characterize the metabolome and on the other hand to study the physiological adaptation to SWL, characteristic of the tolerant breed

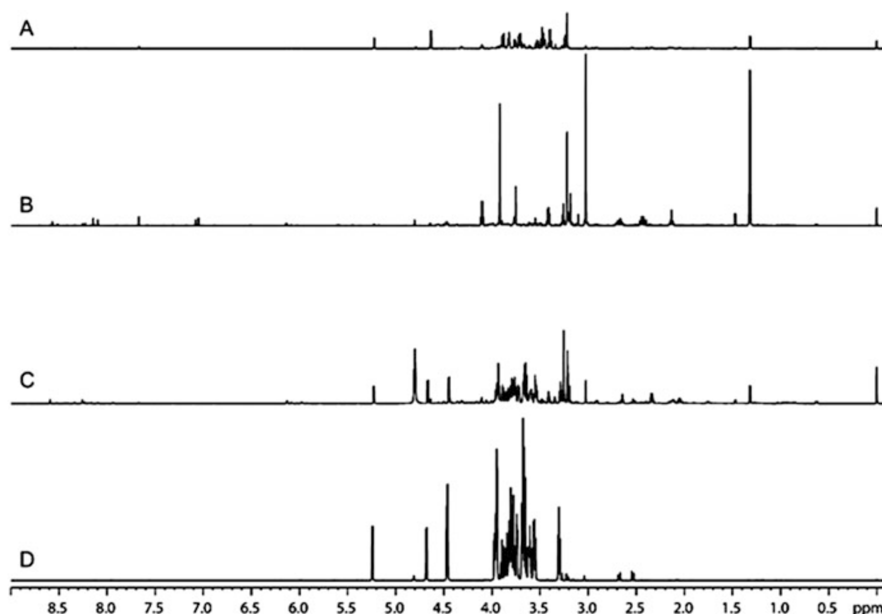

Fig. 3. Representative spectra (obtained at 800 MHz) of the different types of tissues and fluids studied. (**a**) Aqueous fraction of sheep liver, (**b**) aqueous fraction of sheep muscle, (**c**) aqueous fraction of goat mammary, (**d**) goat milk whey

Animals from the three breeds were subjected to nutritional stress over a period of 42 days, leading to relative live weight decreases of about 10–15% (Scanlon et al. 2013), and meat and carcass traits were characterized in depth (Almeida et al. 2013). We have first focused our studies on a lipidomics approach that included the fatty acid profiling of the fat tail tissue in Damara sheep. Its results let us conclude that Damara breed has a distinct lipid metabolism leading to a high concentration of branched chain fatty acids in tissues (Alves et al. 2013). Later, we have also conducted an in-depth characterization of muscle fatty acid profiles (van Harten et al. 2016). Our findings indicated that Damara sheep revealed several differences compared to the other two breeds, namely, a higher concentration of polyunsaturated fatty acids, which can be related to the presence of the fat tail. Furthermore and even in restricted feeding conditions, Damara sheep revealed the highest levels compared to Merino and Dorper sheep of linoleic acid, linolenic acid, eicosapentaenoic acid (EPA), docosapentaenoic acid (DPA) and docosahexaenoic acid (DHA), respectively. Among these, EPA, DPA and DHA, three omega-3 fatty acids, have described beneficial characteristics. This work finally highlighted the potential of the Damara breed as an interesting alternative for animal production in semiarid climates, also for the fatty acid profiles in muscle tissue (van Harten et al. 2016). From these initial studies, research moved to proteomics-based studies. In a first approach, we have conducted a study on the wool proteome on the two experimental merino groups (Almeida et al. 2014). In fact, iTRAQ was used to

study changes in the wool proteome in the Australian Merino animals used in this study. SWL showed a decrease of fibre diameter, increasing the abundance of high sulphur protein KAP13.1 and proteins from the high glycine-tyrosine protein KAP6 family, with implications on wool quality. From there, we used 2DE coupled to MS/MS mass spectrometry to identify proteins differentially expressed in the skeletal muscle (Almeida et al. 2016). This work indicated first of all relevant differences between breeds. Several proteins were suggested as putative biomarkers of tolerance to weight loss: desmin, troponin T, phosphoglucomutase and the histidine triad nucleotide-binding protein 1 (Almeida et al. 2016). Despite its originality and interesting results, the study was however hindered by several limitations characteristic of 2DE, particularly the fact of protein identification and expression which were limited to those spots visible in the gel. As such, we have recently conducted an approach based on shotgun proteomics label-free technique (Ferreira et al. 2017). This approach led to a very complete and reliable characterization of the proteome with over 650 proteins being identified, 95 with differential regulation. Interpreting such data became a daunting task; however, results could be briefly summarized. We observed that the more vulnerable to SWL a breed is, the more differential abundance proteins we find. Protein binding was the most frequently altered molecular function identified. We suggested six putative markers for restricted nutritional conditions independently of breed: ferritin heavy chain, immunoglobulin V lambda chain, transgelin, fatty acid synthase, glutathione S-transferase A2 and dihydrodiol dehydrogenase 3 like. Finally, authors proposed several proteins related to SWL tolerance: S100-A10, serpin A3-5-like and catalase. Overall, results from the proteomics studies will lead to increased stock productivity interesting for animal production, particularly if identified at the muscle level, the tissue of economic importance in meat production (Ferreira et al. 2017).

This study was furthermore completed with a thorough NMR-based metabolomics approach, conducted at the level of the muscle and the liver in these animals (Palma et al. 2016b). Overall, the analysis allowed the identification of 51 metabolites in the muscle and 46 metabolites in the liver, with significant differences between breeds and nutritional groups. Results from this study allowed a unique complementary overview as it involved the two tissues and allowed the contrast and integration with previously obtained proteomics studies. Generally speaking, it is noteworthy to mention that both Dorper and Damara breed seemed to be more adapted to SWL, showing few changes in both tissues when subjected to feed restriction. This tolerance could be a result of the breed selection history. The Merino, probably due to its selection for wool production, showed more marked changes in both tissues and seems to be the less well adapted to SWL. Main variations were observed in the glycogenolysis pathway and antioxidant activity. The Damara presented a specific set of adaptations, such as decrease in muscle development and at the glucose-alanine cycle, reflecting the physiology of its major body characteristic, the fat tail. In the context of the breed selection towards SWL tolerance, our results confirm that the Dorper and Damara breeds have performed better under SWL conditions. Their adaptation seems to be linked to a more

efficient metabolic adaptation to feed restriction, so that a change in the nutritional energy source did not compromise overall muscle structure. A possible adaptation at the rumen level should also be considered in these breeds, since they presented some variations related with rumen microflora composition and activity (Palma et al. 2016b).

4.2 NMR Metabolomics and Tolerance to Seasonal Weight Loss in Dairy Goats

The Canary Islands, off the coast of Africa, is a very interesting archipelago with different types of climate according to the location of the islands. Indeed, if islands located to the East, like Fuerteventura or Lanzarote, have a semiarid climate, islands to the West, like La Palma, have a predominantly rainy weather. Such different climates affect the agricultural practices on the islands as well as the livestock production systems (essentially based on dairy goats). As such, two major dairy goat breeds occur on the islands: the Majorera (from the island of Fuerteventura) and the Palmera (from the island of La Palma). The former display all characteristics of goat breeds of semiarid regions, whereas the latter do not. Majorera animals are therefore well adapted to SWL, contrary to Palmera goats.

We have recently conducted a 21-day trial to contrast these two breeds when subjected to weight loss (Lérias et al. 2013). There were no significant differences between both restricted groups of the two goat breeds regarding live weights or milk yield reductions (live weight reduction of 13% and milk yield reduction of 87% in both breeds). At the same time, there were major differences in the control (adequately fed) groups for both breeds, highlighting the size and milk yield differences between breeds. Interestingly, however, a first analysis of several biochemical and endocrine parameters in the plasma of these animals clearly indicated that both breeds reacted differently to experimentally induced weight loss (Lérias et al. 2015), which was also visible in a PCA (principal component analysis). These results encouraged us to proceed with omics-based studies, particularly through the use of proteomics and metabolomics.

For analysing the proteome, we have first conducted a label-free proteomics approach, in which we identified over 1000 proteins. Of these, 96 showed differential abundance between two of the groups within studied comparisons (Hernandez-Castellano et al. 2016). The results clearly demonstrated that the two breeds used different biochemical pathways to deal with undernutrition. Indeed, Majorera breed showed higher concentrations of immune system-related proteins, whereas (and in contrast) the Palmera breed showed higher abundance of proteins related to apoptosis. Overall, these results indicate that the two goat breeds have a distinct metabolic reaction to SWL and that proteins related to the immune system and apoptosis such as cadherin-13, collagen alpha-1, nidogen-2, clusterin and protein S100-A8 could be considered putative candidates as markers of tolerance

to SWL in dairy goats. The analysis was further complemented with a mitochondrial membrane proteome analysis that included blue native PAGE gels and protein identification using MALDI-TOF/TOF MS (matrix-assisted laser desorption ionization—time of flight—mass spectrometry) (Cugno et al. 2016). The proteomics analysis of mitochondria enabled the resolution of a total of 277 proteins with 148 (53%) identified by MALDI-TOF/TOF mass spectrometry. Some of the proteins were identified as subunits of the glutamate dehydrogenase complex and the respiratory complexes I, II, IV, V from mitochondria, as well as numerous other proteins with functions in metabolism, development, localization, cellular organization and biogenesis, biological regulation and response to stimulus, among others, that were for the first time mapped in both blue native and 2DE gels. Finally, the comparative proteomics analysis enabled the identification of several proteins: NADH-ubiquinone oxidoreductase 75 kDa subunit and lamin B1 mitochondrial (upregulated in the Palmera breed), guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2 (upregulated in the Majorera breed) and cytochrome b-c1 complex subunit 1, mitochondrial and chain D, bovine F1-C8 sub-complex of ATP synthase (downregulated in the Majorera breed).

Finally, the proteomics analysis was combined with an NMR-based metabolomics analysis (Palma et al. 2016a). This study included not only the mammary gland secretory tissue but also an analysis of the milk. The conducted profiling analysis led to the identification of 46 metabolites in the aqueous extract of the mammary gland. Lactose, glutamate, glycine and lactate were found to be the most abundant. Analysis of milk whey allowed the identification of 50 metabolites, the most abundant being lactose, citrate and creatine. Significant differences were observed in mammary gland biopsies and milk whey between control- and restricted-fed groups in both breeds, albeit with no differences between the breeds. Authors determined that the variations may be related to metabolism adaptation to the low-energy diet and are indicative of breed-specific microflora, providing a new insight that had not been possible when using proteomics only. Milk whey showed more metabolites varying between control and restricted groups than the mammary gland. Interestingly, the Majorera breed also displayed more variations between control- and restricted-fed groups than the Palmera breed regarding the milk metabolome profiles. This could finally be an indication of a prompt adaptation to SWL by the Majorera breed, corroborating the results previously mentioned in the proteomics and blood biochemical profiling.

Future Prospects

These illustrative case studies enhance the interest of the technique on a broader scale that involves also an integration with other omics tools, particularly proteomics and transcriptomics. For the moment, applications of NMR-based metabolomics to animal science and, on a broader perspective, to agriculture in general are still rather limited. Several factors contribute to this situation, particularly the lack of knowledge on the technique and its advantages by the majority of animal scientists and the high prices of the instruments used in this research. This is particularly troublesome in present day's context, where research funds seem to be

stagnating during the last decade. Nevertheless, the analytical power of the technique is immense, and as metabolite databases become more and more complete, NMR metabolomics will become an analytical tool in the context of livestock production research. It has numerous advantages, particularly the fact that it is easy, non-destructive, fast and relatively affordable (albeit not including NMR instrument prices) to conduct. It is likely that the outlook will change in the near future and that NMR metabolomics will become a common technique in animal science to be used in conjunction with other post-genomic tools such as transcriptomics or proteomics.

Acknowledgements Authors acknowledge funding from *Fundação para a Ciência e Tecnologia* (FCT, Lisbon, Portugal) through research project PTDC/CVT/116499/2010—Lactation and milk production in Goat (*Capra hircus*): identifying molecular markers underlying adaptation to seasonal weight loss that funded research mentioned in Sect. 4. Authors are members of COST actions FA1002—Proteomics in Farm Animals and FA1308—Dairycare to whom network funding is acknowledged. Author M. Palma acknowledges a postdoctoral grant under project ReNATURE—CENTRO-01-0145-FEDER-000007.

References

Almeida AM (2011) The Damara in the context of Southern Africa fat-tailed sheep breeds. Trop Anim Health Prod 43:1427–1441

Almeida AM, Cardoso L (2008a) Animal production and genetic resources in Guinea Bissau: I – Northern Cacheu Province. Trop Anim Health Prod 40:529–536

Almeida AM, Cardoso L (2008b) Animal production and genetic resources in Guinea Bissau: II– Tombali province. Trop Anim Health Pro 40:537–543

Almeida A, Schwalbach L, Waal H, Greyling J, Cardoso L (2006) The effect of supplementation on productive performance of Boer goat bucks fed winter veld hay. Trop Anim Health Prod 38:443–449

Almeida AM, Schwalbach LMJ, Cardoso LA, Greyling JPC (2007) Scrotal, testicular and semen characteristics of young Boer bucks fed winter veld hay: the effect of nutritional supplementation. Small Rum Res 73:216–220

Almeida AM, Kilminster T, Scanlon T, Araújo SS, Milton J, Oldham C, Greeff JC (2013) Assessing carcass and meat characteristics of Damara, Dorper and Australian Merino lambs under restricted feeding. Trop Anim Health Prod 45:1305–1311

Almeida AM, Plowman JE, Harland DP, Thomas A, Kilminster T, Scanlon T, Milton J, Greeff J, Oldham C, Clerens S (2014) Influence of feed restriction on the wool proteome: a combined iTRAQ and fiber structural study. J Proteomics 103:170–177

Almeida AM, Bassols A, Bendixen E, Bhide M, Ceciliani F, Cristobal S, Eckersall PD, Hollung K, Lisacek F, Mazzucchelli G, McLaughlin M, Miller I, Nally JE, Plowman J, Renaut J, Rodrigues P, Roncada P, Staric J, Turk R (2015) Animal board invited review: advances in proteomics for animal and food sciences. Animal 9:1–17

Almeida AM, Palhinhas RG, Kilminster T, Scanlon T, van Harten S, Milton J, Blache D, Greeff J, Oldham C, Coelho AV, Cardoso LA (2016) The effect of weight loss on the muscle proteome in the Damara, Dorper and Australian Merino Ovine Breeds. PLoS One 11:e0146367

Alves SP, Bessa RJ, Quaresma MA, Kilminster T, Scanlon T, Oldham C, Milton J, Greeff J, Almeida AM (2013) Does the fat tailed Damara ovine breed have a distinct lipid metabolism leading to a high concentration of branched chain fatty acids in tissues? PLoS One 8:e77313

Beauclercq S, Nadal-Desbarats L, Hennequet-Antier C, Collin A, Tesseraud S, Bourin M, Le Bihan-Duval E, Berri C (2016) Serum and muscle metabolomics for the prediction of ultimate pH, a key factor for chicken-meat quality. J Proteome Res 15:1168–1178

Beckonert O, Keun HC, Ebbels TM, Bundy J, Holmes E, Lindon JC, Nicholson JK (2007) Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat Protoc 2:2692–2703

Bernini P, Bertini I, Luchinat C, Nincheri P, Staderini S, Turano P (2011) Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. J Biomol NMR 49:231–243

Bertram H, Stødkilde-Jørgensen H, Karlsson A, Andersen H (2002) Post mortem energy metabolism and meat quality of porcine M. longissimus dorsi as influenced by stunning method - A (31)P NMR spectroscopic study. Meat Sci 62:113–119

Bonnefont C, Guerra A, Theron L, Molette C, Canlet C, Fernandez X (2014) Metabolomic study of fatty livers in ducks: identification by 1H-NMR of metabolic markers associated with technological quality. Poult Sci 93:1542–1552

Cardoso LA, Almeida AM (2013) Seasonal weight loss—an assessment of losses and implications on animal welfare and production in the tropics: Southern Africa and Western Australia as case studies. In: Makkar HPS (ed) Enhancing animal welfare and farmer income through strategic animal feeding. FAO Animal Production and Health Paper No 175 Foof and Agriculture Organization of the United Nations (FAO), Rome, pp 37–45

Castejón D, García-Segura J, Escudero R, Herrera A, Cambero M (2015) Metabolomics of meat exudate: its potential to evaluate beef meat conservation and aging. Anal Chim Acta 901:1–11

Ceccarelli G, Pozzo E, Scorletti F, Benedetti L, Cusella G, Ronzoni FL, Sahakyan V, Zambaiti E, Mimmi MC, Calcaterra V, Deprest J, Sampaolesi M, Pelizzo G (2015) Molecular signature of amniotic fluid derived stem cells in the fetal sheep model of myelomeningocele. J Pediatr Surg 50:1521–1527

Coco L, Papadia P, Pascali S, Bressani G, Storelli C, Zonno V, Fanizzi F (2009) Comparison among Different Gilthead Sea Bream (*Sparus aurata*) farming systems: activity of intestinal and hepatic enzymes and 13C-NMR analysis of lipids. Forum Nutr 1:291–301

Cugno G, Parreira JR, Ferlizza E, Hernández-Castellano LE, Carneiro M, Renaut J, Castro N, Arguello A, Capote J, Campos AM, Almeida AM (2016) The goat (*Capra hircus*) mammary gland mitochondrial proteome: a study on the effect of weight loss using blue-native PAGE and two-dimensional gel electrophoresis. PLoS One 11:e0151599

Dervishi E, Zhang G, Dunn SM, Mandal R, Wishart DS, Ametaj BN (2017) GC–MS metabolomics identifies metabolite alterations that precede subclinical mastitis in the blood of transition dairy cows. J Proteome Res 16:433–446

Ellinger J, Chylla R, Ulrich E, Markley J (2012) Databases and software for NMR-based metabolomics. Curr Metabolomics 1:28–40

Emwas AH (2015) The strengths and weaknesses of NMR spectroscopy and mass spectroscopy with particular focus on metabolomics research. In: Bjerrum JT (ed) Metabonomics. Methods and protocols. Humana Press. Springer Science + Business Media LLC, New York, pp 161–193

Ferreira AM, Grossmann J, Fortes C, Kilminster T, Scanlon T, Milton J, Greeff J, Oldham C, Nanni P, Almeida AM (2017) The sheep (*Ovis aries*) muscle proteome: decoding the mechanisms of tolerance to seasonal weight loss using label-free proteomics. J Proteomics 161:57–67

Gowda GA, Raftery D (2015) Can NMR solve some significant challenges in metabolomics? J Magn Reson 260:144–160

Griffin J, Shockcor J (2004) Metabolic profiles of cancer cells. Nat Rev Cancer 4:551–561

Hernández-Castellano LE, Ferreira AM, Nann IP, Grossmann J, Argüello A, Capote J, Cai G, Lippolis J, Castro N, Almeida AM (2016) The goat (*Capra hircus*) mammary gland secretory tissue proteome as influenced by weight loss: a study using label free proteomics. J Proteomics 145:60–69

Hu F, Furihata K, Kato Y, Tanokura M (2007) Nondestructive quantification of organic compounds in whole milk without pretreatment by two-dimensional NMR spectroscopy. J Agric Food Chem 55:4307–4311

Huber K, Dänicke S, Rehage J, Sauerwein H, Otto W, Rolle-Kampczyk U, von Bergen M (2016) Metabotypes with properly functioning mitochondria and anti-inflammation predict extended productive life span in dairy cows. Sci Rep 6:24642

Jégou M, Gondret F, Lalande-Martin J, Tea I, Baéza E, Louveau I (2016) NMR-based metabolomics highlights differences in plasma metabolites in pigs exhibiting diet-induced differences in adiposity. Eur J Nutr 55:1189–1199

Kadarmideen H (2014) Genomics to systems biology in animal and veterinary sciences: progress, lessons and opportunities. Livest Sci 166:232–248

Klein MS, Almstetter MF, Schlamberger G, Nürnberger N, Dettmer K, Oefner PJ, Meyer HH, Wiedemann S, Gronwald W (2010) Nuclear magnetic resonance and mass spectrometry-based milk metabolomics in dairy cows during early and late lactation. J Dairy Sci 93:1539–1550

Kodani Y, Miyakawa T, Komatsu T, Tanokura M (2017) NMR-based metabolomics for simultaneously evaluating multiple determinants of primary beef quality in Japanese Black cattle. Sci Rep 7:1297

Lamy E, van Harten S, Sales-Baptista E, Guerra M, Almeida AM (2012) Environmental stress and amelioration. In: Sejian V, Naqvi SMK, Ezeji T, Lakritz J, Lal R (eds) Environmental stress and amelioration in livestock production. Springer, Berlin, pp 19–51

Lérias JR, Hernández-Castellano LE, Morales-Delanuez A, Araújo SS, Castro N, Argüello A, Capote J, Almeida AM (2013) Body live weight and milk production parameters in the Majorera and Palmera goat breeds from the Canary Islands: influence of weight loss. Trop Anim Health Prod 45:1731–1736

Lérias JR, Peña R, Hernández-Castellano LE, Capote J, Castro N, Argüello A, Araújo SS, Saco Y, Bassols A, Almeida AM (2015) Establishment of the biochemical and endocrine blood profiles in the Majorera and Palmera dairy goat breeds: the effect of feed restriction. J Dairy Res 82:416–425

Levitt MH (2008) Spin dynamics—basics of nuclear magnetic resonance. Wiley, New York. 740 pp

Li J, Wijffels G, Yu Y, Nielsen L, Niemeyer D, Fisher A, Ferguson DM, Schirra H (2011) Altered fatty acid metabolism in long duration road transport: an NMR-based metabonomics study in sheep. J Proteome Res 10:1073–1087

Lu J, Fernandes E, Cano A, Vinitwatanakhun J, Boeren S, Hooijdonk T, Knegsel A, Vervoort J, Hettinga K (2013) Changes in milk proteome and metabolome associated with dry period length, energy balance, and lactation stage in postparturient dairy cows. J Proteome Res 12:3288–3296

Maher AD, Coles C, White J, Bateman JF, Fuller ES, Burkhardt D, Little CB, Cake M, Read R, McDonagh MB, Rochfort SJ (2012) 1H NMR spectroscopy of serum reveals unique metabolic fingerprints associated with subtypes of surgically induced osteoarthritis in sheep. J Proteome Res 11:4261–4268

Mannina L, Sobolev AP, Capitani D, Iaffaldano N, Rosato MP, Ragni P, Reale A, Sorrentino E, D'Amico I, Coppola R (2008) NMR metabolic profiling of organic and aqueous sea bass extracts: implications in the discrimination of wild and cultured sea bass. Talanta 77:433–444

Mercier P, Lewis M, Chang D, Baker D, Wishart D (2011) Towards automatic metabolomic profiling of high-resolution one-dimensional proton NMR spectra. J Biomol NMR 49:307–323

Mickiewicz B, Heard BJ, Chau JK, Chung M, Hart DA, Shrive NG, Frank CB, Vogel HJ (2015) Metabolic profiling of synovial fluid in a unilateral ovine model of anterior cruciate ligament reconstruction of the knee suggests biomarkers for early osteoarthritis. J Orthop Res 33:71–77

Moestue S, Sitter B, Bathen T, Tessem MB, Gribbestad I (2011) HR MAS MR spectroscopy in metabolic characterization of cancer. Curr Top Med Chem 11:2–26

Nicholson J, Wilson I (2003) Opinion: understanding 'global' systems biology: metabonomics and the continuum of metabolism. Nat Rev Drug Discov 2:668–676

Nyberg N, Nielsen M, Jaroszewski J (2010) Metabolic trajectories based on 1H NMR spectra of urines from sheep exposed to nutritional challenges during prenatal and early postnatal life. Metabolomics 6:489–496

Palma M, Hernández-Castellano LE, Castro N, Arguëllo A, Capote J, Matzapetakis M, Almeida AM (2016a) NMR-metabolomics profiling of mammary gland secretory tissue and milk serum in two goat breeds with different levels of tolerance to seasonal weight loss. Mol BioSyst 12:2094–2107

Palma M, Scanlon T, Kilminster T, Milton J, Oldham C, Greeff J, Matzapetakis M, Almeida AM (2016b) The hepatic and skeletal muscle ovine metabolomes as affected by weight loss: a study in three sheep breeds using NMR-metabolomics. Sci Rep 6:39120

Pears MR, Salek RM, Palmer DN, Kay GW, Mortishire-Smith RJ, Griffin JL (2007) Metabolomic investigation of CLN6 neuronal ceroid lipofuscinosis in affected South Hampshire sheep. J Neurosci Res 85:3494–3504

Rezzi S, Giani I, Héberger K, Axelson D, Moretti V, Reniero F, Guillou C (2007) Classification of gilthead sea bream (*Sparus aurata*) from 1H NMR lipid profiling combined with principal component and linear discriminant analysis. J Agric Food Chem 55:9963–9968

Roth K (1984) Basic principles of NMR spectroscopy in Medicine. Springer, Berlin Heidelberg, pp 4–28

Sacco D, Brescia MA, Sgaramella A, Casiello G, Buccolieri A, Ogrinc N, Sacco A (2009) Discrimination between Southern Italy and foreign milk samples using spectroscopic and analytical data. Food Chem 114:1559–1563

Salek R, Cheng KKK, Griffin J (2011) The study of mammalian metabolism through NMR-based metabolomics. Methods Enzymol 500:337–351

Scanlon TT, Almeida AM, van Burgel A, Kilminster T, Milton J, Greeff JC, Oldham C (2013) Live weight parameters and feed intake in Dorper, Damara and Australian Merino lambs exposed to restricted feeding. Small Rum Res 109:101–106

Scano P, Rosa A, Incani A, Maestrale C, Santucciu C, Perra D, Vasceallari S, Pani A, Ligios C (2015) (1)H NMR brain metabonomics of scrapie exposed sheep. Mol BioSyst 11:2008–2016

Standal I, Praël A, McEvoy L, Axelson D, Aursand M (2008) Discrimination of cod liver oil according to wild/farmed and geographical origins by GC and 13C NMR. J Am Oil Chem Soc 85:105–112

Sun Y, Xu C, Li C, Xia C, Xu C, Wu L, Zhang H (2014) Characterization of the serum metabolic profile of dairy cows with milk fever using 1H-NMR spectroscopy. Vet Q 34:159–163

Sundekilde UK, Poulsen NA, Larsen LB, Bertram HC (2013) Nuclear magnetic resonance metabonomics reveals strong association between milk metabolites and somatic cell count in bovine milk. J Dairy Sci 96:290–299

van Harten S, Kilminster T, Scanlon T, Milton J, Oldham C, Greeff J, Almeida AM (2016) Fatty acid composition of the ovine longissimus dorsi muscle: effect of feed restriction in three breeds of different origin. J Sci Food Agric 96:1777–1782

Viegas I, Mendes VM, Leston S, Jarak I, Carvalho RA, Pardal MA, Manadas B, Jones JG (2011) Analysis of glucose metabolism in farmed European sea bass (*Dicentrarchus labrax* L.) using deuterated water. Comp Biochem Physiol A Mol Integr Physiol 160:341–347

Wang LF, Jia SD, Yang GQ, Liu RY, Yang GY, Li M, Zhu HS, Wang YY, Han LQ (2016) The effects of acute lipopolysaccharide challenge on dairy goat liver metabolism assessed with (1) HNMR metabonomics. J Anim Physiol Anim Nutr (Berl) 101:180–189

Worley B, Powers R (2012) Multivariate analysis in metabolomics. Curr Metabolomics 1:92–107

Xu C, Sun LWW, Xia C, Zhang HYY, Zheng JSS, Wang JSS (2016) (1)H-nuclear magnetic resonance-based plasma metabolic profiling of dairy cows with fatty liver. Asian-Australas J Anim Sci 29:219–229

Yang Y, Zheng N, Zhao X, Zhang Y, Han R, Yang J, Zhao S, Li S, Guo T, Zang C, Wang J (2016) Metabolomic biomarkers identify differences in milk produced by Holstein cows and other minor dairy animals. J Proteomics 136:174–182

Omics and Systems Biology: Integration of Production and Omics Data in Systems Biology

Kasper Hettinga and Lina Zhang

Abstract Omics technologies have become of mainstream use in the study of farm animals, to better understand the physiology of the animal and the quality of the products produced by those animals. Such studies can be done at the level of genes, transcripts, proteins and/or metabolites. An important aspect of doing such omics studies is understanding of variation. For example, in relation to parity, lactation, feeding status and animal health, variation can happen in transcripts, proteins or metabolites found in farm animals and the products produced. This variation can help in better understanding the physiology of the animal. Also variation between individual animals exists, which may assist in better understanding of the animal's physiology. One limitation of the majority of the studies in this area is that they are performed using one specific omics technology. Integrating omics data captured using multiple omics technologies, using a systems biology approach, can shed more light on the biochemistry of the farm animal's physiology. At the end of this chapter, the outlook on such studies and the (software) developments that would be needed for optimal integration of omics data is discussed.

Keywords Genomics • Transcriptomics • Proteomics • Metabolomics • Interactomics • Systems biology • Computation biology • Farm animal • Milk • Biochemistry

K. Hettinga (✉)
Dairy Science & Technology, Wageningen University & Research, Wageningen, The Netherlands
e-mail: kasper.hettinga@wur.nl

L. Zhang
Center for Translational Biomedical Research, University of North Carolina, Greensboro, NC, USA

© Springer International Publishing AG 2018

A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_22

Abbreviations

CN	Casein
DGAT1	Diglyceride acyltransferase 1
DNA	Deoxyribonucleic acid
GWAS	Genome-wide association study
IgA	Immunoglobulin A
miRNA	MicroRNA
mRNA	Messenger RNA
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
SNP	Single nucleotide polymorphisms

1 Introduction to Omics Technologies

In the past decades, many omics technologies have been developed for studying biology on different levels, from genes to metabolites. Figure 1 gives an overview of the different levels of biology with the associated omics technologies available. On the metabolite level, different omics technologies are applied, depending on the target metabolites of interest, with the term metabolomics mainly used for small molecules that are part of the core metabolism of organisms.

Due to continuous technological improvements, more comprehensive technology has been, and is, developed on all these levels. These improvements are leading to a higher resolution of analysis, a higher throughput and all that at a lower cost of analysis. These techniques have become more widely used by scientists around the world in many disciplines, including scientists in animal and food science. In this chapter, the application of the different omics technologies will be explained, to indicate how these can be used to better understand the physiology of farm animals. Most of the omics research focus on better understanding the production characteristics of farm animals (e.g. how much product is produced, what is the quality of the product produced) or animal physiology (e.g. understanding the mammary

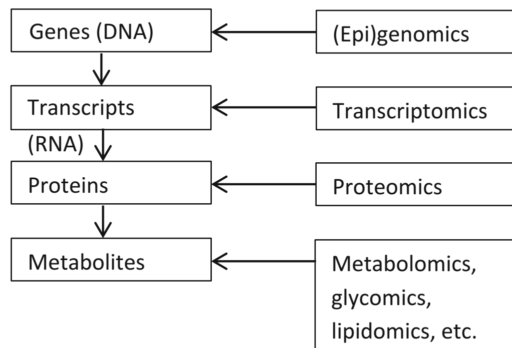

Fig. 1 Overview of biology from gene to metabolite, with the associated omics technologies that can be used for studying the different levels

gland physiology, health of the animal). In this first section, the use of different omics technologies is briefly explained.

However, as already indicated in Fig. 1, biology does not happen on a single level but is an interaction between different levels, from genes, through transcripts and proteins, to metabolites. To perform integrative research on all these levels, combining multiple omics technologies aids in obtaining a better insight in the biology compared to studying the biology on a single level. Such combined approaches can range from studies combining two levels of biology, all the way up to full integration over all these levels. This integrative approach is the basis of systems biology research. In this chapter, the application of omics technologies will be described starting from single omics technologies up till full integration from a system's biology perspective. The chapter will finish with challenges and future developments that are envisaged for further research in this area.

1.1 Genomics

The effect of genes of farm animals on production characteristics and animal physiology has mostly been studied by linking these outcome parameters to variation in gene sequences. Most research on farm animals is done by looking either at a large number of single nucleotide polymorphisms (SNPs, single mutations at specific positions in the genome) or by sequencing target genes that are expected to be involved in the outcome of interest. More recently, whole genome sequencing has become more popular, partly due to a reduction in the cost of performing such analyses. This allows using information about the whole genome sequence, making the translation of findings to the underlying mechanism easier.

When determining many SNPs in the animal's genome, association studies can be done to link these SNPs to specific traits (e.g. milk yield). This is usually conducted in the form of a genome-wide association study (GWAS). Many of such studies have been performed on a large scale in farm animals (Schennink et al. 2009).

One disadvantage of studies based on GWAS is that it is relatively difficult to determine what the causative gene/causative mutation is underlying the found relation. A GWAS will only indicate genomic regions associated with the trait, in which many genes may still be present. Based on screening the genes in the genomic region, and trying to find genes that may be causally associated with the trait, a hypothesis on the role of a specific gene could be made. Once these potentially causative genes are found based on a GWAS, the genes can be sequenced to search for specific mutations (Duchemin et al. 2014; Schennink et al. 2009). This is where whole genome sequencing, as mentioned above, has an advantage, as the full genomic information is collected, so mutations in any gene can be searched for.

A more recent development in the field of genomics is looking at chemical modifications of DNA instead of variations in the sequence, which is called epigenomics. Chemical modifications, such as methylation, of the DNA strand

may alter the propensity of the gene to transcription (Singh et al. 2012). Epigenomics thereby provide a link between the genomics studies, as mentioned above, to transcriptomics as will be discussed in the next section.

1.2 Transcriptomics

Whereas genomics studies the genetic information as such, transcriptomics studies the transcription of genes into mRNA. Before proteins/enzymes are formed, transcription of the genes is the first step towards expressing the activity originating from the genes. In transcriptomics studies, the mRNA molecules present in cells are isolated and characterized. In the past, this was usually done using hybridization-based microarrays. The main disadvantage of such a microarray approach is that it is required that the target sequences are known before the experiment is done. To solve this, RNA-seq has been developed, in which nontargeted RNA sequencing is performed (Kukurba and Montgomery 2015). For studying farm animals, both approaches (microarrays and RNA sequencing) are currently being used.

Transcriptomics is extensively used in farm animal studies. Studies have, for example, focused on differences in the transcriptome during illness (Moran et al. 2016; Younis et al. 2016) or when performing experimental, for example, pharmaceutical, interventions (McCoard et al. 2016). Another area where transcriptomics is frequently used is to compare animals that differ in production characteristics (Bai et al. 2016; Shen et al. 2016; Wall et al. 2013). In all these studies, the objective is usually to determine how animals that differ in specific output parameters differ in the transcription of their genes. This can then be translated back to the known function of genes, to be able to study the underlying mechanism. The data from such experiments can be used to determine whether specific pathways are up- or downregulated in response to the studied contrast in output parameter.

Besides mRNA-based transcriptomics, other types of RNA transcripts may also be studied. One example is microRNA, which are short pieces of non-coding RNA that can influence transcriptional activity, often being involved in reduced transcription. The transcriptional regulation of a specific miRNA can be targeted at many genes simultaneously, thus giving a broad range of possible functions to individual miRNA molecules. Levels of miRNA have been related to animal physiology (Salilew-Wondim et al. 2016; Ioannidis and Donadeu 2016). These miRNAs may not only have local effect in the cell of synthesis but have also been shown to be transported through serum, with exosomes being plausible vectors in which miRNAs are transported (Zhao et al. 2016). Because of the potential broad range of effects on transcription of many genes, as well as their ability to be transported across cells in the organism, changes in miRNA levels can have many consequences, both in local and systemic physiology. These miRNAs can have influences within an individual but can also be transported across individuals, for example, through milk. For milk, transport of miRNA through exosomes has been shown to occur. Such milk-borne miRNAs may function as transcriptional regulators in the newborn (Perge et al. 2016). In addition miRNAs may also play a

role across species; the presence of these components in farm animal (e.g. dairy) products may also be relevant for product quality, as these food-derived miRNAs may have specific transcriptional consequences in the consumer (Benmoussa et al. 2016; Kirchner et al. 2016).

1.3 Proteomics

The next step after transcription of genes into mRNA is the translation of the resulting mRNA into proteins. As mentioned above, this process can be influenced by factors such as miRNA. In this step, highly abundant (food) proteins may be produced, as well as low-abundant proteins which have a wide range of functions across the whole physiology of the farm animal.

Of the highly abundant food productions, much research has been done on the major milk proteins, such as caseins and the major whey proteins (α-lactalbumin and β-lactoglobulin). The genes associated with these proteins have been studied in detail in several dairy animals, showing large genetic variation. The variations have been linked to differences in milk and protein composition (Heck et al. 2009; Buitenhuis et al. 2016). This variation not only exists on the level of protein abundance but also exists on the level of post-translational modifications (phosphorylation and glycosylation). This information is actively used, both directly through genomic selection as well as indirectly through traditional breeding programmes (Hayes et al. 2009) to improve the quality of the farm animal products produced. Besides using it for improving quality through animal breeding, proteomics can also be used to study product quality in more detail. An example of this is meat, in which proteomics has been used for monitoring meat quality throughout the whole production chain (Paredi et al. 2013). It has also been used for cheese, in which proteomics was used to monitor the progress of ripening by looking at the degradation of the major milk proteins (Hinz et al. 2012).

For the low-abundant proteins in farm animals, and their products, these have been studied in relation to many physiological disturbances in such animals (Bendixen et al. 2011; Almeida et al. 2015). Many diseases may occur in farm animals, although most interest has traditionally gone to production-related diseases (Nir Markusfeld 2003), because these occur most frequently and usually have large economic consequences for the farm animal sector. Proteomics has many applications in this area. First, it is often used in relation to studying the physiological effects of problems with animal health. In dairy animals, it has often been used to study the physiological response to mastitis, which also may include the response to specific pathogens, as reviewed by (Boehmer 2011). These researches have shown the up- and downregulation of many proteins, indicating a decrease in milk protein synthesis with a concomitant increase in proteins that are known to support the host defence system.

But proteomics is not limited to the traditional farm animals, it has also been applied in, for example, aquaculture (Rodrigues et al. 2012). As with the traditional farm animals, in aquaculture the focus is on both product quality and animal health.

In aquaculture, diseases can have a large impact on the amount of product that is produced (in other words, the growth rate of the fish). In addition, studying the response of fish to pathogens (as just before described for the case of mastitis in dairy animals) can help unravel the physiology of diseases and thereby lead to a better understanding of the host defence (Zhou et al. 2011). Another topic that has been studied in fish, as with farm animals living on land, is the response to stress, being one of the main determinants of animal welfare but also having an impact on the quality of the product produced (Morzel et al. 2006).

1.4 Metabolomics

As described above, the components of the proteome (protein/enzymes) are involved in the animal's metabolism. Besides studying such enzymes directly using proteomics technology, research can also focus on the metabolites that are produced by the enzymes of interest. Although the metabolites are several steps away from the genes of the cow, research has shown that there are many correlations that can be detected across all levels, from genes to metabolites (Wittenburg et al. 2013).

Many different categories of metabolites exist that can be studied, including lipids, water-soluble metabolites and volatile metabolites. These different categories of metabolites all require their own analytical approach for detection (Wang et al. 2010).

The research areas, in which metabolomics is applied, are for a large part similar to those of transcriptomics and proteomics. The underlying research themes are thus often the same but aiming at components on a different level of the animal's physiology. An example of a specific area of research in metabolomics that is less studied than the other levels is the rumen. Many farm animals are ruminants, and in their rumen a wide range of microorganisms are present to breakdown plant material eaten by the animals. These microorganisms produce a whole range of metabolites that can be detected with metabolomics technologies (Zhao et al. 2014). Although these metabolites are thus not produced by the ruminant itself, they can end up in the body of the animal and thereby in animal food products such as meat and milk. Some of these metabolites in the animal (or its products) can thus be used as a reflection of the metabolic state of the rumen (Antunes-Fernandes et al. 2016). Another area of research in which metabolomics (but also transcriptomics and proteomics) has frequently been used is bovine mastitis. The aim is again to understand the physiology but now on the level of metabolites present (Sundekilde et al. 2013). In poultry research, metabolomics is also used for improving physiological understanding. An example of such research is the production disease ascites syndrome, for which the metabolomic response by chicken has been described (Shen et al. 2014). Although the above examples focus on physiological understanding, the differences in metabolites cannot only be used to explain such metabolic perturbations but also be frequently used for biomarker research. Detection of individual metabolites as biomarker of diseases is often used, and

metabolomics research can contribute to finding such biomarkers in complex samples from farm animal (products).

One specific category of metabolites that is often studied in all different farm animals is the class of lipids. This class includes different categories of lipids and lipid-soluble components [fatty acids, triglycerides, phospholipids, sterols (Sokol et al. 2015)]. Lipids in farm animal products such as milk is an obvious area in which lipid metabolites are studied (Lu et al. 2013, 2015; Sokol et al. 2015; Li et al. 2017), but also blood lipids have been studied in relation to health (Li et al. 2017; Gerspach et al. 2017). In both cases, the lipids are relevant as part of product quality as well as the physiology of the farm animal.

2 How Omics Technologies Can Help in Better Understanding Production Characteristics and Animal Physiology

In the first section of this chapter, many omics technologies are mentioned that have been applied in the study of farm animals. This second section focuses more in detail on how the omics technologies can be applied to better understand the farm animal. This will be done based on three different research directions. First, the combination of variation in the omics data between individual farm animals studied can be used to better understand the animal as such. Capturing the variation requires quantitative omics technologies. Especially in products of farm animals, like milk, there are many examples of capturing variation to better understand the animals producing the product. A second option to better understand the animal and its physiology is the comparison between species. Different farm animal species producing similar products also have similar underlying physiology. Omics technologies applied to study differences among species can help understand these animals better. Finally, omics technologies are often applied to better understand the functioning of specific organs. In this section, especially the morphology of the mammary is used as an example of such research approaches.

2.1 Importance of Capturing Variation

As mentioned above in the introduction of this section, capturing variation is a tool to better understand farm animals and their products. A clear example of such an approach is the study of milk. Milk is a complete and complex food suited to the requirements for the growth and development of the neonate. Milk and dairy products are also central elements in the human diet. The principal function of milk is providing energy and nutrients. For understanding of how farm animal are able to produce these necessary milk components, all levels of physiology can be studied using omics technologies, focusing on capturing variation at all these levels.

Milk yield and milk composition, including milk lactose, protein and fat concentration, for example, have been shown to have large variation with changes in environmental temperature (Alstrup et al. 2016). All casein fractions, except for γ-casein, were present at lower concentration in summer than in winter, whereas immunoglobulins, serum albumin were present at higher concentrations in summer than winter. A consequent worsening of milk coagulation properties was observed in summer season, which may influence cheese production from such milk. In addition, a mild effect of season was observed for milk somatic cell count, with higher values in summer than in the winter and spring (Bernabucci et al. 2015). This data suggests that the risk of mastitis is higher in summer than in other seasons. Linking the variation in milk composition to the genes of the dairy cow, it was shown that the centromeric region of bovine chromosome 14 was strongly associated with test day fat percentage. Several SNPs were associated with eicosapentaenoic acid, docosapentaenoic acid, arachidonic acid, rumenic acid and linolenic acid (Ibeagha-Awemu et al. 2016b). This study also reported some novel potential candidate genes, such as ERCC6, TONSL, NPAS2, ACER3, ITGB4, GGT6, ACOX3, MECR, ADAM12, ACHE, LRRC14, FUK, NPRL3, EVL, SLCO3A1, PSMA4, FTO, ADCK5, PP1R16A and TEP1, which may be involved in complex dairy traits, including milk traits and mammary gland functions (Ibeagha-Awemu et al. 2016b). Another example is that DGAT1 gene mutation is related to milk with changes in saturated, unsaturated and omega-3 fatty acid concentrations. Moreover, milk fat composition also differs between seasons. Summer bovine milk contains higher amounts of unsaturated fatty acids and lower amounts of saturated fatty acids compared with winter bovine milk (Duchemin et al. 2013). In addition, milk fat composition changed with the increase of lactation, for example, C18:1 fatty acid in bovine milk (Samková et al. 2012). This variation in milk fatty acid composition is mainly linked to feed composition. However, part of this variation is also linked to gene x environment interaction (Duchemin et al. 2013). These studies together show that genomics studies are important to better understand the variation in milk metabolites. By this improved understanding, a larger part of the captured variation can be explained.

Proteins in farm animal products not only provide nutrition through the presence of essential amino acids, but many proteins in the farm animal and its products are also involved in the development of the immune system. The investigation of the changes of the proteome provides information on the frequency, onset and progression of different markers (e.g. proteins) due to exogenous (e.g. season, disease) and endogenous (age, lactation) factors, as mentioned in the first section of this chapter. To achieve these objectives, the variation in these proteins needs to be determined using quantitative proteomics technologies. Information from such studies may help farmers to better manage their animals, improving milk yield and providing high-quality milk and milk products for human consumption.

The composition of farm animal (products) shows large variations depending on animal age and for dairy animals especially also over lactation stages. Looking specifically at milk over lactation, the milk host defence-related proteins, such as immunoglobulins, decreased remarkably from colostrum to mature milk in both human and bovine milk (Zhang et al. 2015a, b, 2016b). In addition to these

well-known proteins, many other low-abundant immune-related proteins also decreased from early lactation stage to middle lactation stage, such as complement proteins, lactoferrin, osteopontin, glycosylation-dependent cell adhesion molecule 1, alpha-1-acid glycoprotein 1 and protease inhibitors (Zhang et al. 2013, 2015b, 2016b; Korhonen 2009). On the other hand, lipid transport proteins, including apolipoprotein A-I, A-IV and C-III, were shown to increase from early to middle lactation (Korhonen 2009). In late lactation, proteins related to milk fat synthesis (e.g. adipophilin, fatty acid-binding protein, butyrophilin) and proteins related to lactose synthesis (e.g. α-lactalbumin and β-1,4-galactosyltransferase) were shown to decline (Zhang et al. 2015b, 2016b; Lu et al. 2014), whereas the immune-related proteins increased at this late lactation stage, suggesting the decrease in milk synthesis and a concomitant increase in the protection by immune-related proteins of the mammary gland during involution (Boggs et al. 2015, 2016).

The milk proteins not only show variation over lactation but also show variation depending on breed and genotype (Lu et al. 2015), with also unexplained variation between individual animals (Zhang et al. 2015a, b). Milk from Danish Holstein cows was mainly characterized by higher relative contents of β-casein, α-lactalbumin and β-lactoglobulin and a higher fraction of glycosylated κ-casein, whereas milk from Danish Jersey cows was characterized by higher relative contents of κ-casein, α_{S2}-casein and the less phosphorylated forms of α_{S1}-casein and α_{S2}-casein (Poulsen et al. 2016).

Some of the genetic variability that is known to impact milk composition has been studied from a proteomics perspective. Genotypic variation in the DGAT1 gene (K232A polymorphism) was shown to induce changes in expression of the lipid synthesis-related protein stomatin (Lu et al. 2015). Moreover, genotype variation resulted in the differences in post-translational modifications of milk proteins, which could be related to milk coagulation properties. Poulsen et al. (2016) found that milk from cows with κ-CN BB genotype had relative higher contents of both unglycosylated κ-CN and glycosylated κ-CN compared with that of κ-CN AA (Poulsen et al. 2016).

With respect to the variation between individual animals, this has been studied, for example, in the milk protein of dairy cows. To study this individual variation, we collected proteomics data according to Zhang et al. (2015a) of 17 individual healthy cows in mid-lactation. This unpublished data shows that there was a relative high overlap (80%) in the qualitative milk proteome; however, at the quantitative level, there was a large variation in relative protein concentrations among individual cows (Fig. 2). The variation of relative protein concentration between individual cows in mid-lactation was discussed by Zhang (2015). This quantitative variation in the milk proteins between individual animals is probably due to a multiple factors. Parity/age of cows may result in changes of milk serum proteins. For example, in bovine milk, β-lactoglobulin and immunoglobulins were positively correlated with cow's age, and bovine serum albumin increased from the first to fourth parity followed by a decline as cows became older (Ng-Kwai-Hang et al. 1987).

The variation in the concentration of milk metabolites has also been reported. A metabolomic study in the milk between heat-stress-free and heat-stressed dairy

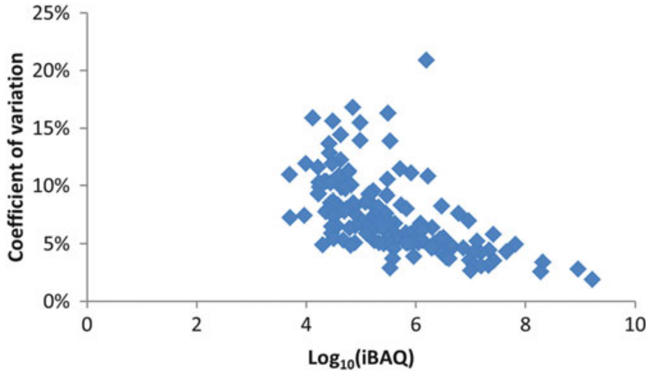

Fig. 2 Quantitative variation of milk proteins among 17 healthy (SCC <250,000) individual cows in mid-lactation (lactation stage from day 112 to day 247). Proteomics analysis was performed according to Zhang et al. (2015a)

cows showed a total of 53 discriminating metabolites that were significantly up- or downregulated in the heat-stressed group compared with the heat-stress-free group, respectively. These metabolic biomarkers were involved in pathways of carbohydrate, amino acid, lipid and gut microbiome-derived metabolism (Tian et al. 2016).

All in all, these previous researches show that by studying the variation in farm animals from the different omics approaches can help better understand the farm animal physiology, as well as how this relates to the products produced from or by the animal.

2.2 *Comparison Between Species*

Apart from variation within species, as discussed above, also differences between species can be used to better understand the farm animal physiology.

When it comes to genetic differences between species, these are relatively broad. Genetic similarity between species depends on the evolutionary origin of these species and how far back in time a common ancestor existed. At the moment, as far as we are aware, these genetic differences between species have not related to physiological differences between species in scientific research. The same is for transcriptomic differences between species.

Similar to capturing variation, comparison between species has also been comprehensively investigated in milk. It was previously shown that milk composition differs between mammalian species (Yangilar 2013). Of the mammals studied, human milk contained the highest amount of lactose and the lowest amount of proteins, especially casein. Sheep milk, on the other hand, is very high in both fat and casein content (Yangilar 2013). Differences between species not only exist on the level of macronutrients but also exist when comparing different omics dataset between species.

The differences in the milk proteome between species have frequently been reported. For example, in milk serum, β-lactoglobulin is the most abundant protein in bovine and caprine milk serum (Tsiplakou and Zervas 2013), whereas it is absent in human and camel milk serum (Zhang et al. 2016a, b). Of the low-abundant host defence proteins, IgG is the predominant antibody in bovine, caprine and camel milk, while it is IgA in human milk (Zhang et al. 2016b; Stelwagen et al. 2009; Sanchez-Macias et al. 2014). IgG can be transferred to the foetus prior to birth in humans but not in several animals, for example, ruminants such as cattle and sheep (Stelwagen et al. 2009). Furthermore, even the lower abundant proteins, such as complement proteins, antibacterial proteins, acute phase proteins, blood coagulation proteins and protease inhibitors, also are present in very different concentrations between species (D'Auria et al. 2005; Hettinga et al. 2011).

In addition to the qualitative and quantitative proteome differences between species, the changes of milk proteins over lactation stages also differ between species. Transport proteins, enzymes and immunologically active proteins are three dominant protein groups which were shown to change differently over lactation between bovine milk and human milk (Zhang 2015). In particular the immunologically active proteins decreased more rapidly in bovine milk than in human milk in early lactation (Fig. 3). The differences in the changes of these three groups of proteins over lactation can be related to the differences in the needs

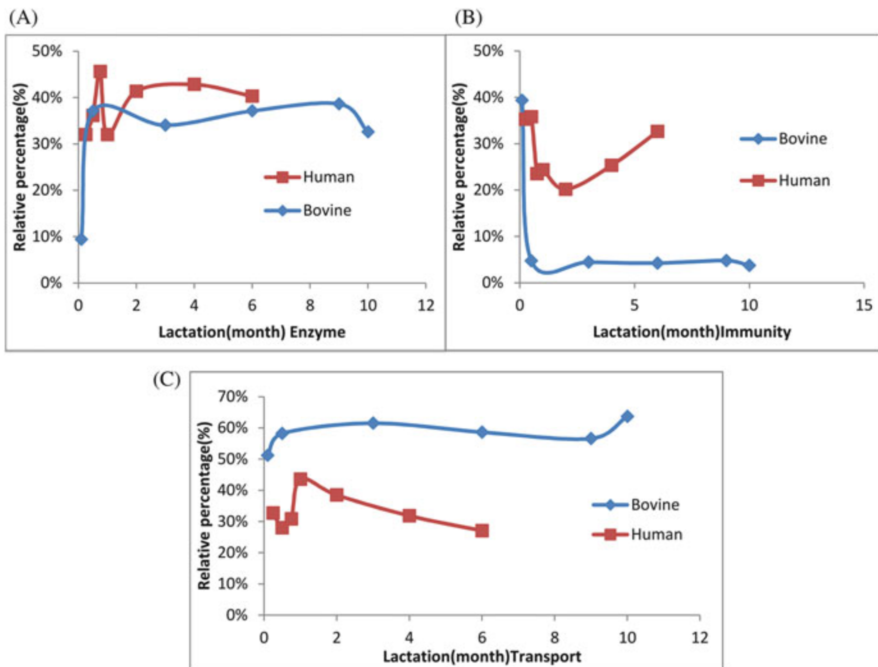

Fig. 3 The changes of enzymes (**a**), immune proteins (**b**) and transport proteins (**c**) in both human and bovine milk over lactation (Zhang et al. 2017)

between infants and calves (Hettinga et al. 2011). This knowledge on variation between species can be used to better understand the role of milk in the development of the newborn mammal. The rapid decrease in immunologically active proteins, for example, is related to the quick production of these proteins by the calf, requiring less support from milk. When designing feed for young animals or infant formula for newborn babies, this knowledge is useful in deciding on the required quality of these milk-based products.

Of the metabolites in milk, the lipids were also shown to differ between species. A comparison of milk fat composition between cow, buffalo, donkey, sheep and camel showed that the total fatty acid composition were quite similar between species. However, the sn-2 fatty acid, triacylglycerol (TAG), phospholipid and phospholipid fatty acid compositions and melting and crystallization profiles were very different between species (Zou et al. 2013).

In addition, other milk metabolites were also shown to differ between species (Qian et al. 2016). These differences in the milk metabolites were mainly clustered into four groups: (1) nonesterified fatty acids, (2) free amino acids, (3) tricarboxylic acid intermediates and (4) free carbohydrates. Metabolic differences between species have been used to distinguish milk from different dairy animals. For example, choline and succinic acid were only identified in milk from Holstein cows but not in milk of Jersey cows, yak, buffalo, goat, camel and horse. Glycerophospholipid metabolism as well as valine, leucine and isoleucine biosynthesis were similar among ruminant animals (Holstein, Jersey, buffalo, yak and goat), and biosynthesis of unsaturated fatty acids was similar among the non-ruminant animals (camel and horse), as shown by Yang et al. (Yang et al. 2016). This indicates that the metabolism of different dairy animals differs, which may be due to differences in the level of milk synthesis and differences in the milk that is produced.

The above studies show how identifying differences between species helps in better understanding different farm animals. Also, the need of the newborn when it comes to nutrition and host defence can be established through such omics studies. These studies also indicate that different dairy animals differ in their metabolic activity. A special case is the study of human milk. Although human milk is not a commercial product, knowledge on its composition can help in designing optimal replacer for human milk (infant formula), which are commonly produced based on farm animal milks (mainly of bovine and caprine origin).

2.3 Omics Studies in the Morphology, Development and Regulation of the Mammary Gland in Health, Disease and Production

Innovative and high-throughput technologies such as genomics, transcriptomics, proteomics and metabolomics can be used to better understand the functioning of

organs in general. One area where this has been applied is in getting much broader and more detailed knowledge on the morphology, development and regulation of the mammary gland. Previous research has studied the mammary gland in a healthy situation, and when it is in a diseased state, giving more information that aids in optimal production for adequate management of dairy farming.

One of the omics technologies applied to study the mammary gland is epigenomics. It was previously shown that epigenetic regulation, by, for instance, DNA methylation or histone modifications (methylation and acetylation), has been addressed as a non-genetic mechanism of regulating mammary function, as explained in the first section of this chapter. A substantial proportion of unexplained phenotypical variation in the dairy cattle has been claimed to be involved in epigenetic regulation, which should also be considered when studying milk production management practices to optimize production (Singh et al. 2010).

Stress and disease of the mammary gland are important factors that can influence milk production and milk quality. Connor et al. (2008) investigated the physiological changes occurring within the mammary gland during stress induced by more frequent milking. Changes in gene expression related to cell proliferation and differentiation, extracellular matrix (ECM) remodelling, metabolism, nutrient transport and immune function were found (Connor et al. 2008). Besides stress, disease is also often studied. Mastitis is the most devastating disease causing staggering economic losses worldwide to the dairy industry (Kaneene and Scott Hurd 1990). It can be caused by a wide range of organisms, including bacteria, fungi and algae. Transcriptomics studies have provided growing evidence that *E. coli*-induced mastitis causes a far higher expression/regulation of TLR genes, especially TLR2 and TLR4 genes, when compared to *S. aureus* (Yang et al. 2008). Proteomics enabled the detection of the increase of immune-related proteins, immunoglobulins, cathelicidins, lactoferrin, lactadherin, alpha-1-acid glycoprotein and serpin A3-8 in the milk from cows with mastitis (Zhang et al. 2015c; Yang et al. 2009). Monitoring the differences in the milk proteome between healthy animals and animals with disease may help to identify disease-related biomarkers. PTGDS was hypothesized to be a biomarker in bulk milk for mastitis, due to its high correlation with the principally accepted indicator for mastitis (somatic cell count) (Zhang et al. 2015c). Recently, metabolomics was also applied in the investigation of the mammary gland's response to infection. Components of bile acid metabolism, linked to the FXR pathway-regulating inflammation, were found to be increased during mammary gland infections. Furthermore, metabolites mapped to carbohydrate and nucleotide metabolism showed a decreasing trend in concentration up to 81 h post-challenge, whereas an increasing trend was found in lipid metabolites and di-, tri- and tetrapeptides up to the same time point, suggesting the degradation of milk proteins during mastitis (Thomas et al. 2016).

3 The Benefit of Combining Data from Different Omics Technologies into a Systems Biology Approach

In research, omics techniques are widely used as single approaches to study changes in animals on the level of genes, transcripts, proteins or metabolites, as described in the first two sections of this chapter. This has led to a better understanding of the physiology of the animal. Combining insights acquired from these different studies have given a better insight in the physiology of farm animals. However, to further our understanding of the physiology of farm animals, integration of these different fields is required based on a more integrated approach. This section will describe the integration of omics technologies, starting from the combination of multiple omics technologies to the full integration from a systems biology approach, all aimed at better understanding of farm animals.

3.1 Interactions Between the Different Omics Techniques

As explained in the first section, biology in general and farm animal physiology in particular can be studied through omics techniques on different levels (Fig. 1). Most studies using omics technologies apply single techniques, because differences are expected on a specific level. However, in real life, biology is often not that easy, and effects on one level will also have influences on other levels. Therefore, studying farm animals on multiple levels using multiple omics technologies simultaneously can be very helpful to better understand the underlying physiology.

Starting at the genetic level, one gene studied extensively in dairy cows is DGAT1. As mentioned in the second section of this chapter, a polymorphism in DGAT1 has been shown to have a wide range of effects on milk synthesis and milk composition (Schennink et al. 2007). To mechanistically study why a single gene polymorphism can have such broad effects, the effect of this polymorphism has been studied on different levels, linking this genetic variation to transcriptomics, proteomics and metabolomics. Transcriptomic studies of cows differing in the DGAT1 K232A polymorphism have been performed to obtain an overview of the gene transcription changes related to this polymorphism. Thereby, that study aimed at better understanding the mechanism underlying the effects this polymorphism has on milk compositional parameters. This microarray-based transcriptomics analysis showed that cows differing in DGAT1 polymorphism had many genes that were differentially regulated, with the largest effects on transcripts related to energy metabolism, although no difference in transcription of the DGAT1 gene itself was found (Mach et al. 2012). This study showed that genes do not necessarily affect the transcript of the gene they encode, but there may be many correlated effects on completely different genes, which may lead to a much broader range of effects. In relation to the same genetic polymorphism, research has also been performed on the proteome and metabolome level (Lu et al. 2015). The results

showed a single protein, stomatin, was differentially regulated. Stomatin is involved in membrane structures in general and of milk fat globule membrane in particular. Simultaneously, lipid metabolites that are also involved in membrane structure were also differentially regulated in the milk fat globule and its membrane. This combination of omics dataset on different levels led to a hypothesis for an underlying mechanism related to differences in membrane structure between cows with different DGAT1 polymorphisms. This shows the benefit of simultaneously collecting omics data on different levels.

The same authors also studied the effect on proteome and metabolome level of another important factor in cow physiology, the negative energy balance in the periparturient period (Lu et al. 2013). In this study, changes in the same protein (stomatin) that is involved in membrane structuring were found. Also enzymes involved in cholesterol synthesis as well as cholesterol itself which is a lipid that is also important in membrane structure itself were changed. In addition, leakage of intracellular water-soluble metabolites was found. This study, by integrating proteomics and water- and lipid-soluble metabolomics, gave rise to the hypothesis that cell integrity and membrane structure were altered in cows in severe negative energy balance in early lactation. The studies mentioned above all linked different omics levels together and thereby reached new hypotheses on mammary gland physiology that could not have been reached if studies would have been limited to only a single level omics research.

To further study hypotheses that have been reached by analysing multiple omics datasets within the same sample, research could also look at multiple samples from the same animal. For example, samples can be taken of both the animal and its products or from different parts of the animal. Studying multiple samples from the same animal may lead to a better understanding of the underlying physiology. One example of such research aimed at better understanding milk synthesis by taking samples from both milk and the mammary gland itself into account. This was done in a study following a feeding intervention consisting of an increased intake of unsaturated fat by dairy cows (Ibeagha-Awemu et al. 2016a). This study showed how lipids in milk that changed after a feeding intervention were correlated with transcriptome alterations in the mammary gland. Such a study gives a more direct indication how interventions in cow management can influence milk synthesis in the mammary gland and thereby the composition of the milk produced. This approach has not only been used to better understand the mammary gland and milk synthesis but also for a better understanding of meat quality, as reviewed elsewhere (Mullen et al. 2006). An important benefit of studying the quality of farm animal products, like milk and meat, on different levels, from genes to metabolites and from animals to products, is that it leads to a better understanding of the underlying mechanisms. And once the mechanism is known, this can lead to better interventions to improve aspects such as animal health and product quality.

3.2 Full Integration of Omics Datasets into Systems Biology

The examples discussed above provide insight in the benefits of combining omics dataset collected on different levels. This can be further extended to looking from a network perspective: genes do not work in isolation but sets of genes encode, through sets of transcripts, sets of proteins that are involved in specific metabolic pathways. Interpretation of omics data can use information on such metabolic pathways, also called network analysis, for its interpretation. Once the genome of an animal is sequenced and annotated, this information can be used to construct metabolic pathways, which describe the integrated picture how different processes in a specie work together. The construction of overviews of metabolic pathways is a first necessary step towards research aiming at a better understanding of the metabolism of an animal. This construction of metabolic pathways can be done using knowledge of known enzymatic reactions and pathways to which the annotated genes can be linked. Different software-assisted approaches exist to perform this task automatically. These pathways by themselves do not directly explain the underlying biology; however, they provide the basic required information for understanding the metabolism, and its regulation, of animals (Seo et al. 2013). Of the farm animals, this has been done almost exclusively for the cow. This approach has been used to study the gene networks involved in lipid (Bionaz and Loor 2008) and protein (Bionaz and Loor 2011) synthesis of dairy cows. More recently, it was also shown that these gene networks are not uniquely associated with either lipid or fat synthesis but also interact (Li et al. 2016). These findings show that milk synthesis pathways for different milk components interact with each other and should not be studied in isolation. This also may explain the earlier mentioned pleiotropic effects of the DGAT1 polymorphism influencing many milk-related parameters, although an analysis of this effect from this perspective has not yet been performed. Furthermore illnesses of the mammary gland have been studied from such a pathways perspective. It was, for example, shown that mastitis was associated with an upregulation of the immune system pathway and a downregulation of the lipid metabolism pathway (Buitenhuis et al. 2011). On the level of proteins, it was also shown than an upregulation of host defence proteins occurred (Boehmer 2011; Boehmer et al. 2010).

This is not only true for the mammary gland, but also for the bovine liver (Khan et al. 2015), where multiple gene networks were shown to be differentially regulated during the periparturient period of negative energy balance. From the perspective of the homeostasis of an organism, balancing many metabolic pathways through many interactions makes sense, but it does complicate research in which often single cause-effects relations are searched for.

Such integrated insights are not only helpful for understanding the physiology of the cow, as discussed above, but also can help us to get a better understanding of the quality of the product. For example, meat has been extensively discussed previously (D'Alessandro and Zolla 2013; Paredi et al. 2012, 2013). By using multiple omics technologies, the development of muscle tissue in the growing animal, also

its degradation post-mortem, as well as its quality (for example, tenderness) can be better understood. Studying this from a systems biology approach by taking into account all relevant metabolic pathways (e.g., pathways related to apoptosis and autophagy) is required for a full understanding of the conversion from muscle into meat (Hollung et al. 2014).

In the end, all these integrated approaches aim at better understanding of the underlying physiology (Davidsen et al. 2016) but, as can also be seen from the relatively small number of published studies in this field, such analyses are not easy to perform. The challenges and outlooks for the systems biology approach to farm animals will be discussed next in the final section of this chapter.

4 Outlook on Future Developments

In the previous sections of this chapter, the current state of development in the area of omics technologies for systems biology application in farm animals is described. In this final section, a discussion is given of the different future developments that could aid the further development in this research area.

4.1 What Are the Challenges of Applying Systems Biology for Farm Animals

One of the main challenges remaining in the computational approaches to systems biology is the software that can be used for combining different omics datasets. Integration of multiple omics dataset, with integrated computer analyses, could be a very powerful tool for understanding the systems biology of healthy animals but is currently limited by software options as discussed previously (Suravajhala et al. 2016). Such studies will as a minimum require an in-depth understanding of all the interactions, both within each level, as well as between each level, as shown in Fig. 1. For bovine this knowledge is to a larger extent available than for other farm animals. This will require more effort in the future to get a better integrated picture of the metabolism of different farm animals. As discussed in Sect. 2, there are many differences between farm animals. However, the basic underlying genes, enzymes and metabolites involved in muscle development and milk synthesis are rather similar between species. The development of these metabolic pathways in other farm animals can thus be based on information already available for the cow.

Another challenge is the availability and format of collected data. Most omics dataset are collected on wide variety of platforms, all with their own structure and encoding of the data. A lot of available software for integrative analysis of omics data is therefore developed for specific types of data. In collaborations, where multiple types of instruments are used, this poses specific challenges in being

able to compare and analyse all data using an integrative perspective. On top of that, discrepancies between datasets (transcripts/proteins/metabolites obtained on one system and not on the other) pose an additional challenge. And finally, mismatches between gene-transcript-protein-metabolites can further complicate systems biology approaches. Due to different underlying detection principles between different analytical approaches, it will not be easy to solve these problems.

Multivariate approaches to analyse integrated omics datasets using, for example, pattern recognition are also not been broadly developed yet. One successful example of such an approach in which support vector machines were applied for biomarker discovery in large omics datasets has been published (Kim et al. 2017). However, the application of such tools in the analysis of very different types of datasets was not shown in practice yet.

4.2 What Can We Learn from Other Fields of Biology

The field of systems biology is of course not exclusive to farm animals. It is frequently, and extensively, used in humans to better understand the physiology of health and disease. For example, in cancer research, multiple omics techniques have previously been combined for diagnosis, prognosis and monitoring treatment (Seeree et al. 2015; Dimitrieva et al. 2016). Also for other human diseases, the combination of multiple omics field data for biomarker discovery has frequently been applied (Kussmann and Blum 2007; Castro-Santos et al. 2015; Mehta et al. 2015).

More generally, knowledge on interaction between different levels of omics, also called interactomics, can be used for better understanding illnesses (Bellay et al. 2012). Specifically on the level of proteins, such an interactomic overview of milk has been previously done (D'Alessandro et al. 2011; Zhang et al. 2017). These studies show how milk proteins are involved in different specific protein-protein interaction networks, leading to a better understanding of the role of milk proteins in the growth and development of the neonate. Further extending such studies to all levels of biology from gene to metabolite could further our understanding of milk and more in general of the farm animal.

References

Almeida AM, Bassols A, Bendixen E, Bhide M, Ceciliani F, Cristobal S, Eckersall PD, Hollung K, Lisacek F, Mazzucchelli G, McLaughlin M, Miller I, Nally JE, Plowman J, Renaut J, Rodrigues P, Roncada P, Staric J, Turk R (2015) Animal board invited review: advances in proteomics for animal and food sciences. Animal 9(1):1–17

Alstrup L, Soegaard K, Weisbjerg MR (2016) Effects of maturity and harvest season of grass-clover silage and of forage-to-concentrate ratio on milk production of dairy cows. J Dairy Sci 99(1):328–340

Antunes-Fernandes EC, van Gastelen S, Dijkstra J, Hettinga KA, Vervoort J (2016) Milk metabolome relates enteric methane emission to milk synthesis and energy metabolism pathways. J Dairy Sci 99(8):6251–6262

Bai X, Zheng Z, Liu B, Ji X, Bai Y, Zhang W (2016) Whole blood transcriptional profiling comparison between different milk yield of Chinese Holstein cows using RNA-seq data. BMC Genomics 17(Suppl 7):512

Bellay J, Michaut M, Kim T, Han S, Colak R, Myers CL, Kim PM (2012) An omics perspective of protein disorder. Mol BioSyst 8(1):185–193

Bendixen E, Danielsen M, Hollung K, Gianazza E, Miller I (2011) Farm animal proteomics—a review. J Proteomics 74(3):282–293

Benmoussa A, Lee CH, Laffont B, Savard P, Laugier J, Boilard E, Gilbert C, Fliss I, Provost P (2016) Commercial dairy cow milk microRNAs resist digestion under simulated gastrointestinal tract conditions. J Nutr 146(11):2206–2215

Bernabucci U, Basirico L, Morera P, Dipasquale D, Vitali A, Piccioli Cappelli F, Calamari L (2015) Effect of summer season on milk protein fractions in Holstein cows. J Dairy Sci 98 (3):1815–1827

Bionaz M, Loor JJ (2008) Gene networks driving bovine milk fat synthesis during the lactation cycle. BMC Genomics 9:366

Bionaz M, Loor JJ (2011) Gene networks driving bovine mammary protein synthesis during the lactation cycle. Bioinform Biol Insights 5:83–98

Boehmer JL (2011) Proteomic analyses of host and pathogen responses during bovine mastitis. J Mammary Gland Biol Neoplasia 16(4):323–338

Boehmer JL, Ward JL, Peters RR, Shefcheck KJ, McFarland MA, Bannerman DD (2010) Proteomic analysis of the temporal expression of bovine milk proteins during coliform mastitis and label-free relative quantification. J Dairy Sci 93(2):593–603

Boggs I, Hine B, Smolenksi G, Hettinga K, Zhang L, Wheeler TT (2016) Proteomics data in support of the quantification of the changes of bovine milk proteins during mammary gland involution. Data Brief 8:52–55

Boggs I, Hine B, Smolenski G, Hettinga K, Zhang L, Wheeler TT (2015) Changes in the repertoire of bovine milk proteins during mammary involution. EuPA Open Proteom 9:65–75

Buitenhuis B, Poulsen NA, Gebreyesus G, Larsen LB (2016) Estimation of genetic parameters and detection of chromosomal regions affecting the major milk proteins and their post translational modifications in Danish Holstein and Danish Jersey cattle. BMC Genet 17:114

Buitenhuis B, Røntved CM, Edwards SM, Ingvartsen KL, Sørensen P (2011) In depth analysis of genes and pathways of the mammary gland involved in the pathogenesis of bovine Escherichia coli-mastitis. BMC Genomics 12(1):130

Castro-Santos P, Laborde CM, Diaz-Pena R (2015) Genomics, proteomics and metabolomics: their emerging roles in the discovery and validation of rheumatoid arthritis biomarkers. Clin Exp Rheumatol 33(2):279–286

Connor EE, Siferd S, Elsasser TH, Evock-Clover CM, Van Tassell CP, Sonstegard TS, Fernandes VM, Capuco AV (2008) Effects of increased milking frequency on gene expression in the bovine mammary gland. BMC Genomics 9:362. https://doi.org/10.1186/1471–2164–9-362

D'Alessandro A, Zolla L (2013) Meat science: from proteomics to integrated omics towards system biology. J Proteomics 78:558–577

D'Alessandro A, Zolla L, Scaloni A (2011) The bovine milk proteome: cherishing, nourishing and fostering molecular complexity. An interactomics and functional overview. Mol BioSyst 7 (3):579–597

D'Auria E, Agostoni C, Giovannini M, Riva E, Zetterstrom R, Fortin R, Greppi GF, Bonizzi L, Roncada P (2005) Proteomic evaluation of milk from different mammalian species as a substitute for breast milk. Acta Paediatrica (Oslo, Norway: 1992) 94(12):1708–1713

Davidsen PK, Turan N, Egginton S, Falciani F (2016) Multilevel functional genomics data integration as a tool for understanding physiology: a network biology perspective. J Appl Physiol (Bethesda, Md: 1985) 120(3):297–309

Dimitrieva S, Schlapbach R, Rehrauer H (2016) Prognostic value of cross-omics screening for kidney clear cell renal cancer survival. Biol Direct 11(1):68. https://doi.org/10.1186/s13062-016-0170-1
Duchemin S, Bovenhuis H, Stoop WM, Bouwman AC, van Arendonk JAM, Visker MHPW (2013) Genetic correlation between composition of bovine milk fat in winter and summer, and DGAT1 and SCD1 by season interactions. J Dairy Sci 96(1):592–604
Duchemin SI, Visker MH, Van Arendonk JA, Bovenhuis H (2014) A quantitative trait locus on Bos taurus autosome 17 explains a large proportion of the genetic variation in de novo synthesized milk fatty acids. J Dairy Sci 97(11):7276–7285
Gerspach C, Imhasly S, Gubler M, Naegeli H, Ruetten M, Laczko E (2017) Altered plasma lipidome profile of dairy cows with fatty liver disease. Res Vet Sci 110:47–59
Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009) Invited review: genomic selection in dairy cattle: progress and challenges. J Dairy Sci 92(2):433–443
Heck JM, Schennink A, van Valenberg HJ, Bovenhuis H, Visker MH, van Arendonk JA, van Hooijdonk AC (2009) Effects of milk protein variants on the protein composition of bovine milk. J Dairy Sci 92(3):1192–1202
Hettinga K, van Valenberg H, de Vries S, Boeren S, van Hooijdonk T, van Arendonk J, Vervoort J (2011) The host defense proteome of human and bovine milk. PLoS ONE 6(4):e19433
Hinz K, O'Connor PM, O'Brien B, Huppertz T, Ross RP, Kelly AL (2012) Proteomic study of proteolysis during ripening of Cheddar cheese made from milk over a lactation cycle. J Dairy Res 79(2):176–184
Hollung K, Timperio AM, Olivan M, Kemp C, Coto-Montes A, Sierra V, Zolla L (2014) Systems biology: a new tool for farm animal science. Curr Protein Pept Sci 15(2):100–117
Ibeagha-Awemu EM, Li R, Ammah AA, Dudemaine PL, Bissonnette N, Benchaar C, Zhao X (2016a) Transcriptome adaptation of the bovine mammary gland to diets rich in unsaturated fatty acids shows greater impact of linseed oil over safflower oil on gene expression and metabolic pathways. BMC Genomics 17:104
Ibeagha-Awemu EM, Peters SO, Akwanji KA, Imumorin IG, Zhao X (2016b) High density genome wide genotyping-by-sequencing and association identifies common and low frequency SNPs, and novel candidate genes influencing cow milk traits. Sci Rep 6:31109
Ioannidis J, Donadeu FX (2016) Circulating microRNA Profiles during the Bovine Oestrous Cycle. PLoS ONE 11(6):e0158160
Kaneene JB, Scott Hurd H (1990) The national animal health monitoring system in Michigan. III. Cost estimates of selected dairy cattle diseases. Prev Vet Med 8(2):127–140
Khan MJ, Jacometo CB, Riboni MV, Trevisi E, Graugnard DE, Correa MN, Loor JJ (2015) Stress and inflammatory gene networks in bovine liver are altered by plane of dietary energy during late pregnancy. Funct Integr Genomics 15(5):563–576
Kim S, Jhong JH, Lee J, Koo JY (2017) Meta-analytic support vector machine for integrating multiple omics data. BioData Mining 10:2
Kirchner B, Pfaffl MW, Dumpler J, von Mutius E, Ege MJ (2016) microRNA in native and processed cow's milk and its implication for the farm milk effect on asthma. J Allergy Clin Immunol 137(6):1893–1895.e1813
Korhonen HJ (2009) Bioactive components in bovine milk. In: Bioactive components in milk and dairy products. Wiley-Blackwell, Ames, IA, pp 13–42
Kukurba KR, Montgomery SB (2015) RNA sequencing and analysis. Cold Spring Harb Protoc 2015(11):951–969
Kussmann M, Blum S (2007) OMICS-derived targets for inflammatory gut disorders: opportunities for the development of nutrition related biomarkers. Endocr Metab Immune Disord Drug Targets 7(4):271–287
Li Q, Zhao Y, Zhu D, Pang X, Liu Y, Frew R, Chen G (2017) Lipidomics profiling of goat milk, soymilk and bovine milk by UPLC-Q-exactive orbitrap mass spectrometry. Food Chem 224:302–309

Li S, Hosseini A, Danes M, Jacometo C, Liu J, Loor JJ (2016) Essential amino acid ratios and mTOR affect lipogenic gene networks and miRNA expression in bovine mammary epithelial cells. J Animal Sci Biotechnol 7:44

Lu J, Antunes Fernandes E, Paez Cano AE, Vinitwatanakhun J, Boeren S, van Hooijdonk T, van Knegsel A, Vervoort J, Hettinga KA (2013) Changes in milk proteome and metabolome associated with dry period length, energy balance, and lactation stage in postparturient dairy cows. J Proteome Res 12(7):3288–3296

Lu J, Boeren S, van Hooijdonk T, Vervoort J, Hettinga K (2015) Effect of the DGAT1 K232A genotype of dairy cows on the milk metabolome and proteome. J Dairy Sci 98(5):3460–3469

Lu J, van Hooijdonk T, Boeren S, Vervoort J, Hettinga K (2014) Identification of lipid synthesis and secretion proteins in bovine milk. J Dairy Res 81(1):65–72

Mach N, Blum Y, Bannink A, Causeur D, Houee-Bigot M, Lagarrigue S, Smits MA (2012) Pleiotropic effects of polymorphism of the gene diacylglycerol-O-transferase 1 (DGAT1) in the mammary gland tissue of dairy cows. J Dairy Sci 95(9):4989–5000

McCoard SA, Hayashi AA, Sciascia Q, Rounce J, Sinclair B, McNabb WC, Roy NC (2016) Mammary transcriptome analysis of lactating dairy cows following administration of bovine growth hormone. Animal 10(12):2008–2017

Mehta SM, Banerjee SM, Chowdhary AS (2015) Postgenomics biomarkers for rabies-the next decade of proteomics. OMICS 19(2):67–79

Moran B, Cummins SB, Creevey CJ, Butler ST (2016) Transcriptomics of liver and muscle in Holstein cows genetically divergent for fertility highlight differences in nutrient partitioning and inflammation processes. BMC Genomics 17(1):603

Morzel M, Chambon C, Lefevre F, Paboeuf G, Laville E (2006) Modifications of trout (Oncorhynchus mykiss) muscle proteins by preslaughter activity. J Agric Food Chem 54 (8):2997–3001

Mullen AM, Stapleton PC, Corcoran D, Hamill RM, White A (2006) Understanding meat quality through the application of genomic and proteomic approaches. Meat Sci 74(1):3–16

Ng-Kwai-Hang KF, Hayes JF, Moxley JE, Monardes HG (1987) Variation in milk protein concentrations associated with genetic polymorphism and environmental factors. J Dairy Sci 70(3):563–570

Nir Markusfeld O (2003) What are production diseases, and how do we manage them? Acta Vet Scand Suppl 98:21–32

Paredi G, Raboni S, Bendixen E, de Almeida AM, Mozzarelli A (2012) “Muscle to meat” molecular events and technological transformations: the proteomics insight. J Proteomics 75 (14):4275–4289

Paredi G, Sentandreu MA, Mozzarelli A, Fadda S, Hollung K, de Almeida AM (2013) Muscle and meat: new horizons and applications for proteomics on a farm to fork perspective. J Proteomics 88:58–82

Perge P, Nagy Z, Decmann A, Igaz I, Igaz P (2016) Potential relevance of microRNAs in interspecies epigenetic communication, and implications for disease pathogenesis. RNA Biol 14:1–11

Poulsen NA, Jensen HB, Larsen LB (2016) Factors influencing degree of glycosylation and phosphorylation of caseins in individual cow milk samples. J Dairy Sci 99(5):3325–3333

Qian L, Zhao A, Zhang Y, Chen T, Zeisel SH, Jia W, Cai W (2016) Metabolomic approaches to explore chemical diversity of human breast-milk, formula milk and bovine milk. Int J Mol Sci 17(12):2128

Rodrigues PM, Silva TS, Dias J, Jessen F (2012) PROTEOMICS in aquaculture: applications and trends. J Proteomics 75(14):4325–4345

Salilew-Wondim D, Ibrahim S, Gebremedhn S, Tesfaye D, Heppelmann M, Bollwein H, Pfarrer C, Tholen E, Neuhoff C, Schellander K, Hoelker M (2016) Clinical and subclinical endometritis induced alterations in bovine endometrial transcriptome and miRNome profile. BMC Genomics 17:218

Samková E, Spicka J, Pesek M, Pelikánová T, Hanus O (2012) Animal factors affecting fatty acid composition of cow milk fat: a review. S Afr J Anim Sci 42:83–100
Sanchez-Macias D, Moreno-Indias I, Castro N, Morales-Delanuez A, Arguello A (2014) From goat colostrum to milk: physical, chemical, and immune evolution from partum to 90 days postpartum. J Dairy Sci 97(1):10–16
Schennink A, Bovenhuis H, Leon-Kloosterziel KM, van Arendonk JA, Visker MH (2009) Effect of polymorphisms in the FASN, OLR1, PPARGC1A, PRL and STAT5A genes on bovine milk-fat composition. Anim Genet 40(6):909–916
Schennink A, Stoop WM, Visker MH, Heck JM, Bovenhuis H, van der Poel JJ, van Valenberg HJ, van Arendonk JA (2007) DGAT1 underlies large genetic variation in milk-fat composition of dairy cows. Anim Genet 38(5):467–473
Seeree P, Pearngam P, Kumkate S, Janvilisri T (2015) An omics perspective on molecular biomarkers for diagnosis, prognosis, and therapeutics of cholangiocarcinoma. Int J Genomics 2015:179528
Seo S, Larkin DM, Loor JJ (2013) Cattle genomics and its implications for future nutritional strategies for dairy cattle. Animal 7(Suppl 1):172–183
Shen B, Zhang L, Lian C, Lu C, Zhang Y, Pan Q, Yang R, Zhao Z (2016) Deep sequencing and screening of differentially expressed MicroRNAs related to milk fat metabolism in bovine primary mammary epithelial cells. Int J Mol Sci 17(2):200
Shen Y, Shi S, Tong H, Guo Y, Zou J (2014) Metabolomics analysis reveals that bile acids and phospholipids contribute to variable responses to low-temperature-induced ascites syndrome. Mol BioSyst 10(6):1557–1567
Singh K, Erdman RA, Swanson KM, Molenaar AJ, Maqbool NJ, Wheeler TT, Arias JA, Quinn-Walsh EC, Stelwagen K (2010) Epigenetic regulation of milk production in dairy cows. J Mammary Gland Biol Neoplasia 15(1):101–112
Singh K, Molenaar AJ, Swanson KM, Gudex B, Arias JA, Erdman RA, Stelwagen K (2012) Epigenetics: a possible role in acute and transgenerational regulation of dairy cow milk production. Animal 6(3):375–381
Sokol E, Ulven T, Færgeman NJ, Ejsing CS (2015) Comprehensive and quantitative profiling of lipid species in human milk, cow milk and a phospholipid-enriched milk formula by GC and MS/MSALL. Eur J Lipid Sci Technol 117(6):751–759
Stelwagen K, Carpenter E, Haigh B, Hodgkinson A, Wheeler TT (2009) Immune components of bovine colostrum and milk. J Anim Sci 87(13 Suppl):3–9
Sundekilde UK, Larsen LB, Bertram HC (2013) NMR-based milk metabolomics. Metabolites 3 (2):204–222
Suravajhala P, Kogelman LJA, Kadarmideen HN (2016) Multi-omic data integration and analysis using systems genomics approaches: methods and applications in animal production, health and welfare. Genet Sel Evol 48(1):38
Thomas FC, Mudaliar M, Tassi R, McNeilly TN, Burchmore R, Burgess K, Herzyk P, Zadoks RN, Eckersall PD (2016) Mastitomics, the integrated omics of bovine milk in an experimental model of Streptococcus uberis mastitis: 3. Untargeted metabolomics. Mol BioSyst 12 (9):2762–2769
Tian H, Zheng N, Wang W, Cheng J, Li S, Zhang Y, Wang J (2016) Integrated metabolomics study of the milk of heat-stressed lactating dairy cows. Sci Rep 6:24208
Tsiplakou E, Zervas G (2013) Changes in milk and plasma fatty acid profile in response to fish and soybean oil supplementation in dairy sheep. J Dairy Res 80(2):205–213
Wall EH, Bond JP, McFadden TB (2013) Milk yield responses to changes in milking frequency during early lactation are associated with coordinated and persistent changes in mammary gene expression. BMC Genomics 14:296
Wang JH, Byun J, Pennathur S (2010) Analytical approaches to metabolomics and applications to systems biology. Semin Nephrol 30(5):500–511
Wittenburg D, Melzer N, Willmitzer L, Lisec J, Kesting U, Reinsch N, Repsilber D (2013) Milk metabolites and their genetic variability. J Dairy Sci 96(4):2557–2569

Yang W, Zerbe H, Petzl W, Brunner RM, Gunther J, Draing C, von Aulock S, Schuberth HJ, Seyfert HM (2008) Bovine TLR2 and TLR4 properly transduce signals from Staphylococcus aureus and E. coli, but S. aureus fails to both activate NF-kappaB in mammary epithelial cells and to quickly induce TNFalpha and interleukin-8 (CXCL8) expression in the udder. Mol Immunol 45(5):1385–1397

Yang Y, Zheng N, Zhao X, Zhang Y, Han R, Yang J, Zhao S, Li S, Guo T, Zang C, Wang J (2016) Metabolomic biomarkers identify differences in milk produced by Holstein cows and other minor dairy animals. J Proteomics 136:174–182

Yang YX, Zhao XX, Zhang Y (2009) Proteomic analysis of mammary tissues from healthy cows and clinical mastitic cows for identification of disease-related proteins. Vet Res Commun 33 (4):295–303

Yangilar F (2013) As a potentially functional food: goats' milk and products. J Food Nutr Res 1 (4):68–81

Younis S, Javed Q, Blumenberg M (2016) Meta-analysis of transcriptional responses to mastitis-causing Escherichia coli. PLoS ONE 11(3):e0148562. https://doi.org/10.1371/journal.pone.0148562

Zhang L (2015) Dynamics of the proteome in human and farm animal milk. Wageningen University, Wageningen

Zhang L, Boeren S, Hageman JA, van Hooijdonk T, Vervoort J, Hettinga K (2015b) Bovine milk proteome in the first 9 days: protein interactions in maturation of the immune and digestive system of the newborn. PLoS ONE 10(2):e0116710

Zhang L, Boeren S, Hageman JA, van Hooijdonk T, Vervoort J, Hettinga K (2015c) Perspective on calf and mammary gland development through changes in the bovine milk proteome over a complete lactation. J Dairy Sci 98(8):5362–5373

Zhang L, Boeren S, Smits M, van Hooijdonk T, Vervoort J, Hettinga K (2016a) Proteomic study on the stability of proteins in bovine, camel, and caprine milk sera after processing. Food Res Int 82:104–111

Zhang L, Boeren S, van Hooijdonk AC, Vervoort JM, Hettinga KA (2015a) A proteomic perspective on the changes in milk proteins due to high somatic cell count. J Dairy Sci 98 (8):5339–5351

Zhang L, de Waard M, Verheijen H, Boeren S, Hageman JA, van Hooijdonk T, Vervoort J, van Goudoever JB, Hettinga K (2016b) Changes over lactation in breast milk serum proteins involved in the maturation of immune and digestive system of the infant. J Proteomics 147:40–47

Zhang L, van Dijk AD, Hettinga K (2017) An interactomics overview of the human and bovine milk proteome over lactation. Proteome Sci 15:1

Zhang Q, Cundiff J, Maria S, McMahon R, Woo J, Davidson B, Morrow A (2013) Quantitative analysis of the human milk whey proteome reveals developing milk and mammary-gland functions across the first year of lactation. Proteomes 1(2):128

Zhao K, Liang G, Sun X, Guan le L (2016) Comparative miRNAome analysis revealed different miRNA expression profiles in bovine sera and exosomes. BMC Genomics 17(1):630

Zhao S, Zhao J, Bu D, Sun P, Wang J, Dong Z (2014) Metabolomics analysis reveals large effect of roughage types on rumen microbial metabolic profile in dairy cows. Lett Appl Microbiol 59 (1):79–85

Zhou S, Wan Q, Huang Y, Huang X, Cao J, Ye L, Lim TK, Lin Q, Qin Q (2011) Proteomic analysis of Singapore grouper iridovirus envelope proteins and characterization of a novel envelope protein VP088. Proteomics 11(11):2236–2248

Zou X, Huang J, Jin Q, Guo Z, Liu Y, Cheong L, Xu X, Wang X (2013) Lipid composition analysis of milk fats from different mammalian species: potential for use as human milk fat substitutes. J Agric Food Chem 61(29):7070–7080

Proteomics in Domestic Animals on a Farm to Systems Biology Perspective: Final Remarks and Future Prospects

André M. de Almeida, Ingrid Miller, and P. David Eckersall

In this book, we have explored the fascinating world of proteomics and the developments in the field that may enhance research, specifically in the area of animal science and production, as well as veterinary medicine. Furthermore, we have shown how proteomic studies may successfully be integrated with other omics, namely, transcriptomics and metabolomics. At this stage, we hope that the reader who has never used proteomics and associated methods may now consider using it in future research projects in order to answer the many challenges that animal production has to address in the years to come. As for the reader that already has a history of farm animal proteomic research, we hope that this book may serve as an inspiration for additional or alternative uses of the technologies available and perhaps stimulate future collaborations.

In the last decades, proteomics has experienced numerous developments, as have analytical sciences in general. The trend in proteomics has moved from a gel-based technology that could be available in almost every laboratory at relatively affordable costs to a mass spectrometry-based technology that implies the use of multimillion euro/dollar instruments that in turn require very specialised staff and have to run 24/7, 365 days a year in order to compensate for the initial investment. Such instruments cannot be available to the majority of institutions that conduct

A.M. de Almeida (✉)
Instituto Superior de Agronomia, Universidade de Lisboa, Tapada da Ajuda, Lisbon, Portugal
e-mail: aalmeida@isa.ulisboa.pt

I. Miller
University of Veterinary Medicine Vienna, Vienna, Austria
e-mail: Ingrid.Miller@vetmeduni.ac.at

P.D. Eckersall
School of Veterinary Medicine, University of Glasgow, Glasgow, Scotland, UK

Faculty of Veterinary Medicine, University of Zagreb, Zagreb, Croatia
e-mail: david.eckersall@glasgow.ac.uk

© Springer International Publishing AG 2018
A.M. de Almeida et al. (eds.), *Proteomics in Domestic Animals: from Farm to Systems Biology*, https://doi.org/10.1007/978-3-319-69682-9_23

farm animal research. They may however be available through centres with hi-tech proteomic platforms that analyse samples from different countries, regions or even the whole world but in very close collaboration with the institutes that have conducted the animal trials and which would have the task of sample collection and handling and at a later stage data treatment and biological interpretation. This is perhaps one of the most important challenges to researchers in animal sciences. Indeed, a successful connection to a proteomic platform may be difficult to achieve for a researcher that is not familiar with proteomic principles. We hope that this book may be helpful in filling that gap. Another difficulty relates to the high requirements for funds that are necessary to access such platforms. In fact, farm animal research project budgets seldom include the amounts required to analyse all samples in a trial. As such, a compromise between statistical significance and budgetary limitations may have to be found.

Proteomics in farm animals has also to struggle with some technical limitations. Probably the most significant is related to the multispecies nature of the research. Indeed, if proteomic research in sequenced species like cattle (*Bos taurus*), pig (*Sus scrofa*) or chicken (*Gallus gallus*) tends to lead to very complete results, research with less known species like goats (*Capra hircus*), horses (*Equus caballus*) or turkeys (*Meleagris gallopavo*) and geese (*Anser anser*), not to mention game species, tends to lead to very incomplete results due to the lower representation of the latter in protein (and gene) databases. It is likely that in the future sequencing of such species will become a reality. The lowering of sequencing costs experienced in the last decade does point that way. To contribute to solving this problem and from an individual research initiative perspective, it is important to supply sequencing data obtained and insert it in public databases, so it may be accessible to all researchers in the field. Other technical limitations have to do with the nature of the proteinaceous material studied. Indeed, it is relatively straightforward to find protein extraction protocols for certain tissues such as the muscle or liver; for others like wool or honey, adequate protocols are needed. It is therefore important that researchers publish their findings in international peer-reviewed journals, highlighting protocol aspects in detail so that other researchers may adopt (and likely cite) such works contributing to strengthen the farm animal proteomic research community. It is equally important that journal editors and reviewers recognise the importance of comparative proteomic data from multiple species so that the knowledge base on the animal species used for livestock and aquaculture production keeps pace in dissemination with the advances in technology. While there are disadvantages to undertaking proteomic analysis in domestic animal species, there are also several advantages that come from an ability to plan experimental procedures that allow scientific hypothesis-based research to be undertaken. Experimental proteomic study that could not ethically be performed in humans is possible with appropriate approval in farm animals, and these large species can provide multiple samples in volumes that are not achievable in laboratory rodents. In just one example, bacterial infection of a dairy cow's udder quarter leading to mastitis is an almost perfect in vivo incubator for the study of host-pathogen interaction leading to changes to the proteome during infection, recovery and therapy with

millilitre volumes of milk being readily available throughout for detailed time-series studies.

The farm animal proteomic community has reached an important stage, as may be seen from this book. Indeed, the number of teams dedicated to this issue seems to be growing worldwide, as has the number of publications and the overall quality of manuscript published in both animal/veterinary sciences and proteomics journals. Regardless of the challenge, the best way to face it seems to be through collaboration and cooperation between all players in the area: farmers, veterinarians, researchers, technicians, retailers, consumers and regulatory bodies. International cooperation, especially among scientists, is also an equally important aspect. Indeed, international cooperation may render accessible new technologies and instruments that otherwise may not be available. Nevertheless, much more important are the new avenues that may be opened by international cooperation. These include, for instance, access to international consortia and funds ultimately improving research quality and significance. As a final note on this last aspect, we would like to highlight the importance of projects such as the now finished COST action FA1002—Proteomics in Farm Animals (http://www.cost.eu/COST_Actions/fa/FA1002) or Prime-XS (www.Primexs.eu) that allowed, for the first time, an effective interaction of classical animal science researchers and proteomic platforms, allowing the former to access top-quality instrumentation. The development of similar enterprises with research funding capacity dedicated solely to agriculture and animal sciences would be a very important initiative to leverage the development of proteomics in farm animals.

We finally hope that this book may be an important step in the growing internationalisation of farm animal proteomic research and the research quality deriving from this vital resource for maintaining the global supply of edible protein for the growing human population.

Lisbon, Vienna and Zagreb 1st of August 2017

André M. de Almeida

Ingrid Miller

David Eckersall