Richard J. Morris *Editor*

Mathematical Modelling in Plant Biology

Springer

Mathematical Modelling in Plant Biology

Richard J. Morris
Editor

Mathematical Modelling in Plant Biology

Springer

Editor
Richard J. Morris
Computational and Systems Biology
John Innes Centre
Norwich, Norfolk, UK

ISBN 978-3-319-99069-9 ISBN 978-3-319-99070-5 (eBook)
https://doi.org/10.1007/978-3-319-99070-5

Library of Congress Control Number: 2018960264

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

Preface

Plant biology is a rich, fascinating and rewarding playground for those with an interest in modelling and simulation.

Plants are constantly processing information, computing and adapting to their surroundings. This notion of information processing is becoming increasingly important in biology as is the appreciation of physical and engineering approaches for understanding these processes and how they are manifested in form and function. Our basic premise is that whatever plants are doing, they are doing within the laws of physics. Physical approaches using the established language of mathematics and computation are therefore key research tools for unravelling the biology of plants. For instance, plant growth can be viewed as a mechanical problem in which plants exploit hydraulics and material properties to determine their shape. Photosynthesis can be treated as a quantum mechanical problem, whereby photons are captured and their energy used to catalyse chemical processes that convert carbon dioxide to sugar. Plants use diffusible particles and electrical waves for transmitting information. Plants generate pressure gradients that drive fluid flow through small elastic tubes to transport nutrients. The list goes on.

Perhaps with the exception of the theory of evolution, biology may seem to currently lack the unifying laws of physics, the axiomatic nature of mathematics or the abstractions of computer science, thus giving rise to the usual clichés. The ‘hard’ sciences are typically viewed as abstract, reductionist, mathematical and quantitative. For some theories, such as quantum electrodynamics or special relativity, astonishing levels of precision have been reported, achieving better than ten significant digits. Such precision (and accuracy) can induce the idea of these theories being ‘exact’. On the other hand, ‘soft’ sciences have a reputation of being more descriptive and qualitative. Biological systems perform tasks such as self-repair and reproduction that fit less readily within existing physical and engineering frameworks. The science of living systems can appear ‘messy’ and therein lies the challenge. It is ‘easy’ to achieve robust computation using well-characterised, virtually error-free components, but how does biology with its noisy, fluctuating systems manage to carry out tasks such as reproducing cells and whole organisms so robustly? Biology is full of such ‘hard’ problems.

More important than this artificial division into disciplines is the approach. Unlike in physics where the question of 'function' or 'utility' of something (why questions) makes little sense, in biology thinking about function can help place observations in an evolutionary context and enhance our understanding. Thinking about problems in terms of cause and effect (how questions) is what leads to mechanistic insights which in turn helps understand evolutionary innovation, and this is the spirit of the chapters presented in this book. Modelling and simulation have a key role to play in unravelling mechanism and trying to discern cause and effect but also in making sense of evolutionary changes. Efficient models can compress a lot of data and knowledge into a few rules or equations. Despite all models being wrong at some level, they offer powerful tools to synthesise data and ideas, to generate and evaluate hypotheses and to make predictions. Importantly, useful models are falsifiable (which is when we learn the most about a system). Predictions can be used to guide new experiments and to validate, falsify or define the application boundaries of a model.

This book aims to provide a mix of introductory chapters and latest state-of-the-art research overviews that place key questions in plant biology within a cause-and-effect framework, thereby drawing on relevant physical, mathematical and computational approaches. Modern plant biology requires an increasing and diverse set of skills and seamlessly blends them into an integrative, interdisciplinary approach. All chapters are written by leading experts who are driving such interdisciplinary developments. The chapters are not meant to be exhaustive but to give a flavour of some of the current problems and to provide some background upon which can be built to develop a solid foundation for research in the area. The book is aimed at physicists, mathematicians, computer scientists and engineers, whom we hope to excite with the challenges and opportunities in plant biology but also the increasing number of mathematically skilled biologists with an interest in modelling and simulation as a means to understand biology. For others there are likely better and more suitable introductions.

For those without much biology background, there are many truly excellent text books. A fantastic and stimulating up-to-date classic is *Molecular Biology of the Cell* by Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts and Peter Walter. Some basic understanding of DNA, genes and proteins can readily be developed from online resources. For the current book, the prerequisites in terms of plant biology are limited, many of which are explained in the individual chapters. Important concepts include the following: plants consist of cells that are surrounded by a cell wall and are therefore not mobile; plant cells can build substantial pressures within them through osmosis, which is the exchange of water driven by the chemical potential that arises from the concentration difference on solutes; plant cells have various components (proteins) that can transport ions across membranes, giving rise to concentration differences and therefore voltages which can be exploited for signalling. Personal favourites for more in-depth studies include *Plant Biomechanics* by Karl Niklas, *An Introduction to Systems Biology* by Uri Alon and *Information Theory, Inference and Learning Algorithms* by David MacKay.

Given that human life is sustained by plants in that they influence our atmospheric composition, providing us with oxygen to breath, form important ecosystems, contribute a substantial part of our diet and calorific intake and provide natural products and medicines, understanding the biology, physics and computation of plants is perhaps one of the most relevant challenges of our time.

As I hope will become apparent from this book, plant biology is a 'hard' science. It extensively uses quantitative data, physical theories and mathematical and computational modelling and is becoming increasingly predictive. Furthermore, plant biology is great fun.

Chapter 1 introduces physical models of plant morphogenesis. The theory of forces, stress and strain is explained with selected case studies. Chapter 2 summarises the basic mathematical approaches to fluid transport in plants. Fluid dynamics plays an important role in the transport of nutrients, growth and long-distance signalling. Chapter 3 describes how we can use physical models to understand ion channels. Much of the information processing carried out by plants uses changes in ionic concentrations to transmit signals that activate responses to environmental challenges. Chapter 4 introduces plant microtubules and mathematical and computational techniques for modelling their behaviour. Microtubules are dynamic entities that play a key role in determining plant cell shape and function. From here onwards we are exposed to problems of organisation over multiple spatial and temporal scales—a reoccurring theme and challenge in approaches to modelling in biology. Chapter 5 takes a closer look at how cell shape changes as a function of ion channel activity on the example of guard cells, thereby integrating transport processes with macroscopic function. In Chap. 6, an overview is provided of the most recent developments in single-cell approaches for understanding morphogenesis, particularly in terms of image processing, quantitative data analysis and computational modelling techniques. Chapter 7 goes beyond single cells and tackles approaches for describing collections of cells, tissues, their interactions, growth and division. Chapter 8 takes an abstract computational approach to plant development with the development of L-systems. L-systems offer powerful tools for studying plant development at different levels from reactions to whole plant behaviour. Chapter 9 describes recent results on the important trait of flowering time to move up further in scales of synthesising knowledge. This chapter considers gene networks, phenology and evolution. Chapter 10 takes the scale of modelling one important step further and investigates the lifestyle strategy of plants in natural environment on the example of seed banks. Together these chapters are exemplars of how plant science is developing and the inherent challenges of bridging scales between micro-mechanisms through cells to whole plant behaviour and populations of plants in a changing environment. There is clearly no shortage of really exciting and highly relevant challenges ahead for which computational approaches will have a key role to play.

There are many exciting computational developments in plant biology, and hopefully future editions of this book can be extended to include further chapters that display the power of interdisciplinary journeys into the processes and mechanisms of plants.

Norwich, UK Richard J. Morris
April 2018

Contents

Contributors

Milad Adibi Max Planck Institute for Plant Breeding Research, Köln, Germany

Michael R. Blatt Laboratory of Plant Physiology and Biophysics, University of Glasgow, Glasgow, UK

M. G. Blyth School of Mathematics, University of East Anglia, Norwich, UK

Arezki Boudaoud Laboratoire Reproduction et Développement des Plantes, Univ Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA, Lyon, France

Mikolaj Cieslak Department of Computer Science, University of Calgary, Calgary, AB, Canada

Eva E. Deinum Biometris, Wageningen University & Research, Wageningen, The Netherlands

Mathilde Dumond Laboratoire Reproduction et Développement des Plantes, Univ Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA, Lyon, France

Pascal Ferraro Department of Computer Science, University of Calgary, Calgary, AB, Canada

Pau Formosa-Jordan Sainsbury Laboratory, University of Cambridge, Cambridge, UK

Jim Hanan Centre for Horticultural Science, University of Queensland, Brisbane, QLD, Australia

Adrian Hills Laboratory of Plant Physiology and Biophysics, University of Glasgow, Glasgow, UK

Henrik Jönsson Sainsbury Laboratory, University of Cambridge, Cambridge, UK

Computational Biology and Biological Physics, Lund University, Lund, Sweden

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK

R. J. Morris Computational and Systems Biology, John Innes Centre, Norwich, Norfolk, UK

Gabriella Mosca Max Planck Institute for Plant Breeding Research, Köln, Germany

Bela M. Mulder Living Matter Department, Institute AMOLF, Amsterdam, The Netherlands

Department of Plant Science, Wageningen University & Research, Wageningen, The Netherlands

Maria Papanatsiou Laboratory of Plant Physiology and Biophysics, University of Glasgow, Glasgow, UK

Przemyslaw Prusinkiewicz Department of Computer Science, University of Calgary, Calgary, AB, Canada

Adam Runions Max Planck Institute for Plant Breeding Research, Köln, Germany

Aleksandra Sapala Max Planck Institute for Plant Breeding Research, Köln, Germany

Akiko Satake Department of Biology, Faculty of Science, Kyushu University, Fukuoka, Japan

Richard S. Smith Max Planck Institute for Plant Breeding Research, Köln, Germany

Soeren Strauss Max Planck Institute for Plant Breeding Research, Köln, Germany

J. Tabak University of Exeter Medical School, University of Exeter, Exeter, UK

José Teles Sainsbury Laboratory, University of Cambridge, Cambridge, UK

Aurélien Tellier Section of Population Genetics, Technical University of Munich, Freising, Germany

K. Tsaneva-Atanasova Department of Mathematics and Living Systems Institute, University of Exeter, Exeter, UK

K. C. A. Wedgwood Department of Mathematics and Living Systems Institute, University of Exeter, Exeter, UK

Daniel Živković Section of Population Genetics, Technical University of Munich, Freising, Germany

About the Editor

Richard J. Morris Richard's research aims to shed light on the physics of information processing in plants. He completed a degree in Mechanical Engineering at the age of 19 before obtaining a BSc in Physics and then an MSc in Theoretical Physics in 1996 from the Erzherzog University of Graz, Austria. He won an EMBL fellowship to carry out his PhD research at the European Molecular Biology Laboratory (EMBL) in the field of computational protein crystallography with Dr Victor Lamzin. After completing his PhD in 2000, Richard joined the group of Dr Gerard Bricogne (MRC-LMB Cambridge and Global Phasing Ltd.) to work on Bayesian approaches for protein structure solution. Richard then joined the group of Prof Dame Janet Thornton, FRS, at the European Bioinformatics Institute (EMBL-EBI) in 2002, where he developed novel shape mathematics for protein function prediction. In 2005, Richard was recruited to the bioinformatics group at the John Innes Centre (JIC) as a tenure-track project leader. Richard played a key role in building up computational biology at JIC. He became Head of the Department for Computational and Systems Biology in 2010. In 2013 he took on the role of institute strategic programme leader as an associate director. He is active in promoting quantitative, and in particular physical, approaches to plant biology and in training the next generation in mathematical modelling and computational methods.

Chapter 1
Physical Models of Plant Morphogenesis

Mathilde Dumond and Arezki Boudaoud

Abstract Biological form is closely associated with function. Yet, despite much progress in developmental biology, we are still far from understanding how organs grow and reach their final size and shape, through a process known as morphogenesis. Morphogenesis is associated with a variety of cellular scale phenomena such as cell expansion, cell proliferation, and cell differentiation. These processes occur within the thousands to billions of cells that yield a well-defined organ. How these phenomena are coordinated over time and space to shape a consistent and reproducible organ or organism is still an open question. In this chapter, we focus on physical models of morphogenesis. We first introduce quantitative descriptions of growth. We then expand on mechanical models of growth; we review types of models and we discuss case studies where such models were used.

1.1 Describing Morphogenesis

To better understand morphogenesis and reliably compare models to experiments, qualitative observations are not sufficient and quantitative measurements are necessary. From an analytical viewpoint, morphogenesis can be dissected as the combination of a small set of elementary transformations. The final shape of an organ results from the integration throughout time of growth. Growth can be decomposed into three parameters: growth rate (differential of area over time), growth anisotropy (ratio between the maximal and the minimal principal directions of growth), and maximal growth direction (see Fig. 1.1) [12, 18, 27, 65]. Formally, growth is a tensor that can be defined similarly to the strain tensor in continuum mechanics. Consider a generic material point of coordinates (x_1, x_2, x_3); it is displaced by growth to $(x_1 + u_1, x_2 + u_2, x_3 + u_3)$, (u_1, u_2, u_3) being the displacement field. The growth tensor is then:

M. Dumond · A. Boudaoud (✉)
Laboratoire Reproduction et Développement des Plantes, Univ Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRA, Lyon, France
e-mail: arezki.boudaoud@ens-lyon.fr

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_1

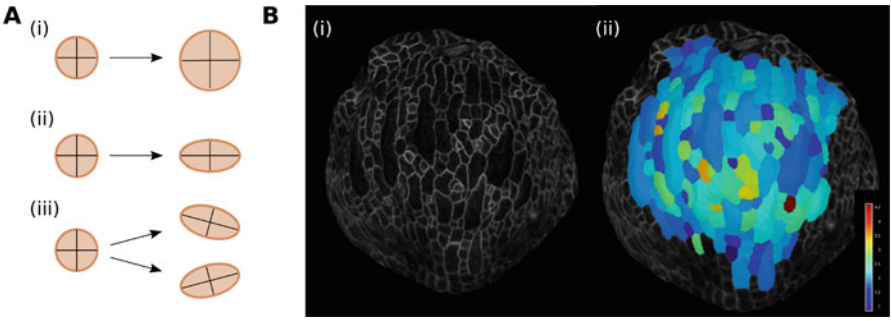


Fig. 1.1 Quantifying morphogenesis. (**a**) Growth can be assessed by monitoring a circle drawn on the tissue (a sphere in 3D): (i) if the circle remains a circle, growth is isotropic; (ii) if the circle becomes an ellipse, growth is anisotropic (growth rate can be deduced from the ratio of surfaces between the two time points; growth anisotropy derives from the ratio in length of the ellipse axes); (iii) growth direction corresponds to the direction of the great axis of the ellipse. (**b**) Example of a software package developed to quantify growth: (i) image of an *Arabidopsis thaliana* sepal with the membrane tagged with a fluorescent molecule; (ii) growth rates quantified over each cell for a 24h interval using MorphoGraphX [5] (the color scale corresponds to the ratio in cell area between the two time points)

$$g_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_i}{\partial x_j} + \sum_m \frac{\partial u_i}{\partial x_m}\frac{\partial u_j}{\partial x_m}\right). \tag{1.1}$$

Note that this definition implicitly assumes that the elastic strain tensor (due to internal or external forces exerted on the growing body) is negligible. Other slightly different definitions, or sometimes rates (time derivatives), are also used. Such quantifications enable the conversion of successive images of organs into quantitative data to perform statistical analyses and to compare models to experiments in a systematic manner.

1.1.1 Quantifying Cell Growth

The starting point is two- or three-dimensional images of tissues or organs showing cell contours, for instance tagged with a fluorescent protein when using confocal microscopy. Quantitative measurements have been facilitated by the development of software that segment cells from such images and measure their growth parameters semi-automatically, in 2D [5], or in 3D [29]. Such software has been used to extract and characterize cell shapes [47, 53, 64], cell growth [15, 37, 42, 66], and to compare mutants to wild-type growth [40, 69]. These cell-based quantifications have been performed so far for rather small organs (less than a thousand of cells). The growth patterns of bigger organs such as older leafs and flowers are often measured at the supracellular level using a continuous description.

1.1.2 Quantifying Organ Growth

The growth patterns of large organs are measured with methods such as landmark analysis and clonal analysis. It is possible to use landmarks on an organ, and measure the relative displacements of the landmarks over time. Early studies considered either natural landmarks such as the vein intersections in leaves [48] or ink-drawn landmarks—grids [3] or set of points [33, 55]—restricting these approaches to relatively older leaves. More recent studies used fluorescent microparticles deposited on leaves [57, 60].

It is also possible to perform clonal analysis, which consists in labelling groups of cells or single cells by expressing a specific heritable marker, and observing their descendants. When a notable fraction of the cells is labelled without any neighbor marked, this enables to define growth rate, anisotropy, and direction at the supracellular level [56]. Because the tissue deforms during growth, the interpretation of clonal analysis requires the use of a model to account for the advection of material points by growth [59].

Finally, two studies have shown that measuring the leaf contour change over time was sufficient to use conformal maps to roughly predict the displacement field of all points inside the leaf [1, 52], because the leaf remains flat all over its development, and because its growth is roughly isotropic at later stages of leaf morphogenesis.

1.2 Forces in Plants

Plant cells are surrounded by a stiff extracellular matrix called the cell wall, put under tension by the internal hydrostatic pressure known as turgor that typically ranges from 0.1 to 1 MPa [8]. Growth is achieved by modulating turgor pressure and cell wall mechanical properties. How can we relate growth to cell mechanics? What is the mechanical status of plant tissues? We here focus on the evidence for forces, as measurements of mechanical properties were reviewed elsewhere [8, 51, 62, 68].

1.2.1 Forces in Tissues and Cells

It has been observed that when peeling a stem, the outer tissue shrinks whereas the internal tissue expands, suggesting that internal cell layers are in compression while the outer layers are under tension [67]. Similarly, cutting a plant tissue leads to a deformation—the cut opens or remains closed according to whether the tissue is in tension or not. Measuring such deformations thus yields information on mechanical stress patterns. For instance, the epidermis of the early sunflower capitulum is under tension in its center and under circumferential compression in the concave region that surrounds the center [23].

Turgor pressure results from osmotic pressure, and so is determined by the difference in osmolyte concentration between the cell and the outer medium. Hence, it is possible to plasmolyze cells by increasing the outer concentration of osmolytes, removing turgor pressure and the tension in the cell wall. Comparing plasmolyzed and turgid cells yields the elastic strain in turgid cell walls and gives information on the stress pattern in cell walls [63].

1.2.2 Forces and Growth

The interplay between turgor pressure and cell wall mechanical properties is at the core of our current understanding of plant cell growth. The growth of the plant cell is due to the yielding of the cell wall under the tension generated by turgor pressure, in addition to the synthesis and export of new wall materials by the cell. From the mechanical point of view, cell wall tension induces deformations that depend on cell wall rheology.

In the simplest rheological model, the cell wall is considered purely elastic: it behaves like a spring. In this case, the deformation, or strain ϵ, depends only on the stress σ applied on the spring, or the cell wall:

$$\epsilon = \sigma / E \tag{1.2}$$

where E the stiffness modulus of the cell wall. The principal limitation of this model is that the spring reverts to its rest shape when the force is released: such a cell wall does not grow. This issue is dealt with within the framework of incremental growth: The spring is loaded and the equilibrium length is taken as the next rest length of the spring. This new rest length initiates the next loading step. Thus growth is modeled as a succession of steps of loading and updating of the rest length.

A more realistic representation of the cell wall is a viscoelastic material. In this case, it behaves like the association of a spring and a damper (see Fig. 1.2):

$$\frac{d\epsilon}{dt} = \sigma/\mu + \frac{d\sigma}{dt}/E \tag{1.3}$$

where μ is the dynamic viscosity. In a few models of growth [16], elasticity is fully neglected and only the damper is accounted for. Note that the incremental approach to growth is equivalent to a viscoelastic model if a timescale is associated with step increments [11]. Here, when the force is released, the material does not revert to its original configuration: this rheology allows cell wall growth.

One of the most elaborate rheological models of the cell wall behavior was introduced in [54], and proposes that the cell wall behaves as a visco-elasto-plastic material (see Fig. 1.2): the cell wall behaves as an elastic material when the stress is lower than a threshold σ_0, but as a viscoelastic material if the stress is larger ($|f|^{(+)} = 0$ if $f < 0$ and else $|f|^{(+)} = f$):

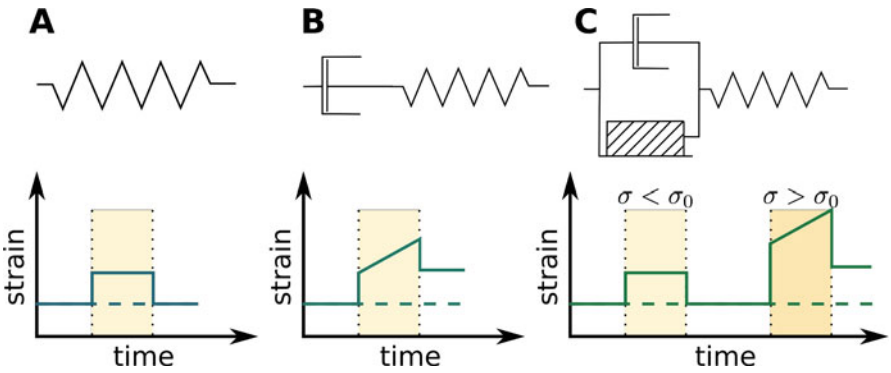


Fig. 1.2 Simple rheological behaviors and plant growth. (**a**) A purely elastic material behaves like a spring: it returns to its original state when applied stress is released. (**b**) A viscoelastic material (of Maxwell type) behaves like the association of a spring and a damper; rate of damper elongation depends on the applied stress. (**c**) The behavior of a visco-elasto-plastic material depends on the value of the force applied: if the stress is smaller than the material-specific threshold σ_0, it behaves elastically, otherwise it behaves viscoelastically. The bottom plots show strain as a function of time, with force applied during the period highlighted in yellow

$$\frac{d\epsilon}{dt} = |\sigma - \sigma_0|^{(+)}/\mu + \frac{d\sigma}{dt}/E \tag{1.4}$$

Cell wall rheological parameters can be measured, using techniques reviewed in [51, 63, 68]. Depending on the question addressed, one can use one or the other of these rheological models. Note that the rheological parameters of the cell wall can be heterogeneous and/or anisotropic. Full models require the generalization of these simple rheological models to 2 or 3 dimensions and the assembly of simple bricks to account for cell and/or tissue geometry and for links between cellular processes and cell wall mechanics.

1.3 Modeling Morphogenesis

1.3.1 Different Types of Models

Models of morphogenesis fall into two main categories: models considering a continuous growing medium and models individualizing each cell.

Continuous models are usually used for large organs, comprising thousands of cells, where cell size is very small compared to organ size. Most studies considered flat organs such as leaves, petals, or sepals (see, e.g., [40, 41]). The surface of the organ is modeled as a 2D surface embedded in 3D space, assuming the thickness is small with respect to other dimensions. 3D models are less common, one example being the morphogenesis of fruits [17]. A widespread assumption for other organs

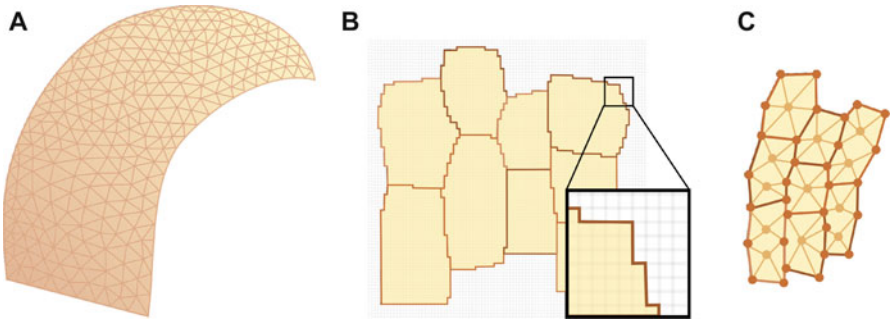


Fig. 1.3 Main types of models: Examples in two dimensions. (**a**) Continuous model: the surface of the organ is continuous, for instance represented by a triangulated mesh. (**b**) Cellular Potts model, where cells are defined on a non-deformable grid. (**c**) Vertex-based model, the cells are defined by vertices (circles) and their edges with the neighboring cells

such as stems is that the epidermis is dominant in the control of growth, because the surface cell wall is thicker and stiffer than internal walls (see, e.g., [9] for a discussion). This enables modeling the dynamics of the surface of the organ [39]. Formally, these models come in the form of partial differential equations, often obtained by accounting for cell wall rheology and mechanical equilibrium. The finite element method is often used to solve them because the triangular meshes are well-suited to domains of arbitrary geometry.

Cells are individualized in other types of models, such as cellular Potts models or vertex-based models (see Fig. 1.3). In the cellular Potts model, cells are defined on a discrete fine grid, and the status of each point of the grid is updated depending on a set of rules, leading to the movement of the edges of the cells. This framework was originally developed for physical systems such as foams. It is widely used in the animal field, and can have a finer subcellular resolution than the vertex-based models discussed hereafter. Nevertheless, the cellular Potts model seems to have been used only once for plants, in the context of auxin concentration dynamics in a growing root [35]. Indeed, this framework is not well-suited for an elastic material and is commonly used to model purely viscous materials such as animal cells.

Vertex models are broadly used to investigate plant development [25, 49]. Cells are often assumed to be polygonal, so that cell shape is defined by the position of vertices and their dynamics in space. Sometimes, edges are allowed to be curved, for instance assumed to be arc of circles [20]. The mechanical elements can be placed at cell edges (e.g. spring-damper systems) accounting for anticlinal cell walls [20, 25], or on the whole surface of the cell accounting for periclinal cell walls [22, 49, 61]. Such models can also incorporate gene regulatory networks, cell–cell communication, or cell division [25].

More recently, 2D models started to combine vertex-based and continuous approaches, with cells individualized and their cell walls represented as continuous structures [14, 30]. This enables to simultaneously model cell scale behavior (cell division, cell shape) and consider the mechanical properties of the periclinal cell

wall at a subcellular resolution. This approach was extended to 3D, enabling for the first time to model morphogenesis a 3D tissue at cellular resolution [12]: by accounting for epidermis and for internal layers, the authors investigated the relative role of cell layers in the outgrowth of organ primordia in the shoot apical meristem.

1.3.2 Implementation of Growth

So far, growth has been modeled using two different approaches. In the first type of description, growth rate, direction, and anisotropy are specified or inferred from a gene regulatory network at each point in time and space of the simulation. The tissue is assumed to be elastic and the equilibrium state is computed from force balance [41]. This amounts to prescribing the rest length of an assembly of springs and computing their equilibrium lengths.

In the second type of description, the cell wall's mechanical properties such as elastic modulus or viscosity (these properties may be anisotropic) are specified or inferred from regulatory networks, and the mechanical equilibrium under loading by turgor pressure defines the current growth rate. In this framework, the tissue deformation depends on the tissue rheology, which can be one of the types previously presented: elastic, viscoelastic, or visco-elasto-plastic.

1.4 Case Studies: The Use of Models to Understand Plant Morphogenesis

We now illustrate the concepts outlined above by discussing representative physical models of morphogenesis.

1.4.1 Morphogenesis of an Isolated Plant Cell

A first step towards understanding organ morphogenesis is to study cell morphogenesis. A classical system of interest is the pollen tube, which is one of the model systems for tip growth: elongation of the tube by expansion of the cell wall localized at the cap of the cylinder. During pollination, the pollen grain lands on the summit of the carpel and germinates. The pollen tube emerges and subsequently grows into the carpel reaching the ovule. How a cell can grow in such directional manner has been extensively investigated.

Considering the cell wall as a hyperelastic (extension of linear elasticity to large deformation) membrane and using an incremental approach to growth, [32] showed that a lower elastic modulus at the tip of the tube was sufficient to produce

self-similar tip growth. A fully viscous model led to similar conclusions [16]. In the visco-elasto-plastic model developed in [24], the authors also needed a softer tip; they found, however, that the cell wall anisotropy was required to retrieve self-similar tip growth, except in very few specific cases. In these three studies, the equations were numerically solved based on the circumferential symmetry of the pollen tube, making this modeling framework difficult to extend to other systems. The study in [28] released this assumption of axisymmetry and used the finite element method for numerical solutions. They implemented an incremental approach to growth, considering the cell wall as an elastic material. The influence of anisotropy and of the steepness of the gradient of stiffness over the edges of the tube were tested, and the interaction between these two quantities allowed self-similar growth depending on the parameters: a steeper cell wall stiffness gradient was associated with a more isotropic cell wall to produce self-similar growth. Finally, [58] introduced a model coupling cell wall chemistry with mechanics, assuming a viscoelastic rheology in which the viscosity depends on the concentration of cross-links in the wall. In addition to self-similar tip growth, they retrieved the oscillations in growth rate and tube diameter observed in fast growing tubes.

1.4.2 Growth Motion of an Elongated Organ

Plants cannot move, but they react to their environment: for instance, the main shoot and the main root bend towards the gravity vector. The molecular mechanisms involved are relatively well described, and the integration of these mechanisms during growth has been investigated using models coupling biomechanical and biochemical processes [31, 70]. The differential localization of auxin transporters from the PIN-FORMED family leads to differential concentrations of auxin, which in turn cause differential growth rates along the transverse axis of roots and hypocotyl, ultimately inducing a bending of the organ. These studies focused on the relationship between auxin, cell differential growth, and bending initiation, but they did not fully investigate how the vertical orientation of the organ is reached. Actually, when only gravity-sensing is taken into account, the shoot oscillates around the axis of gravity whereas real shoots reach this orientation [6]. Sensing the local curvature (proprioception) needs to be included to reproduce the observed dynamics of stem curving, in the case of both gravitropism [6] and phototropism [7].

1.4.3 Shaping a Sheet-Like Organ

Volvox is a green algae in the form of a spherical sheet of cells with an aperture. One major even in the morphogenesis of Volvox is its inversion: the organism turns inside out. The inversion of the sheet of cells is associated with a sequence of deformations where cells firstly circularly invaginate at the equator, accompanied by the posterior

hemisphere which moves into the anterior and inverts as well [38]. Finally, the aperture stretches out over the posterior hemisphere. The deformation of the sheet is associated with cell shape changes, but their role in the sheet inversion remained poorly understood. Modeling the sheet as an elastic material in which the rest state is actively controlled (this is formally similar to the models prescribing growth), the authors showed that the shift in the sheet curvature at the equator combined with the contraction of the posterior hemisphere were sufficient to trigger this major morphogenetic event [38].

In angiosperms, the main sheet-like organs are leaves and petals. The shapes of such organs are very diverse and can change drastically from one species to another. The underlying regulatory networks are very intricate and it is difficult to relate them to final organ shapes. The Snapdragon corolla, in particular, has a very elaborated, asymmetric shape. The wild-type and several morphogenetic mutants have been successfully modeled using a continuous approach based on incremental elastic growth [34]. This model removes mechanical stress at each step when using the equilibrium configuration to define the following rest configuration. This stress is induced by spatial gradients in growth rates (e.g. fast growing regions exert pressure on other regions) and is called residual stress.

Models prescribing mechanical properties instead of growth rates generally account for residual stress [12]. Such residual stress may have significant effect on morphogenesis, for instance in the case of thin organs. Larger growth rates at the edge of these organs induces compressive stress there, which leads to the buckling of the edges of the organ into a wavy shape [2]. Such waviness of edges is observed in leaves and petals of many species, such as in Lily [44, 45]. Based on this, it is likely that a specific regulation of growth rates is required for leaves or petals to remain flat [1, 52].

1.4.4 Feedback Through Mechanical Signals

Mechanics are at the core of morphogenesis: growing cells interact mechanically during morphogenesis, relaxing and generating residual stress. Can this mechanical stress have an impact on cell behavior? It has been shown that plant cells can sense and react to mechanical stress by orientating cortical microtubule networks in the direction of maximal principal stress [36], leading to the synthesis of cellulose microfibrils in this direction and to the mechanical reinforcement of the cell wall along mechanical stress. The consequences on organ morphogenesis were investigated in Arabidopsis thaliana sepals [37]. The authors used an incremental model, with a prescribed elastic modulus, and a mechanical anisotropy imposed by the anisotropy of stress and following the same orientation. They obtained a gradient in growth rates with a slowly growing tip and a fast growing base, leading to a transverse tension in the tip and a mechanical reinforcement there. By comparing simulations with mutants affected in sensing mechanical stress, they showed that this mechanical feedback enabled the modulation of organ shape. Similarly, a vertex

model accounting for cell divisions and viscoelastic periclinal cell walls helped showing that cell divisions follow the direction of maximal stress in the shoot apical meristem [46]. In all these studies, it was assumed that cells sensed mechanical stress.

However, cells could sense either stress or strain. The two are usually correlated, but the direction of maximal strain and maximal stress may differ when the material is mechanically anisotropic. [13] investigated which of the two sensing mechanisms was more plausible in the case of microtubule orientation and cellulose synthesis. They modeled the plant epidermis as an elastic 2D surface embedded in 3D and pressurized from inside by turgor, increasing elastic modulus along the direction of the maximal stress or along maximal strain, and they examined the subsequent elastic strain, considered as a proxy for growth. Simulations where the cellulose orientation followed stress were in accordance with experimental observations, whereas simulations where cellulose oriented depending on strain were less stable and disagreed with observations.

Mechanical cues can also guide differentiation [26]. Models have started addressing how this may pattern growing organs. For instance, the epidermis of leaves is stiffer than internal tissues, so that the effect of turgor is tension in the epidermis and compression in internal tissues. Models assumed that such compression, when above a threshold, leads to the differentiation of ground cells into pro-vascular cells [21, 43]. This mechanism is sufficient to produce venation patterns that are similar to the patterns observed in dicotyledon leaves [21, 43]. However, it also established that biochemical patterning by auxin flow is crucial for venation [10]. It may well be that the combination chemical signals and mechanical signals provides robustness to vascular patterning.

1.4.5 Variability and Morphogenesis

All the models discussed so far are deterministic: They describe the expected average behavior of the system. However, cells in an organ are variable [50]. We here consider three examples of mechanical variability in flat organs.

Modeling the wheat leaf using a cell-based model which takes turgor pressure and water movements into account, [71] showed that turgor pressure was variable between the cells of the wheat leaf, and that this variability correlated with cell identity.

In Arabidopsis sepals, a mutant showing more variability of shape was less heterogenous spatially than wild-type concerning growth rates and mechanical properties [40]. An incremental organ growth model with random mechanical properties showed that increasing the correlation length of elastic modulus allowed to retrieve the observed changes from wild type to mutant [40].

The growth of a leaf depends on its venation pattern, because veins are stiffer than ground tissue. Areas surrounded by veins (areoles) grow at different rates, depending on their geometry and on the thickness of neighboring veins, which

likely determines their stiffness [19]. In order to disentangle the links between global leaf growth and venation, [4] investigated the effect of external stress on venation networks. They found that the venation network reorients towards the direction of the applied stress. Using a vein-based model with veins represented as viscoelastic rods and growth driven by turgor, they showed that external stress increased the variability in growth, opposite to experimental observations. In order to reconcile model with experiments, they needed to include randomness in vein thickness and a threshold in tension for growth to occur, considering an elasto-visco-plastic rheology for veins.

1.5 Conclusion

As developmental biology is becoming more and more quantitative, mechanical models are increasingly used to test hypotheses and help understand the mechanisms behind morphogenesis. Many approaches and types of models are now available. However, a number of questions remain open. For instance, is there a unifying framework to model growth in plants? What is the exact rheology of plant cells? These issues are exemplified by models for pollen tubes: 4 different types of rheology all yield self-similar growth. A continuous dialogue between experiments and theory will be instrumental in future progress.

References

1. Alim K, Armon S, Shraiman BI, Boudaoud A (2016) Leaf growth is conformal. Phys Biol 13(5):05LT01
2. Audoly B, Boudaoud A (2003) Self-similar structures near boundaries in strained systems. Phys Rev Lett 91(8):86105
3. Avery GS (1933) Structure and development of the tobacco leaf. Am J Bot 20(9):565–592
4. Bar-Sinai Y, Julien J-D, Sharon E, Armon S, Nakayama N, Adda-Bedia M, Boudaoud A (2016) Mechanical stress induces remodeling of vascular networks in growing leaves. PLOS Comput Biol 12(4):e1004819
5. Barbier de Reuille P, Routier-Kierzkowska A-L, Kierzkowski D, Bassel GW, Schüpbach T, Tauriello G, Bajpai N, Strauss S, Weber A, Kiss A, Burian A, Hofhuis H, Sapala A, Lipowczan M, Heimlicher MB, Robinson S, Bayer EM, Basler K, Koumoutsakos P, Roeder AHK, Aegerter-Wilmsen T, Nakayama N, Tsiantis M, Hay A, Kwiatkowska D, Xenarios I, Kuhlemeier C, Smith RS (2015) MorphoGraphX: a platform for quantifying morphogenesis in 4D. eLife 4:1–20
6. Bastien R, Bohr T, Moulia B, Douady S (2013) Unifying model of shoot gravitropism reveals proprioception as a central feature of posture control in plants. Proc. Natl. Acad. Sci. U. S. A. 110(2):755–760
7. Bastien R, Douady S, Moulia B (2015) A unified model of shoot tropism in plants: photo-, gravi- and propio-ception. PLoS Comput Biol 11(2):1–30
8. Beauzamy L, Nakayama N, Boudaoud A (2014) Flowers under pressure: ins and outs of turgor regulation in development. Ann Bot 114(7):1517–1533

9. Beauzamy L, Louveaux M, Hamant O, Boudaoud A (2015) Mechanically, the shoot apical meristem of arabidopsis behaves like a shell inflated by a pressure of about 1 MPa. Front Plant Sci 6:1038
10. Berleth T, Scarpella E, Prusinkiewicz P (2007) Towards the systems biology of auxin-transport-mediated patterning. Trends Plant Sci 12(4):151–159
11. Bonazzi D, Julien JD, Romao M, Seddiki R, Piel M, Boudaoud A, Minc N (2014) Symmetry breaking in spore germination relies on an interplay between polar cap stability and spore wall mechanics. Dev Cell 28(5):534–546
12. Boudon F, Chopard J, Ali O, Gilles B, Hamant O, Boudaoud A, Traas J, Godin C (2015) A computational framework for 3D mechanical modeling of plant morphogenesis with cellular resolution. PLoS Comput Biol 11(1):e1003950
13. Bozorg B, Krupinski P, Jönsson H (2014) Stress and strain provide positional and directional cues in development. PLoS Comput Biol 10(1):e1003410
14. Bozorg B, Krupinski P, Jönsson H (2016) A continuous growth model for plant tissue. Phys Biol 13(6):065002
15. Bringmann M, Bergmann DC (2017) Tissue-wide mechanical forces influence the polarity of stomatal stem cells in arabidopsis. Curr Biol 27(6):1–7
16. Campàs O, Mahadevan L (2009) Shape and dynamics of tip-growing cells. Curr Biol 19(24):2102–2107
17. Cieslak M, Cheddadi I, Boudon F, Baldazzi V, Génard M, Godin C, Bertin N (2016) Integrating physiology and architecture in models of fruit expansion. Front Plant Sci 7:1–19
18. Coen E, Rolland-Lagan A-G, Matthews M, Bangham JA, Prusinkiewicz P (2004) The genetics of geometry. Proc Natl Acad Sci 101(14):4728–4735
19. Corson F, Adda-Bedia M, Boudaoud A (2009) In silico leaf venation networks: growth and reorganization driven by mechanical forces. J Theor Biol 259(3):440–448
20. Corson F, Hamant O, Bohn S, Traas J, Boudaoud A, Couder Y (2009) Turning a plant tissue into a living cell froth through isotropic growth. Proc Natl Acad Sci 106(21): 8453–8458
21. Corson F, Henry H, Adda-Bedia M (2010) A model for hierarchical patterns under mechanical stresses. Phil Mag 90(1–4):357–373
22. De Vos D, Vissenberg K, Broeckhove J, Beemster Gerrit TS (2014) Putting theory to the test: which regulatory mechanisms can drive realistic growth of a root? PLoS Comput Biol 10(10):e1003910
23. Dumais J, Steele CR (2000) New evidence for the role of mechanical forces in the shoot apical meristem. J Plant Growth Regul 19:7–18
24. Dumais J, Shaw SL, Steele CR, Long SR, Ray PM (2006) An anisotropic-viscoplastic model of plant cell morphogenesis by tip growth. Int. J. Dev. Biol. 50:209–222
25. Dupuy L, MacKenzie J, Rudge T, Haseloff J (2008) A system for modelling cell-cell interactions during plant morphogenesis. Ann Bot 101(8):1255–1265
26. Engler AJ, Sen S, Sweeney HL, Discher DE (2006) Matrix elasticity directs stem cell lineage specification. Cell 126(4):677–689
27. Erickson RO (1976) Modeling of plant growth. Annu Rev Plant Physiol 27:407–434
28. Fayant P, Girlanda O, Chebli Y, Aubin C-E, Villemure I, Geitmann A (2010) Finite element model of polar growth in pollen tubes. Plant Cell 22(8):2579–2593
29. Fernandez R, Das P, Mirabet V, Moscardi E, Traas J, Verdeil J-L, Malandain G, Godin C (2010) Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution. Nat Methods 7(7):547–53
30. Fozard JA, Lucas M, King JR, Jensen OE (2013) Vertex-element models for anisotropic growth of elongated plant organs. Front Plant Sci 4:233
31. Fozard JA, Bennett MJ, King JR, Jensen OE (2016) Hybrid vertex-midline modelling of elongated plant organs. Interface Focus 6(5):20160043
32. Goriely A, Tabor M (2003) Self-similar tip growth in filamentary organisms. Phys Rev Lett 90(10):108101 (2003)

33. Granier C, Tardieu F (1998) Spatial and temporal analyses of expansion and cell cycle in sunflower leaves. A common pattern of development for all zones of a leaf and different leaves of a plant. Plant Physiol 116(3):991–1001
34. Green AA, Kennaway JR, Hanna AI, Bangham JA, Coen E (2010) Genetic control of organ shape and tissue polarity. PLoS Biol 8(11):e1000537
35. Grieneisen VA, Xu J, Marée AFM, Hogeweg P, Scheres B (2007) Auxin transport is sufficient to generate a maximum and gradient guiding root growth. Nature 449(7165):1008–1013
36. Hamant O, Heisler MG, Jonsson H, Krupinski P, Uyttewaal M, Bokov P, Corson F, Sahlin P, Boudaoud A, Meyerowitz EM, Couder Y, Traas J (2008) Developmental patterning by mechanical signals in arabidopsis. Science 322(5908):1650–1655
37. Hervieux N, Dumond M, Sapala A, Routier-Kierzkowska A-L, Kierzkowski D, Roeder AHK, Smith RS, Boudaoud A, Hamant O (2016) A mechanical feedback restricts sepal growth and shape in arabidopsis. Curr Biol 26(8):1019–1028
38. Höhn S, Honerkamp-Smith AR, Haas PA, Trong PK, Goldstein RE (2015) Dynamics of a volvox embryo turning itself inside out. Phys Rev Lett 114(17):1–5
39. Holloway DM, Harrison LG (2008) Pattern selection in plants: coupling chemical dynamics to surface growth in three dimensions. Ann Bot 101(3):361–374
40. Hong L, Dumond M, Tsugawa S, Sapala A, Routier-Kierzkowska A-L, Zhou Y, Chen C, Kiss A, Zhu M, Hamant O, Smith RS, Komatsuzaki T, Li C-B, Boudaoud A, Roeder AHK (2016) Variable cell growth yields reproducible organdevelopment through spatiotemporal averaging. Dev Cell 38(1):15–32
41. Kennaway R, Coen E, Green A, Bangham JA (2011) Generation of diverse biological forms through combinatorial interactions between tissue polarity and growth. PLoS Comput Biol 7(6):e1002071
42. Kierzkowski D, Nakayama N, Routier-Kierzkowska A-L, Weber A, Bayer EM, Schorderet M, Reinhardt D, Kuhlemeier C, Smith RS (2012) Elastic domains regulate growth and organogenesis in the plant shoot apical meristem. Science 335(6072):1096–1099
43. Laguna MF, Bohn S, Jagla EA (2008) The role of elastic stresses on leaf venation morphogenesis. PLOS Comput Biol 4(4):e1000055
44. Liang H, Mahadevan L (2009) The shape of a long leaf. Proc Natl Acad Sci U. S. A. 106(52):22049–54
45. Liang H, Mahadevan L (2011) Growth, geometry, and mechanics of a blooming lily. Proc Natl Acad Sci U.S.A 108(14):5516–5521
46. Louveaux M, Julien J-D, Mirabet V, Boudaoud A, Hamant O (2016) Cell division plane orientation based on tensile stress in Arabidopsis thaliana. Proc Natl Acad Sci 113(30):E4294–E4303
47. Lucas M, Kenobi K, von Wangenheim D, Voss U, Swarup K, De Smet I, Van Damme D, Lawrence T, Peret B, Moscardi E, Barbeau D, Godin C, Salt D, Guyomarc'h S, Stelzer EHK, Maizel A, Laplaze L, Bennett MJ (2013) Lateral root morphogenesis is dependent on the mechanical properties of the overlaying tissues. Proc Natl Acad Sci 110(13):5229–5234
48. Maksymowych R (1959) Quantitative analysis of leaf development in xanthium pensylvanicum. Am J Bot 46(9):635–644
49. Merks RM, Guravage M, Inzé D, Beemster GTS (2011) VirtualLeaf: an open source framework for cell-based modeling of plant tissue growth and development. Plant Phys 155(2): 656–666
50. Meyer HM, Roeder AHK (2014) Stochasticity in plant cellular growth and patterning. Front Plant Sci 5:420
51. Milani P, Braybrook SA, Boudaoud A (2013) Shrinking the hammer: micromechanical approaches to morphogenesis. J Exp Bot 64(15):4651–4662
52. Mitchison G (2016) Conformal growth of Arabidopsis leaves. J Theor Biol 408:155–166
53. Montenegro-Johnson TD, Stamm P, Strauss S, Topham AT, Tsagris M, Wood ATA, Smith RS, Bassel GW (2015) Digital single-cell analysis of plant organ development using 3DCellAtlas. Plant Cell 27(4):1018–1033

54. Ortega JK (1985) Augmented growth equation for cell wall expansion. Plant Physiol 79(1):318–320
55. Poethig RS, Sussex IM (1985) The developmental morphology and growth dynamics of the tobacco leaf. Planta 165(2):158–169
56. R S L B P B S Record: 2290 Poethig (1987) Clonal analysis of cell lineage patterns in plant development. Am J Bot 74(4):581–594
57. Remmler L, Rolland-Lagan AG (2012) Computational method for quantifying growth patterns at the adaxial leaf surface in three dimensions. Plant Physiol 159:27–39
58. Rojas ER, Hotton S, Dumais J (2011) Chemically mediated mechanical expansion of the pollen tube cell wall. Biophys J 101(8):1844–1853
59. Rolland-Lagan A-G, Bangham JA, Coen E (2003) Growth dynamics underlying petal shape and asymmetry. Nature 422(6928):161–163
60. Rolland-Lagan A-G, Remmler L, Girard-Bock C (2014) Quantifying shape changes and tissue deformation in leaf development. Plant Physiol 165:496–505
61. Romero-Arias JR, Hernández-Hernández V, Benítez M, Alvarez-Buylla ER, Barrio RA (2017) Model of polar auxin transport coupled to mechanical forces retrieves robust morphogenesis along the <math> <mi mathvariant="italic">Arabidopsis</mi> </math> root. Phys Rev E 95(3):032410
62. Routier-Kierzkowska A-L, Smith RS (2013) Measuring the mechanics of morphogenesis. Curr Opin Plant Biol 16(1):25–32
63. Routier-Kierzkowska A-L, Weber A, Kochova P, Felekis D, Nelson BJ, Kuhlemeier C, Smith RS, Breakthrough Technologies (2012) Cellular force microscopy for in vivo measurements of plant tissue mechanics. Plant Physiol 158:1514–1522
64. Sampathkumar A, Gutierrez R, McFarlane HE, Bringmann M, Lindeboom J, Emons A-M, Samuels L, Ketelaar T, Ehrhardt DW, Persson S (2013) Patterning and lifetime of plasma membrane-localized cellulose synthase is dependent on actin organization in Arabidopsis interphase cells. Plant Physiol 162(2):675–688
65. Silk WK, Erickson RO (1979) Kinematics of plant growth. J Theor Biol 76(4):481–501
66. Tauriello G, Meyer HM, Smith RS, Koumoutsakos P, Roeder AHK (2015) Variability and constancy in cellular growth of Arabidopsis sepals. Plant Physiol 169:2342–2358
67. Vandiver R, Goriely A (2008) Tissue tension and axial growth of cylindrical structures in plants and elastic tissues. Europhys Lett 84, 58004
68. Vogler H, Felekis D, Nelson B, Grossniklaus U, Measuring the mechanical properties of plant cell walls. Plants 4(2):167–182
69. Yang W, Schuster C, Beahan CT, Doblin MS, Wightman R, Meyerowitz EM, Yang W, Schuster C, Beahan CT, Charoensawan V, Peaucelle A, Bacic A (2016) Regulation of meristem morphogenesis by cell wall synthases in arabidopsis article regulation of meristem morphogenesis by cell wall synthases in arabidopsis. Curr Biol 26(11):1404–1415
70. Žádníková P, Wabnik K, Abuzeineh A, Gallemi M, Van Der Straeten D, Smith RS, Inzé D, Friml J, Prusinkiewicz P, Benková E (2016) A model of differential growth-guided apical hook formation in plants. Plant Cell 28(10):2464–2477
71. Zubairova U, Nikolaev S, Penenko A, Podkolodnyy N, Golushko S, Afonnikov D, Kolchanov N (2016) Mechanical behavior of cells within a cell-based model of wheat leaf growth. Front Plant Scie 7:1–15

Chapter 2
Fluid Transport in Plants

M. G. Blyth and R. J. Morris

Abstract Fluid motion is of fundamental importance for plant survival, growth and development. This distribution of water and nutrients is achieved by hydraulics. Fluid flow also plays a key role in long-distance signalling, allowing plants to adapt to environmental challenges. Fluid dynamics thus maintains plant vitality and health. In this chapter we derive the basic governing equations for fluid motion from first principles and describe the pertinent boundary conditions. Pressure-driven flow in a tube is discussed as a conceptualised model of fluid transport in the plant's vasculature system. We also discuss solute transport with particular reference to the individual roles played by convection and diffusion and the enhanced dispersive effect that can be achieved when these two effects work in unison.

2.1 An Overview of Plant Hydraulics

Typically around 70% of a plant cell is water. Water pressure within cells is required for plants to maintain their mechanical properties, as described in other chapters in this book. Water influx into cells is achieved by osmosis. Osmosis is the flow of a solvent across a semipermeable membrane, such as the plasma membrane which surrounds a cell, driven by the chemical potential arising from a solute concentration gradient. Water influx is key for cell expansion and growth. Most of the water used by land plants is absorbed from the soil by the roots. Plants thus need a water distribution system to supply water to other tissues. Likewise, numerous micronutrients are taken up from the soil by roots and need to be transported throughout the plant. Water is also used for temperature control by cooling through vapour loss to the surrounding air. Water transport from root to

M. G. Blyth (✉)
School of Mathematics, University of East Anglia, Norwich, UK
e-mail: m.blyth@uea.ac.uk

R. J. Morris
Computational and Systems Biology, John Innes Centre, Norwich, Norfolk, UK
e-mail: Richard.Morris@jic.ac.uk

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_2

shoot is enabled by the *xylem*. The process of photosynthesis, in which carbon dioxide is converted to sugars using light energy, also requires water. The carbon that is taken from atmospheric gas and placed within biologically accessible sugar molecules is termed *fixed*. This fixed carbon is a key building block for numerous complex biological molecules and as such needs to be transported to those cells that cannot photosynthesise to provide their own source material. The transport of sugars occurs in the *phloem*. The main transport routes for fluids are depicted in Fig. 2.1. These structures have complex geometries that hinder mathematical analysis and, for this reason, it is traditional to model them as cylindrical tubes. The fluids themselves are complex solutions of water, ions, small molecules and biological polymers. In the interest of simplicity, we will assume that they can be treated as Newtonian fluids, meaning that the internal stress is a linear function of the strain rate, as will be discussed in more detail below.

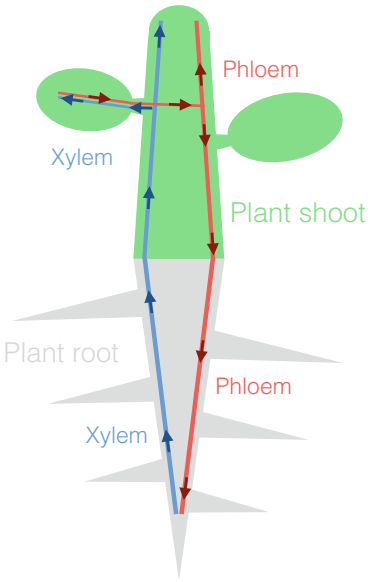


Fig. 2.1 Main hydraulic transport routes within a plant. Plants use waterways for long-distance transport: sugars are transported from where they are produced in mature leaves to other tissues through the phloem; water and minerals are transported from root to shoot through the xylem. These waterways operate through different dedicated transport networks using distinct physical processes to drive them. Flow through the phloem is driven by osmolarity-induced pressure differences, whereas flow through the xylem is driven by transpiration. In addition to resource allocation, effective communication is key for the success of multicellular organisms and plants use their long-distance transport streams to coordinate their activities, for instance to inform distal tissues of locally perceived challenges. Xylem and phloem are known to play important roles in signalling. In particular the phloem has been shown to contain a wide range of signalling molecules. So, not only does the phloem play a key role in delivering nutrients but also in transporting messages over long distances and thus coordinating growth, development and various physiological responses

2.2 Physical Principles and Basic Equations

In this section we introduce the fundamental physical principles and mathematical equations describing the motion of a Newtonian fluid. A Newtonian fluid represents an idealisation in which the shear stress within the fluid is assumed to be proportional to the rate of strain, the constant of proportionality being the fluid viscosity. This is known as Newton's law of viscosity. So-called non-Newtonian fluids obey different laws and may change their viscosity under stress. The following derivations proceed on the tacit assumption that we can treat a fluid as being effectively infinitely divisible such that at any point we can unambiguously ascribe a unique value to physical quantities such as density, mass, pressure and so on. This assumption is known as the *continuum hypothesis*.

2.2.1 Conservation of Mass

Consider an infinite fluid in motion in three-dimensional space, and imagine within the fluid an arbitrary closed volume V with surface S which is fixed in both space and time, and which has a unit outward-pointing surface normal vector $\boldsymbol{n}$, as is illustrated in Fig. 2.2. It should be emphasised that the volume V is purely notional and its surface is not a physical boundary; the fluid is free to flow into and out of the volume unhindered. Our aim is to write down a statement expressing the idea that, in the absence of any source of fluid within V, the total fluid mass inside V changes at a rate equal to the net inflow (or outflow) of fluid particles across S. If $\boldsymbol{u}(\boldsymbol{x}, t)$ and $\rho(\boldsymbol{x}, t)$ are, respectively, the fluid velocity and density at a general point in space $\boldsymbol{x}$ at time t, then this statement takes the form,

$$\frac{\mathrm{d}}{\mathrm{d}t} \iiint_V \rho \, \mathrm{d}V = - \iint_S \rho \boldsymbol{u} \cdot \boldsymbol{n} \, \mathrm{d}S. \tag{2.1}$$

The first term on the left-hand side of (2.1) represents the rate of change of the total fluid mass within V. The term on the right-hand side represents the rate at which fluid mass is leaving V by passing through the surface S (note that the dot product of $\boldsymbol{u}$ with $\boldsymbol{n}$ is taken since it is only motion normal to the boundary S which

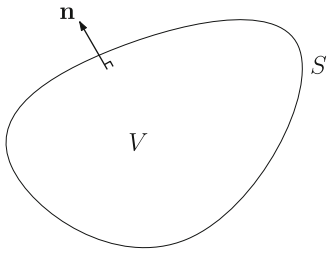


Fig. 2.2 Sketch of a notional, closed volume of fluid V with surface S and unit outward-pointing normal $\boldsymbol{n}$

is relevant). If the term on the right-hand side is positive/negative, then the total mass in V increases/decreases.

Since V is fixed we can shift the time derivative on the left-hand side inside the integral. Moving the term on the right-hand side across the equals sign, and making use of the divergence theorem (e.g. Schey et al. [19]), we obtain

$$\iiint_V \big(\rho_t + \nabla \cdot (\rho \boldsymbol{u})\big)\, \mathrm{d}V = 0. \tag{2.2}$$

Appealing to the fact that the choice of volume V was arbitrary, we conclude that the integrand in (2.2) must itself vanish, and therefore

$$\rho_t + \nabla \cdot (\rho \boldsymbol{u}) = 0. \tag{2.3}$$

This is the equation of *conservation of mass*.

2.2.2 The Navier-Stokes Equation

Consider now the momentum balance within the fixed volume V. Newton's second law of motion (e.g. Goldstein et al. [6]) requires that the total force acting on the volume must be equal to the rate of change of momentum within the volume. We express this idea mathematically by writing the balance

$$\frac{\mathrm{d}}{\mathrm{d}t} \iiint_V \rho \boldsymbol{u}\, \mathrm{d}V + \iint_S (\rho \boldsymbol{u})(\boldsymbol{u} \cdot \boldsymbol{n})\, \mathrm{d}S = \iiint_V \rho \boldsymbol{F}\, \mathrm{d}V + \iint_S \boldsymbol{\sigma} \cdot \boldsymbol{n}\, \mathrm{d}S. \tag{2.4}$$

The first term on the left-hand side of (2.4) represents the rate of change of the total momentum within V. The second term on the left-hand side expresses the rate of departure of momentum, carried by fluid particles, out of the fixed volume V through the surface S. The first term on the right-hand side is the total body force acting on the fluid inside V, for example due to gravity; here, $\boldsymbol{F}$ is the body force per unit mass. The second term on the right-hand side represents the viscous stress imposed on the surface S by the surrounding fluid; here, $\boldsymbol{\sigma}$ is the stress tensor, which for a Newtonian fluid may be written in the form (e.g. Batchelor [2]), $\boldsymbol{\sigma} = -p\boldsymbol{I} + \mu(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T)$, where p is the fluid pressure, $\boldsymbol{I}$ is the identity matrix and μ is the dynamic viscosity of the fluid.

Applying the divergence theorem to the second term on the left-hand side of (2.4) and to the second term on the right-hand side of (2.4), moving the time derivative on the first term on the left-hand side into the integrand, making use of the statement of conservation of mass (2.3), and finally appealing to the fact that the volume V is arbitrary, we arrive at the Navier-Stokes equation,

$$\rho(\boldsymbol{u}_t + \boldsymbol{u} \cdot \nabla \boldsymbol{u}) = -\nabla p + \rho \boldsymbol{F} + \mu \nabla^2 \boldsymbol{u}. \tag{2.5}$$

This equation is named after *Claude-Louis Navier* (1785–1836) and *George Gabriel Stokes* (1819–1903). It is complemented by the statement of conservation of mass (2.3), which, as has been already noted, was also used to derive it. In the case of an incompressible flow, which occurs, for example, when the fluid density ρ is constant, the latter reduces to the requirement that the velocity field be solenoidal, namely

$$\nabla \cdot \boldsymbol{u} = 0. \tag{2.6}$$

Equations (2.5) and (2.6) govern the motion of an incompressible Newtonian viscous fluid, and will be used in this chapter as the basic model equations for studying fluid motion in plants.

2.2.3 Boundary Conditions

To complete the mathematical description, boundary conditions must be appended to the momentum equation (2.5) and the conservation of mass equation (2.6). Most flows of practical interest are confined, or partially confined, by physical boundaries. For example, from a simplistic viewpoint, flow in the interior of a plant cell is confined by the cell wall and by the surface of the vacuole. It is therefore appropriate to consider what conditions should be imposed on a fluid motion at a solid surface.

One such condition follows from the intuitively obvious statement that a fluid cannot penetrate a solid, impermeable surface. Referring to a boundary Σ which is moving at velocity $\boldsymbol{U}(\boldsymbol{x}, t)$, where $\boldsymbol{x}$ is a point on the boundary and t is time, this condition is expressed mathematically as

$$\boldsymbol{u} \cdot \boldsymbol{n} = \boldsymbol{U} \cdot \boldsymbol{n} \quad \text{on } \Sigma, \tag{2.7}$$

where $\boldsymbol{n}$ is a unit normal to Σ. Physically it stipulates that the fluid particles located on Σ move in a direction normal to Σ with the normal component of velocity of Σ. This is usually referred to as the *no penetration* or *no normal flow* condition.

A second boundary condition requires that the tangential velocity of the fluid particles on Σ should match the tangential velocity of Σ itself. This demands that

$$\boldsymbol{P} \cdot \boldsymbol{u} = \boldsymbol{P} \cdot \boldsymbol{U} \quad \text{on } \Sigma, \tag{2.8}$$

where $\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{nn}$ is a projection matrix,[1] which has the effect of removing the normal component of the vector on which it is operating. This is usually referred to

[1] In the definition of $\boldsymbol{P}$, the term $\boldsymbol{nn}$ is a *dyadic product*. This can be thought of as a matrix so that in index notation $\boldsymbol{nn}$ is interpreted as the matrix with elements $n_i n_j$. Note that some authors write this product as $\boldsymbol{n} \otimes \boldsymbol{n}$.

as the *no-slip* condition. In the case of two-dimensional flow, (2.8) simplifies to the requirement that $\boldsymbol{u} \cdot \boldsymbol{t} = \boldsymbol{U} \cdot \boldsymbol{t}$ on Σ, where $\boldsymbol{t}$ is a vector tangent to Σ. The boundary condition (2.8) is less intuitive than the no penetration condition; nevertheless, it is appropriate for most fluid motions and is supported by a wealth of experimental observations (e.g. Lauga et al. [12]).

Taken together, conditions (2.7) and (2.8) demand that the fluid velocity at a solid boundary Σ matches the velocity of Σ itself. If $\boldsymbol{U} = \boldsymbol{0}$, so that the boundary is stationary, these conditions require that $\boldsymbol{u} = \boldsymbol{0}$ on Σ. Solving the flow problem (2.5), (2.6) in an arbitrary geometry with boundary conditions (2.7), (2.8) is in general a formidable task. Much of the difficulty stems from the presence of the nonlinear term, $\boldsymbol{u} \cdot \nabla \boldsymbol{u}$, in the Navier-Stokes equation. Fortunately, substantial simplifications occur in particular geometries or under certain flow conditions. A classical example of considerable interest in the fluid mechanics of plants is the flow in a straight tube to be considered in Sect. 2.3.1.

2.2.4 The Reynolds Number

In the cgs system of units the fluid density, viscosity, and velocity have units of $\mathrm{g\,cm^{-3}}$, $\mathrm{g\,cm^{-1}s^{-1}}$, and $\mathrm{cm\,s^{-1}}$, respectively. In the flow of a fluid of density ρ and dynamic viscosity μ, which has a characteristic speed U and a characteristic length scale L, the dimensionless grouping known as the Reynolds number, named after *Osborne Reynolds* (1842–1912),

$$Re = \frac{\rho U L}{\mu}, \tag{2.9}$$

is free of units and, consequently, provides a very convenient means of categorising the flow. A flow with a particular value of Re is dynamically similar to another flow at the same Reynolds number, and information about the one may be inferred from information about the other via a simple rescaling. For example, the velocity on the axis of a straight tube of radius $1\mu\mathrm{m}$ carrying a purely axial flow at Reynolds number $Re = 0.1$ is exactly twice the axial velocity in a purely axial flow of a fluid of the same density and dynamic viscosity in a straight tube of radius $2\mu\mathrm{m}$ at the same Reynolds number.

Physically speaking, the Reynolds number indicates the relative importance of inertia to viscous forces. If Re is large, as is typical in aerodynamics for example, inertia tends to dominate over viscous effects. If Re is small, then viscous effects are predominant and inertia may be neglected to a leading order approximation. In fact, assuming steady flow, so that $\boldsymbol{u}_t = \boldsymbol{0}$, and formally taking the limit $Re \to 0$, the inertia term, $\boldsymbol{u} \cdot \nabla \boldsymbol{u}$, on the left-hand side of the Navier-Stokes equation (2.5) disappears, and we are left with the linear form

$$\boldsymbol{0} = -\nabla p + \rho \boldsymbol{F} + \mu \nabla^2 \boldsymbol{u}. \tag{2.10}$$

Equation (2.10) is formally valid when $Re = 0$ and the flow it describes is sometimes referred to as *zero Reynolds number flow* or, more commonly, as *Stokes flow*, or occasionally as *creeping flow*.

2.3 Plant Hydraulics: Flow in Tubes

The rise of water through a plant from root to leaf, against gravity, occurs as a result of transpiration (see, for example, Niklas [15, chapter 4]). Evaporation in the photosynthetic leaf tissues lowers the local water potential, defined as the chemical potential per unit volume, and thereby essentially establishes a pressure gradient to draw water up through the xylem. The transportation of food products, namely the outputs of photosynthesis, occurs in the phloem. The Münch hypothesis proposed in 1930 has been the traditionally accepted explanation for the driving transport mechanism here, and support for this hypothesis has recently been bolstered by strong experimental evidence (Knoblauch et al. [9]). The hypothesis asserts that the motion is pressure-driven from, say, leaf to root, with the pressures at the leaf and root being set by the local osmotic movement of water from the xylem into the phloem cells; furthermore, no additional motive force between the two locations is required.

The water-conducting chambers in the xylem consist of tracheary elements comprising elongated cells known as tracheids and vessels (e.g. Evert [4, chapter 10]). Generally speaking, vessels are joined end to end to form extended tubes, and perforation plates at the end-to-end connections permit relatively easy progress of water from one vessel to another. Water may move from one tracheid cell to another at so-called pits, where the secondary cell wall comes away exposing the primary wall. Water can pass through this part of the primary cell wall, which is known as the pit membrane, but does so under heavy resistance (e.g. Niklas [15, chapter 4]). Further geometrical complications which may arise include helical thickenings, which manifest as spiralling ridges along the inside of the cell wall (e.g. Evert [4, chapter 10]).

Despite the complex morphology, water transport has traditionally been modelled by unidirectional fluid motion in a straight tube, and this topic forms the focus of the remainder of this chapter.

2.3.1 Flow in a Straight Tube

To describe the flow, we assume that the tube centreline coincides with the x axis of a Cartesian or other suitable coordinate system, and seek a description in which the fluid velocity $\boldsymbol{u} = u\boldsymbol{e}_x$, where $\boldsymbol{e}_x$ is the unit vector in the x direction, so that the flow is unidirectional and purely parallel to the tube centreline. Keeping in mind the application to flow in a vertically-oriented xylem tube, we assume that the x axis

is vertical and aligned with the direction of gravitational acceleration. Accordingly we assume that the body force in the momentum equation (2.5) takes the form $\boldsymbol{F} = -g\boldsymbol{e}_x$, where $g = 9.8\,\mathrm{m\,s^{-2}}$. In what follows it is convenient to effectively eliminate the body force term from the momentum equation by defining the modified pressure $\tilde{p} = p + \rho g x$.

The Navier-Stokes equation (2.5) is considerably simplified in the case of unidirectional flow. The conservation of mass equation (2.6) implies that $u_x = 0$ so that u is independent of the axial coordinate, x. It follows that $\boldsymbol{u} \cdot \nabla \boldsymbol{u} = \boldsymbol{0}$ and so the problematic nonlinear term on the left-hand side of Eq. (2.5) vanishes identically. Assuming steady flow, Eq. (2.5) reduces to the linear form,

$$\boldsymbol{0} = -\nabla \tilde{p} + \mu \nabla^2 \boldsymbol{u}, \tag{2.11}$$

which we observe to coincide with Eq. (2.10), which governs steady Stokes flow. Since $\boldsymbol{u} = u\boldsymbol{e}_x$, taking the dot product of (2.11) with any vector $\boldsymbol{v}$ which is orthogonal to $\boldsymbol{e}_x$, we deduce that $\boldsymbol{v} \cdot \nabla \tilde{p} = 0$ meaning that the pressure gradient in the direction of $\boldsymbol{v}$ is zero. Consequently the tube pressure is a function of x and t only so that $\tilde{p} = \tilde{p}(x, t) = p(x, t) + \rho g x$.

We assume that the flow is driven by a constant axial pressure gradient $-G$, where $G > 0$. For vertical xylem transport the pressure gradient must be sufficiently large to overcome the force of gravity. For a tube of length L with entrance and exit pressures p_1 and p_2, respectively (see Fig. 2.3a), the pressure gradient along the tube is

$$-G = \frac{\tilde{p}_2 - \tilde{p}_1}{L} = \frac{p_2 - p_1}{L} + \rho g. \tag{2.12}$$

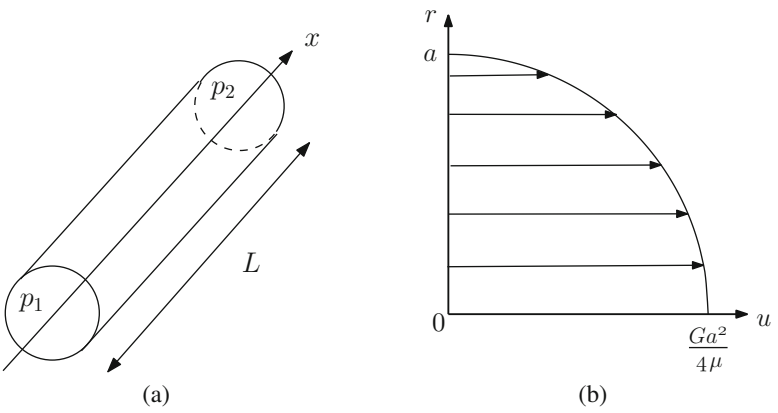


Fig. 2.3 (**a**) A pressure gradient drives flow through a tube in the positive x direction if the exit pressure is less than the entrance pressure, $p_2 < p_1$. (**b**) The Poiseuille velocity profile (2.15) in a straight circular tube. The arrows indicate the direction of flow

To ensure that $-G < 0$ and so that a flow is driven through the tube in the positive x direction, we need $p_2 < p_1 - \rho g L$. For a tree of height, say, $L = 10\,\mathrm{m}$, taking $\rho = 10^3\,\mathrm{kg\,m^{-3}}$ as the density of water, assuming normal atmospheric pressure at the root, so that $p_1 = 1\mathrm{atm}$, we find that an exit pressure of $p_2 < 0.02$ atm is required for water to be drawn upwards through the tube, that is $\frac{1}{50}$th normal atmospheric pressure. Kramer and Boyer [10] demonstrate that the water potentials measured in tall trees are sufficient for the water to rise to substantial heights.

To compute the flow itself it remains to solve the x component of (2.11), namely

$$0 = G + \mu \nabla^2 u, \tag{2.13}$$

subject to the no-slip condition that $u = 0$ at the tube wall (see Sect. 2.2.3). Since the flow is everywhere parallel to the wall, the no normal flow condition (2.7) is automatically satisfied. We start with a discussion of flow in a tube of circular cross-section of radius a.

2.3.1.1 Circular Tube Flow

Cylindrical polar coordinates (x, r, θ) are the most natural coordinate system to use to describe flow in a circular tube. We take the axis of the tube to be located at $r = 0$. Assuming axisymmetry, so that the velocity field does not depend on θ, Eq. (2.13) takes the form

$$0 = G + \mu \left(u_{rr} + \frac{u_r}{r} \right). \tag{2.14}$$

Noting that the term in brackets can be expressed more succinctly in the form $(r u_r)_r / r$, it is straightforward to integrate twice to obtain the velocity field in terms of two arbitrary constants. These constants are determined by enforcing the no-slip condition at the wall, $u(r = a) = 0$, as well as a regularity condition $u_r(r = 0) = 0$ at the tube axis.[2] Ultimately, we obtain

$$u(r) = \frac{G}{4\mu}(a^2 - r^2). \tag{2.15}$$

The motion described by this velocity field is usually referred to as *Poiseuille flow*, or sometimes *Hagen-Poiseuille* flow, after *Gotthilf Hagen* (1797–1884) and *Jean Léonard Marie Poiseuille* (1797–1869). The axial velocity profile (2.15) is sketched in Fig. 2.3b.

The volumetric flow rate Q through the tube is obtained by integrating (2.15) over the tube's cross section, yielding

[2] This condition is required to avoid a coordinate singularity at $r = 0$ which would otherwise result in an unphysical, infinite velocity at the tube axis.

$$Q = \int_0^{2\pi} \int_0^a u(r)\, r \, \mathrm{d}r \, \mathrm{d}\theta = \sigma G, \qquad \sigma = \frac{a^4 \pi}{8\mu}. \tag{2.16}$$

Here σ is the *hydraulic conductivity* of the tube, which provides a measure of the difficulty of driving fluid through the tube against viscous resistance. It is notable that σ scales with the fourth power of the tube radius, so that increasing the radius of a tube rapidly increases its conductivity. The maximum speed occurs on the pipe centreline and is given by $u_{\max} \equiv u(r = 0) = Ga^2/4\,\mu$. The average speed over a cross-section is

$$\overline{u} = \frac{1}{\pi a^2} \int_0^{2\pi} \int_0^a u(r)\, r \, \mathrm{d}r \, \mathrm{d}\theta = \frac{Q}{\pi a^2} = \frac{Ga^2}{8\mu}. \tag{2.17}$$

Evidently $\overline{u} = u_{\max}/2$ so that the average flow speed is equal to one half the maximum flow speed.

In the present case it makes sense to base the Reynolds number on the mean flow speed and the radius of the tube, and to set $U = \overline{u}$ and $L = a$ in (2.9) so that

$$Re = \frac{\rho a \overline{u}}{\mu}. \tag{2.18}$$

In the xylem the tube diameter is around 50 μm. A typical measured flow speed is around $0.1\,\mathrm{cm\,s^{-1}}$ [3]. Assuming that the fluid in the xylem is essentially water with dynamic viscosity $\mu = 0.01\,\mathrm{g\,cm^{-1}\,s^{-1}} = 1\mathrm{cP}$ and density $\rho = 1\,\mathrm{g\,cm^{-3}}$, we obtain $Re = 0.1$, suggesting that we are in the regime of low Reynolds number Stokes flow discussed in Sect. 2.2.4.

2.3.1.2 Non-Circular Tube Flow

As has been noted by Tyree and Ewers [21], the tracheary elements are rarely circular in cross-section and, in fact, a rectangular or elliptical cross-sectional shape is more appropriate. For an elliptical tube whose cross-sectional boundary C is described by the equation $y^2/a^2 + z^2/b^2 = 1$, where a is the semi-major axis and b is the semi-minor axis of the ellipse, we obtain the solution to (2.13),

$$u(y, z) = \frac{Ga^2b^2}{2\mu(a^2 + b^2)} \left(1 - \frac{y^2}{a^2} - \frac{z^2}{b^2}\right), \tag{2.19}$$

which satisfies the requirement that $u = 0$ on C. The maximum flow speed $u_{\max} = Ga^2b^2/2\mu(a^2 + b^2)$ occurs on the tube axis, and the hydraulic conductivity is

$$\sigma = \frac{Q}{G} = \frac{a^3b^3\pi}{4\mu(a^2 + b^2)}. \tag{2.20}$$

Notice that these results reduce to those for circular tube flow when $b = a$. If the cross-section is only a slight elliptical deviation from a circle, we may set $b = a(1+\epsilon)$, where $\epsilon \ll 1$, and expand the hydraulic conductivity formula as a Taylor series to obtain[3]

$$\sigma = \frac{a^4\pi}{8\mu}\left(1 + 2\epsilon + O(\epsilon^2)\right). \tag{2.21}$$

Note that the term preceding the bracket coincides with the hydraulic conductivity for a circular tube given in (2.16).

For a tube of rectangular cross-section, the solution for the axial velocity may be found in the form of an infinite series (e.g. Batchelor [2]). The hydraulic conductivity for such a tube with sides of length $2a$ and $2b$ is calculated to be

$$\sigma = \frac{Q}{G} = \frac{4ab^3}{3\mu}\left[1 - \frac{6b}{a}\sum_{k=1}^{\infty}\frac{1}{\beta_k}\tanh\left(\frac{a\beta_k}{b}\right)\right], \tag{2.22}$$

where $\beta_n = (k - 1/2)\pi$.

Analytic expressions for Poiseuille flow in a tube may be obtained for a few other cross-sectional shapes, including an equilateral triangle (see, for example, Batchelor [2]), or a small deviation from a regular shape such as a rectangle or a trapezium (Navardi et al. [14]). For an arbitrary cross-section the solution can be computed numerically via a complex variable formulation (e.g. Langlois & Deville [11]), or using the boundary element method (e.g. Pozrikidis [17]).

2.3.1.3 Tube Collapse and Cavitation

Since the water pressure decreases in the direction of the flow, a potentially large difference can develop between the interior and exterior pressures, creating a substantial transmural pressure across the tube wall. If this transmural pressure exceeds a critical value, the tube will tend to buckle and collapse[4]; and, at the cellular level, the tracheids and vessels need to be strong enough to resist this. Furthermore, if the pressure is lowered significantly below the water vapour pressure, then the formation of cavitation bubbles is expected (e.g. Batchelor [2]). Such bubbles, which in plants are known as embolisms, have the potential to cause a blockage and obstruct the flow of water. The architecture of the tracheid cells helps to mitigate against this since it is difficult for vapour or air bubbles to squeeze through the small pores in the pit membrane (see Evert [4, chapter 10]).

[3] The $O(\epsilon^2)$ term in (2.21) may be interpreted to mean "terms of typical size ϵ^2 and smaller".

[4] This has been demonstrated for elastic tubes by Flaherty et al. [5], who also calculated the buckled tube states.

2.3.2 Flow in a Tube with Varying Diameter

The diameter of the lumen aperture in mature tracheary vessels is not expected to be uniform throughout a given plant. Vessel diameter may tend to decrease with stem height, for example, and vessel cross-section may change due to cell wall thickening. In this section we account for these effects in a simplified manner by considering flow in a tube whose radius is a slowly-varying function of axial distance. Since we expect the Reynolds number to be small in the xylem (see Sect. 2.3.1.1), it is reasonable to neglect the inertia terms in the Navier-Stokes equation and work instead with the linear equation of Stokes flow (2.10).

As before, we assume that the flow is driven by a constant axial pressure gradient $-G$, with $G > 0$. Suppose that L and a^* are typical length scales in the axial and radial directions, respectively. Our analysis will be predicated on the assumption that L is much larger than a^* so that the tube slenderness parameter $\delta \equiv a^*/L$ is small. Given the separation of scales, it is appropriate to introduce the new variables

$$X \equiv x/L, \qquad R \equiv r/a^*, \qquad \tilde{\boldsymbol{u}} \equiv \left(a^{*2}G/\mu\right)^{-1} \boldsymbol{u}, \qquad \tilde{p} \equiv (LG)^{-1}\delta^2 p.$$

Writing the Stokes equation (2.10) and the conservation of mass equation (2.6) in component form, we have

$$\begin{aligned} 0 &= -\tilde{p}_X + \left(\tilde{u}_{RR} + \tilde{u}_R/R + \delta^2 \tilde{u}_{XX}\right), \\ 0 &= -\delta^{-2}\tilde{p}_R + \left(\tilde{v}_{RR} + \tilde{v}_R/R - \tilde{v}/R^2 + \delta^2 \tilde{v}_{XX}\right), \\ 0 &= \delta \tilde{u}_X + \tilde{v}_R + \tilde{v}/R, \end{aligned} \tag{2.23}$$

where we have taken $\tilde{\boldsymbol{u}} = \tilde{u}\boldsymbol{e}_x + \tilde{v}\boldsymbol{e}_r$, where $\boldsymbol{e}_r$ is the unit vector in the radial direction.

At this stage we drop the tilde decorations to simplify the notation. Considering the limit $\delta \to 0$, to leading order approximation the second equation in (2.23) yields $p_R = 0$ so that the pressure does not vary over a cross-section, and therefore $p = p(X)$. This simplification means that the first equation in (2.23) can be integrated directly (having dropped the small $\delta^2 u_{XX}$ term). Integrating twice with respect to R, and applying the no-slip condition $u(R = a) = 0$ at the wall, where $a(X)$ is the local tube radius, and the regularising condition[5] $u_R(R = 0) = 0$, we obtain

$$u = -\frac{1}{4}\frac{\mathrm{d}p}{\mathrm{d}X}\left(a^2(X) - R^2\right). \tag{2.24}$$

[5] This is required to avoid a singularity on the pipe axis. Compare with the calculation for Poiseuille flow in a circular tube in Sect. 2.3.1.1.

Integrating (2.24) over the tube cross-section, we compute the volumetric flow rate, $Q = -(\pi a^4/8)\mathrm{d}p/\mathrm{d}X$.

The third equation in (2.23), which represents conservation of mass, suggests that $v = O(\delta)$ so that the flow is almost unidirectional. Integrating this third equation and applying the physically intuitive condition $v(R = 0) = 0$, we find that the no penetration condition at the tube wall, $v(R = a) = 0$, can only be satisfied if the pressure satisfies the ordinary differential equation

$$\frac{\mathrm{d}^2 p}{\mathrm{d}X^2} + \left(\frac{4}{a}\frac{\mathrm{d}a}{\mathrm{d}X}\right)\frac{\mathrm{d}p}{\mathrm{d}X} = 0, \tag{2.25}$$

where we have assumed that $a \neq 0$ so that the tube does not become completely constricted at any point. Integrating this equation once, we may deduce that

$$\frac{\mathrm{d}p}{\mathrm{d}X} = -\frac{8Q}{\pi a^4}. \tag{2.26}$$

Thus the pressure gradient along the tube decreases as the inverse fourth power of the tube radius, $a(X)$.

Integrating (2.26) we obtain the following formula for the tube pressure at any axial position for a given tube shape $a(X)$,

$$p(X) = p_0 - \frac{8Q_0}{\pi} I(X), \qquad I(X) = \int_0^X \frac{1}{a^4(\xi)}\,\mathrm{d}\xi, \tag{2.27}$$

where p_0 is the pressure at the tube entrance at $X = 0$. For a general choice of shape function $a(X)$, the integral in (2.27) can be calculated using a suitable numerical quadrature (e.g. Hildebrand [7]). Figure 2.4 shows how p varies according to (2.27) for a tube whose radius is decreasing according to one of the two forms:

$$\text{(i)}\quad a(X) = 1 - \epsilon X, \qquad \text{(ii)}\quad a(X) = 1 + \epsilon\cos(X), \tag{2.28}$$

for constant ϵ. Option (i) might be considered as a simple model designed to capture the trend of decreasing vessel diameter with stem height, while option (ii) works as a simple model to capture the fact that water flowing from tracheid to tracheid must pass through narrow constrictions at the pits. For both options in (2.28), the integral I in (2.27) can be calculated exactly to obtain a closed formula for the pressure. Assuming $\epsilon \ll 1$, we obtain the following approximations:

$$\text{(i)}\quad I(X) = X + 2\epsilon X^2 + O(\epsilon^2), \qquad \text{(ii)}\quad I(X) = X - 4\epsilon\sin X + O(\epsilon^2). \tag{2.29}$$

The pressures $p(X)$ computed using these approximations are shown in the plots in Fig. 2.4 with broken lines. Evidently the approximations become less accurate with distance along the tube.

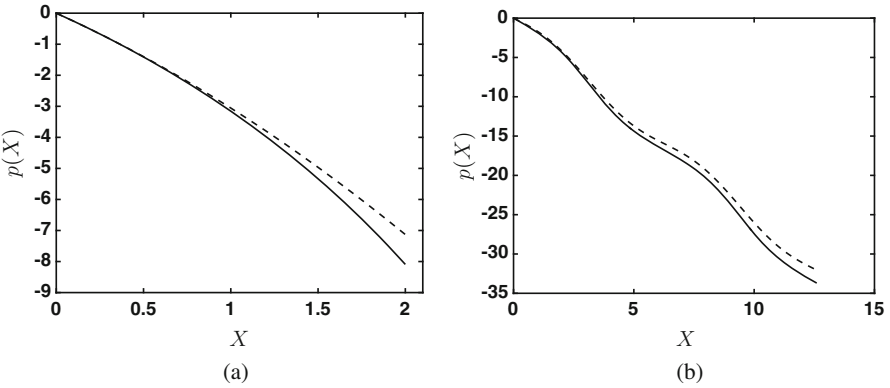


Fig. 2.4 Pressure p versus X according to (2.27) for the tube with radius (*a*) $a(X) = 1-0.1X$, and (*b*) $a(X) = 1 + 0.1\cos X$, corresponding, respectively, to choices (i) and (ii) in (2.28), both with $\epsilon = 0.1$ and with $Q_0 = 1$ and $p_0 = 0$. The broken lines indicate the approximations computed using (2.29)

If the assumption of a slowly-varying tube radius is relaxed, the analysis required to determine the flow velocity and the tube pressure is substantially more involved. Numerical calculations of pressure-driven flow at zero Reynolds number through a channel with sinusoidal walls were carried out by Pozrikidis [16]. These computations revealed that recirculating eddies are present in the wall troughs provided that the amplitude of the wall is sufficiently large; and, indeed, this is expected on the basis of the corner-flow analysis of Moffatt [13]. Roth [18] provided an axisymmetric model of flow in a tracheary vessel with helical wall thickenings, and also observed the presence of near-wall eddies. Such thickenings may result in a highly contorted flow, which provides an efficient environment for mixing. This has knock-on benefits for the transport of solutes. This topic forms the focus of the next section.

2.4 Solute Transport

The transport of solutes is of considerable interest in the fluid mechanics of plants. In this section, the fundamentals of solute transport by convection or diffusion are discussed, and the enhanced mixing effect obtained when these two effects act in unison, usually referred to as *Taylor-Aris* dispersion, is examined.

2.4.1 The Transport Equation

To derive the governing equation for solute transport, we return to the notional volume V which is fixed in time and space within an infinite fluid envisaged

in the derivation of the basic equations in Sects. 2.2.1 and 2.2.2. Consider now the transport of a passive solute, such as a nutrient or a chemical, through this volume due to the dual processes of convection by the carrier fluid and molecular diffusion. If $c(\boldsymbol{x}, t)$ is the solute concentration at a point $\boldsymbol{x}$ at time t, then we may write, assuming that there is no continuous production of the solute from a source within V,

$$\frac{\mathrm{d}}{\mathrm{d}t}\iiint_V c\,\mathrm{d}V = -\iint_S c\boldsymbol{u}\cdot\boldsymbol{n}\,\mathrm{d}S - \iint_S \boldsymbol{q}\cdot\boldsymbol{n}\,\mathrm{d}S, \tag{2.30}$$

where S is the surface of V, $\boldsymbol{u}$ is the velocity of the fluid, $\boldsymbol{n}$ is the outward-pointing normal to S, and $\boldsymbol{q}$ is the diffusive flux. The first term in (2.30) expresses the rate of change of the total amount of solute contained within V. The first term on the right-hand side represents the rate of transport of solute through the surface S by convection. The second term on the right-hand side quantifies the rate of transport of solute across S by diffusion.

According to Fick's law, solute is spread by diffusion down its concentration gradient, so that $\boldsymbol{q} = -D\nabla c$, where D is the diffusivity. Substituting this form for $\boldsymbol{q}$ into (2.30), applying the divergence theorem to both terms on the right-hand side, and moving the time derivative on the first term into the integrand on the basis that V is fixed, and finally moving all terms to the left-hand side, we have

$$\iiint_V \Big(c_t + \nabla\cdot(c\boldsymbol{u}) - \nabla\cdot(D\nabla c)\Big)\,\mathrm{d}V = 0. \tag{2.31}$$

Appealing to the fact that the volume V is arbitrary, we conclude that the integrand in (2.31) must vanish, yielding a partial differential equation governing the dispersal of the solute. If the flow is incompressible, so that the solenoidal velocity field condition (2.6) holds, we may expand and simplify the second term in the integrand of (2.31) using the product rule of differentiation. If we also assume that the diffusivity D is constant, we arrive at the *convection-diffusion equation*,

$$c_t + \boldsymbol{u}\cdot\nabla c = D\nabla^2 c. \tag{2.32}$$

Prior to considering the combined effects of convection and diffusion, it is informative to first discuss the effect of each in isolation.

2.4.2 Transport by Diffusion

In the absence of convection, the transport equation (2.32) reduces to the *diffusion equation*,

$$c_t = D\nabla^2 c. \tag{2.33}$$

To illustrate the basic process of diffusion, it is convenient to work in one spatial dimension and assume that $c = c(x, t)$. In this case (2.33) takes the simpler form $c_t = Dc_{xx}$. It has the *fundamental solution*,

$$c(x, t) = \frac{c_0}{(4\pi D\, t)^{1/2}}\, \mathrm{e}^{-x^2/4Dt}, \qquad -\infty < x < \infty. \tag{2.34}$$

This solution has been constructed so that the total mass of solute, c_0 say, remains the same for all time t,

$$\int_{-\infty}^{\infty} c(x, t)\, \mathrm{d}x = c_0. \tag{2.35}$$

Figure 2.5a shows how the solute profile given by (2.34) spreads out along the x axis as time increases. Formally, at $t = 0$ the solute is all concentrated at the single point $x = 0$, where c takes an infinite value. In any case, it is clear that a concentration of solute will spread outwards over time under the action of diffusion.

The question of how rapidly the solute spreads is of fundamental importance. Suppose that $x = \chi(t)$ marks the boundary such that half of the solute is contained in the region $|x| < \chi$. Using (2.34) and (2.35), the mass of solute in the bounded zone is

$$\int_{-\chi}^{\chi} c(x, t)\, \mathrm{d}x = \frac{c_0}{(4\pi D\, t)^{1/2}} \int_{-\chi}^{\chi} \mathrm{e}^{-x^2/4Dt}\, \mathrm{d}x = \frac{c_0}{2}. \tag{2.36}$$

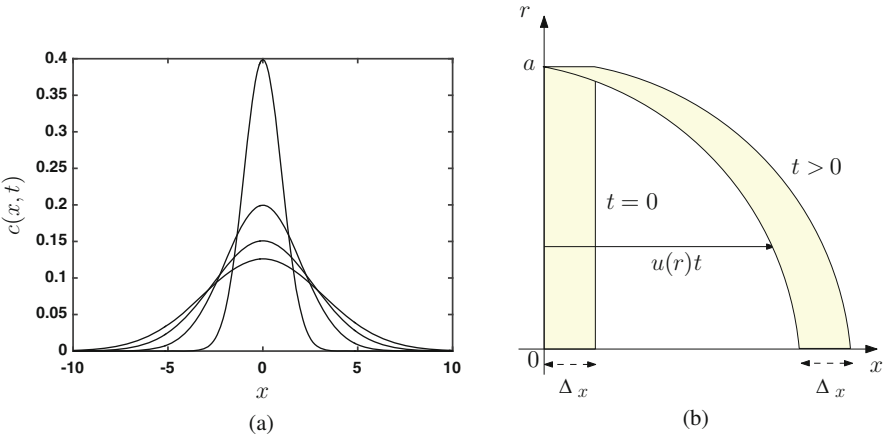


Fig. 2.5 (**a**) Solute dispersal under pure diffusion: profiles $c(x, t)$ versus x according to the fundamental solution (2.34) with $c_0 = D = 1$ at times $t = 0.5\ (0.40),\ 2.0\ (0.20),\ 3.5\ (0.15),\ 5.0\ (0.13)$, with the values of $c(0, t)$ in brackets. The area under each curve is equal to c_0 according to (2.35). (**b**) Deformation of an initially rectangular distribution of solute of axial length Δ_x under pure convection in circular Poiseuille flow in a tube of radius a

Making the change of variable $x = (4Dt)^{1/2}\xi$ in the integral, we have that

$$\int_{-\chi^*}^{\chi^*} \mathrm{e}^{-\xi^2}\,\mathrm{d}\xi = \frac{\pi^{1/2}}{2}, \tag{2.37}$$

where $\chi^* = \chi/(4Dt)^{1/2}$. Although we cannot obtain an analytic formula for the value of χ^* which satisfies (2.37), it is clear that χ^* must be constant, and hence $\chi(t) \propto (D\,t)^{1/2}$. So the boundary in question advances along the x axis as the square root of time, that is to say rather slowly. Indeed, its speed $\mathrm{d}\chi/\mathrm{d}t \propto (D/t)^{1/2}$ decreases with time and so the rate of advance slows as time progresses.

We conclude that diffusion is a *slow* process and, therefore, taken on its own it offers a poor mechanism for transport when a relatively fast delivery is required. Also of note is that diffusion acts to smear out spatial gradients, preferring a more uniformly-distributed concentration profile.

2.4.3 Transport by Convection

In the absence of diffusion, the transport of solute is dictated purely by convection and (2.32) reduces to the form

$$c_t + \boldsymbol{u}\cdot\nabla c = 0. \tag{2.38}$$

Suppose that an individual fluid particle moving in a three-dimensional flow has time-dependent position $\boldsymbol{x}(t) = x(t)\boldsymbol{e}_x + y(t)\boldsymbol{e}_y + z(t)\boldsymbol{e}_z$. Since the particle is moving at the local fluid velocity, we must have

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{u}(\boldsymbol{x}, t), \tag{2.39}$$

where $\boldsymbol{u}(\boldsymbol{x}, t) = u\boldsymbol{e}_x + v\boldsymbol{e}_y + w\boldsymbol{e}_z$ is the fluid velocity. Consider now the solute concentration associated with *this particular fluid particle*, $c(\boldsymbol{x}(t), t)$. Applying the chain rule, the rate of change of the concentration following this particle is given by

$$\frac{\mathrm{d}c}{\mathrm{d}t} = c_t + uc_x + vc_y + wc_z = c_t + \boldsymbol{u}\cdot\nabla c. \tag{2.40}$$

It follows that the diffusion-free transport equation (2.38) can be rewritten simply as $\mathrm{d}c/\mathrm{d}t = 0$, meaning that the concentration *associated with an individual fluid particle* does not change as the particle moves with the flow.

By way of example, consider the transport of solute by pure convection in the pressure-driven Poiseuille flow in a circular tube described in Sect. 2.3.1.1. The axial velocity is given by formula (2.15). Following the preceding observation that the solute is simply carried with the flow, it is necessary only to describe the motion of

the fluid particles. We therefore aim to solve (2.39) for the time-dependent position $\boldsymbol{x}(t)$ of a general fluid particle. Splitting the equation into components, we have

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \lambda(a^2 - r^2), \qquad \frac{\mathrm{d}r}{\mathrm{d}t} = 0, \tag{2.41}$$

where $\lambda = G/4\mu$, and where x measures axial distance along the tube and $r = (y^2 + z^2)^{1/2}$. The second equation in (2.41) implies that the radial distance of a particle from the tube axis does not change during the motion. Integrating the two equations in (2.41), we obtain

$$x(t) = \lambda(a^2 - r_0^2)\, t + x_0, \qquad r(t) = r_0, \tag{2.42}$$

assuming that the particle started at $x = x_0$ and $r = r_0$ at $t = 0$. Notice that particles on the tube axis, $r = 0$, move the fastest and particles on the wall $r = a$ do not move at all, in accordance with the no-slip condition (see Sect. 2.2.3).

Labelling individual fluid particles by their initial position $(x_0,\ r_0)$, we can infer knowledge of the solute concentration at any time $t > 0$ from knowledge of the concentration profile at $t = 0$ since we know the trajectories of the fluid particles. For example, a distribution of solute which is initially shaped like a flat circular disk will deform into a parachute-type shape, as is illustrated in Fig. 2.5b. This distortion in the concentration profile in convective transport underscores the importance of *shear*, which results from a spatially-dependent velocity field. If the convection velocity in this example were constant, the disk-shaped region at $t = 0$ would simply move downstream like a rigid body, maintaining its original shape, size and orientation.

2.4.4 Shear-Enhanced Diffusion

When both convection and diffusion are in play, the situation is considerably more complicated. While analytical progress is difficult, there exist numerous computational algorithms for solving the convection-diffusion equation (see, for example, Hoffman [8]). Broadly speaking, convection carries solute along with fluid particles while diffusion acts to smear it out across fluid particles. Shear tends to distort the solute profile, creating potentially strong spatial concentration gradients, as we saw in the previous subsection. But diffusion dislikes spatial gradients, and it tries to smooth them out. This competition between convective distortion and diffusive smoothing can produce an efficient mechanism for solute dispersal.

In a celebrated piece of analysis, Taylor [20], and later Aris [1] demonstrated how streamwise diffusion in a circular tube flow, driven by a constant pressure gradient, is substantially enhanced by the shear in the basic velocity profile. The analysis is predicated on the following two assumptions following the release of a solute into the flow: (i) sufficient time has passed for the solute to have diffused

radially outwards a distance equal to the tube radius, so that any deviation in concentration in a cross-sectional plane from its average value across this plane is small; and, moreover, this time scale is short compared to that over which appreciable changes due to convective transport are observed; and (ii) the radial diffusion of this deviation dominates over its axial diffusion.

In following this analysis, our goal is to derive an approximate evolution equation for the cross-sectional average of the concentration field. To achieve this aim, it will be convenient to decompose the concentration field by writing $c = \overline{c}(x,t) + C(x,r,t)$, where $\overline{c}$ is the cross-sectional average,

$$\overline{c}(x,t) = \frac{1}{\pi a^2}\int_0^{2\pi}\int_0^a c(x,r,t)\, r\, \mathrm{d}r\, \mathrm{d}\theta, \tag{2.43}$$

and C is the deviation from this average; note that the deviation has zero mean, so that $\overline{C} = 0$. From the discussion in Sect. 2.4.2 we can infer that assumption (i) will broadly hold true when t is on the order of the radial diffusion timescale t_D, where $t_D = a^2/D$, and where a is the tube radius and D is the diffusivity. Accordingly we introduce the new variables

$$X \equiv x/L, \qquad R \equiv r/a, \qquad \tau \equiv t/t_D, \qquad U(R) \equiv u(r,t)/\overline{u},$$

where L is a representative axial length scale. We rewrite the circular Poiseuille velocity profile (2.15) in the form $u = \overline{u}\, U(R)$, where $U = 2(1-R^2)$, and where the cross-sectional average speed $\overline{u} = Ga^2/8\mu$ was given in (2.17). The convection-diffusion equation (2.32) then takes the form

$$\overline{c}_\tau + C_\tau + (\delta\, Pe)\, U(\overline{c}_X + C_X) = \delta^2 \overline{c}_{XX} + \left(\delta^2 C_{XX} + C_{RR} + \frac{C_R}{R}\right), \tag{2.44}$$

where $\delta = a/L$ and the Péclet number is defined as $Pe = \overline{u}a/D = Ga^3/(8\mu D)$.

We assume that the solute cannot escape through the tube wall. Appealing to Fick's law (see Sect. 2.4.1), this requires that

$$c_R(R=1) = 0. \tag{2.45}$$

Taking the cross-sectional average of (2.44), and enforcing the condition (2.45), we obtain

$$\overline{c}_\tau + (\delta\, Pe)\left(\overline{U}\,\overline{c}_X + \overline{UC_X}\right) = \delta^2 \overline{c}_{XX}. \tag{2.46}$$

Subtracting (2.46) from (2.44), we arrive at the evolution equation for the deviation,

$$C_\tau + (\delta\, Pe)\left(\left(U - \overline{U}\right)\overline{c}_X + UC_X - \overline{UC_X}\right) = \delta^2 C_{XX} + C_{RR} + \frac{C_R}{R}. \tag{2.47}$$

Up to this point our analysis is exact. To proceed we now take the limit $\delta \to 0$. In this limit it is clear from the right-hand side of (2.47) that the axial diffusion of the deviation concentration is much smaller than its radial diffusion; this corresponds to assumption (ii). Invoking assumption (i), namely that $\overline{c} \gg C$, and retaining the leading order terms on the left and right-hand sides of (2.47), we arrive at the approximate equation,

$$\delta\, Pe \left(U - \overline{U}\right) \overline{c}_X = C_{RR} + \frac{C_R}{R}, \tag{2.48}$$

where we have assumed that $\delta\, Pe \ll 1$. This latter condition ensures that the time scale for radial diffusion is much shorter than that for convective transport, in accordance with assumption (i). Substituting the form for U given above, (2.48) can be integrated directly. Demanding that C be finite on the tube axis, and enforcing the condition that $\overline{C} = 0$, we find that

$$C = \frac{1}{4}\delta\, Pe\, \overline{c}_X \left(R^2 - \frac{R^4}{2} - \frac{1}{3} \right). \tag{2.49}$$

It is now straightforward to compute the term $\overline{UC_X}$. Substituting the result into (2.44), and rearranging terms, we arrive at the desired evolution equation for the cross-sectional mean concentration. To aid physical interpretation, it is convenient to restore the original variables, whereupon the equation takes the form

$$\overline{c}_t + \overline{u}\, \overline{c}_x = D_e\, \overline{c}_{xx}, \tag{2.50}$$

where D_e is an *effective diffusivity* given by

$$D_e = D + \frac{\kappa}{D}, \qquad \kappa = \frac{a^2 \overline{u}^2}{48}. \tag{2.51}$$

Equation (2.50) can be simplified by making the Galilean transformation $x \mapsto x + \overline{u}\, t$ to remove the second term on the left-hand side, leaving the pure diffusion equation, $\overline{c}_t = D_e\, \overline{c}_{xx}$. Physically, this means that in a frame of reference which is travelling in the x direction at speed $\overline{u}$, the averaged concentration field $\overline{c}$ spreads out purely under the action of axial diffusion with effective diffusivity D_e. Referring to Sect. 2.4.2, we can then say that after a time t the solute will have spread axially in this moving frame of reference over a distance on the order of $(D_e\, t)^{1/2}$; this is illustrated in Fig. 2.6a.

The tracheary vessels in the xylem, and the conduits in the phloem, tend to be long and thin, and consequently the assumption that δ is small is reasonable. Mastro et al. (1984) have estimated that $D \approx 10^{-6}\,\mathrm{cm^2\,s^{-1}}$ for small molecules in cellular environments. Using this value for the diffusivity, and taking $a \approx 25\,\mu\mathrm{m}$ as a reasonable value for the tube radius, we find that $t_D = 6.25\,\mathrm{s}$. Assuming that the mean convection speed $\overline{u} \approx 0.1\,\mathrm{cm\,s^{-1}}$ [3], after time t_D the solute has been

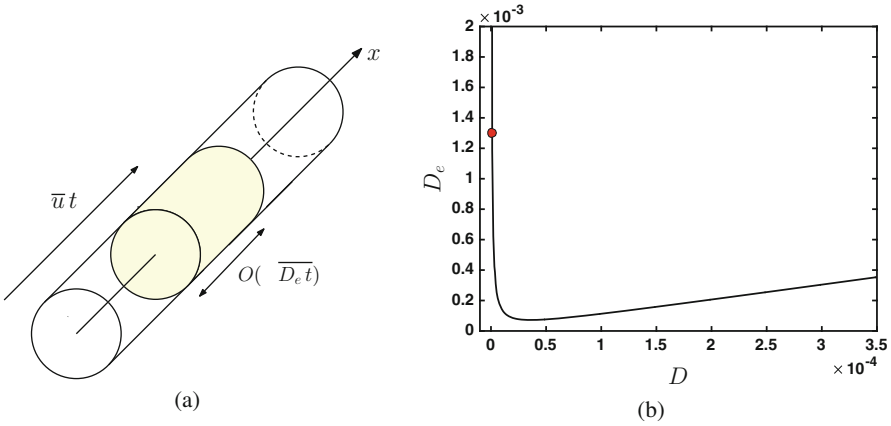


Fig. 2.6 (**a**) Solute dispersal in a circular tube, radius a, after time $t \approx t_D = a^2/D$. Convection has carried a solute bolus (shaded) a distance $\overline{u}\,t$ along the tube, radial diffusion has smeared it out across the tube, and enhanced axial diffusion has spread it out over a distance $O(\sqrt{D_e\,t})$ along the tube axis. (**b**) Variation of the effective diffusivity with D according to formula (2.51). The filled circle shows the data point for $D = 10^{-6}\text{cm}^2\,\text{s}^{-1}$

convected through a distance of roughly $\overline{u}\,t_D = 0.63\,\text{cm}$. This distance is likely to be substantially shorter than the total distance to be travelled along a plant stem. Computing the Péclet number using these values, we obtain the fairly large value $Pe = 250$. Nevertheless if $L = 10\,\text{cm}$, say, then $\delta\,Pe = 0.063$. We therefore expect both of the conditions (i) and (ii) to be fulfilled in a plant flow.

Figure 2.6b shows the dependence of D_e on D according to formula (2.51). The curve has a minimum at $D = \sqrt{\kappa}$, and D_e increases sharply as D falls below value. Using the preceding physical estimates, we compute $\sqrt{\kappa} \approx 3.6 \times 10^{-5}\,\text{cm}^2\,\text{s}^{-1}$. The filled circle in the figure corresponds to the value $D = 10^{-6}\,\text{cm}^2\,\text{s}^{-1}$ suggested by Mastro et al. (1984). According to the graph, the effective diffusivity at this value is given by $D_e = 1.3 \times 10^{-3}\,\text{cm}^2\,\text{s}^{-1}$, which is about a thousand times larger than D! This strongly suggests that shear-enhanced dispersion will be an active mechanism in the transport of solutes in plant hydraulics.

2.5 Conclusions

Plant water relationships are critical for plant growth and survival. Plant hydraulics is of fundamental importance both to plant form and to plant function, and is the subject of ongoing and active research. Such research demands a solid understanding both of the plant biology and of the underlying physics. In this chapter we have discussed some of the elementary principles underlying fluid motion in plants. We have introduced the basic mathematical equations and discussed some simple

model solutions, and presented an elementary description of solute transport. Much remains to be done to understand the intricacies of fluid motion in the complex geometry of a real plant. It is hoped that this chapter has provided a suitable introduction for scientists looking to delve into this fascinating interdisciplinary area of research.

References

1. Aris R (1956) On the dispersion of a solute in a fluid flowing through a tube. Proc Roy Soc Lond A 235:67–77
2. Batchelor GK (1967) An introduction to fluid dynamics. Cambridge University Press, Cambridge
3. Choi WG, Hilleary R, Swanson SJ, Kim SH, Gilroy S (2016) Rapid, long-distance electrical and calcium signaling in plants. Annu Rev Plant Biol 67(1):043015–112130
4. Evert RF (2006) Esau's plant anatomy. Wiley, Hoboken
5. Flaherty JE, Keller JB, Rubinow SI (1972) Post buckling behavior of elastic tubes and rings with opposite sides in contact. SIAM J Appl Math 23:446–455
6. Goldstein H, Poole CP, Safko JL (2014) Classical mechanics. Pearson Higher Ed, Upper Saddle River
7. Hildebrand FB (1987) Introduction to numerical analysis. Courier Corporation, North Chelmsford
8. Hoffman JD (1992) Numerical methods for engineers and scientists. McGraw Hill, New York
9. Knoblauch M, Knoblauch J, Mullendore DL, Savage JA, Babst BA, Beecher SD, Dodgen AC, Jensen KH, Holbrook NM (2016) Testing the münch hypothesis of long distance phloem transport in plants. Elife 5:e15341
10. Kramer PJ, Boyer JS (1995) Water relations of plants and soils. Academic Press, Cambridge (1995)
11. Langlois WE, Deville MO (2014) Slow viscous flow. Springer, Berlin (2014)
12. Lauga E, Brenner M, Stone H (2007) Microfluidics: the no-slip boundary condition. In: Springer handbook of experimental fluid mechanics. Springer, Berlin, pp 1219–1240
13. Moffatt HK (1964) Viscous and resistive eddies near a sharp corner. J Fluid Mech **18**:1–18
14. Navardi S, Bhattacharya S, Azese M (2016) Analytical expression for velocity profiles and flow resistance in channels with a general class of noncircular cross sections. J Eng Math 99:103–118
15. Niklas KJ (1992) Plant biomechanics. University of Chicago Press, Chicago
16. Pozrikidis C (1987) Creeping flow in two-dimensional channels. J Fluid Mech 180:495–514
17. Pozrikidis C (2002) A practical guide to boundary element methods with the software library BEMLIB. Chapman & Hall, London
18. Roth A (1996) Water transport in xylem conduits with ring thickenings. Plant, Cell Environ. 19(5):622–629
19. Schey HM (2005) Div, grad, curl, and all that. WW Norton, New York (2005)
20. Taylor GI (1953) Dispersion of soluble matter in solvent flowing slowly through a tube. Proc Roy Soc Lond A 219:186–203
21. Tyree MT, Ewers FW (1991) The hydraulic architecture of trees and other woody plants. New Phytol 119:345–360

Chapter 3
Modelling Ion Channels

K. C. A. Wedgwood, J. Tabak, and K. Tsaneva-Atanasova

Abstract Plant adaption and survival relies on signalling, much of which is achieved through concentration changes in ions. Furthermore, plants can influence their growth and shape via changes in hydraulic pressure which in turn can be modulated by changes in ionic concentrations that drive osmosis. We present an introduction to mathematical modelling of ionic currents and transmembrane voltages, both intracellular and intercellular. We introduce the modelling techniques used to describe the physical processes involved in ion channel dynamics and illustrate their application using generic examples. We begin by discussing modelling of individual ion channels. Next, we present computational algorithms most commonly employed in simulating ionic currents passing via a single as well as an ensemble of the same ion channel type. We then discuss modelling of ionic current flow across cellular membrane that could involve different ion channel species. We end with an overview of modelling action potentials and their propagation resulting from interactions between different ion channels within as well as between cells. We illustrate this using a simplified example of plant action potential.

3.1 Introduction

One of the most interesting aspects of ion channel dynamics is that local opening and closing events can be organised into complex spatio-temporal patterns such as action potentials that can propagate from cell to cell. Not only are these phenomena physiologically important, they are also mathematically interesting, and challenging to grasp. Like in all living systems, ion channels in plants are involved in signal transduction and are vital for normal plant function, being associated with rapid

K. C. A. Wedgwood · K. Tsaneva-Atanasova
Department of Mathematics and Living Systems Institute, University of Exeter, Exeter, UK
e-mail: K.C.A.Wedgwood@exeter.ac.uk; k.tsaneva-atanasova@exeter.ac.uk

J. Tabak (✉)
University of Exeter Medical School, University of Exeter, Exeter, UK
e-mail: j.tabak@exeter.ac.uk

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_3

responses to environmental stimuli [9, 26]. Reviews of the experimental literature on ion channels in plants are given by Tester [24], Gradmann [11], and Hedrich [14]. The mathematics behind ion channel modelling has a lot in common with the study of action potential propagation in excitable models such as the Hodgkin-Huxley or FitzHugh-Nagumo models [7, 18, 19]. However, useful insights into the mechanisms of plants signal propagation would require the use of more than simple generic excitable models. Therefore, the biophysical framework described in this chapter could be used a starting point for construction of models that are tailor-made to the particular type of cell or electrical activity under consideration.

3.2 Modelling Single Ion Channel Dynamics

Single ion channels are proteins that make a (intra)cellular membrane permeable to specific ions by allowing them to pass through a transmembrane pore. By changing conformation, or other mechanisms, the pore is able to `'open'` or `'close'` in each case permitting, or not, the passage of ions [17]. Under voltage-clamp conditions, when permeable ions are present in solution surrounding the channel, steps in current amplitude can be seen as the channel opens or closes [17]. In this section, we introduce the modelling steps involved in mathematically describing single ion channel currents.

3.2.1 Diffusion of Ionic Species

The movement of charged ionic particles down a concentration gradient creates an electric current. The cellular membrane provides a barrier to the movement of charged ionic particles and thus induces a charge separation and establishes a transmembrane potential difference. Along the cellular membrane, there exist ion channels that facilitate the diffusion of ionic species across the membrane. In addition to these channels are pumps that actively move ionic species against the concentration gradient to maintain a transmembrane concentration difference.

We can model the average movement across the cellular membranes through passive diffusion. The general form describing the passive diffusion of a particular ionic species, Y, obeys the Fickian flux rules:

$$\frac{\partial Y}{\partial t} = -D\frac{\partial Y}{\partial x}, \tag{3.1}$$

where $x \in \mathbb{R}$ represents the distance across the membrane and $D > 0$ is the diffusion coefficient, which describes how easily the particles move across the membrane. If we assume that the cellular membranes are thin compared to the cell and its surroundings, we can replace (3.1) with

$$\frac{\mathrm{d}Y_\mathrm{i}}{\mathrm{d}t} = -D[Y_\mathrm{i} - Y_\mathrm{o}], \tag{3.2}$$

where Y_i and Y_o, respectively, represent the ionic concentrations on the inside and outside of the membrane. The diffusion coefficient is dependent on the permeability of the membrane to that particular ionic species, the surface area of the cell, the density of channels, and the thickness of the cellular membranes. The membrane permeability is then controlled in a dynamic fashion through the opening and closing of ion channels, which we refer to as the gating. It is the aim of this chapter to mathematically describe this gating and its impact on the transmembrane potential difference.

3.2.2 Channel Dynamics

Ion channels are transmembrane proteins that connect the intracellular as well as intercellular compartments. These proteins can exist in conformations which either allow or prevent the diffusion of particles across cellular membranes. Conformational changes that switch between these states depend on a variety of factors, including the transmembrane potential difference, hereby referred to as the voltage, the binding to specific ligands, temperature or mechanical force. We will first describe the scenario in which transitions between states are purely random, that is, they depend on no external factors. We note that the modelling paradigm we describe is applicable to all types of channels, but that care must be taken to ensure that the correct dependencies on extrinsic factors are included when modelling specific ion channels.

Consider a generic ion channel that can exist in one of two states: an open state (O) in which ions are free to pass through; and a closed state (C) in which they are not. Transitions between these states are random in nature, though the probabilities of moving from open to closed states and vice versa do not have to be the same. Stable conformations of proteins representing the closed and open states are achieved by minimising the Gibb's free energy. The open and closed states thus correspond to minima of the free energy landscape. Conformational changes are then represented as transitions between these minima induced by noise caused by thermal fluctuations.

Given the considerations above, our two state channel can be represented by the stochastic differential equation:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -U_x(x) + \xi(t), \quad x \in \mathbb{R}, \tag{3.3}$$

where x is a coordinate that represents the conformational state of the channel, ξ is a Gaussian noise process with mean $\langle \xi(t) \rangle = 0$ and autocorrelation $\langle \xi(t)\xi(t') \rangle = 2k_\mathrm{B}T\,\delta(t - t')$, where k_B is the Boltzmann constant and T is the temperature in

Kelvin. The function $U(x)$ represents the energy landscape, and for exposition, we will here choose this to be a quartic of the form

$$U(x) = \frac{1}{4}x^4 - \frac{\mu}{2}x^2 - \eta x. \tag{3.4}$$

This energy landscape has two minima, located at $x_\pm$, separated by a maximum at x_0. Associated with the two minima are 'wells', which are separated by the 'barrier' at x_0. When $x(t)$ is in the well corresponding to x_-, we say that the channel is in the closed conformation, and when it is in the well associated with x_+, it is in the open conformation. The parameters $\mu > 0$ and $\eta \in \mathbb{R}$, respectively, control the height of the barrier at x_0 and the difference between the depths of the two wells. Thus, increasing μ makes transitions between states harder, whilst varying η alters the relative probabilities of transition.

If the barrier separating the two minima is sufficiently high compared to the thermal fluctuations, transitions between states will be rare, and we here consider the expected transition times between them. The magnitude of the energy barrier relative to the closed (open) state is equal to $\Delta U_\pm = U(x_0) - U(x_\pm)$. Writing $\sigma = (k_B T)^{-1}$, the expected transition time, $\tau_\pm$, from the closed (open) state is then

$$\tau_\pm = \frac{\mathrm{e}^{\sigma \Delta U_\pm}}{\nu_\pm}. \tag{3.5}$$

We note that the functional form of the transition rate, (3.5) was derived first by Arrhenius [1], who left $\nu_\pm$ as undetermined parameters. The specific forms for $\nu_\pm$ was later derived using a number of different methods; our presentation here is based on the exposition by Eyring [4, 5].

As a representative example, we show in Fig. 3.1 the double-well potential system according to (3.3)–(3.4). The energy landscape is depicted in Fig. 3.1a, with an example trajectory shown in Fig. 3.1b. Superimposed on this trajectory is an indicator showing the state of the channel at time t.

3.2.3 Markov Processes

Whilst the potential well system defined by (3.3) is a good descriptor of channel dynamics, it is impractical for large scale channel modelling, since it requires computation of sample paths. Instead of considering the full system, considerable simplification can be made by noting that trajectories will spend most of the time near the minima of the system, making infrequent excursions over the barrier and transitioning into the well associated with the other minima. This leads us to think of simplifying the system to only consider these transition times, which are formally known as *passage times* or *escape times*.

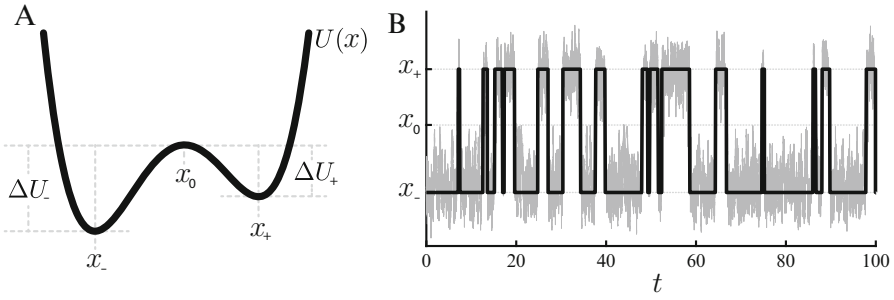


Fig. 3.1 (**a**) Double well potential landscape given by (3.4) with $\mu = 10$ and $\eta = -2$. (**b**) A sample path of (3.3)–(3.4) with $\sigma = 0.05$. Superimposed on this sample path is a bold line indicating which state the channel is in

We can take the reciprocal of the transition times given in (3.5) to provide *rate constants* that describe the probability per unit time of transitioning to the other state. In what follows, we shall denote the probability per unit time of transitioning from the closed state to the open state by $\alpha = 1/\tau_-$, and the probability per unit time of making the opposite transition as $\beta = 1/\tau_+$. Schematically, the system can now be represented as

$$\mathrm{C} \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} \mathrm{O}. \tag{3.6}$$

This system is Markovian in nature, meaning that the evolution of the state is dependent only on the current state of the system and not on previous states. This system can then be easily simulated as follows. Initialise the system with the channel being either in the closed or the open state and pick a time discretisation $\mathrm{d}t \ll 1$. At each time point, draw a random number ξ from the standard uniform distribution. If the channel is closed, transition to the open state if $\xi < \alpha\,\mathrm{d}t$, else remain closed. If the channel is open, transition to the closed state if $\xi < \beta\mathrm{d}t$, else remain open. After completing this step, advance time by $\mathrm{d}t$ and repeat as much as necessary.

3.3 Modelling the Dynamics of an Ensemble of Ion Channels

For a single channel, the method described in the previous section is a simple and convenient way to simulate its gating dynamics. However, cells may have many channels of the same type. Using this method to describe a cell with N channels thus requires drawing N samples per time step. Furthermore, many of these samplings will ultimately result in no transition being made, particularly if α and β are small. An alternative way to simulate the system is offered by the *Gillespie algorithm* [10], which exactly simulates a network of Markov processes under the assumption that all channels act independently of one another and are identical in nature.

3.3.1 Gillespie Algorithm

Instead of assessing at each step whether each channel in the system has transitioned, the Gillespie algorithm estimates a time over which an event has occurred, then computes which event has occurred and updates the system as appropriate. Thus, the Gillespie algorithm only requires drawing two random numbers per time step and only considers times at which events have occurred. In our network of N channels, let us assume that $N_c(t)$ of these are closed and that $N_o(t)$ are open at time t, so that $N_c(t) + N_o(t) = N$.

Each step of the Gillespie algorithm then involves drawing $\xi_{1,2}$. The probability density function describing the time to the next transition, τ, is given by

$$P(\tau = t) = \lambda e^{-\lambda t}, \tag{3.7}$$

where $\lambda = N_c\alpha + N_o\beta$ is the average rate of the next transition. The time to the next transition is then sampled as

$$\tau = \frac{-\ln \xi_1}{\lambda}. \tag{3.8}$$

Between times t and $t + \tau$, we then know that a transition must have occurred. If $\lambda\xi_2 < N_c(t)\alpha$, we say that one of the channels has opened, and we thus set $N_c(t + \tau) = N_c(t) - 1$ and $N_o(t + \tau) = N_o(t) + 1$. Otherwise, if $\lambda\xi_2 > N_c(t)\alpha$, then one of the channels has closed and we thus set $N_c(t + \tau) = N_c(t) + 1$ and $N_o(t + \tau) = N_o(t) - 1$. Time is then updated by setting $t = t + \tau$ and the process is continued. We can see the rationale behind these choices by noting that $N_c(t)\alpha/\lambda$ is the probability that a channel opening has occurred and $N_o(t)\beta/\lambda$ is the probability that a channel closing has occurred, conditioned on the fact that one of these events has happened. Note that during the simulation of the system, the Gillespie algorithm automatically selects the timestep τ over which to evolve the system. Whilst this is an advantage in terms of computational efficiency, since the system is only sampled at event times, one must take care that any extrinsic factors that affect the transition probabilities α and β are approximately constant over the interval $[t, t + \tau)$.

3.3.2 Transition Probabilities

Thus far, we have considered the case where transition probabilities are constant. In general, however, the gating of ion channels may be dependent on a variety of factors, including the transmembrane voltage, see, e.g., [14] in plants; ligands such as intracellular Ca^{2+} and nucleotides, see, e.g., [3, 13, 15] in plants. We can capture this dependence by making the transition probabilities depend on one or more of these factors. For example, for voltage-gated ion channels, we can write $\alpha = \alpha(V)$, $\beta = \beta(V)$, where

$$\alpha(V) = A\mathrm{e}^{k_\alpha(V-V_\alpha)}, \quad \beta(V) = B\mathrm{e}^{k_\beta(V-V_\beta)}, \tag{3.9}$$

is an often used functional form.

A simple way to incorporate Ca^{2+} dependence is via the following schematic

$$\mathrm{C} \underset{\beta}{\overset{\alpha\,[\mathrm{Ca}^{2+}]}{\rightleftharpoons}} \mathrm{O} \tag{3.10}$$

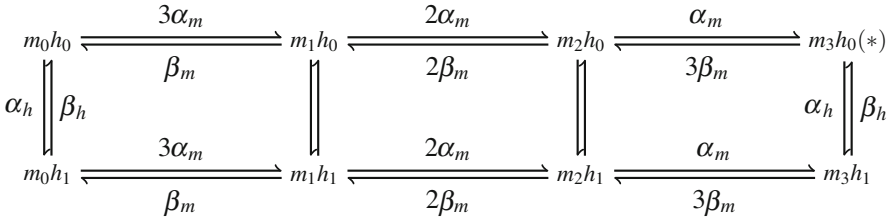
so that the transition rate from the closed to the open state has a linear dependence on the intracellular Ca^{2+} concentration. Modelling the dependence on other factors, such as ligands and nucleotides, follows in a similar fashion, with functional forms chosen to either match the physiology of the channel, or more commonly, fitted to empirical data.

Making transition probabilities depend explicitly on factors such as voltage does not significantly alter the approach to simulating channel dynamics; the transition probabilities α and β are simply altered to reflect this dependence. However, as we shall later see, the movement of ions across cellular membranes impacts upon the transmembrane voltage. In addition, this can induce further downstream effects, for example by opening Ca^{2+} channels so that intracellular Ca^{2+} concentrations vary. As such, channel gating dynamics are tightly coupled to such factors and the approximation in the Gillespie algorithm that these are unaltered in the interval $[t, t+\tau)$ may not hold. In this case, however, a simple fix is to impose a maximum timestep that can be taken by the algorithm.

3.3.3 Many Gates

We have thus far considered channels with only one (activating) gate. However, many ion channels have a multitude of these gates and the channel will only be permeable to ions if all gates are in the open state. In this case, there exist a number of closed states that correspond to the case where at least one of the gates is closed. In addition to the so-called 'activating gates' that we have thus far discussed, ion channels may also possess 'inactivating gates'. Inactivating gates can close ion channels through mechanisms independent of the closing of activating gates and thus a channel will only allow particles to pass through when the activating gates are open and the channel is not inactivated. This gives rise to the four mechanisms that govern channel gating: activation, inactivation, deactivation and reactivation.

Activation describes the process whereby channels move from closed to open states; inactivation occurs when the inactivating gates close; deactivation arises when activating gates close; and reactivation (or recovery from inactivation) describes the process by which inactivating gates open. Consider the simplified model of Cl^- channel in the plasmalemma of cells of *C. corallina.*, which has three activating gates and one inactivating gate [21]. The schematic for this channel can be represented aswhere m_i, $i = 0, 1, 2, 3$, is the number of channels with i



open activating gates and h_j, $j = 0, 1$ is the number of channels with j closed inactivated gates. Recall that the only state in which the channel is permeable is the one in which the inactivating gate and all three activating gates are open, which is indicated by the asterisk. Also note the transitions between the different closed and open states are now scaled by a factor that captures how many gates must be open/closed in that state and that this factor assumes that the opening and closing of gates are independent of the state of the other gates.

The inclusion of multiple activating and inactivating gates does not affect the method of simulation, as long as the assumption that the opening and closing of channels are independent events. In cases where channels have 'memory', that is, that transition probabilities are dependent on the history of the channel, or the length of time the channel has been in a certain state, this does not hold. In the latter case, where transition probabilities are dependent on the 'dwell times' in specific states, a modified Markov chain description that respects this behaviour can be constructed, so that the Gillespie algorithm can be used [20].

3.3.4 Master Equation

Since the processes that we are describing are Markovian, we can instead represent the dynamics of the system through evolution of the probability, $P(S, t)$ of being in state S at time t. For the simple two state system represented by (3.6), the evolution of P is given by:

$$\frac{\mathrm{d}}{\mathrm{d}t}P(C,t) = -\alpha P(C,t) + \beta P(O,t) \tag{3.11}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}P(O,t) = \alpha P(C,t) - \beta P(O,t). \tag{3.12}$$

The linear equation (3.12) is known as the *master equation*, and is often the starting point for many considerations of biochemical reactions. General schema describing the states of channels with M states can be written in the form

$$S_i \underset{q_{ji}}{\overset{q_{ij}}{\rightleftharpoons}} S_j \tag{3.13}$$

where S_i and S_j, $(i, j = 1, \ldots, M)$, are two distinct states of the system and q_{ij} represents the probability (which may be zero) of entering state S_j from S_i. The master equation for system (3.13) can then be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}P(S_i, t) = \sum_{j=1}^{M} P(S_j, t)q_{ji} - \sum_{j=1}^{M} P(S_i, t)q_{ij}, \quad i = 1, \ldots M. \tag{3.14}$$

The term on the left represents a *source* term describing all the routes to state S_i, whilst the right-hand *sink* term captures all the transition away from state S_i. Since (3.14) is linear, we can write it in matrix form as

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{P} = \mathbf{P}Q, \quad \mathbf{P} \in [0, 1]^M, \tag{3.15}$$

where the $M \times M$ matrix Q contains all of the transition probabilities between the distinct states.

Whilst the master equation is a succinct way to describe the probability density function of the system, it is often impractical when simulating a system since it captures dynamics in a *distributional* sense, and does not provide a stochastic representation of a specific *realisation* of the process. However, when the number of channels becomes large, we can use a similar density based representation to describe the average behaviour across the whole cell.

3.3.5 Averaging

When the number of channels is large, instead of considering the probability of the system being in state S_i at time t, we can instead construct an equation to describe the *fraction of channels in state* S_i, which we shall denote s_i. Under this approximation, we replace (3.14) with

$$\frac{\mathrm{d}s_i}{\mathrm{d}t} = \sum_{j=1}^{M} s_j r_{ji} - \sum_{j=1}^{M} s_i r_{ij}, \quad i = 1, \ldots M, \tag{3.16}$$

where r_{ij} now represents the *rate* at which channels transition from state S_i to state S_j. Since channels are only permeable when the channel is open, we can simplify (3.16) by only considering the fraction of channels that are in the open state. Letting m represent the fraction of open channels, we replace (3.16) with

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \alpha_m(1 - m) - \beta_m m, \quad m \in [0, 1], \tag{3.17}$$

where α_m is the rate of channels entering the open state and β_m is the rate of channels exiting the open state. Note that these rates are the same as the probabilities of transitioning to the open state that we considered earlier. Thus, in the limit of infinitely many channels, transition probabilities are replaced with transition rates. Of course, cells do not have infinitely many channels and since (3.17) is deterministic, one may wonder how we are supposed to reflect the fluctuations arising from having finitely many channels. One way to incorporate these fluctuations, for example in a system with N_m channels, is to add an additive noise term to (3.17) [8]

$$\frac{\mathrm{d}m}{\mathrm{d}t} = \alpha_m(1-m) - \beta_m m + \xi_m(t), \tag{3.18}$$

where $\xi(t)$ is a Gaussian noise term with moments given by $\langle \xi_m(t) \rangle = 0$ and

$$\left\langle \xi_m(t)\xi_m(t') \right\rangle = \frac{\alpha_m(1-m) + \beta_m m}{N_m}\,\delta(t-t').$$

When incorporating this noise term in, care must be taken to ensure that m remains in the interval $[0, 1]$. For practical reasons, these finite size fluctuations are very often ignored in most ion channel based models of cell electrophysiology.

It is very common to see (3.17) written in the form

$$\tau_m \frac{\mathrm{d}m}{\mathrm{d}t} = m_\infty - m, \tag{3.19}$$

where

$$\tau_m = \frac{1}{\alpha_m + \beta_m}, \quad m_\infty = \frac{\alpha_m}{\alpha_m + \beta_m}. \tag{3.20}$$

If α and β are given as in (3.9), with $A = B$, then we can write

$$m_\infty(V) = \frac{1}{1 + \mathrm{e}^{-(V-V_0)/S_0}}, \tag{3.21}$$

where

$$S_0 = \frac{1}{k_\beta - k_\alpha}, \quad V_0 = \frac{k_\beta d_\beta - k_\alpha d_\alpha}{k_\beta - k_\alpha}. \tag{3.22}$$

Thus, according to (3.19), m evolves towards m_∞ at a rate determined by τ_m. The form of m_∞ given by (3.21) suggests that as the cell depolarises, the fraction of open channels increases or decreases in a sigmoidal fashion, which is often observed in experiments. In particular, the form given in (3.19) is often preferred over (3.17) since the parameters V_0 and S_0 in (3.21) can be fit directly to electrophysiological recordings.

3.4 Modelling Cellular Membrane Excitability

Now that we have arrived at a deterministic description for the channel dynamics, we are ready to start describing the currents that flow into and out of the cellular compartments as a result of channel opening.

3.4.1 Nernst Potential

In order to understand the current flow induced by the movement of ions across a cellular membrane, we need to first provide an equation that describes the electrochemical gradient across it. This can be done by equating the chemical potentials across the membrane. For space reasons, we here omit the derivation and simply state that, at equilibrium, the voltage across the cell membrane for ionic species X *independent of all other species* is given by the Nernst potential:

$$V_X = \frac{RT}{z_X F} \ln\left(\frac{[X]_\mathrm{o}}{[X]_\mathrm{i}}\right), \tag{3.23}$$

where R is the ideal gas constant, T is the temperature in Kelvin, z_X is the valence of species X, F is the Faraday constant and $[X]_{\mathrm{i,o}}$ are the concentrations of species X on the inside and outside of the membrane, respectively. Similar calculations can be used to find the voltage across cellular compartments at equilibrium accounting for all ionic species of interest.

3.4.2 Membrane Currents

We are now in a position to describe current flow across the cellular membranes, such as plasma, vacuolar or nuclear envelope membranes. Under the assumption that the concentrations of ions in the outside and inside of the membrane remain constant, the opening of channels permeable to ionic species X will push the transmembrane voltage towards the Nernst potential V_X. The current I_X induced by the flow of this species across cellular membranes is then given by Ohm's law, which can be written as

$$I_X = g_X(V - V_X), \tag{3.24}$$

where g_X is the summed channel conductance over the whole cell.

This conductance is proportional to the fraction of open channels, and thus we can rewrite (3.24) as

$$I_X = \overline{g}_X m^a h^b (V - V_X), \tag{3.25}$$

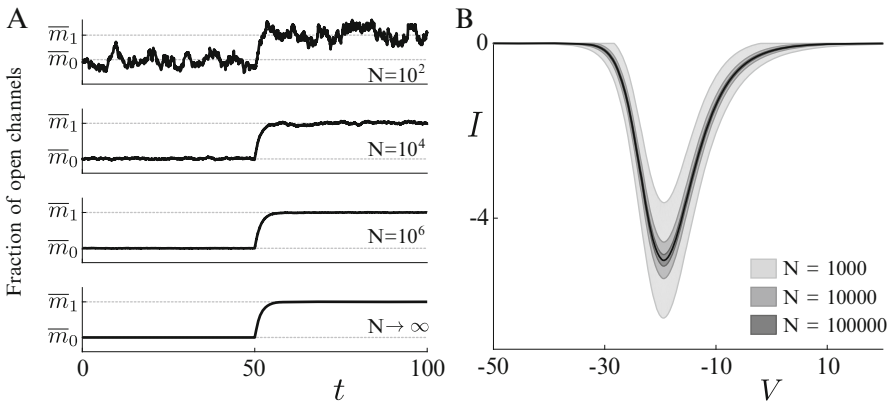


Fig. 3.2 (**a**) Sample paths representing the fraction of open channels for different values of the total channel number N. At $t = 50$, the transition rates for the channels are instantaneously changed as per the text. (**b**) Prototypical steady state currents under variation of voltage. The bold line represents the deterministic $I - V$ curve in the limit as $N \to \infty$. The shaded regions represent the mean $\pm$ standard error corresponding to the indicated number of channels. For further details, please refer to the text

in which there are a activating gates and b inactivating gates per channel, m and h, respectively, represent the fraction of open activating gates and inactivating gates and $\overline{g}_X$ is the total conductance when all channels are open. The dynamics of m and h obey the equation given by (3.19), though we note that the sigmoid for h_∞ is reversed compared to m_∞.

In Fig. 3.2, we demonstrate how the number of channels in the cell impacts the variability of the cell response. Figure 3.2a illustrates how the fraction of open activating channels varies as the transition rates are instantaneously changed from $\alpha_0 = 0.1$, $\beta_0 = 0.5$ to $\alpha_1 = 0.2$, $\beta_1 = 0.4$ at $t = 50$. In this subfigure, we also vary the total number of channels, N, in the cell and observe a reduction in the variability around the steady states, which are at $\overline{m}_i = \alpha_i/(\alpha_i + \beta_i)$, $i = 0, 1$. The lowermost plot shows the system in the thermodynamic limit as $N \to \infty$, and we observe that as N is increased, the sample paths approach this deterministic limit.

To assess the variability of channel opening and closing on current flow, we illustrate in Fig. 3.2b a prototypical $I - V$ curve, relating the steady-state current to the transmembrane voltage, assuming that the channels exhibit voltage dependence. The bold line indicates the current flow in the limit as $N \to \infty$, and we clearly see that as N increases, the variability around the deterministic profile decreases. In this example, we assume that the current is modelled according to (3.25), with $\overline{g}_X = 10$, $a = 3$, $b = 1$, $V_x = 50$ using steady state representations for m and h in the form given by (3.20) with α and β in the form (3.9). Specific parameters used are $A = B = 1$, $k_\alpha^m = 0.1$, $V_\alpha^m = -30$, $k_\beta^m = -0.4$, $V_\beta^m = -25$, $k_\alpha^h = 0.1$, $V_\alpha^h = -40$, $k_\beta^h = 0.3$ and $V_\beta^h = -40$.

In the following, we assume that the number of channels in plant cells is always high, so that we can use the deterministic limit to describe the ionic currents flowing through the channels.

3.4.3 Action Potentials

Equation (3.24) reflects the notion that we can represent the electrophysiology of the cell by making comparisons to electrical circuits. The Nernst potential provides an electromotive force similar to that produced by an electrical battery (cell), whilst the ion channels collectively behave as a variable conductance obeying nonlinear dynamics. By extending the equivalence of the biological system with an electrical circuit, and considering the evolution of currents associated with all ionic species to which a cellular membrane is permeable, we can now describe the evolution of the voltage across such a cell membrane.

By providing a barrier to the movement of ions across it, the cell membrane effectively separates charge between the inside and outside of the cell and thus acts as a capacitor. This capacitor is in parallel with all the ion channels on the membrane. Upon applying Kirchoff's second law, we can thus describe the evolution of the transmembrane voltage via

$$C\dot{V} = -\sum_X I_X, \tag{3.26}$$

where $C = 2 \times 10^{-2}$ F/cm^2 is the capacitance of the cell and $C\dot{V}$ is the current flowing through the capacitance. This equation, together with the equations that describe how the gating variables of each channel type vary with V and time (3.17), constitutes the Hodgkin-Huxley formalism for modelling action potentials [18].

As an example, we present a simple model of action potential in guard cells [11, 12]. This model incorporates three types of ion channels: one chloride channel that provides a fast activating and slowly inactivating inward current driving the action potential $I_{\mathrm{Cl}} = \overline{g}_{\mathrm{Cl}} m_{\mathrm{Cl}} h_{\mathrm{Cl}} (V - V_{Cl})$ with a Nernst potential at $V_{\mathrm{Cl}} = 100$ mV, one slowly activity outward rectifier potassium channel that provides an outward current that helps terminate action potentials $I_{\mathrm{Ko}} = \overline{g}_{\mathrm{Ko}} m_{\mathrm{Ko}} (V - V_{\mathrm{K}})$, and one inward rectifier potassium channel responsible for potassium influx into the cytoplasm at negative resting membrane potential $I_{\mathrm{Ki}} = \overline{g}_{\mathrm{Ki}} h_{\mathrm{Ki}} (V - V_{\mathrm{K}})$. Both K^+ currents have Nernst potentials at $V_K = -100$ mV. Two other ion transporters also create ionic currents. One is the proton pump, extruding H^+ ions at negative membrane potentials and thus generating an outward current $I_{\mathrm{pu}} = \overline{g}_{\mathrm{pu}} h_{\mathrm{pu}} (V - V_{\mathrm{pu}})$ with Nernst potential $V_{\mathrm{pu}} = -400$ mV, which is responsible for the resting membrane potential being more negative than V_K. The second one is a $2\mathrm{H}^+/\mathrm{Cl}^-$ symporter that brings in chloride at negative membrane potentials and creates a net inward current $I_{\mathrm{sy}} = \overline{g}_{\mathrm{sy}} h_{\mathrm{sy}} (V - V_{\mathrm{sy}})$ that reverses at $V_{\mathrm{sy}} = 20$ mV. Note that this model assumes there is no long-term change in ionic concentration that would be large enough to affect the Nernst potentials.

The steady state value of each current, i.e. the current observed at any given membrane potential V when the gating variables have reached their steady state values, is shown in Fig. 3.3a. The bottom panel shows the sum of all the currents, revealing three membrane potential values for which the total steady state current is 0. The lowest (leftmost) and highest (rightmost) values are stable steady states of V. This bistability provides a simple explanation for the observation that steady state membrane potential of many plant cells can be either at a hyperpolarised level well below V_K or at a level above V_K. The middle value acts as a threshold, if V is above/below that value it will converge toward the high/low steady state.

To simulate the variations of V with time, we numerically integrate the differential equations for V (3.26) together with the equation for each gating variable (3.17). Results of several simulations are displayed in Fig. 3.3b. For each simulation

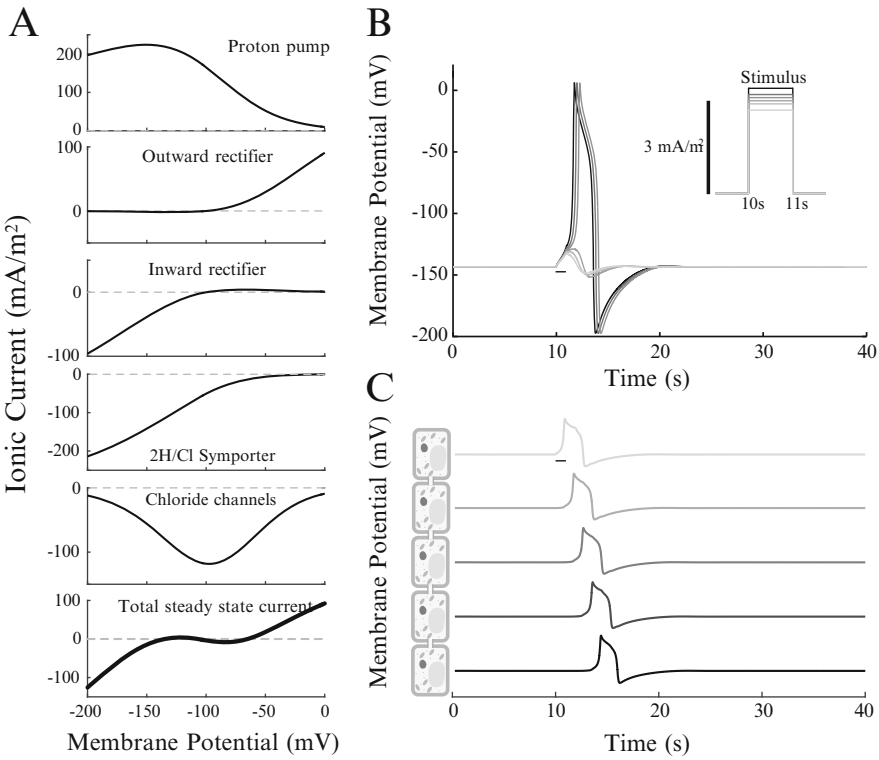


Fig. 3.3 (**a**) Steady state current of each ionic current. Bottom panel shows the sum of all the currents. Values for each conductance are: $\overline{g}_{pu} = \overline{g}_{Ko} = \overline{g}_{Ki} = \overline{g}_{sy} = \overline{g}_{Cl} = 1\ \mathrm{S/m^2}$. Other parameter values are given in text or in [12]. (**b**) Action potential are generated in response to a stimulating current if the current is larger than a threshold value. All parameters as in A, except $\overline{g}_{Cl} = 0.9\,\mathrm{S/m^2}$. Stimulating current of 2.8, 2.9, 3 (subthreshold), 3.1, 3.2, 3.4 $\mathrm{mA/m^2}$ is applied for one second (inset). (**c**) Propagation of an action potential wave. The Cl^- conductance is increased by 0.1 $\mathrm{S/m^2}$ between 10 and 11 s in the top cell. This triggers an action potential in this cell, which triggers an action potential in the next cell, and so on. The conductance of each plasmodesma is $g_p = 0.05\,\mathrm{S/m^2}$

we inject a brief (one second) stimulating current to the model cell, which creates a rapid rise in membrane potential. This results in a rapid return of membrane potential towards its equilibrium value, or in a large amplitude action potential before going back to equilibrium. The action potential is triggered only if the depolarizing current is sufficiently large, illustrating the threshold behaviour. Once the current is above its threshold value, a full action potential occurs. The shape of the action potential varies little with further increases in stimulating current, but the latency between stimulus and action potential onset decreases. Thus, this simple model of plant action potential exhibits features also found in action potentials produced by animal cells.

Plant action potentials can also propagate from cell to cell, thanks to plasmodesmata that create electrical connections between cells and allow exchange of ions [6]. To illustrate this action potential propagation [16], we model the electrical connection, which is similar to gap junctions that electrically connect excitable cells in animals [2, 25]. Thus, the current I_{pab} flowing through a plasmodesma from cell a with membrane potential V_a to cell b with membrane potential V_b is simply given by

$$I_{pab} = g_p(V_a - V_b) \tag{3.27}$$

where g_p is the conductance of the plasmodesma [22, 23]. In Fig. 3.3c, an action potential is initiated in the first cell by a brief increase in the Cl^- conductance. This sequentially triggers action potentials in the neighbouring cells, thus propagating a wave of electrical excitation.

3.5 Conclusions

Plants (and animals) very effectively use concentration differences of ions across membranes (inside vs outside the cell and between cellular compartments) to regulate numerous processes. Calcium ions, for instance, are key information carriers in plants that regulate responses to abiotic and biotic stresses from their environment. In this chapter, we have introduced the fundamental physical principles that underpin ionic changes via the function of ion channels and provided some useful mathematical and computational approaches for modelling such systems. Many exciting questions in plant biology converge on ion channels and transporters and we hope that the tools offered here will provide a foundation on which interested researchers can build to develop specific biophysical models to address specific questions.

Acknowledgements KW was generously supported by the Wellcome Trust Institutional Strategic Support Award (WT105618MA). KT-A gratefully acknowledges the financial support of the EPSRC via grant EP/N014391/1.

References

1. Arrhenius SA (1889) Über die dissociationswärme und den einfluß der temperatur auf den dissociationsgrad der elektrolyte. Zeitschrift für Physikalische Chemie 4:96–116
2. Brink PR, Cronin K, Ramanan S (1996) Gap junctions in excitable cells. J Bioenerg Biomembr 28(4):351–358
3. Dietrich P, Anschütz U, Kugler A, Becker D (2010) Physiology and biophysics of plant ligand-gated ion channels. Plant Biol 12(s1):80–93
4. Eyring H (1935) The activated complex and the absolute rate of chemical reactions. Chem Rev 17:65–77
5. Eyrin H (1935) The activated complex in chemical reactions. Chem Rev 3:107–115
6. Faulkner C (2013) Receptor-mediated signaling at plasmodesmata. Front Plant Sci 4:521
7. FitzHugh R (1961) Impulses and physiological states in theoretical models of nerve membrane. Biophys J 1(6):445–466
8. Fox R (1997) Stochastic versions of the hodgkin-huxley equations. Biophys J 72:2068–2074
9. Fromm J, Lautner S (2007) Electrical signals and their physiological significance in plants. Plant Cell Environ 30(3):249–257
10. Gillespie D (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81:2340–2361
11. Gradmann D (2001) Models for oscillations in plants. Funct Plant Biol 28(7):577–590
12. Gradmann D, Blatt M, Thiel G (1993) Electrocoupling of ion transporters in plants. J Membr Biol 136(3):327–332
13. Granqvist E, Wysham D, Hazledine S, Kozlowski W, Sun J, Charpentier M, Vaz Martins T, Haleux P, Tsaneva-Atanasova K, Downie JA, Oldroyd GE, Morris RJ (2012) Buffering capacity explains signal variation in symbiotic calcium oscillations. Plant Physiol 160(4):2300–2310
14. Hedrich R (2012) Ion channels in plants. Physiol Rev 92(4):1777–1811
15. Hedrich R, Neher E (1987) Cytoplasmic calcium regulates voltage-dependent ion channels in plant vacuoles. Nature 329(6142):833–836
16. Hedrich R, Salvador-Recatalà V, Dreyer I (2016) Electrical wiring and long-distance plant communication. Trends Plant Sci 21(5):376–387
17. Hille B et al. (2001) Ion channels of excitable membranes, vol 507. Sinauer, Sunderland
18. Hodgkin AL, Huxley AF (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol 117(4):500
19. Keener JP, Sneyd J (2009) Mathematical physiology, vol 1. Springer, Berlin
20. Lowen SB, Liebovitch LS, White JA (1999) Fractal ion-channel behavior generates fractal firing patterns in neuronal models. Phys Rev E 59(5):5970–5980
21. Spalding EP, Slayman CL, Goldsmith MHM, Gradmann D, Bertl A (1992) Ion channels in arabidopsis plasma membrane transport characteristics and involvement in light-induced voltage changes. Plant Physiol 99(1):96–102
22. Sukhov V, Nerush V, Orlova L, Vodeneev V (2011) Simulation of action potential propagation in plants. J Theor Biol 291:47–55
23. Sukhov V, Vodeneev V (2009) A mathematical model of action potential in cells of vascular plants. J Membr Biol 232(1–3):59
24. Tester M (1990) Tansley review no. 21 plant ion channels: whole-cell and single channel studies. New Phytol 114(3):305–340
25. Unwin N (1989) The structure of ion channels in membranes of excitable cells. Neuron 3(6):665–676
26. Ward JM, Schroeder J (1997) Roles of ion channels in initiation of signal transduction in higher plants. In: Signal transduction in plants. Springer, Berlin, pp 1–22

Chapter 4
Modelling the Plant Microtubule Cytoskeleton

Eva E. Deinum and Bela M. Mulder

Abstract The physical shape and structure of plants are manifestations of the actions of gene products and their concerted responses to their environment. In this chapter we introduce the plant cortical microtubule array. This structure is both a nexus in the control of plant cell shape and function, and a fascinating out-of-equilibrium system for state-of-the-art physics research. We describe how analytical and computational approaches complement each other in the study of the array, and highlight some recent results and open research questions.

4.1 Microtubules: Dynamic Controllers of Plant Cell Growth and Development

Plant cells distinguish themselves from other eukaryotic cells by being encased in a rigid cell wall. These walls are composed of long cellulose microfibrils embedded in a matrix consisting of a complex mixture of other polysaccharides [1]. The cells typically adopt strongly anisotropic shapes requiring a significant degree of control over the way they grow. Moreover, the typical dimensions of plant cells in the range of 10–100 μm exceed those of most mammalian cells. This implies that plant cells require the means to sense and control their geometry on the scale of 10s of μm's. A major molecular workhorse in providing these key services is the microtubule (MT) cytoskeleton. It is composed of long and highly dynamic filamentous protein polymers [2] that, aided-and-abetted by a class of microtubule associated proteins (MAPs) [3], are able to self-organize into a number

E. E. Deinum
Biometris, Wageningen University & Research, Wageningen, The Netherlands
e-mail: eva.deinum@wur.nl

B. M. Mulder (✉)
Living Matter Department, Institute AMOLF, Amsterdam, The Netherlands

Department of Plant Science, Wageningen University & Research, Wageningen, The Netherlands
e-mail: mulder@amolf.nl

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_4

of cell-scale functional structures. In order of appearance during the cell cycle they are the cortical array (CA), the pre-prophase band, the mitotic spindle and the phragmoplast. Of these four, the CA [4], the pre-prophase band [5] and the phragmoplast [6] are plant specific. Even the plant mitotic spindle, although overall similar in shape to its mammalian counterpart, does have significant structural differences to the latter [7]. MTs in mammalian cells are associated to so-called centrosomes [8], which act as hubs for their nucleation and typically dictate a radial organization. Plant cells do not have centrosomes, which implies that other organizational principles must be at play [9]. Of the four MT structures mentioned above, the CA has to date received the most interest. The CA invariably appears in growing plant cells and is composed of MTs roughly homogeneously dispersed over the entire cell cortex, which are oriented in a direction transverse to the growth axis of the cell.

4.2 Dynamic Self-organization: Exciting Biology and Challenging Physics/Mathematics

What makes structures like the CA doubly interesting is that they are apparently self-organizing [10]. In root epidermal cells, the CA develops in approximately 1 h after cytokinesis from an initial state in which the cortex which is devoid of any MTs. More strikingly, while a fully developed CA can be totally disrupted by adding MT-depolymerizing drugs such as oryzalin, it returns to its original state within an hour after these drugs are washed out are washed out [30]. This obviously speaks for their robustness from the point of view of biological functionality. At the same time this means that the CA poses the challenging question of which mechanism(s) drive and control its formation. The CA is composed of $10^2 - 10^4$ of MTs, each of which is a stochastic dynamical system in its own right. Somehow, through interactions between these dynamical actors a state is reached which has well-defined and stable global properties, in spite of the unceasing stochastic activity of its elements. From the point of view of statistical physics, this characterizes the CA as a far out of equilibrium steady state of an ensemble of interacting spatially extended particles. Such systems represent the cutting edge of our physical understanding [11].

4.3 MT Behaviour at Different Levels of Description

4.3.1 At the MT Level: Dynamic Instability

Individual MTs are highly dynamic protein structures. They are long (up to many μm), thin (about 25 nm in diameter) and polar. Growth and shrinkage of MTs occurs at their ends, which are called “plus” and “minus”. MT behaviour can be described at many different levels, with consequences for the kinds of questions that can be tackled with a certain description. The basic building blocks of MTs

are α, β-tubulin heterodimers. These subunits form chains, and typically 13 of these so-called protofilaments together form an MT. The plus end of the MT is highly dynamic: it may grow through polymerization, the incorporation of new subunits, shrink through depolymerization or pause. Switches between growth and shrinkage are called catastrophes and rescues. The minus end is also dynamic, but with different kinetic parameters, and typically has a tendency towards net depolymerization.

If you are interested in understanding or predicting the parameters that describe this so-called dynamic instability, you have to consider many molecular details. Tubulin subunits are incorporated in a GTP-bound state, and stochastically get hydrolyzed into a GDP-state. Protofilaments of GTP-tubulin are almost straight, but GDP-tubulin protofilaments have a tendency to curve. Consequently, a growing MT is stabilized by a cap of GTP-tubulin. If this cap disappears, however, due to depletion of free subunits, delayed growth caused by some obstruction, or whatever other reason, the protofilaments start curving outward and the plus end starts shrinking fast: a catastrophe [12]. To understand the process of a rescue, you moreover have to consider the many kinds of proteins that interact with MTs. Detailed molecular models may help predict how mutations in tubulin or other relevant proteins affect MT dynamics.

4.3.2 At the Cell Level: Interactions

Understanding the behaviour of the entire cortical array, on the other hand, is easier with a much simplified model of MT dynamics (Fig. 4.1). Disregarding all molecular detail, growth and shrinkage may be described with one constant (average) velocity each, and the transitions may be described with constant rates. For many questions, we may entirely forget about the pause state and map a 3-state (growth/shrinkage/pause) description onto a 2-state (growth/shrinkage description) [13].

Similarly, nucleation of new MTs occurs mostly from existing MTs and with specific distributions of relative angles between parent and child [14]. These distributions, in turn, can be tuned by the plant. Nucleation complexes, moreover, are about 10 times more active when associated with an MT than when not [15]. Depending on the questions you are interested in, it may be important to incorporate many details of the nucleation process, a caricature of the process that can be described with one or few parameters, or simple isotropic nucleation may even suffice.

And what about the cell itself? Is it sufficient to describe the cell cortex as a simple periodic surface of an appropriate size, is it essential to approximate the shape of real plant cells as close as possible, or would some simple geometry like a box or cylinder be better? That, too, depends on the question.

The MTs comprising the CA are attached to the cell membrane. This essentially reduces the CA to an (almost) two-dimensional system. Consequently, growing

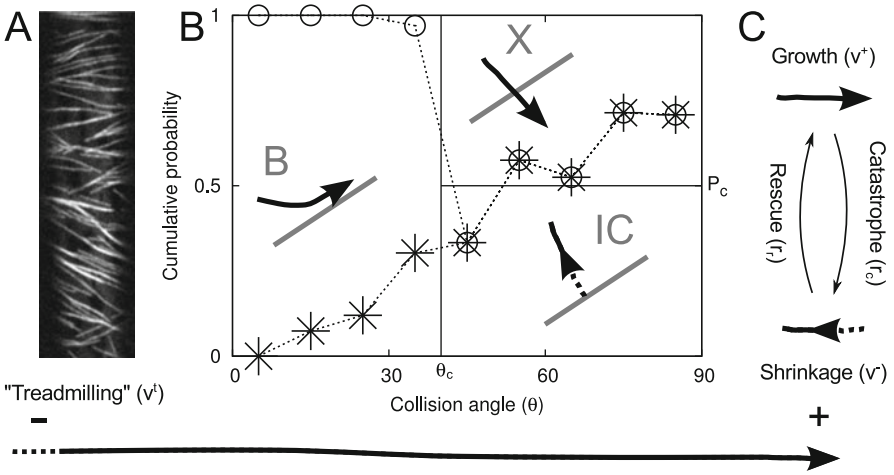


Fig. 4.1 (**a**) Cortical array in an *A. thaliana* hypocotyl cell (confocal microscope image courtesy J.J. Lindeboom, Stanford). (**b**) Interaction between MTs. Depending on the relative angle θ, a collision may result in bundling (grey B), or an induced catastrophe (grey IC, with probability P_c) or crossover (grey X). (**c**) Dynamic instability at the MT plus end

MTs will frequently collide into other, obstructing, MTs. Modelling studies of the CA were sparked by the observations by Dixit and Cyr [16] that the outcomes of such collisions strongly depend on the relative angle (θ) between incoming and obstructing MT (Fig. 4.1). For shallow angle interactions ($\theta \lesssim 40°$), the incoming MT will typically continue growing along the obstructing MT, a process referred to as bundling (or sometimes as zippering or entrainment). For larger angle interactions, the MT may either continue growing in its original direction ("crossover") or, typically after a short pause, start shrinking ("induced catastrophe").

The observation of these angle dependent collision outcomes immediately raised the question: are these interactions, combined with the standard MT dynamic instability, sufficient for spontaneous self-organization of the CA? Several groups approached this question using models, both analytical [17–19] and computational [13, 20–22].These approaches complement each other, for example, see Sect. 4.6.

4.3.3 At the Tissue Level: Coarse Graining

Going up to the tissue level, cell expansion properties must be coordinated to ensure tissue coherence, especially during growth. Such coordination may be achieved through simultaneous responses to the same external cue, for example in blue light induced reorientation of the array [23], with mobile developmental signals such as a plant hormones [24], or in response to tissue spanning stresses [25]. With regard to the latter: a correlation between stresses in the cell wall and the net orientation of the

CA is observed in certain cell types [26, 27]. Current models of this phenomenon ignore the mechanistic basis of this correlation and explore its consequences in mechanical models which describe the entire CA with just two numbers per cell [26, 28, 29].

4.4 The Consensus Model of MT Dynamics

Following the discussion in 4.3.2, we focus on what has become the generally accepted model for describing MT dynamics in the CA. In this model MTs live and die in a two-dimensional domain which represents (part of the) the cell cortex. They are born through nucleation events that can occur either at random locations within this domain, in which case their initial growth direction is also random, or from a random position on the body of an existing MT, in which case their initial orientation may depend on that of the parent MT. MTs are taken to be composed of straight line segments, where on curved surfaces "straight" should be interpreted as following a geodetic path. The plus-end tip of a growing MT moves with the growth speed v^+, while if applicable the minus-end retracts with the treadmilling speed v^t. A growing MT can switch to the shrinking state with the constant catastrophe rate r_c. The plus-end of a shrinking MT retracts with a speed v^- and can switch to the growing with the rescue rate r_r. When the tip of a growing MT collides with another MT, the outcome depends on the angle of incidence θ. Whenever this angle is below the critical angle $\theta_c = 40°$, a bundling event occurs. At a bundling event, a new segment of the colliding MT is created along the obstructing MT. In this way, an MT can consist of multiple segments, each with its own orientation and length. Only the front segment is *active* and either in the growing or the shrinking state. The remaining segments are *inactive*, except in the case of minus-end retraction in which the rear segment is always shrinking. Inactive segments can thus be *reactivated* when they start shrinking either at their plus- or their minus-end. When at a collision the incidence angle is above θ_c, the MT undergoes an induced catastrophe with probability $0 < P_c < 1$, where $P_c = 0.5$ is here chosen as default. Finally, when an MT shrinks to length 0 it simply disappears.

4.5 Simulations: Time Steps Versus Event Driven

As described above, simulation studies addressing the question of array alignment model individual MTs as connected sets of line segments that switch between growth, shrinkage and possibly pausing at the plus end, and have either stable or steadily retracting minus ends. There are essentially two ways of implementing these phenomena in a simulation: using discrete time steps or with an event-driven algorithm.

With *discrete time steps*, all plus (and minus) ends are updated every time step (Δt). Simple cases are when they grow a certain length ($v^+\Delta t$) without obstruction, or shrink ($v^-\Delta t$) during the whole time step: the tip's is propagated by the specified length and that's it. Additionally, tips may switch state during the time step with certain probability (r_c, r_r, ...). This may sound simple enough, but you have to think carefully about the length change of the MT during this time step: depending on when during the time step it presumptively has switched, it may show net growth, shrinkage, or no change of length. The difference between (the extremes of) these options, of course, decreases with shorter time steps, but that has a computational cost.

When the length increase of a growing MT brings it across an obstructing MT, the resulting interaction must be handled: for example, if this resulted in a bundling event, the new position of the plus tip will be along the obstructing MT, at a distance beyond the intersection the length of its initial overshoot.

Handling these interactions one by one works as long as it is safe to assume that they are all independent. With increasing Δt, however, the probability increases that MT changes become interdependent. For example, an early updated plus end may undergo bundling onto an MT that itself undergoes a (spontaneous or induced) catastrophe in the same time step—effectively resulting in bundling onto empty space. As you can see, this method requires a lot of careful thinking and bookkeeping that is not apparent from the simple idea behind the algorithm. The problem of interdependent interactions/state changes not only aggravates with longer time steps, but also with increasing array density. Consequently, the maximum allowable time step size may turn out to be so small that it becomes hard to obtain good simulation statistics of large or high density arrays.

In *event-driven* simulations, the problems of event interdependence and the uncertainty of MT length after a state switch are solved. The essence of event-driven simulations is that events (interactions among MTs, spontaneous state changes, and even measurements) are scheduled against a continuous wall clock. These events are handled one by one and, where necessary, the probabilities of particular event types are adjusted after each event. Calculating future event times requires forward integrating the equations of tip motion (for growth and shrinkage, which are *deterministic events*), or drawing random numbers based on the probability density of *stochastic events*.

If done in a smart way, such event-driven simulations can be very computationally efficient. We review some concepts that help keep the simulations efficient [13, 30]. First of all, as MT are modelled using straight line segments, it is possible to compute exactly when a growing MT may interact with another one by extrapolating MT paths. The intersections between such lines ("trajectories" in [13]) can be precomputed and will be the same for all MTs within a bundle. To keep the computational cost of computing trajectory intersections sufficiently low, the simulation domain may be subdivided into smaller parts.

For stochastic events (rescues, spontaneous catastrophes, MT nucleation, ...), we only need to know the event time of the first event. If the first stochastic event is scheduled before the first deterministic event, it is executed, and the first

deterministic event otherwise. In the former case, the type of stochastic event and its specifics are selected proportional to the respective rates.

In general, the time interval Δt_s until the next stochastic event can be calculated from

$$\int_t^{t+\Delta t_s} R(\tau)\mathrm{d}\tau = -\log u \tag{4.1}$$

with $R(t)$ the sum of all instantaneous rates $r_i(t)$ for all different stochastic event types and u a uniform random variate on the interval (0,1). If all event types have constant rates *between two events*, then

$$R(t) = R = \sum_i r_i \tag{4.2}$$

and

$$\Delta t_s = -\frac{\log u}{R}. \tag{4.3}$$

For examples of non-constant rates, we refer to [13].

After an event, it is likely that the rates of some event types have changed slightly, for example, because the number of growing MTs has changed. This shows that one must only calculate a single stochastic event in advance, as the waiting time distributions may change at every time step.

Scheduling events with a constant rate is easy, as their waiting times all follow exponential distributions. Some rates, however, may not be constant over time. In that case there are two major options. The first is to calculate the correct waiting time distribution of the variable rate and sample according to that. The second is to use a rejection method: schedule events with a constant rate corresponding to the highest possible instantaneous rate and if the event is selected for execution, only execute it with probability instantaneous rate / maximum rate. The validity of the latter option depends on the Markov property of the model: the behaviour of the system is fully determined by its current state. Which of these two options is best depends on how hard it is to compute the exact waiting time distribution—mathematically but also computationally—and the difference between maximum rate and average instantaneous rate.

Finally, considerable computational savings are possible by only updating the necessary information. The positions of all MT ends must be known exactly for measuring the state of the array, but not for the correct scheduling and execution of events. The number of growing and shrinking MTs is enough to calculate, for example, how the total MT length and density change between events, which in turn is sufficient to correctly schedule all events that depend on these quantities.

4.6 Understanding Order

Before discussing how order comes about, we first need to define a *quantitative measure* of the degree of order. To that end we employ the notion of an *order parameter*. Order parameters have a long history in physics [31], and are widely used to distinguish different macroscopic states of physical systems. The one appropriate to describe the orientational order of line-like objects independent of their polarity is

$$S_2 = \sqrt{\langle \cos(2\theta) \rangle^2 + \langle \sin(2\theta) \rangle^2} \tag{4.4}$$

where θ is the angle of the MT segments with respect to an arbitrary, but fixed, axis in the plane, and the angle brackets $\langle \cdots \rangle$ denote taking the length weighted average over many microstates in the system. If all the directions are equally abundant, we find $\langle \cos(2\theta) \rangle = \langle \sin(2\theta) \rangle = 0$ and hence $S_2 = 0$, which characterizes the disordered state. On the other hand, if all segments have exactly the same orientation θ_0, then $S_2 = \sqrt{\cos^2(2\theta_0) + \sin^2(2\theta_0)} = 1$, which marks the fully ordered state. In this way S_2 creates a natural scale measuring the degree of order between the limits 0 and 1.

In order to understand how ordering could in principle arise in a system of dynamical MTs interacting in the manner described in Sect. 4.4, an analytical approach is indispensable. First of all, such a theory allows one to establish a direct link between the assumptions of the underlying model and the outcomes. Moreover, through a combination of structural and dimensional analysis of the formulated equations the "true" parameters that determine the behaviour of the system are revealed. Finally, the results of theory can be used to provide an independent validation of simulation results. In our analytical approach we forgo the description of individual MTs, but rather focus on probability densities per unit of their properties in a large statistical ensemble of equivalent systems. The basic dependent variables we consider are $m_i^+(l, \theta, t)$, $m_i^-(l, \theta, t)$ and $m_i^0(l, \theta, t)$ being respectively the density of growing, shrinking and dormant segments of length l, orientation θ at time t that were created after their parent MT underwent the bundling event. The evolution equations for these densities are schematically of the form

$$\frac{\partial}{\partial t} m_i^+(l, \theta, t) = \Phi_{\text{growth}} + \Phi_{\text{rescue}} - \Phi_{\text{cat}} - \Phi_{\text{indcat}} - \Phi_{\text{bundle}} \tag{4.5}$$

$$\frac{\partial}{\partial t} m_i^-(l, \theta, t) = \Phi_{\text{shrink}} - \Phi_{\text{rescue}} + \Phi_{\text{cat}} + \Phi_{\text{indcat}} + \Phi_{\text{reactivate}} \tag{4.6}$$

$$\frac{\partial}{\partial t} m_i^0(l, \theta, t) = \Phi_{\text{bundle}} - \Phi_{\text{reactivate}} \tag{4.7}$$

The fluxes on the right denote the effects of the different process that contribute to the changes of the densities, with the sign indicative of whether it is effectively a gain or a loss term. In addition there is a boundary condition that describes the nucleation of new MTs, which in the simplest case of uniform MT-independent nucleation is $v^+ m_0^+(0,\theta,t) = r_n/2\pi$, where r_n is the nucleation rate per unit area. The set of Eqs. (4.5)–(4.7) unfortunately is intractable in the general case. However, if we assume that a steady is reached, things simplify enormously. The first observation is that the length dependence of all densities is exponential, with a single orientation dependent length scale

$$m_i^+(l,\theta) = m_i^+(\theta) e^{-l/l(\theta)} \tag{4.8}$$

$$m_i^-(l,\theta) = \frac{v^-}{v^+} m_i^+(l,\theta) \tag{4.9}$$

$$m_i^0(l,\theta) = (1+\frac{v^-}{v^+}) Q(\theta) m_i^+(l,\theta) \tag{4.10}$$

The mean length of segments $l(\theta)$ in a given direction depends on the total intensity of collisions that lead to either an induced catastrophe or a bundling event

$$\frac{1}{l(\theta)} = \frac{1}{\bar{l}} + \int \mathrm{d}\theta' \, \sin(|\theta-\theta'|)\{H(\theta_c - |\theta-\theta'|) + P_c H(|\theta-\theta'|-\theta_c)\} k(\theta') \tag{4.11}$$

Here $k(\theta'')$ is the total length density of segments with orientation θ'' that forms the "target" which growing segments can hit, $\bar{l}$ is the mean length of an MT in the absence of any collisions, and $H(x)$ is the Heaviside step function. Clearly, the more collisions in a certain direction the shorter the filaments in that direction will become. This reveals the basic mechanism by which an ordered state can maintain itself: an MT segment aligned with the majority will suffer fewer state changing collisions and is therefore likely to be longer than less fortunate segments who are oriented differently from the majority. Without going into the details, there are three more equations needed for the full description, which together fix, next to the average segment length $l(\theta)$ and total segment length density $k(\theta)$ already mentioned, the ratio between active and passive segments $Q(\theta)$, which already appeared in (4.10), and $t(\theta)$ the density of active tips. While in principle the nature of the steady states depends on all five MT dynamical parameters r_n,v^+,v^-,r_c and r_r, an appropriate choice of the units length shows that only the ratio

$$G = -\frac{l_0}{\bar{l}} = -\left(\frac{2v^+v^-}{r_n(v^+ + v^-)}\right)^{\frac{1}{3}} \left(\frac{r_r}{v^-} - \frac{r_c}{v^+}\right) \tag{4.12}$$

is important [17]. The perhaps surprising choice of the minus sign ensures that the larger G, the larger the average length of the MTs. In fact, we limit ourselves

to the regime $G < 0$, where the MTs are in the bounded growth regime. For $G > 0$ MTs could in principle grow unbounded in length, and a steady state is no longer possible. The utility of the control parameter G extends far beyond the theory itself. As we have also validated by simulation, it predicts that systems that have very different dynamical parameters will still show the same self-organization if their G value is the same. The strategy to deal with this still arguably complex set of coupled equations is the following. One can show that the assumption that all four unknowns $l(\theta)$,$k(\theta)$,$Q(\theta)$ and $t(\theta)$ do *not* depend on θ, works for any value G. This yields the so-called isotropic solution to the equations. However, because the equations are non-linear this need not be the only solution. To probe for the existence other non-isotropic solutions, we use an approach called bifurcation theory. In bifurcation theory one perturbs the original isotropic solution with an appropriate small anisotropic component, e.g.

$$l(\theta) = l_0 + \epsilon l_2 \cos(2\theta) \tag{4.13}$$

where l_0 is the isotropic value of the mean segment length and $\epsilon << 1$ a small quantity. Note that at the outset we assume that although MTs are polarized structures, the collisions between them do not depend on the relative polarity. We therefore expect that inverting the direction of all MTs, i.e. letting θ go to $\pi - \theta$ should not change the solution, hence the choice of perturbation with period π. The perturbed values of the unknowns are then entered into the equations, which are subsequently expanded in orders of ϵ. Requiring that the perturbed functions are a solution to first order in ϵ leads to a soluble set of *linear* equations in the amplitudes l_2, k_2, Q_2 and t_2. This set of so-called bifurcation equations only has a solution for a specific value of the control parameter, which identifies the bifurcation points G^*. Thus, for all values of $G > G^*$ the isotropic solution is no longer stable, and an ordered state will spontaneously develop. If we work out G^* for the consensus model, we find

$$G^* = \left(\frac{2}{\pi} 2P_c \sin(2\theta_c)\right)^{1/3} \left(\frac{\pi - \theta_c}{\sin(2\theta_c)} - 1\right) \tag{4.14}$$

The most surprising aspect of this result is that if the probability of induced catastrophes vanishes ($P_c = 0$), the critical control parameter goes to the effectively unattainable value $G^* = 0$, implying that you need either infinitely long MTs or infinitely many ($r_n \to \infty$) to achieve ordering. It also implies that bundling by itself cannot be responsible for the ordering process. Thus the induced catastrophes are the driver of the ordering process. How can this be understood? The answer is what we have dubbed the "survival of the aligned" principle. When MTs are (almost) aligned they have few collisions among themselves, and if they do these are likely to lead to bundling rather than induced catastrophes. These aligned MTs will therefore be likely to live for their full lifetime. On the other hand, a discordant MT segment that encounters a set of aligned MTs will suffer many collisions. If at these collisions it has a finite probability of suffering an induced catastrophe, it will likely be switched

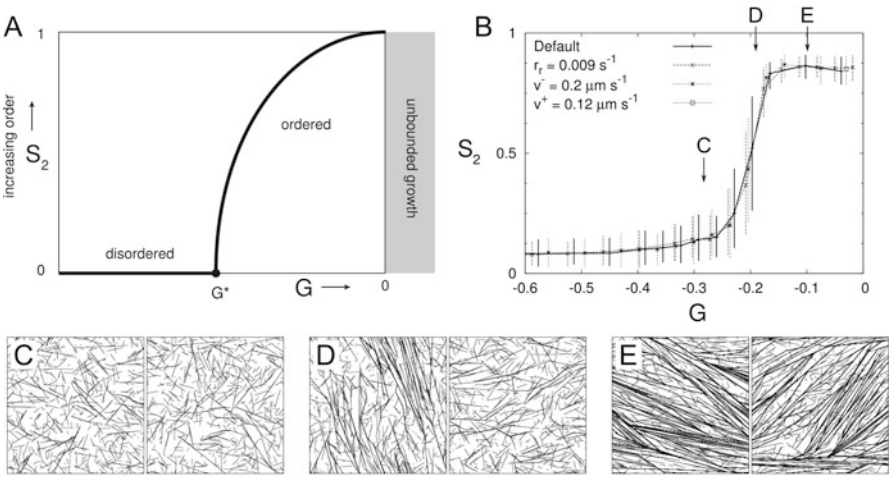


Fig. 4.2 Understanding order in the CA. (**a**) Theoretically predicted bifurcation diagram. Order arises if $G > G^*$. (**b**) Simulation results: the same G-value can be reached in many different ways. The transition between order and disorder always occurs around the same G-value. (**c**)–(**e**) representative snapshots of two simulated CAs per parameter combination as indicated with arrows in B along the curve with "Default" parameters, $r_n = 0.001\ \mathrm{s}^{-1}$, $r_r = 0.007\ \mathrm{s}^{-1}$, $v^t = 0.01\ \mu\mathrm{m\,s}^{-1}$, $v^- = 0.16\ \mu\mathrm{m\,s}^{-1}$, $v^+ = 0.08\ \mu\mathrm{m\,s}^{-1}$ and $r_c = 0.0075\ \mathrm{s}^{-1}$ (C), $0.006\,\mathrm{s}^{-1}$ (D) and $0.0045\,\mathrm{s}^{-1}$ (E)

into the shrinking state. A shrinking MT has a significantly shorter expected lifetime than a growing MT. Hence a discordant MT will likely disappear more quickly from the system than the aligned ones, leaving the latter as the stable majority. Likewise an MT nucleated in the majority direction survives longer than one that is nucleated in another direction, hence the name of the principle.

Armed with these results, we can show the generic behavior our theory of CA ordering predicted as a function of the control parameter G, and how this compares to the results of simulations described in Sect. 4.5 (Fig. 4.2).

4.7 Understanding Orientation

The orientation of the CA changes in response to environmental signals [23] as well as internal cues, such as mechanical stresses [26]. How this orientation is controlled is still an open area of research. In fact, two different processes play a role: first the *de novo* establishment of the array and the concurrent "first" choice of orientation, and second the reorientation of the array.

As stated, the CA is newly established after cell division and also after washout of MT depolymerizing drugs. During both processes a transient population of obliquely oriented MTs is observed, which disappears when the CA adopts its final

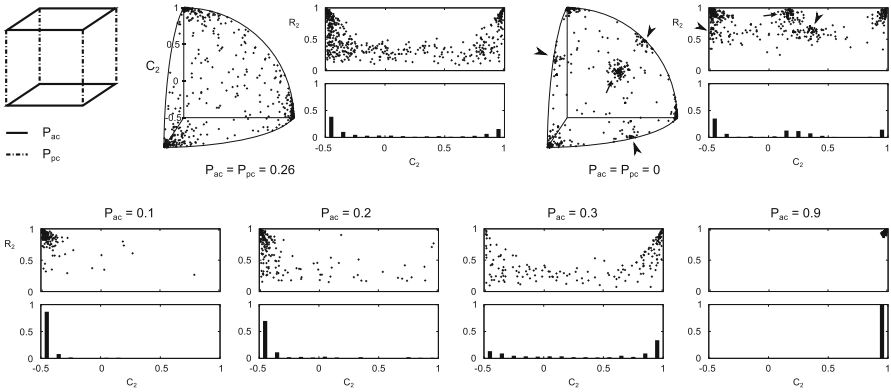


Fig. 4.3 Cell geometry favours certain orientations. The oblique orientations (arrows and arrow heads) only appear when the penalties for crossing edges (induced catastrophe probabilities P_{ac} and P_{pc}) are 0 or very low. Differences between P_{ac} and P_{pc} may help select between transverse ($C_2 = 1$) and longitudinal ($C_2 = -0.5$) orientations. Cubes are $15 \times 15 \times 15\,\mu$m. $n = 500$ simulations per parameter combination. Simulated time $T = 40000$ s. R_2: order parameter between 0 (isotropic) and 1 (perfectly aligned) [13]

transverse orientation [30]. The functional importance of this population remains unclear. The current hypothesis is that the oblique MTs disappear because the cell favours only certain orientations. It has been suggested that the sharp corners between cell faces—particularly those to the cell face that was formed during the last cell division—are hard to cross for growing MTs, and that this crossing may be facilitated by the protein CLASP [32]. This suggests a mechanism in which the cell selects a particular orientation through selective facilitation of edge crossing (Fig. 4.3). Even without (controllable) edge crossing penalties, the geometry of the cell itself typically allows only a few orientations [33].

The different orientations allowed by the cell geometry could be seen as local minima in a potential energy landscape. On simple geometries such as rectangular boxes and cylinders, these potential wells correspond to closed paths on the geometry surface that minimize curvature along the path because MTs are stiff. In this interpretation, the establishment of the initial orientation is like finding one of the potential wells and if the difference between different wells is large, most cells are expected to find the global minimum (deepest well). To change orientation, then, would require that the CA is actively "pushed out" of its current well. Making another state more favourable (deeper well) alone would be insufficient for reorienting the array. Such active reorientation has indeed been observed in hypocotyl cells that change from transverse to longitudinal CA orientation in response to blue light [23]. During the initial stages of this reorientation process, severing of discordant (oblique and longitudinal) MTs followed by immediate rescues quickly amplifies the population of longitudinal MTs [23].

4.8 The Role of Severing

Various experiments have indicated the MT severing enzyme katanin as a key player in inducing alignment in the CA as well as enabling rapid reorientation of the CA in response to environmental signals and cell wall stresses [23, 26] At first sight, this poses a theoretical problem: by the "survival of the aligned" principle, alignment critically depends on having sufficient interactions per MT life time to effectively "weed out" discordant MTs. Random severing along the MT lattice—with new MT plus ends starting in the shrinking state—results in shorter MTs, and hence fewer interactions per MT life time and less alignment [19]. Severing, however, preferentially occurs at MT crossovers and predominantly affects the crossing MT [23, 34]. As the crossing MT is more likely to be a discordant one, this makes crossover severing a delayed punishment of such MTs, effectively contributing to alignment like a weaker form of induced catastrophes [35]. Notably, severing at crossovers only contributes to alignment if shallow angle interactions are protected from severing. Bundling normally offers this protection, because it prevents the formation of crossovers with small relative angles. This for the first time indicates that bundling *is* important for alignment *in planta* [35]. Understanding how severing can boost alignment also gives a foundation for understanding how severing is important for changes in orientation. Reorientation requires the breakdown of MTs following the old orientation as well as amplification of MTs following the new orientation. The latter can be achieved if part of the novel plus ends start in the growing rather than the shrinking state, turning a suppression of the deviant orientation into an amplification [23].

4.9 Outlook

The CA is a very important structure that bridges developmental signaling and the mechanical execution of growth responses in plants. Modelling studies have played a key role in understanding how the CA self-organizes and will continue to play an important role in understanding CA behaviour. The relative simplicity of the system allows for a high degree of similarity between models and experimental observations, making it possible to quantitatively assess the impact of experimental findings. Currently pressing issues about CA behaviour include its reorientation in response to various internal and external cues. There might be different mechanisms at play in the different reorientation processes, although it seems that katanin often plays a key role [23, 29]. Here, progress in modelling will strongly depend on additional mechanistic insights from cellular observations.

In addition, CA modelling studies so far have focused on homogeneous arrays. For several critical functions and cell types, the array must form inhomogeneous

patterns. Striking examples are the condensation of the array into the pre-prophase band prior to cell division [36], and the range of patterns found in xylem—where complex CA patterns translate to structured cell walls along a trade-off between element extensibility and resistance against vessel collapse [37].

References

1. Cosgrove DJ (2005) Nat Rev Mol Cell Biol 6(11):850. https://doi.org/10.1038/nrm1746. http://www.nature.com/doifinder/10.1038/nrm1746
2. Desai A, Mitchison TJ (1997) Annu Rev Cell Dev Biol 13(1):83. https://doi.org/10.1146/annurev.cellbio.13.1.83. http://www.annualreviews.org/doi/10.1146/annurev.cellbio.13.1.83
3. Mandelkow E, Mandelkow EM (1995) Curr Opin Cell Biol 7(1):72. https://doi.org/10.1016/0955-0674(95)80047-6. http://linkinghub.elsevier.com/retrieve/pii/0955067495800476
4. Ehrhardt DW, Shaw SL (2006) Annu Rev Plant Biol 57(1):859. https://doi.org/10.1146/annurev.arplant.57.032905.105329. http://www.annualreviews.org/doi/10.1146/annurev.arplant.57.032905.105329
5. Mineyuki Y (1999), pp. 1–49. https://doi.org/10.1016/S0074-7696(08)62415-8. http://linkinghub.elsevier.com/retrieve/pii/S0074769608624158
6. Smith LG (2001) Nat Rev Mol Cell Biol 2(1):33. https://doi.org/10.1038/35048050. http://www.nature.com/doifinder/10.1038/35048050
7. Zhang H, Dawe RK (2011) Chromosom Res 19(3):335. https://doi.org/10.1007/s10577-011-9190-y. http://link.springer.com/10.1007/s10577-011-9190-y
8. Bornens M (2002) Curr Opin Cell Biol 14(1):25. https://doi.org/10.1016/S0955-0674(01)00290-3. http://www.sciencedirect.com/science/article/pii/S0955067401002903
9. Ehrhardt DW (2008) Curr Opin Cell Biol 20(1):107. https://doi.org/10.1016/j.ceb.2007.12.004. http://www.sciencedirect.com/science/article/pii/S0955067407001937
10. Isaeva VV (2012) Biol Bull 39(2):110. https://doi.org/10.1134/S1062359012020069. http://link.springer.com/10.1134/S1062359012020069
11. Grzybowski BA, Wilmer CE, Kim J, Browne KP, Bishop KJM (2009) Soft Matter 5(6):1110. https://doi.org/10.1039/b819321p. http://xlink.rsc.org/?DOI=b819321p
12. Howard J, Hyman AA (2003) Nature 422:753–758. http://dx.doi.org/10.1038/nature01600.
13. Tindemans SH, Deinum EE, Lindeboom JJ, Mulder B (2014) Front Physiol 2(19):9. https://doi.org/10.3389/fphy.2014.00019. http://www.frontiersin.org/biophysics/10.3389/fphy.2014.00019/abstract
14. Chan J, Sambade A, Calder G, Lloyd C (2009) Plant Cell 12(8):2298. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?cmd=prlinks&dbfrom=pubmed&retmode=ref&id=19706794
15. Nakamura M, Ehrhardt DW, Hashimoto T (2010) Nat Cell Biol 12(11):1064
16. Dixit R, Cyr R (2004) Plant Cell Online 16(12):3274
17. Tindemans SH, Hawkins RJ, Mulder BM (2010) Phys Rev Lett 104(5):058103. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?cmd=prlinks&dbfrom=pubmed&retmode=ref&id=20366797
18. Hawkins RJ, Tindemans SH, Mulder BM, Phys Rev E Stat Nonlinear Soft Matter Phys (2010) 82(1 Pt 1):011911
19. Tindemans SH, Mulder BM (2010) Phys Rev E Stat Nonlinear Soft Matter Phys 81(3 Pt 1):031910. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?cmd=prlinks&dbfrom=pubmed&retmode=ref&id=20365773
20. Allard JF, Wasteneys GO, Cytrynbaum EN (2010) Mol Biol Cell 21(2):278. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?cmd=prlinks&dbfrom=pubmed&retmode=ref&id=19910489

21. Eren EC, Dixit R, Gautam N (2010) Mol Biol Cell 21(15):2674. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?cmd=prlinks&dbfrom=pubmed&retmode=ref&id=20519434
22. Deinum E, Tindemans S, Mulder B (2011) Phys Biol 8(5):056002. https://doi.org/10.1088/1478-3975/8/5/056002. http://dx.doi.org/10.1088/1478-3975/8/5/056002
23. Lindeboom JJ, Nakamura M, Hibbel A, Shundyak K, Gutierrez R, Ketelaar T, Emons AMC, Mulder BM, Kirik V, Ehrhardt DW (2013) Science 342(6163):1245533. https://doi.org/10.1126/science.1245533. http://dx.doi.org/10.1126/science.1245533
24. Vineyard L, Elliott A, Dhingra S, Lucas JR, Shaw SL (2013) Plant Cell 25(2):662. https://doi.org/10.1105/tpc.112.107326. http://dx.doi.org/10.1105/tpc.112.107326
25. Hervieux N, Dumond M, Sapala A, Routier-Kierzkowska AL, Kierzkowski D, Roeder AH, Smith RS, Boudaoud A, Hamant O (2016) Curr Biol 26(8):1019
26. Hamant O, Heisler MG, Jonsson H, Krupinski P, Uyttewaal M, Bokov P, Corson F, Sahlin P, Boudaoud A, Meyerowitz EM, Couder Y, Traas J (2008) Science 322(5908):1650. https://doi.org/10.1126/science.1165594. http://www.sciencemag.org/cgi/content/abstract/322/5908/1650
27. Uyttewaal M, Burian A, Alim K, Landrein B, Borowska-Wykrt D, Dedieu A, Peaucelle A, Ludynia M, Traas J, Boudaoud A, Kwiatkowska D, Hamant O (2012) Cell 149(2):439. https://doi.org/10.1016/j.cell.2012.02.048. http://dx.doi.org/10.1016/j.cell.2012.02.048
28. Heisler MG, Hamant O, Krupinski P, Uyttewaal M, Ohno C, Jönsson H, Traas J, Meyerowitz EM (2010) PLoS Biol 8(10):e1000516. https://doi.org/10.1371/journal.pbio.1000516. http://dx.doi.org/10.1371/journal.pbio.1000516
29. Sampathkumar A, Krupinski P, Wightman R, Milani P, Berquand A, Boudaoud A, Hamant O, Jönsson H, Meyerowitz EM (2014) Elife 3:e01967. https://doi.org/10.7554/eLife.01967. http://dx.doi.org/10.7554/eLife.01967
30. Lindeboom JJ, Lioutas A, Deinum EE, Tindemans SH, Ehrhardt DW, Emons AMC, Vos JW, Mulder BM (2013) Plant Physiol 161(3):1189. https://doi.org/10.1104/pp.112.204057. http://dx.doi.org/10.1104/pp.112.204057
31. Sethna JP (1992). In: Nagel L, Stein D (eds) 1991 lectures in complex systems. Santa Fe Institute studies in sciences of complexity, vol 15. Addison-Wesley, Reading
32. Ambrose C, Allard JF, Cytrynbaum EN, Wasteneys GO (2011) Nat Commun 2:430. https://doi.org/10.1038/ncomms1444. http://dx.doi.org/10.1038/ncomms1444
33. Deinum EE (2013) Simple models for complex questions on plant development. Ph.D. thesis
34. Zhang Q, Fishel E, Bertroche T, Dixit R (2013) Curr Biol 23(21):2191. https://doi.org/10.1016/j.cub.2013.09.018. http://dx.doi.org/10.1016/j.cub.2013.09.018
35. Deinum EE, Tindemans SH, Lindeboom JJ, Mulder BM (2017) Proc Natl Acad Sci. https://doi.org/10.1073/pnas.1702650114. http://www.pnas.org/content/early/2017/06/15/1702650114.abstract
36. de Keijzer J, Mulder BM, Janson ME (2014) Syst Synth Biol 8:187. https://doi.org/10.1007/s11693-014-9142-x.
37. Oda Y, Fukuda H (2012) Science 337(6100):1333. https://doi.org/10.1126/science.1222597. http://dx.doi.org/10.1126/science.1222597

Chapter 5
Bridging Scales from Protein Function to Whole-Plant Water Relations with the OnGuard Platform

Maria Papanatsiou, Adrian Hills, and Michael R. Blatt

Abstract The characteristics of transport across plant membranes, like all eukaryotic membranes, is highly non-linear. The complexity inherent to such characteristics defies intuitive understanding and, in these circumstances, quantitative mathematical modelling is essential as a tool, both to integrate the detailed knowledge of individual transporters and to extract the properties emerging from their interactions. As the first, fully-integrated and quantitative modelling environment for the study of ion transport dynamics in a plant cell, the OnGuard platform offers a unique tool for examining such emergent properties associated with guard cell metabolism and ion transport at the plasma membrane and tonoplast. The OnGuard platform has already yielded details guiding phenotypic and mutational studies. These advances represent key steps towards 'reverse-engineering' of stomatal physiology to improve water use efficiency and carbon assimilation, based on rational design and testing in simulation. The newly expanded platform, OnGuard2, bridges the micro-macro gap in stomatal models, coupling whole-plant transpiration to the molecular functionalities of the guard cell. Here we set out guidelines for use of OnGuard2 and outline a standardized approach that will enable users to advance quickly in applying the platform in classroom and laboratory situations.

5.1 Introduction

Though it often goes unacknowledged, the network of interactions between metabolism, ion transport, and solute partitioning among cellular compartments poses a major barrier to a quantitative understanding of guard cell physiology and of stomatal regulation mediated by gas exchange of the leaf. Cellular physiology is dictated by the characteristics intrinsic to each process, whether of ion transport, organic solute synthesis, or its catabolism. Each of these processes incorporates

M. Papanatsiou · A. Hills · M. R. Blatt (✉)
Laboratory of Plant Physiology and Biophysics, University of Glasgow, Glasgow, UK
e-mail: Michael.Blatt@glasgow.ac.uk

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_5

kinetics that are highly non-linear, often with respect to multiple parameter inputs, including substrate concentrations and membrane voltage, as well as regulatory inputs that engage ligand and other post-translational controls. Such non-linearities underpin much of physiology that is seemingly counterintuitive, that is the 'emergent' behaviours of the guard cells. Thus, even without considering transcriptional and translational regulation, addressing how guard cells respond to environmental inputs, and their coupling to the macroscopic properties of the whole leaf, demands a full and quantitative accounting of the characteristics for each transport, metabolic and buffering reaction.

Understanding and predicting stomatal behaviour will inform agricultural practices and potentially anticipate plant responses to climatic changes [1]. Stomata are pores found on the leaf epidermis that form between pairs of specialized cells, the guard cells. Stomata allow the uptake of CO_2, required for photosynthesis at the expense of water loss via transpiration from the tissues within the leaf. Gas exchange is regulated by exogenous signals, such as CO_2, water availability, and light [2–14], each of which affects stomatal pore size. Changes in the size and shape of stomatal pores arise from water fluxes that are driven by the accumulation and depletion of osmotically-active ions in guard cells [10]. Therefore, guard cell solute transport and metabolism have a substantial impact on plant fitness as they feed directly into the photosynthetic and water status of the plant. To better predict and engineer plant physiology, one has to consider the changes occurring at the guard cell level and how these relate to the tissue-wide responses of the leaf. In other words, there is a need to bridge the gap between the microscopic events of cellular ion transport and metabolism to the macroscopic properties of gas exchange and its regulation in the leaf and whole plant.

5.2 Modelling Stomata

A wealth of information exists to address mechanism of guard cell physiology and stomatal function, both in relation to solute transport [10] and metabolism [15]. Many of the molecular components are known and their kinetic properties defined with quantitative detail. Additionally, their cellular interactions have been identified in many cases, establishing the associations if not the detail of their dependencies. There is also a very large body of information that relates stomatal behaviour, as expressed in gas exchange of the whole leaf, with environmental inputs of temperature, relative atmospheric humidity, and light and the consequences for photosynthetic carbon assimilation. With this information in hand, it is possible to expand our knowledge through holistic approaches that use mathematical techniques to simulate stomatal behaviour at the systems level. Effective modelling efforts rely on the interactions between component elements of the system to yield information on the system as a whole, information which may then be validated through experiment.

Of guard cells, for which stomatal function impacts directly on gas exchange and thereby on plant water relations and photosynthesis, in general two different approaches have been pursued that divide across scales. At the macroscopic scale, the first approach has been to treat stomata as discrete, phenomenological components, each stoma serving as a pathway for transpirational water loss and CO_2 uptake [16–20]. These macroscopic approaches consider stomatal gas exchange in the context of discrete conductance units that add up to yield stomatal conductance of the whole plant (g_s). These conductance units vary in quasi-linear fashion in response to environmental inputs of light, water availability, and atmospheric water vapour and CO_2. Indeed, robust models of gas exchange have been used to scale stomatal transpiration and CO_2 exchange from leaf to canopy [21–24]. However, such phenomenological models do not capture the cellular components needed to translate g_s to the molecular mechanics of the guard cell and its functioning. In short, these models lack the necessary links essential to translate the outputs through 'reverse-engineering' to the molecular processes of the guard cell, the key connections needed for predictive in silico modelling.

At the microscopic level, efforts to model stomata have focused on the sub-cellular components, notably the transporters and their associated signal cascades in the guard cells, that facilitate ion, solute and water flux for stomatal opening and closing [10, 12, 25]. Here, one approach, borrowed from the methods of logic circuit design, describes the guard cell in terms of Boolean nodes and links that connect these nodes. The power of Boolean models lies in its use as a tool to analyse networks for which there are a large number of components and possible connections between them, but little quantitative information [26–28]. Its most common application is to identify critical nodes and their relative connections to one another, effectively mapping the predominant causalities within a network. Boolean models are defined through logic gates that can only be 'on' or 'off', which simplifies analysis but precludes kinetics relationships that are essential to understanding dynamic interactions and their temporal associations. Li et al. [29] and Sun et al. [30] have applied pseudo-temporal characteristics to Boolean networks of guard cell signalling, but these outputs are disconnected from any meaningful physiological mechanisms or their kinetics.

True mechanistic models allow the kinetic properties of the individual components within the model to be encoded with parameters defining their operation as functions of the relevant physiological inputs [16, 25]. A so-called hydromechanical model [17] represented a step in this direction; it proposed a simple hyperbolic relation between the ATP concentration of the guard cells and their osmotic content. Even so, the model lacked explicit detail for the mechanics of guard cell solute transport and, thus, the intrinsic regulatory processes that are known to determine much of stomatal function [10]. It is important to note, too, that the hydromechanical model and almost all other modelling efforts have sought analytic solutions for endpoint or stationary states only. They fail to address the wealth of information available relating to the temporal kinetics for stomatal movements and transpiration.

To date, only the OnGuard platform [31–33] has encapsulated the mechanistic components and their parameters for solute transport and metabolism sufficient

to accurately simulate guard cell physiology and the stomatal movements that it engenders. The latest revision of the OnGuard platform, OnGuard2, now bridges the gap between the macroscopic characteristics of transpirational water relations in the whole plant and the microscopic behaviour of guard cells in solute and water transport that drives stomatal movements. Most important to modelling guard cells, the OnGuard platform incorporates the biophysical and kinetic features of each of the transporters that drive solute flux at the two membranes underpinning stomatal movements, the guard cell plasma membrane and tonoplast. Furthermore, it connects the activities of these transporters, as in vivo, through the essential kinetic variables of membrane voltage as well as substrate and regulatory ligand concentrations. The importance of membrane voltage, especially, lies in its role in feedback between transporters and even with respect to a single transporter.

Consider K^+ transport out of the guard cell. The guard cell outward-rectifying K^+ channel—in Arabidopsis, the GORK channel—is strongly voltage-dependent, activating only at voltages positive of the prevailing K^+ equilibrium voltage, E_K [10, 34–36]. Depolarizing the membrane activates the channel for K^+ efflux from the cell, but this activity is countered by the K^+ flux itself. As K^+ passes outward through the channel, across the plasma membrane, it carries charge to repolarize the membrane. In other words, even without any effects from other regulatory processes, the activity of the channel counteracts its own flux as determined by intrinsic kinetic relations of the channel gate.

Of course, the membrane voltage will be affected by every other transporter that moves a net charge across the membrane which, in turn, will also affect the channel-mediated K^+ flux. The consequence is that, in the steady-state, ion transport is a highly non-linear process that is not solely determined by the gating characteristics of an individual transporter such as GORK per se, but by the balance of ion and charge transport across the membrane as a whole. There are platforms available, such as the Virtual Cell [37], E-Cell [38], and the Cellerator [39] with which similar reaction-diffusion processes can be modelled. These platforms are less flexible in enabling parameter modifications during simulations, however. Most important, however, they are poorly adapted to accommodating the feedback that arises from the parallel activities of membrane transporters and the associated variable of membrane voltage.

5.3 The Elements of the OnGuard Platform

Stomatal movement arises from changes in cell volume and turgor driven by the transport, accumulation and release of osmotically-active solutes, primarily K^+, Cl^-, sucrose and the organic anion malate^{2-} (Mal) [10, 40, 41]. Such movements are ideally suited to a ‘bottom-up’ approach in mathematical modelling. They are governed by quantitative relations that describe mass and charge conservation, ion and water flux. The voltage across each membrane, in turn, is linked to the relevant ion gradients and permeabilities across the membrane. Together, these

physico-chemical relations are easy to incorporate mathematically and they constrain all homeostatic interactions within any model. For plant cells, and especially for the guard cell, the important output variables are the cell volume and osmolality, water potential and turgor, the voltages across each membrane, the predominant ion concentrations, especially K^+ and Cl^-, as well as the total and free Ca^{2+} and H^+ concentrations, and the corresponding ion fluxes through each transporter. These variables are subject to buffering for intracellular H^+ and Ca^{2+}, the osmotic contributions from impermeant solutes, mostly of protein, their charge and pH dependencies, all of which are available or can be estimated from experimental data [42–45].

For membrane transport, the biophysical relations are all well-defined and for several plant cell types, including the guard cell, have been studied in depth sufficient for quantitative mathematical description. For example, H^+ transport via ATP-driven pumps and coupled transporters as well as the transport of other ions via channels at the plasma membrane has been characterized, with detailed information on stoichiometry and mechanism [10]. Thus, their operation can be described quantitatively within sets of kinetic equations fully constrained by experimental results. Even if our knowledge of individual transporters at the tonoplast is less well developed, in these circumstances there is ample data from experiments to define the vacuolar ion contents and fluxes [10, 46–50], so constraining any modelling effort by minimizing the range of parameters needed to comply with experimental results. Essential kinetic data are available also for transport regulation, notably its control by cytosolic-free $[Ca^{2+}]$ ($[Ca^{2+}]_i$) and pH, and in many cases by reactive oxygen species (ROS), and protein phosphorylation [10, 51–53].

Of course, there are gaps in our knowledge of many of these transporters, at least their molecular identities and some details of their regulation. However, in general it is sufficient to know the kinetic relationships that describe a process, even if the structural gene products are not known or the precise mechanisms of their regulation have not been resolved. For the modeller seeking to understand how a system responds to physiological or experimental perturbation, the only relevant biology is encapsulated in how one model variable is connected to another. For example, we do not know the proportions of inward-rectifying K^+ channels that are composed of KAT1 and of KAT2 subunits in Arabidopsis, nor whether KC1 might also assemble together with these subunits to form the functional channels [54, 55]. Nevertheless, we know how the ensemble K^+ current is gated by voltage and that its amplitude depends on extracellular $[K^+]$, and we can describe these dependencies in quantitative terms [56–59]; so we are able to describe both dependencies with sufficient detail to model the behaviour of the current in vivo. This description can be expanded to incorporate the different subunits when their unique characteristics are known and these become the focus of the modelling effort. Similarly, we know that $[Ca^{2+}]_i$ inactivates the inward-rectifying K^+ channels of the guard cell [60, 61], and surmise that this may occur via phosphorylation by one or more Ca^{2+}-dependent protein kinases [53, 62–64]. Quantitative kinetic information is still lacking to model the steps between Ca^{2+} binding, the kinase cascades and their ultimate phosphorylation of the K^+ channels. Nevertheless, we know the

relationship between $[Ca^{2+}]_i$ and K^+ channel activity, and we can safely place the mechanistic details in a mathematical description that subsumes the intermediate kinetics. In effect, such a phenomenological approach introduces modules with adjustable levels of resolution that may be expanded if, and when, studies come to focus on a specific module [65].

5.4 Expanding the OnGuard Platform to Define and Model Plant Water Relations

To bridge scales from the guard cell to whole-plant water relations in OnGuard2, we integrate three additional sets of variables and parameters associated with the water relations of the leaf [14]. First, we note that water in the guard cell wall must equilibrate with water vapour in the substomatal cavity. Water vapour equilibration affects the osmotic potential of the guard cell wall and, hence, the microscopic variables of the effective ionic activities in the apoplast and the ion and water flux across the guard cell plasma membrane. We calculate the partial pressure of water vapour in the substomatal cavity from the gradient in its partial pressures between the sites of evaporation in the leaf and the atmosphere outside. Second, to accommodate water delivery to the leaf, we define water flux through the xylem with a pseudo-linear relationship that describes the evaporative surface area within the leaf in relation to the area of the stomatal pore. Finally, we introduce a finite hydraulic permeability of the guard cell plasma membrane in order to place water flux under control of relevant cellular signal cascades, notably $[Ca^{2+}]_i$ and pH [66–69]. These descriptors are sufficient to simulate the behaviour of stomata with changes in atmospheric relative humidity (%RH) and with temperature, which will also affect the rate of evaporation within the leaf. They lead to models that accurately predict the effects of transpiration on guard cell membrane transport as well as the effects of manipulating membrane transport on stomatal transpiration in the whole leaf [14].

While iterative computational modelling, such as used by the OnGuard platform, does not pre-define a final endpoint, it does require a starting point—a reference state—from which time increments may be calculated and the dynamics of the model can evolve base on the constraining physical laws, the equations and their parameters that define the components of the model and, most importantly, the interactions that arise from their functioning over time. The best place to start in any modelling effort of this kind is a state in which no net change in solute flux or content occurs, for example that of the guard cell of the closed stoma in the dark. To aid in resolving such a starting (or reference) state, the OnGuard platform incorporates a Reference State Wizard. The Wizard allows the user to specify the underlying biophysical status of the system and then query the model for solute and metabolic fluxes in total and through each of the model processes. With little practice, the user should be able to balance each of component flux, starting with

H^+ transport and concluding with sucrose and Mal metabolism, by adjusting the populations of transporters and, if necessary, their underlying kinetic descriptors in order to comply with the requirements for a starting reference state.

Obviously establishing a Reference State implies prior knowledge of the probable unit densities, or at least the typical amplitudes of each current, and the characteristic parameters defining the kinetics for each transporter. The biological validity of a model is first judged by this knowledge, and such knowledge is critical if a model is to avoid indetermination and yield true predictive power. There is no absolute rule that will ensure predictive power in a model, but with experience the user will come to recognize some basic guidelines. Critically, there must exist at least two pathways for flux of each solute species if flux balance is to be achieved across a membrane. For example, any model with a pathway for the influx of cation M^+ across the plasma membrane must include at least one other pathway for its efflux, otherwise flux balance within the model as a whole over time is not possible. Similarly, to ensure charge balance, complementary pathways for oppositely charged species must exist across each membrane. If these conditions are met, and kinetic detail is available to define quantitatively at least 80–85% of the total flux of all species in both directions and between all compartments, then parameters for the remaining fluxes generally will be constrained sufficiently to render a model with true predictive power.

Models based on Reference States and diurnal Reference Cycles for guard cells of *Vicia* [31, 32] and Arabidopsis [14, 33, 70] are available for download with the OnGuard platform (www.psrg.org.uk; note that to access full functionality users must register to obtain and install a passcode). Full descriptions of transporters and their parameter sets will be found in these publications and in Jezek and Blatt [10]. These models offer good starting points for users new to the OnGuard platform, as they circumvent the tasks of defining the reference state. Of course, each of these 'pre-packaged' models come with the standard proviso of all working systems: the models offer good approximations to experimental data within the bounds of the conditions and data used for their validation. As new experimental data becomes available for one or more of the processes within a model, corresponding refinements to the model parameters are likely to be needed. We welcome communication with users, and we are always open to suggestions for refinements and new implementations to the OnGuard platform.

5.5 Simulating Stomatal Responses to Humidity

In practice, simulating stomatal physiology and its response to experimental challenges is straightforward using the OnGuard platform. A 20-min video is available for viewing at the software download site. The video introduces the basic operation of the platform and describes the first-generation OnGuard software. Nonetheless, it covers material relevant to OnGuard2, including how to access and adjust user controls for iteration and sampling frequencies, how to modify transport

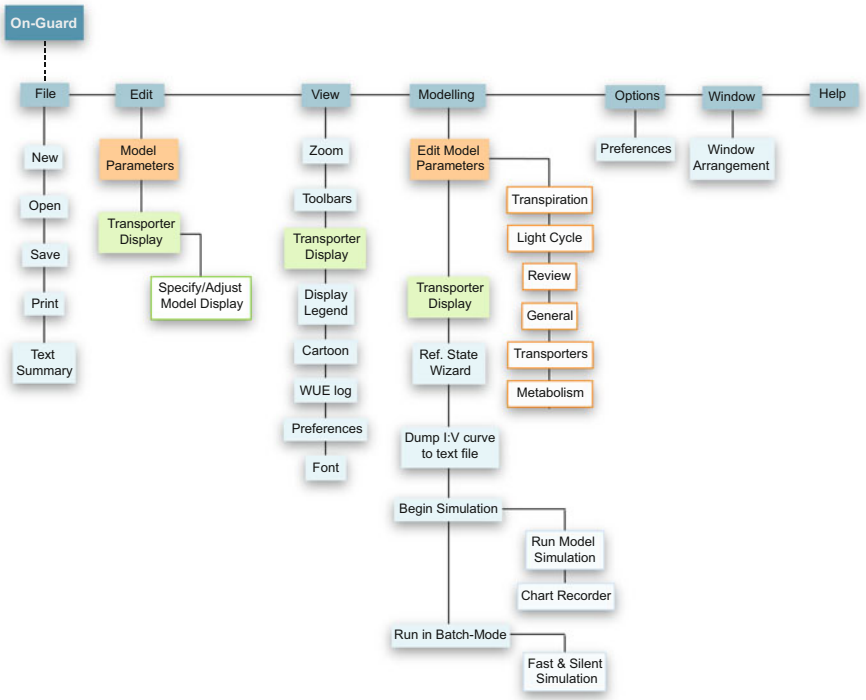


Fig. 5.1 A software map of the OnGuard platform highlighting the general platform options and logic flow. Once a model (*.ogb) file is loaded, key operations are to edit model parameters and transporter display options (orange and green boxes), accessed through the 'Edit' and 'Modelling' dropdowns, and to run the model ('Begin simulation' and 'Run in batch mode'), accessed through the 'Modelling' dropdown. Platform runtime 'Preferences' are accessed through both the 'View' and 'Options' dropdowns. The 'Text Summary' option exports the full set of model parameters specified in a model file. The 'Dump I:V curve to text file' option exports all current-voltage curves visible in the main window currently active at the time. The current-voltage curves export in a format compatible with all spreadsheet software. The 'Chart Recorder' window is also available with the 'Run in batch mode' option, but the user has access to tab through the displays only after exiting from a batch operation cycle. A 'Cartoon' display is available from the 'View' dropdown for real-time visualization of ion transport across the plasma membrane and tonoplast as well as the observation of dynamic changes in physiological outputs such as stomatal aperture and conductance. The 'Cartoon' display does not operate with the batch operation mode

and metabolic parameters, how to set ion and solute concentrations, and how to interrogate and log platform outputs. OnGuard2 has expanded on these controls, as noted above, to incorporate parameters defining aquaporins at the guard cell plasma membrane, xylem water feed to the leaf, stoma and leaf geometries, and ambient temperature and atmospheric humidity relevant to transpiration. Figure 5.1 summarizes OnGuard2 in a platform map with logic flow. We focus here on the newer elements of OnGuard2 and direct the reader to a previous guide on the use of OnGuard for examples of manipulations relevant to the guard cell in isolation [71].

Before starting OnGuard2 for the first time, select 'View' in the main window and choose 'Preferences' in the drop-down list. Access the 'Run-Time Limits' tab and check that the 'Minimum T-inc' and 'Maximum T-inc' values are set to 0.001 and 20 s, respectively. We recommend that the 'Max %age change' is set to 2 or 3%, the 'Max Iterns' is set to 1000, the 'Default dV' is 1×10^{-9}, and finally the 'I tolerance' is set to 1×10^{-21}. These parameters determine the variable time interval limits, the range of permissible variable changes per time interval, and the tolerances for free-running voltage and current estimates. Additionally, you can adjust the 'O/P Efficiency' settings for the best compromise between speed and quality in graphical display: increasing the number of points plotted in each curve adds to the display burden on the processor; increasing the 'Wait-sleep time' slows the computational cycle and allows time for display functions. You need only add to the 'Wait-sleep time' if values in the flux window fail to update regularly during a run. Note that 'O/P Efficiency' and 'Wait-sleep time' have no effect on platform operation when a model is run in the much faster, batch mode (see below).

Once a model is loaded, you have access to 'Edit Model Parameters' from the 'Modelling' tab in the main window. The 'Edit Model Parameters' gives you access to define parameters for the various ion concentrations, transporters, sugar and malate metabolism, and their regulatory characteristics. You can even add new transporters and remove existing ones. For now, we recommend using the preset parameter settings that come with the published models. Later, when you are comfortable running the software and interrogating its outputs, you may want to explore the effects of eliminating one or more transporters, such as has been described before [71], and we describe below how to examine the effects of changes in atmospheric relative humidity.

From the 'Modelling' tab, you also have access to 'Begin Simulation', and to 'Run in Batch-Mode'. The second of these options will generate outputs rapidly, but does not provide the depth of display and control during simulations. You may find this rapid simulation mode useful once you become familiar with the platform, and we recommend you explore its operation only then. Use the 'Begin Simulation' option to open the chart recorder, with tabs for many of the major outputs of the model which can be followed during a simulation, and to open a flux window which reports details of the various ion fluxes and membrane voltages (Fig. 5.2). The flux window provides run-time controls, including those for spreadsheet data logging. Within the flux window, we recommend activating the 'Auto-increment' tickbox at the lower left, and setting the 'Min Log Interval' to 20 s for a suitable temporal resolution in the data logged for spreadsheet access and in the chart recorder. For comparative purposes, run one of the models supplied with the software through three diurnal cycles (72 h) as a control by activating the 'Stop at Time' tick box and entering 0003:00:00:00 (days:hours:minutes:seconds) in the time window below and then clicking on 'Run'. You may minimize windows, but do not close any windows, as this will terminate the simulation and close the data log.

After this first 3-day period is complete, use the 'Modelling' drop-down menu at the top of the main window to 'Edit Model Parameters' and then select the 'Transpiration' tab to access the plant water relations (Fig. 5.3). Ensure the radio

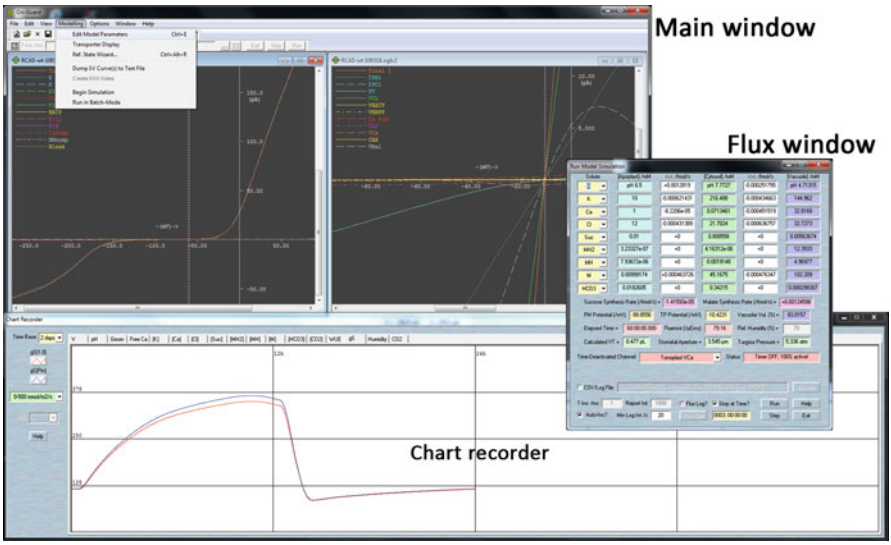


Fig. 5.2 OnGuard2 with the Modelling dropdown activated to show access for editing model parameters and running models. Also visible are the main and flux windows, and the chart recorder. Within the main window, two graphic windows are visible showing the current-voltage curves all of the transporters at the plasma membrane (left) and the tonoplast (right)

button for 'Use Ws/Wp to calculate Ciso' is selected. Here you have a readout of the external relative humidity and the calculated internal relative humidity. You can set the relative water feed (RWF, see Wang et al. [14]), defined as the ratio of the evaporative surface inside the leaf to the cross-sectional area of the open stomatal pore. Reducing the RWF leads to stomata that are more sensitive to changes in external relative humidity and simulates water stress conditions. Set the RWF to 5 or 10, then Click on the 'Edit Humidity Protocol' button to call up an editable, 24-h protocol. A window displaying the external relative humidity opens with hooks and connectors that can be moved, added, or removed to set the relative humidity time sequence. Use these hooks to create a protocol with a step to 30%RH for a period of 2 h during the daylight period (Fig. 5.3). Also shown on the 'Transpiration' tab are settings to define the leaf geometry, including the stomatal pore length, depth and density, as well as the mesophyll depth and percentage of air space within the leaf. These latter parameters are editable but are species-specific and act as scalars in the calculations of stomatal conductance and foliar transpiration.

Again, we recommend that OnGuard2 run with these new settings for three diurnal cycles. For this purpose, simply ensure the 'Stop at Time' tick box is active and enter 0006:00:00:00 (days:hours:minutes:seconds) in the time window below. Once you are finished, terminate the simulation by clicking the 'Stop' button, if still running, then close the flux window. Closing the flux window closes the *.csv file log. If you wish to save your model with any new parameter settings that have been entered, first call up 'Edit Model Parameters' in the 'Modelling' tab of the main

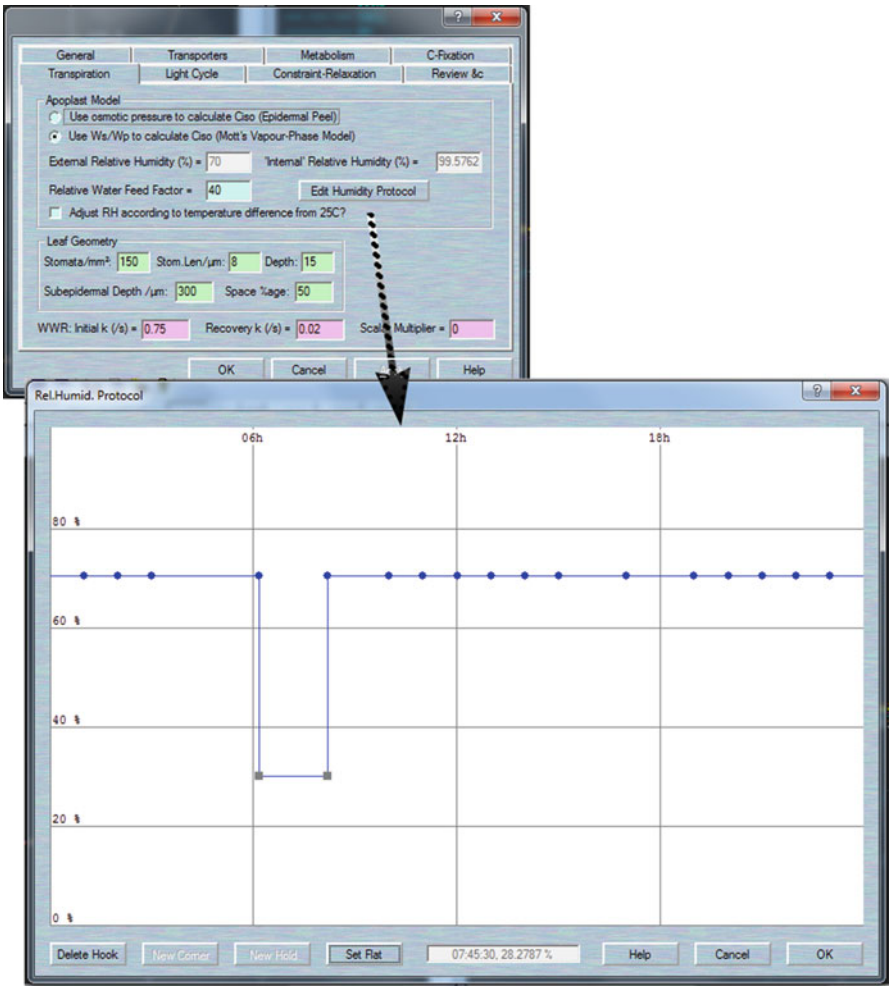


Fig. 5.3 OnGuard2 popup window (above) with tabbed access to all of the model parameters. Shown active is the tab for Transpiration control with the button activating (arrow) the Relative Humidity Protocol window (below) as described in the text

window, choose the 'Review &c' tab, and reset the 'Time of day (hh:mm:ss:ms)' to 00:00:00:00 to reset the start time of the simulation, then use the 'File' tab in the main window to access the 'Save as . . . ' option and give the model (*.ogb) file a new name to avoid overwriting the original model. Note that you can continue running the model at any time, also without saving a file, but if you have reset the start time you must close the flux and chart recorder windows and then re-open them by clicking on the 'Modelling' tab and selecting 'Begin Simulation'.

OnGuard2 normally logs the apoplast, cytosol, and vacuole contents and the net fluxes across the plasma membrane and tonoplast for each ion and solute. It will also

log the fluxes through the individual transporters, $[Ca^{2+}]_i$, cytosolic and vacuolar pH, the rates of sucrose and malic acid synthesis, the membrane voltages, and a number of derived variables such as stomatal conductance and transpiration rate. Many of these variables are displayed during simulations in one of the tabs of the on-screen chart recorder. You may choose to log the fluxes of each ionic species by activating the 'Flux Log?' tick box and selecting one or more of the outputs from the 'Flux Opts' button. These data are saved in *.csv format readable by almost all spreadsheet programs. For a first review, you may plot the outputs using Microsoft Exel, but for a more detailed review of platform outputs, we recommend using a publication-quality graphics package such as SigmaPlot (Systat).

5.6 Interrogating OnGuard Platform Outputs

Interpreting the output of an OnGuard2 simulation requires interrogating the model variables. Changes in each of these variables—for example, solute concentration, the associated rates of ion and solute flux through each of the transporters, the membrane voltages, cytosolic-free $[Ca^{2+}]$ and pH—depend on interactions between the transporters, cytosolic and vacuolar buffering for pH and $[Ca^{2+}]_i$, and metabolism, just as they do in vivo. So, in general, the task of interpretation reduces to one of tracing the sequence of events following a trigger, or change in a specific model parameter, and then tracing the consequences through the network of interrelated homeostatic processes. Here, the output variables, their kinetics, flux, and metabolic origins are most useful and will help in identifying emergent behaviours of the system that are often the most informative aspects of any modelling effort.

A review of the OnGuard2 output in this case shows that it predicts stomatal aperture and g_s changes that reproduce experiments as reported in the literature (Fig. 5.4; see also Wang, et al. [14]), with step decreases in %RH outside reducing aperture and g_s. As Wang et al. [14] demonstrated, when RWF is set below 40, the stomatal sensitivity to %RH is strongly enhanced much as it is in vivo when soil water content is reduced. The OnGuard2 outputs, both the online chart recorder and the logged data, show that decreasing aperture is accompanied by an increase in osmotic solute concentration as water is drawn from the guard cells and their volume decreases. The concentrations of K^+, Cl^- and other solutes rises in proportion but, after an initial rise, cytosolic and vacuolar Cl^- decline to values near those in the absence of the %RH step. Other features of the response include oscillations in $[Ca^{2+}]_I$, which are associated with stomatal closure and facilitate solute efflux [60, 70] as well as an alkaline shift in cytosolic pH, both of which are known to affect many of the transporters at both the plasma membrane and tonoplast.

How are the connections made between transport in the guard cell and transpiration in the whole leaf? The key to the effects on transport with the step decrease in %RH outside is reduction in the partial pressure of water vapour in the substomatal cavity. Over subsequent time increments, the rate of transpiration through the pore

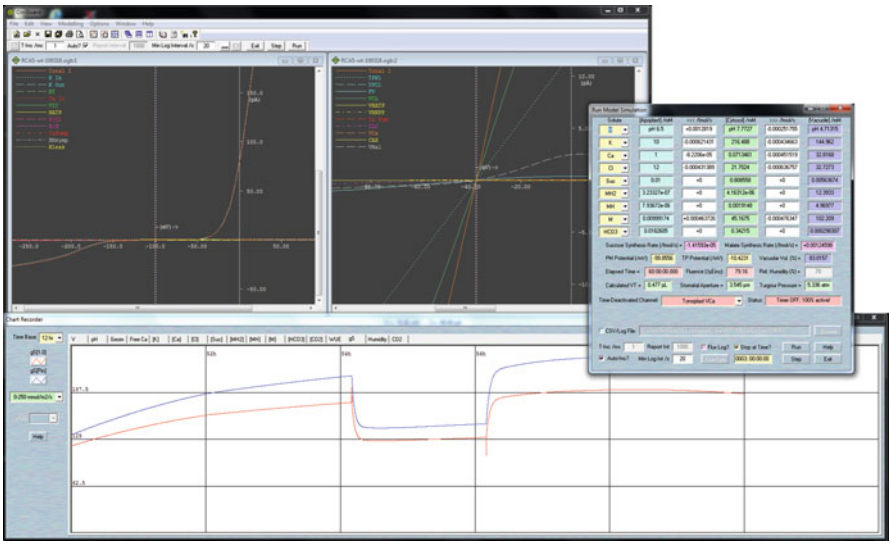


Fig. 5.4 OnGuard2 simulation windows with the Chart Recorder tab for stomatal conductance (g_s) active and showing the result of the 2-h step to 30%RH programmed Fig. 5.3. Note that the red trace for g_s uses the driving force for water vapour diffusion assuming 100%RH in the leaf; the blue trace shows g_s calculated assuming a driving force across the stomatal pore alone with a reduction in water vapour pressure just inside the substomatal cavity resulting from transpiration to the outer atmosphere (see Wang et al. [14])

rises with the increase in the driving force for water vapour diffusion. This decline in substomatal water vapour leads to an increase in osmotic potential of the guard cell wall and, consequently, a decrease in total guard cell volume as water is drawn out of the guard cell. As the guard cell volume decreases, there is a corresponding increase in solute concentration, including that of $[Ca^{2+}]_i$, which leads to $[Ca^{2+}]_i$ oscillations as the change in the concentration of Ca^{2+} in the vacuole enhances the driving force for Ca^{2+} influx across the tonoplast [14].

Just focusing on Ca^{2+} for now, the rise in $[Ca^{2+}]_i$ will act on all fluxes mediated by $[Ca^{2+}]_i$-sensitive transporters—some 70% of the transport activities at the two membranes [72]—including the several pumps, K^+, Cl^- and Mal channels at both the plasma membrane and tonoplast, and, consequently, on the total osmotic solute content of the guard cell. The rise in $[Ca^{2+}]_i$ also affects water flux through its regulation of water transport [14]. Finally, as $[Ca^{2+}]_i$ reduces the rate of water flux, it also affects the rate at which the stoma closes and, hence, stomatal conductance, transpiration, and the partial pressure of water vapour within the leaf.

Most important, the connections above complete the circle with feedback between macro- and microscopic processes: The macroscopic effects of external %RH, transpiration and changes in substomatal water vapour influence the microscopic processes of ion and water flux, and the guard cell volume and osmotic solute content. In turn, guard cell volume and osmotic content determine the macroscopic

variables of stomatal aperture and conductance, the rate of transpiration and, hence, the water vapour pressure within the air space internal to the leaf. Of course, all of the fluxes are rooted in time, so the system evolves governed only by the equations and parameters that define the ensemble of underlying processes, microscopic and macroscopic. In short, the computational approach of the OnGuard platform bridges scales through an evolution of the temporal behaviours encapsulated within the ensemble of flux equations, their relationships to osmotic water and solute content, and the exchange between water in the cell wall and the partial pressure of water vapour within the leaf.

Of course, no model is useful unless it has the capacity to yield predictions that can be tested experimentally, not simply to reproduce known behaviours. There are a number of predictions that can be drawn from OnGuard2 simulations, several of which have been validated experimentally as outlined by Wang et al. [14]. Of these, the OnGuard2 Arabidopsis model predicted a substantial retention of Mal relative to Cl^-, consistent with observations that Mal is retained under osmotic stress [73, 74]. It also predicted transient enhancements in the activity of the inward-rectifying K^+ channels that arose from suppressions in $[Ca^{2+}]_i$ on recovery from %RH steps, a slowing in the rates of stomatal closure and of g_s decreases in the *slac1* Cl^- channel and *ost2* H^+-ATPase mutants, an acceleration in aperture and g_s recovery in the *slac1* mutant following %RH steps, and a slowing in these rates for the *ost2* mutant. Each of these predictions was confirmed in vivo. Other predictions still remain to be tested experimentally, notably the effects of changes in %RH on anion channel and Ca^{2+} activities as well as selected transport across the tonoplast. You may well find some other outputs that are worth pursuing through experiments; we have no doubt that there are further revelations to come.

5.7 Conclusion and Outlook

OnGuard2 offers users an unprecedented tool with which to explore the mechanics of stomatal transpiration across scales from the molecule to the whole plant. Until now, research into stomatal physiology and efforts to model stomatal behaviours have been divided across scales. The mechanics of guard cell membrane transport and metabolism are defined by a wealth of knowledge, including the identities of the key transporters, their biophysical properties and regulation, and have been modelled successfully with quantitative kinetic detail at the cellular level. By contrast, foliar transpiration has been described empirically through quasi-linear relations with atmospheric humidity, CO_2, and light, but without connection to guard cell mechanics. OnGuard2 bridges this historic and conceptual gap seamlessly with a systems platform that defines foliar water relations and transpiration in the context of the molecular mechanics of guard cell ion transport, metabolism and signalling. To date, it has been shown to reproduce faithfully the kinetics of whole plant transpiration with stomatal conductance, its dependence on VPD, and

water feed to the leaf as integral components of OnGuard2 models. We encourage researchers and educators alike to adopt OnGuard2 for their own applications as they relate to guard cell homeostasis, stomatal dynamics and transpiration from plants.

Acknowledgements MP, AH and MRB, and their associated publications, were supported by BBSRC grants BB/L001276/1, BB/L019025/1, BB/M001601/1, and BB/N01832X/1.

References

1. Hetherington AM, Woodward FI (2003) The role of stomata in sensing and driving environmental change. Nature 424:901–908
2. Allaway WG (1973) Accumulation of malate in guard cells of Vicia faba during stomatal opening. Planta 110:63–70
3. Allen GJ, Murata Y, Chu SP, Nafisi M, Schroeder JI (2002) Hypersensitivity of abscisic acid-induced cytosolic calcium increases in the Arabidopsis farnesyltransferase mutant era1-2. Plant Cell 14:1649–1662
4. Assmann SM (1999) The cellular basis of guard cell sensing of rising CO_2. Plant Cell Environ 22:629–637
5. Assmann SM, Jegla T (2016) Guard cell sensory systems: recent insights on stomatal responses to light, abscisic acid, and CO_2. Curr Opin Plant Biol 33:157–167
6. Blatt MR (1990) Potassium channel currents in intact stomatal guard cells: rapid enhancement by abscisic acid. Planta 180:445–455
7. Blatt MR (2000) Cellular signaling and volume control in stomatal movements in plants. Annu Rev Cell Dev Biol 16:221–241
8. Blatt MR, Clint GM (1989) Mechanisms of fusicoccin action kinetic modification and inactivation of potassium channels in guard cells. Planta 178:509–523
9. Hetherington AM (2001) Guard cell signaling. Cell 107:711–714
10. Jezek M, Blatt MR (2017) The membrane transport system of the guard cell and its integration for stomatal dynamics. Plant Physiol 174:487–519
11. Kim TH, Bohmer M, Hu HH, Nishimura N, Schroeder JI (2010) Guard cell signal transduction network: advances in understanding Abscisic acid, CO_2, and Ca^{2+} signaling. Annu Rev Plant Biol 61:561–591
12. Lawson T, Blatt MR (2014) Stomatal size, speed, and responsiveness impact on photosynthesis and water use efficiency. Plant Physiol 164:1556–1570
13. Shimazaki KI, Doi M, Assmann SM, Kinoshita T (2007) Light regulation of stomatal movement. Annu Rev Plant Biol 58:219–247
14. Wang Y, Hills A, Vialet-Chabrand SR, Papanatsiou M, Griffiths H, Rogers S, Lawson T, Lew V, Blatt MR (2017) Unexpected connections between humidity and ion transport discovered using a model to bridge guard cell-to-leaf scales. Plant Cell 29:2921–2139
15. Santelia D, Lawson T (2016) Rethinking guard cell metabolism. Plant Physiol 172:1371–1392
16. Buckley TN (2017) Modeling stomatal conductance. Plant Physiol 174:572–582
17. Buckley TN, Mott KA, Farquhar GD (2003) A hydromechanical and biochemical model of stomatal conductance. Plant Cell Environ 26:1767–1785
18. Farquhar GD, Sharkey TD (1982) Stomatal conductance and photosynthesis. Annu Rev Plant Physiol Plant Mol Biol 33:317–345
19. Farquhar GD, von Caemmerer S, Berry JA (2001) Models of photosynthesis. Plant Physiol 125:42–45
20. McAdam SAM, Brodribb TJ (2016) Linking turgor with ABA biosynthesis: implications for stomatal responses to vapor pressure deficit across land plants. Plant Physiol 171:2008–2016

21. Ball JT, Woodrow IE, Berry JA (1987) A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In: Biggens J (ed) Progress in photosynthesis research. Martinus-Nijhoff, Dordrecht, pp 221–224
22. Pieruschka R, Huber G, Berry JA (2010) Control of transpiration by radiation. Proc Natl Acad Sci U S A 107:13372–13377
23. Vialet-Chabrand S, Matthews JSA, Brendel O, Blatt MR, Wang Y, Hills A, Griffiths H, Rogers S, Lawson T (2016) Modelling water use efficiency in a dynamic environment: an example using Arabidopsis thaliana. Plant Sci 251:65–74
24. Von Caemmerer S, Farquhar GD (1981) Some relationships between the biochemistry of photosynthesis and the gas exchange of leaves. Planta 153:376–387
25. Buckley TN, Mott KA (2013) Modelling stomatal conductance in response to environmental factors. Plant Cell Environ 36:1691–1699
26. Laschov D, Margaliot M (2011) A maximum principle for single-input Boolean control networks. IEEE Trans Autom Control 56:913–917
27. Pawlak Z, Skowron A (2007) Rudiments of rough sets. Inf Sci 177:3–27
28. Siuti P, Yazbek J, Lu TK (2013) Synthetic circuits integrating logic and memory in living cells. Nat Biotechnol 31:448–452
29. Li S, Assmann SM, Albert R (2006) Predicting essential components of signal transduction networks: a dynamic model of guard cell abscisic acid signaling. PLoS Biol 4:1732–1748
30. Sun Z, Jin X, Albert R, Assmann SM (2014) Multi-level modeling of light-induced Stomatal opening offers new insights into its regulation by drought. PLoS Comput Biol 10:e1003930
31. Chen ZH, Hills A, Baetz U, Amtmann A, Lew VL, Blatt MR (2012) Systems dynamic modeling of the stomatal guard cell predicts emergent behaviors in transport, signaling, and volume control. Plant Physiol 159:1235–1251
32. Hills A, Chen ZH, Amtmann A, Blatt MR, Lew VL (2012) OnGuard, a computational platform for quantitative kinetic modeling of guard cell physiology. Plant Physiol 159:1026–1042
33. Wang Y, Papanatsiou M, Eisenach C, Karnik R, Williams M, Hills A, Lew VL, Blatt MR (2012) Systems dynamic modelling of a guard cell Cl^- channel mutant uncovers an emergent homeostatic network regulating stomatal transpiration. Plant Physiol 160:1956–1972
34. Blatt MR (1988) Potassium-dependent bipolar gating of potassium channels in guard cells. J Membr Biol 102:235–246
35. Blatt MR, Gradmann D (1997) K^+ −sensitive gating of the K^+ outward rectifier in Vicia guard cells. J Membr Biol 158:241–256
36. Hosy E, Vavasseur A, Mouline K, Dreyer I, Gaymard F, Poree F, Boucherez J, Lebaudy A, Bouchez D, Very AA, Simonneau T, Thibaud JB, Sentenac H (2003) The Arabidopsis outward K^+ channel GORK is involved in regulation of stomatal movements and plant transpiration. Proc Natl Acad Sci U S A 100:5549–5554
37. Loew LM, Schaff JC (2001) The virtual cell: a software environment for computational cell biology. Trends Biotechnol 19:401–406
38. Tomita M, Hashimoto K, Takahashi K, Shimizu TS, Matsuzaki Y, Miyoshi F, Saito K, Tanida S, Yugi K, Venter JC, Hutchison CA (1999) E-CELL: software environment for whole-cell simulation. Bioinformatics 15:72–84
39. Shapiro BE, Levchenko A, Meyerowitz EM, Wold BJ, Mjolsness ED (2003) Cellerator: extending a computer algebra system to include biochemical arrows for signal transduction simulations. Bioinformatics 19:677–678
40. McAinsh MR, Pittman JK (2009) Shaping the calcium signature. New Phytol 181:275–294
41. Willmer C, Fricker MD (1996) Stomata. Chapman and Hall, London, pp 1–375
42. Grabov A, Blatt MR (1997) Parallel control of the inward-rectifier K^+ channel by cytosolic-free Ca^{2+} and pH in Vicia guard cells. Planta 201:84–95
43. Tiffert T, Lew VL (1997) Cytoplasmic calcium buffers in intact human red cells. J Physiol 500:139–154
44. Tsien RW, Tsien RY (1990) Calcium channels, stores and oscillations. Annu Rev Cell Biol 6:715–760

45. Wang Y, Blatt MR (2011) Anion channel sensitivity to cytosolic organic acids implicates a central role for oxaloacetate in integrating ion flux with metabolism in stomatal guard cells. Biochem J 439:161–170
46. Gobert A, Isayenkov S, Voelker C, Czempinski K, Maathuis FJM (2007) The two-pore channel TPK1 gene encodes the vacuolar K + conductance and plays a role in K + homeostasis. Proc Natl Acad Sci U S A 104:10726–10731
47. MacRobbie EAC (1995) Effects of ABA on 86 Rb + fluxes at plasmalemma and tonoplast of stomatal guard cells. Plant J 7:835–843
48. MacRobbie EAC (2000) ABA activates multiple Ca 2+ fluxes in stomatal guard cells, triggering vacuolar K + (Rb +) release. Proc Natl Acad Sci U S A 97:12361–12368
49. MacRobbie EAC (2002) Evidence for a role for protein tyrosine phosphatase in the control of ion release from the guard cell vacuole in stomatal closure. Proc Natl Acad Sci U S A 99:11963–11968
50. MacRobbie EAC (2006) Osmotic effects on vacuolar ion release in guard cells. Proc Natl Acad Sci U S A 103:1135–1140
51. Blatt MR, Garcia-Mata C, Sokolovski S (2007) Membrane transport and Ca^{2+} oscillations in guard cells. In: Mancuso S, Shabala S (eds) Rhythms in plants. Springer, Berlin, pp 115–134
52. Wang PT, Song CP (2008) Guard-cell signalling for hydrogen peroxide and abscisic acid. New Phytol 178:703–718
53. Zou JJ, Li XD, Ratnasekera D, Wang C, Liu WX, Song LF, Zhang WZ, Wu WH (2015) Arabidopsis Calcium-Dependent Protein KINASE8 and CATALASE3 function in Abscisic acid-mediated signaling and H2O2 homeostasis in stomatal guard cells under drought stress. Plant Cell 27:1445–1460
54. Jeanguenin L, Alcon C, Duby G, Boeglin M, Cherel I, Gaillard I, Zimmermann S, Sentenac H, Very A-A (2011) AtKC1 is a general modulator of Arabidopsis inward shaker channel activity. Plant J 67:570–582
55. Pilot G, Lacombe B, Gaymard F, Cherel I, Boucherez J, Thibaud JB, Sentenac H (2001) Guard cell inward K + channel activity in Arabidopsis involves expression of the twin channel subunits KAT1 and KAT2. J Biol Chem 276:3215–3221
56. Blatt MR (1992) K^+ channels of stomatal guard cells: characteristics of the inward rectifier and its control by pH. J Gen Physiol 99:615–644
57. Blatt MR, Thiel G, Trentham DR (1990) Reversible inactivation of K^+ channels of Vicia stomatal guard cells following the photolysis of caged inositol 1,4,5- trisphosphate. Nature 346:766–769
58. Gajdanowicz P, Garcia-Mata C, Sharma T, Gonzalez W, Morales-Navarro SE, Gonzalez-Nilo FD, Gutowicz J, Mueller-Roeber B, Blatt MR, Dreyer I (2009) Distributed structures determine K^+ and voltage dependent gating of the K_{in} channel KAT1 and the K_{out} channel SKOR. New Phytol 182:380–391
59. Roelfsema MG, Prins HA (1997) Ion channels in guard cells of Arabidopsis thaliana (L) Heynh. Planta 202:18–27
60. Grabov A, Blatt MR (1998) Membrane voltage initiates Ca2+ waves and potentiates Ca2+ increases with abscisic acid in stomatal guard cells. Proc Natl Acad Sci U S A 95:4778–4783
61. Grabov A, Blatt MR (1999) A steep dependence of inward-rectifying potassium channels on cytosolic free calcium concentration increase evoked by hyperpolarization in guard cells. Plant Physiol 119:277–287
62. Acharya BR, Jeon BW, Zhang W, Assmann SM (2013) Open stomata 1 (OST1) is limiting in abscisic acid responses of Arabidopsis guard cells. New Phytol 200:1049–1063
63. Li JX, Lee YRJ, Assmann SM (1998) Guard cells possess a calcium-dependent protein kinase that phosphorylates the KAT1 potassium channel. Plant Physiol 116:785–795
64. Ronzier E, Corratge-Faillie C, Sanchez F, Prado K, Briere C, Leonhardt N, Thibaud JB, Xiong TC (2014) CPK13, a noncanonical Ca2+−dependent protein kinase, specifically inhibits KAT2 and KAT1 shaker K+ channels and reduces stomatal opening. Plant Physiol 166: 314–326

65. Endy D, Brent R (2001) Modelling cellular behaviour. Nature 409:391–395
66. Alleva K, Niemietz CM, Maurel C, Parisi M, Tyerman SD, Amodeo G (2006) Plasma membrane of Beta vulgaris storage root shows high water channel activity regulated by cytoplasmic pH and a dual range of calcium concentrations. J Exp Bot 57:609–621
67. Chaumont F, Tyerman SD (2014) Aquaporins: highly regulated channels controlling plant water relations. Plant Physiol 164:1600–1618
68. Verdoucq L, Grondin A, Maurel C (2008) Structure–function analysis of plant aquaporin AtPIP2;1 gating by divalent cations and protons. Biochem J 415:409–416
69. Yang HM, Zhang XY, Wang GX, Zhang JH (2006) Water channels are involved in stomatal oscillations encoded by parameter-specific cytosolic calcium oscillations. J Integr Plant Biol 48:790–799
70. Minguet-Parramona C, Wang Y, Hills A, Vialet-Chabrand S, Griffiths H, Rogers S, Lawson T, Lew VL, Blatt MR (2016) An optimal frequency in Ca^{2+} oscillations for stomatal closure is an emergent property of ion transport in guard cells. Plant Physiol 170:32–45
71. Blatt MR, Wang Y, Leonhardt N, Hills A (2014) Exploring emergent properties in cellular homeostasis using OnGuard to model K^+ and other ion transport in guard cells. J Plant Physiol 171:770–778
72. Vialet-Chabrand S, Hills A, Wang Y, Griffiths H, Lew VL, Lawson T, Blatt MR, Rogers S (2017) Global sensitivity analysis of OnGuard models identifies key hubs for transport interaction in stomatal dynamics. Plant Physiol 174:680–688
73. Asai N, Nakajima N, Kondo N, Kamada H (1999) The effect of osmotic stress on the solutes in guard cells of Vicia faba L. Plant Cell Physiol 40:843–849
74. Asai N, Nakajima N, Tamaoki M, Kamada H, Kondo N (2000) Role of malate synthesis mediated by phosphoenolpyruvate carboxylase in guard cells in the regulation of stomatal movement. Plant Cell Physiol 41:10–15

Chapter 6
Single-Cell Approaches for Understanding Morphogenesis Using Computational Morphodynamics

Pau Formosa-Jordan, José Teles, and Henrik Jönsson

Abstract In multicellular organisms cells grow, divide and adopt different fates, resulting in tissues and organs with specific functions. In recent years, a number of studies have brought quantitative knowledge about how these processes are orchestrated, shedding new light on cells as active and central players in morphogenesis. We explore recent advances in understanding plant morphogenesis from a quantitative perspective, defining the research field of Computational Morphodynamics. The focus is on studies combining theoretical and experimental approaches integrating hypotheses of how molecular and mechanical regulation at the cellular level lead to tissue behaviour. Finally, we discuss some of the main challenges for future work.

6.1 Introduction

Morphogenesis, the spatial development of tissues with specialized structure and function from populations of undifferentiated cells, is one of the most fascinating and complex problems in nature. Formation of functional tissues requires a strict spatiotemporal control of cellular morphology and gene expression. To achieve this, individual cells can read a multitude of microenvironmental cues, respond via their gene regulatory networks, and ultimately, undertake the appropriate fate

Pau Formosa-Jordan and José Teles contributed equally.

P. Formosa-Jordan · J. Teles
Sainsbury Laboratory, University of Cambridge, Cambridge, UK

H. Jönsson (✉)
Sainsbury Laboratory, University of Cambridge, Cambridge, UK

Computational Biology and Biological Physics, Lund University, Lund, Sweden

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK
e-mail: Henrik.jonsson@slcu.cam.ac.uk

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_6

decisions and morphological changes [1]. A detailed characterisation of the factors influencing cellular fate decision, the mechanisms by which they are executed at the cellular level, and how they are coordinated in multicellular tissues, is paramount to understanding the dynamic pattern formation driving development.

Single cell studies in unicellular organisms and cell cultures have made great progress in providing a quantitative description of the signalling and gene regulatory mechanisms that underlie cell fate decisions [2–5]. These experiments typically involve time-lapse microscopy methods, in which the cells of interest can be tracked as they grow in controlled microenvironments. Expression dynamics of key regulators can be characterised by computational quantification of fluorescent reporters [6, 7] and correlated to cellular variables such as size, division and death [8].

Progress in the development of reporters, microscopy technologies and computational methods now allows for expanding these approaches to the study of developing tissues in multicellular organisms in vivo. In plant biology in particular, morphogenesis has been the object of multiple studies over the years and the role of short-range biochemical signals, long-range hormone transport, and mechanical forces in shaping tissue formation is well documented [9–11]. The ability to quantify the spatiotemporal dynamics of gene expression and cellular behaviour at high resolution in single cells enables a finer quantitative appreciation of how signalling factors ultimately lead to specialised three-dimensional patterns. This level of detailed analysis requires powerful experimental and computational tools in order to identify, track and quantify variables of individual cells in a three-dimensional tissue as it grows in vivo. As these tools become increasingly available, it has been possible to combine high-resolution time-lapse microscopy, image processing, quantification tools and computational models in multidisciplinary efforts to better understand plant development. This integrative approach is the cornerstone of the emerging field of Computational Morphodynamics [12–14] (or Systems Morphodynamics [15]), which, by iterating between experimental design, quantitative data analysis and computer simulations, aims to understand the factors that shape tissues and how they affect individual cells (Fig. 6.1).

In this chapter, we will provide an overview of the most recent developments in the Computational Morphodynamics field in plants, particularly in terms of image processing, quantitative data analysis, and computational modelling techniques. We will review recent studies that applied these methods and technologies, with strong emphasis on work performed at the single-cell level in plant systems, taking advantage of plant-specific properties such as lack of cell migration and relatively slow growth. To illustrate some of the techniques, we will refer to several examples using *Arabidopsis thaliana* as biological model, particularly where single cell descriptions lead to behaviour at the scale of small tissues. We will also discuss the main hurdles and challenges in the field, as well as new methods and technologies that open new and exciting avenues for further understanding the dynamical principles and mechanisms underlying tissue formation in plants.

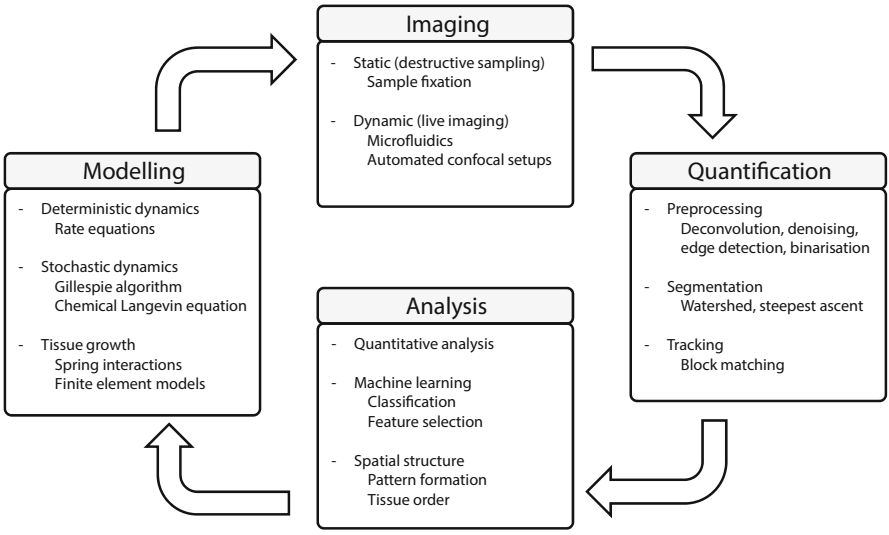


Fig. 6.1 Typical Computational Morphodynamics workflow. An integrative approach iterating between image collection, signal quantification, quantitative analysis and computational modelling. For each step, representative examples for methods and techniques are given and further detailed in the text. Note, this is not meant to be an exhaustive list of all available possibilities

6.2 Capturing Single Cell Dynamics in Space and Time

The accurate characterisation of how single cells are affected by microenvironmental signals and respond in accordance during plant development requires a detailed spatiotemporal description of the tissue as it grows. This description is typically based on time-lapse confocal microscopy of plant tissue in vivo, or live imaging, which allows the non-destructive sampling of the tissue in three dimensions over time. This approach, particularly in combination with other cell and molecular biology tools such as fluorescently tagged reporters, can allow for the quantification of a wealth of variables related to the dynamics of tissue mechanics, cell geometry and cell division, as well as gene expression. Such experiments require high resolution imaging data both spatially and temporally. Ideally, the number of confocal slices spanning the tissue should be as high as possible (thus covering as small a distance as possible) for individual time points, with time intervals being kept to a period as small as possible. A limiting factor in both cases is the phototoxicity that arrives from prolonged and repeated exposure of the tissue to the laser, and so a suitable compromise needs to be found. Sampling times may also need to be adjusted depending on tissue growth rates, such that individual cells can be unequivocally tracked.

The questions of how and when cells divide in plant tissues have attracted much attention. The first studies discussing cell division rules date back to the late 1800s, in which it was proposed that the geometry of the cell would determine the

orientation of the cell division plane (more details on this topic in [16, 17]). In the last decades, improvements of microscopy technologies connected to fluorescent markers have enabled revisiting these questions in more detail. In a pioneering study, Laufs et al. used propidium iodide to stain DNA and visualise individual nuclei in *Arabidopsis thaliana* inflorescence shoot apical meristems (SAMs) [18]. This method allowed for a detailed morphometric analysis that included the quantification of size, spatial distribution and mitotic index of individual cells in wild-type and mutant lines. Propidium iodide staining effectively arrests growth, making it impossible to collect dynamical data. Reddy et al. circumvented this issue by using fluorescent reporters for components of the plasma membrane, histone markers and mitotic cyclins, thus being able to observe the dynamics of cell and nuclear division events, as well as to capture cells about to or in the process of division [19].

As imaging technologies and computational power increase, higher resolution data can be generated which affords automated identification and tracking of individual cells, as well as higher precision in the quantification of variables of interest [20–25]. The implementation of a powerful automated 4D imaging pipeline allowed Willis et al. to observe in great detail the dynamics of cell growth and division in Arabidopsis SAMs [20]. By sampling meristematic growth every 4 h over the course of 2–3 days, and tracing each individual epidermal cell in space and time, the authors could show that cell size regulation does not strictly follow the so-called sizer nor adder models that had been previously described [26], but instead is an intermediate between these two paradigms. Jones et al. also approached this question with single-cell quantification of cell division coupled with a mechanistic model of cyclin-dependent kinase (CDK) activity [22]. The study pointed to the existence of regulatory mechanisms that dynamically regulate cell size as a function of a number of factors including growth rates. Both these studies suggest that cell size in the meristem is not an intrinsically defined (or measured) property. Beyond the timing of division, it is also of interest to understand how cells define the orientation of division planes and localisation of new cell walls. Factors such as cell size, geometry and mechanical forces have all been deemed of potential relevance over the years. Recent quantitative studies have supported the view that cells divide along local minima of plane area leading to equal-sized daughter cells [23], and the alternative view that new division planes establish along directions of maximal tensile stress [24]. Shapiro et al. implemented a quantitative model that predicts the localisation of new division planes at minima of a potential function that incorporates these and other division rules [25].

Despite all its potential, quantitative image analysis at the single cell level is experimentally and computationally demanding and requires careful planning in order to avoid misleading conclusions due to technical artefacts at all stages, from image capture to quantification and analysis [27]. Continuous collection of 4D data in vivo is extremely challenging in plants, due to plant growth and movement, particularly in tissues such as the SAM where plants have to be kept in light- and temperature-controlled chambers between sampling times, and where the imaging is usually done with water-dipping lenses [19, 20]. These challenges can in themselves

be the source of artefacts, and the development of imaging technologies that circumvent them by allowing automated image capture with minimal disturbance can be of paramount importance. To this end, von Wangenheim et al. developed a confocal microscope setup that allows vertical imaging of plant root tip growth with automated adjustment of sampling positions between time points, facilitating the tracking of objects of interest [28]. This platform allowed continuous observation of root growth even when inducing rapid changes in gravity or light, and was successfully tested in zebrafish embryos, showing its general applicability.

The integration of microfluidic technologies for tissue live imaging holds great potential and, indeed, has already proved of importance in different tissues [29]. One prominent example is the RootChip [30], which allows continuous imaging of different growing plant tissues while being exposed to different environmental perturbations [30–32]. This system has been used for studying, for instance, gibberellin response in growing hypocotyls [31] and root–bacteria interactions [32]. Microfluidic technologies afford working in very small volumes as well as the minute control over cellular microenvironment, and hence they have also become an attractive system to observe and manipulate plant cell cultures, either maintained in stable cultures [33] or freshly extracted from the tissue by cell wall digestion [34]. Recently, microfluidic chips have been used to explore mechanisms of regulation of cellular geometry by applying directional auxin gradients to single BY-2 tobacco cells [35]. In the same cell system, optical tweezers had previously been used to create cytoplasmic protrusions in single cells, thus allowing the characterisation of actin and myosin dynamics and their impact on cytoplasm stiffness [36].

For some tissues, it is technically very challenging to perform single-cell quantification while maintaining the plant alive, in which case fixing tissues can be a suitable option [37–39]. Although this makes it impossible to collect dynamical data, it allows for detailed imaging, and statistical methods can then be applied to extract information on growth and division properties from these well-defined cell patterns (Fig. 6.4c, d) [37]. By design, confocal microscopy limits the size of the biological tissue that can be imaged at high resolution. In that regard, whole tissue and even whole organism techniques have been deployed to quantitatively characterize 3D large-scale phenomena in plant development [40, 41].

While the challenges of imaging multicellular and slow growing tissues are many, several of these challenges have recently been overcome and plant development is now approaching a stage where it can be studied with live imaging tools similar to those used for bacteria and yeast, both in vivo and in cell cultures.

6.3 Quantitative Image Analysis

Once image acquisition has been achieved, ideally resulting in a high spatiotemporal resolution time lapse of tissue growth, quantification of cellular variables of interest is performed by deploying a set of computational methods that broadly allow the identification and tracking of each individual cell in space and time. It should be

stressed, however, that image capture and analysis are not independent steps and that particular requirements of image processing should be taken into account when planning time-lapse experiments. Microscopy parameters such as laser intensity, voxel size (in particular, defining the thickness of each confocal slice) and resolution can be optimised by iterating with the image analysis software in single images before performing the full time course. Time sampling should also be carefully defined such that cell tracking can be performed with maximal efficiency. A number of preprocessing steps can also be crucial to improve image quality for the purpose of quantification. Deconvolution may be of importance when processing 3D images, in order to circumvent artificial stretching in the Z-direction caused by the point-spread function [42, 43]. Another relevant factor is plant growth that occurs during image acquisition outside the region of interest. In this case, a fast confocal scan with few slices covering the whole tissue may be acquired first, and used to correct potential growth effects in the longer acquisition, by correcting z-direction slice thickness [20]. In addition to this, other operations may be beneficial such as denoising, edge detection or binarisation of the images to facilitate computational identification of regions of interest (Fig. 6.2a).

The identification of individual cells in a tissue is achieved by applying computational algorithms that make use of specific features of the imaging data to delimitate each region of interest (e.g. a cell, nucleus or other cellular subcompartment) in a process called segmentation. One of the most popular algorithms for cell segmentation is the watershed algorithm [48]. This is particularly true for cases where fluorescence intensities are maximal at the cell membrane, with the inside of each cell having intensity close to zero. Different variants of the watershed algorithm have been successfully applied to the identification of single cells in plant tissues, from embryos [37], to SAMs [45, 49] (Fig. 6.2b), and root meristems [45]. When the fluorescent signal accumulates in a subcellular compartment such as the nucleus, other algorithms may be more suitable. An implementation of the steepest gradient ascent in particular has been used to segment individual meristem cells [50, 51], and nuclei in time lapses of sepal growth [44] (Fig. 6.2a). Once a region of interest has been segmented, its size and morphological parameters can be quantified. The same is true for fluorescence intensity within each region, which is normally used as a proxy for mRNA or protein levels for each cell (Fig. 6.2c). Although this is often not the case, ideal experimental design should include a second fluorescent reporter for a constitutive gene, which can be used to normalise fluorescence intensities in individual cells.

Once individual cells are identified, characterised and computationally catalogued for each time point, temporal correspondence must be resolved by tracking each cell, as well as their progeny, through the time course. This process thus needs to consider cell divisions as well as tissue growth and orientation changes between time points. Due to the existence of rigid cell walls, cell movement is not as much of a factor in plants as it can be in animal tissues. Optimal cell–cell pairing between consecutive time points can be achieved through a process of registration, where one of the images is linearly or non-linearly transformed to maximise matching with the other image of the pair [52]. Different registration

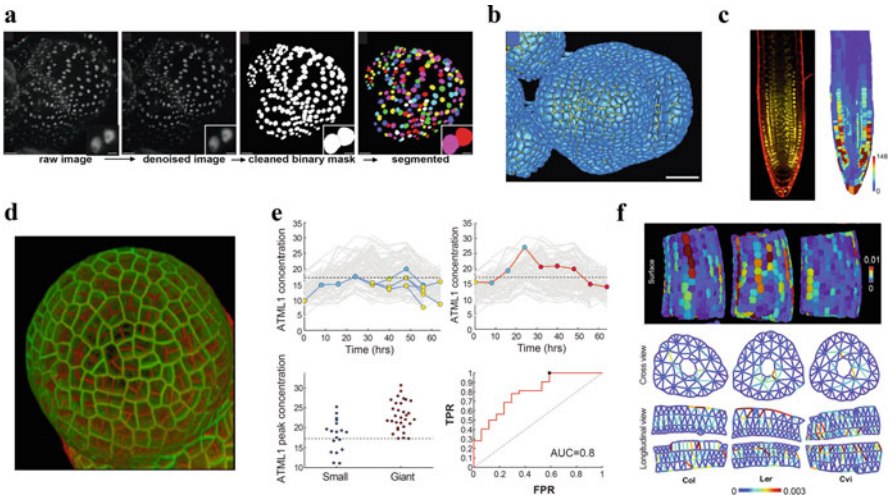


Fig. 6.2 Quantitative image analysis at the single cell level. (**a**) Image analysis pipeline for 3D nuclear segmentation from confocal images of developing sepals in a plant expressing the mCitrine-ATML1 fluorescent reporter [44]. (**b**) 3D visualisation of a shoot apical meristem (SAM) segmented using the watershed algorithm as implemented by Fernandez et al. [45]. (**c**) Single cell quantification of gene expression in the *Arabidopsis* root [46]. Left: DII-VENUS (yellow) and cell geometries given by propidium iodide staining (red). Right: measured DII-VENUS levels quantified from the raw intensity image. (**d**) Single cell tracking of a SAM, using the block matching algorithm [20]. Red lines represent newly formed walls within a period of 24 h of meristematic growth. (**e**) Classification analysis of live imaging mCitrine-ATML1 expression data from sepal growth experiments (Fig. 6.2a) [44]. Top left: ATML1 concentration time course of a small cell lineage. Top right: ATML1 concentration time course of a giant cell lineage. In both cases, coloured circles represent ploidy: yellow – 2C; blue – 4C; red – 8C and above, and the grey lines show all cell linages. Bottom left: each circle represents the ATML1 concentration maximum recorded in 4C cells (i.e. during the G2 stage of the cell cycle) separately for small and giant cell lineages. Bottom right: performance of classification of cells (as either small or giant) based exclusively on ATML1 concentration maxima in 4C cells evaluated by area under the ROC curve (AUC). The red line is the ROC curve; diagonal dashed line represents AUC = 0.5 for comparison. An AUC of 0.8 suggests a good classifier (AUC = 0.5 corresponds to a random classifier; AUC = 1 to perfect classification) and led to the hypothesis that a threshold-based mechanism for cell fate decision is at play involving a combination of high ATML1 concentration during a specific stage of the cell cycle. Horizontal black dashed lines in the top panels as well as bottom left panel represent the inferred ideal ATML1 concentration threshold. (**f**) quantification of topological features of *Arabidopsis* hypocotyls in different ecotypes [47]. Top panel: tissue segmentation meshes with single cell segmentations; heatmap represents scale of betweenness centrality values. Bottom panel: virtual cross and longitudinal sections of extracted cellular networks for each ecotype. Heatmap represents scale of edge betweenness centrality values. Panels A,E have been extracted from [44], panel B from [45], panel C from [46], panel D from [20] and panel F from [47]. Panels (**a**, **e** and **f**) are subject to the CC BY license (https://creativecommons.org/licenses/by/4.0/). Panels (**b** and **c**) have been added with permission, copyright by Springer Nature and the American Society of Plant Biologists, respectively

methods can be applied with reasonable success, particularly when taking into account the specificities of the tissue and objects of interest. These algorithms are available through implementations in different image analysis packages, often with convenient graphical user interfaces and the possibility for manual correction of segmentation and tracking errors [49, 53–56]. In plant tissues specifically, block matching registration algorithms [57–59] have recently been used to successfully perform tracking of individual cells in growing SAMs (Fig. 6.2d) [20], and individual nuclei in growing sepals [44].

The current multitude of image processing tools provide a large pool of options that in principle should cover the exploration of a wide range of biological questions. On the other hand, such diversity also brings about the fragmentation of the community regarding tools of choice and, more importantly, file and data formats. Advances in the state of the art allow us to observe biological phenomena at different scales and complexities from single molecule to whole organs [60, 61]. It is of paramount importance to agree on file and data standards within the community that maximise transferability of both raw image and processed data, such that we can increase power, accuracy and efficiency of quantification protocols.

The vast volumes of spatiotemporal single cell data collected through quantification of time-lapse microscopy experiments can be explored by computational methods in a considerable number of ways. An increasingly used approach is the application of machine learning methods to extract interesting features or behaviours [62]. In supervised machine learning, objects of interest (e.g. cells) can be manually classified within a set of classes a priori (e.g. cell types) given a set of variables (e.g. size or morphological parameters). This dataset is then used to train a model that, once applied to a new dataset, can be able to predict with high accuracy the class of a given object by taking into account its set of variables. There are a number of models that can be used for classification such as regression methods, support vector machines, artificial neural networks, decision trees or random forests [62, 63]. In all these methods, training involves minimizing a function that quantifies the error between prediction and known outcome. A good model should be able to learn from the known data structure and successfully generalise to large sets of new data without overfitting. During the image processing stages, classification methods that take into account cellular morphology and tissue growth can be used to automatically perform and improve segmentation and tracking [63, 64]. Cellular size and morphology parameters have indeed been used for classifying cells in a number of ways, from cell cycle stage [65] to cancer activity [66]. Gene expression data are also a fertile ground for the application of classification methods in order to gain biological insight. Single cell gene expression differences between related cellular populations have been used to predict the identity of genes [67] or the expression thresholds for a specific gene [68] likely to be involved in the decision to choose a particular cellular fate. In plants, a similar approach was recently followed to infer a mechanism by which expression maxima of the transcription factor ATML1 above a certain threshold during the G2 but not the G1 stage of the cell cycle predicts with high accuracy the giant cell fate in growing sepals (Fig. 6.2e) [44]. Unsupervised machine learning methods allow exploring features of interest in the

data structure itself without having to a priori determine the different classes. These include clustering as well as dimensionality reduction methods such as principal component analysis or multidimensional scaling [62].

Beyond gene expression and single cell morphologies, it is of great interest to quantify spatial cellular patterns within a specialised tissue, as a means to further understand its morphodynamics. Methods that allow quantification of tissue order as a function of the spatial distribution of the cells that compose it can be of interest to understand its development and function [69]. In plants, this approach has been followed to understand pattern formation in leafs, in particular the variability of trichome patterns on its surface [70]. More generally, a recent study provided a topological characterisation of complex plant organs by applying quantitative network analysis methods to high resolution descriptions of cellular interactions (Fig. 6.2f) [47]. As the volume of data accumulates, it becomes crucial to integrate these measurements in order to understand how spatial cellular patterns correlate with the underlying gene expression and cellular division patterns during morphogenesis [39].

Great progress has recently been made in the ability to segment, track and quantify single cell information from 3D live imaging data. The continuous development of imaging technologies will predictably lead to data with higher spatiotemporal resolution gathered at quicker rates. The expected magnitude of data produced by these technologies will benefit from current and improved computational methods, but will also require more automatised (and preferably standardised) analysis and storage protocols in order not to become a major bottleneck in future studies.

6.4 Modelling Tissue Morphogenesis from the Bottom-Up

Tissue patterning results from the interplay of intracellular regulatory networks, cell-to-cell interactions, cell growth and division, and mechanical forces [1]. How can mathematical and computational models take these different processes into account? And how can models integrate and be compared with experimental data?

Regulatory networks have mostly been modelled using deterministic rate equations of cellular concentrations of key components. In this context, it can be assumed that, for a given set of model parameter values, the dynamics of regulatory networks will achieve certain steady states, which can be associated with different cell states or fates [50, 71–75]. For example, different regulatory networks have been modelled for understanding epidermal patterning of hair and non-hair fates in both roots and leaves [71, 72]. Also, a model for brassinosteroids signalling in the root has proposed two alternative states that would represent a quiescent and a division state of the quiescent centre cells [74].

In different plant developing tissues, stochasticity in the key regulators has been proposed to play a role in cell fate decisions [76, 77]. In multicellular models, stochasticity has mostly been implemented as small initial cell-to-cell differences in the signalling regulators, which are deterministically propagated through the

regulatory network. Though, due to the probabilistic nature of chemical reactions within cells, certain regulatory networks can show stochastic effects throughout time, and hence, stochasticity needs to be modelled in a dynamic manner. The Gillespie algorithm is an exact and discrete method that can be used for modelling the stochastic nature of regulatory networks [78, 79]. Yet, in the context of multicellular modelling, discrete methods that simulate the number of molecules can become computationally very expensive, and the chemical Langevin equation, which is a continuous concentration-based approach, seems more applicable [80, 81] (Fig. 6.3). Although not being exact, this approach can give quite good approximations to the exact Gillespie method when the number of molecules is not too low [80]. The chemical Langevin equation contains a deterministic contribution and a stochastic contribution, and the stochastic contribution becomes smaller when both the number of molecules and the compartment volume where reactions happen tend to be high [80]. This model formulation facilitates the possibility to use different analytical tools that are used in deterministic systems, as nullcline and bifurcation analyses [82].

Throughout a developmental process, both deterministic and stochastic processes may simultaneously take place. Hence, to model this, hybrid algorithms are necessary, in which both stochastic and deterministic variables are integrated. In the context of giant cell fate commitment in the developing sepal, ATML1 fluctuations were simulated using chemical Langevin equations for ATML1 itself and its downstream target together with a timer variable, while cell growth was modelled deterministically [44].

Often, the patterning process happens while cells are growing and dividing. In the root, the cellular growth and division patterns seem deterministically controlled, leading to stereotypical files of cells that robustly conform the root (Fig. 6.2c) [46, 74]. The root shows a distinct behaviour; growth is mainly anisotropic, cell divisions are most importantly happening in a single dimension along the length of the root. In contrast, the SAM epidermis undergoes isotropic growth and cell divisions are occurring in multiple, but anticlinal, directions, preserving the 2D layer (Fig. 6.2d) [20]. In the case of the developing sepal, a pattern of giant cells gradually forms while cells anisotropically grow and divide [44]. Also, in leaves, the stomata pattern emerges as a result of an interplay of growth, cell divisions and cell fate decisions [83]. Therefore, incorporating cell growth and cell division in the gene regulatory network models becomes fundamental for reproducing the different behaviours and spatial organisations found during morphogenesis of different tissues.

How can we incorporate growth into the models? Cell growth is the result of the internal turgor pressure and mechanical cues [9, 84]. In turn, signalling — and hence, regulatory networks — can modify cell wall properties, which feed back into how cells respond to mechanical cues. Turgor pressure can be effectively included into models in different ways; for instance, by imposing exponential tissue growth [44] (Fig. 6.3), or it can be included in the equations accounting for mechanical cues [84, 85]. In simulations, mechanical cues can be effectively included between vertices through spring interactions acting according to effective potentials [16] or through energies that need to be minimised [86]. Alternatively, stress-strain

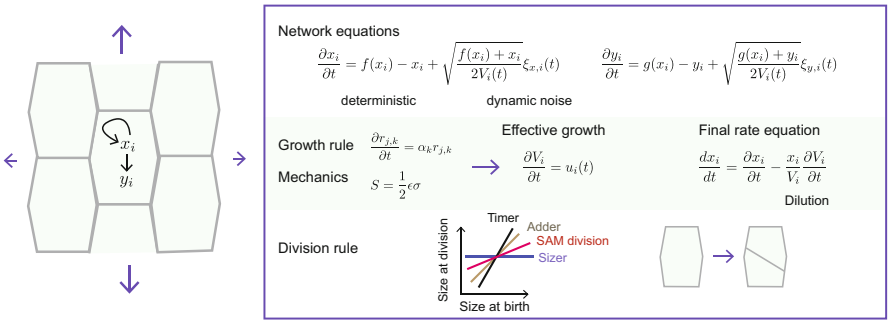


Fig. 6.3 Modelling cell fate decisions in a growing tissue. (Left) Cartoon illustrating a growing tissue. A regulatory network is represented just in the central cell i for simplicity, in which a gene x autoactivates itself and activates a downstream target y. (Right) Possible schematic modelling pipeline. Top: Regulatory networks can be modelled either deterministically or stochastically, depending on the nature of the studied problem. In this scheme, a chemical Langevin equation is shown describing the x and y gene dynamics in cell i. $V_i(t)$ is the volume of the cell i, $f(x_i)$ and $g(x_i)$ are generic functions describing the autoregulation of gene x_i and the activation of y_i by x_i, respectively. $\xi_{z,i}$ is a Gaussian random number with zero mean fulfilling $\langle \xi_{zi}(t) \cdot \xi_{z'j}(t') \rangle = \delta(t - t')\delta_{zz'}\delta_{ij}$, where i and j are cell indices, z and z' the modelled variables (x and y), $\delta_{zz'}$ and δ_{ij} are Kronecker deltas and $\delta(t-t')$ is the Dirac delta [44]. Tissue growth can be implemented by displacing the vertices out the tissue centre of mass, e.g., by following the proposed growth rule, where $r_{j,k}$ refers to the k-th coordinate of the j-th vertex in the tissue, and α_k is the exponential growing rate along k-th direction. On top of the exponential growth, other mechanical cues can be implemented. In this case, we exemplify the mechanical contribution using a strain energy, S, which is a function of the strain (ε) and stress (σ) tensors. The integration of the growth rule explained above with the mechanical cues results into a more complex tissue growth, $u_i(t)$, where i refers to the i-th cell. This resulting growth needs to be taken into account when computing the rate of change of each concentration variable over time. When the resulting growth is more complex than an exponential, its dilution effects can be easily computed given the volumetric changes after each integration step. Finally, different cell division rules can be taken into account to determine the moment at which a cell divides and how the cell division plane will be positioned. Willis et al. [20] showed that the shoot apical meristem (SAM) divides using a rule that is in between the cell adder and sizer paradigms. The division plane can be set with the use of different algorithms [16]

relations (Fig. 6.3) can be included in triangulated cell walls, using finite element method based approaches in 3D [87]. The inclusion of mechanical cues in the models and interactions between joint walls will lead to different growth rates for different cells. Indeed, heterogeneity in cellular growth rates due to mechanical factors and in asymmetric cell divisions have recently been reported [20, 86]. Incorporation of cell growth in the models should drive dilution effects in the different modelled concentrations, and distribution of individual molecules between daughters at cell division. If cell growth is exponential, dilution can be included as effective degradation rates of the modelled concentrations (Fig. 6.3). Though, given the heterogeneity and time-dependent cellular growth rates due to mechanical cues, it becomes more convenient to compute dilution effects given the actual volumetric change [44] (Fig. 6.3).

Taking into account the multitude of cell division patterns found in plants, another challenge is to include cell division in plant tissues models. As previously mentioned, a recent study showed cells in the SAM seemingly divide following a rule that is in between the sizer and adder paradigms [20]. However, in many cases, assumptions and simplifications of the division process will need to be made due to lack of data. Most of the works for plant tissue modelling have made use of sizers (see, *e.g.,* [16, 17, 51, 88]), but there are several models that use timers as well [21, 44]. Although a timer will not allow cell size homeostasis over time and might trigger cell instabilities in the growing tissue [26], it can work as a mechanism for determining cell division times in a differentiating tissue with a countable number of cell cycles before cell division arrests.

In plant tissues, cell-to-cell interactions underpin the transport of key regulators and hormones that are fundamental at the tissue or at the whole plant scale. This includes, for instance, the directed transport of the phytohormone auxin. There are different theoretical proposals of how auxin is transported [89–92], and current studies are still evaluating through a combination of experiments and modelling whether such different competing existing models can explain patterning arising in different tissues [87, 90, 93]. Other diffusible factors including cytokinins, proteins like the stem cell activator WUSCHEL, and microRNAs seem to be fundamental for patterning [94]. Hence, in the study of certain patterning process, diffusible and directed transport between cells may need to be included into the models.

Modelling morphogenetic processes as a whole requires the simultaneous integration of growth, cell division, and signalling. An example where growth, cell division and intercellular signalling are important is the stem cell regulation in shoot meristems. In the SAM, a central pool of cells grows exponentially and the cells undergo division while the stem cell marker CLAVATA3 and WUSCHEL expression domains are continuously maintained. In this context, deterministic models including cell division have been able to predict the different expression domains and could also explain the variability seen following variability in tissue size [95].

Ideally, a good (i.e. useful) model should at a first instance be sufficient to robustly describe the already existing data, and then, it should have some predictive power [14] (Fig. 6.4). To achieve the first stage, one should define equivalent descriptors or observables in both experimental and modelled datasets that can easily be compared. A possible descriptor could be the expression pattern of a fluorescent reporter, or the time series of a certain variable of interest (Fig. 6.4). A more quantitative comparison can be performed by comparing experimental and theoretical concentration histograms and statistical properties of the different datasets. Classification analyses applied to both experimental and theoretical datasets can also bring quantitative ways to execute these comparisons (Fig. 6.4a, b, cf. Fig. 6.4e). Ideally, models should present explorations in the parameter space to establish whether the model can recapitulate the experimental data in a significant region of the parameter space, either by investigating large parameter regions [44, 72, 96, 97] or, if the model is more complex, by performing multiple optimisation runs [75, 95]. This will determine whether the model hypotheses are sufficient to robustly describe

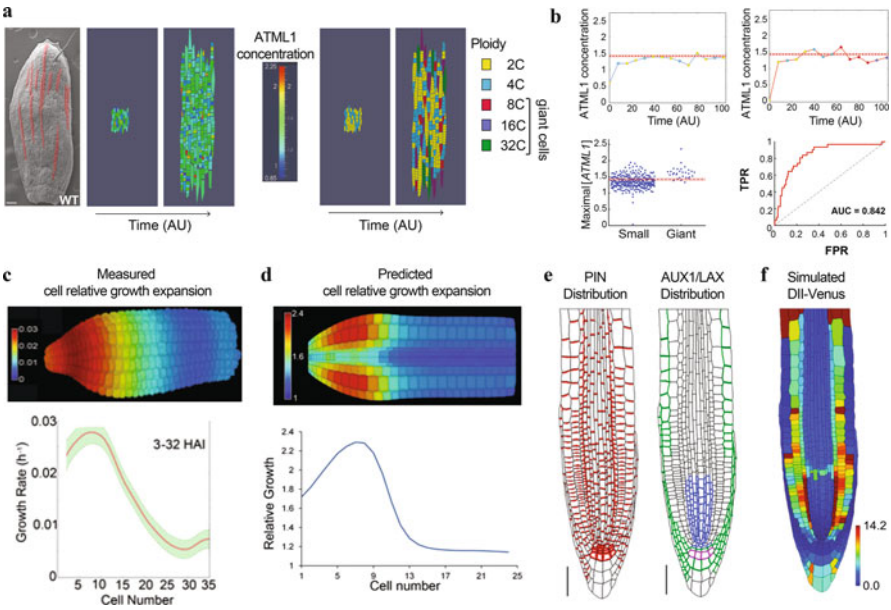


Fig. 6.4 Bridging simulated data with experimental data. (**a** and **b**) Giant cell formation as a case study [44]. (**a**) (Left) Scanning electron microscope image of a wild-type adult sepal, where giant cells are coloured in red. Scale bar: 100 μm. (Right) Simulation example of a developing sepal emulating the experimental data (cf. Fig. 6.2e). The tissue grows anisotropically, while the ATML1 transcription factor dynamically fluctuates. Cells having ATML1 above a certain threshold at the 4C stage (i.e. being in G2) are likely to endoreduplicate and become giant. AU refers to arbitrary units. (**b**) Data analysis of simulation shown in (**a**) (cf. Fig. 6.2e). Data analysis is performed on lower time resolution than the simulated time resolution in order to facilitate comparisons with the experimental data. (Top) Example of ATML1 simulated time-courses for a (top left) normally dividing cell and (top right) a cell becoming a giant cell. Colour refers to the ploidy of the cells as in panel (**a**) on the right. The red dashed horizontal lines refer to the soft threshold for giant cell fate commitment. (Bottom left) Spread plot showing the ATML1 maxima at 4C and the predicted ATML1 threshold. (Bottom right) ROC curve. (**c** and **d**) Mechanical study of the *Arabidopsis* embryo during seed germination [37]. (**c**) Experimental data showing the relative cell expansion in the seed. (**d**) Mechanical growth simulations through finite element methods. Cell colours in top panels in **c** and **d** show relative cell expansions. Growth rates depend on a combination of gene expression input and cell sizes [37]. (**e** and **f**) Auxin patterning in the *Arabidopsis* root [46]. (**e**) Auxin transporter localisations extracted from data. (Left) PIN auxin efflux exporters are shown in red. (Right) AUX1 (green and purple), LAX2 (blue), and LAX3 (purple) auxin influx importers. (Right) Scale bars = 50 μm. (**f**) Simulated DII-Venus levels (compare it with experimental data shown in Fig. 6.2c) given the auxin transporter localisations shown in (**e**). Panels (**a** and **b**) have been adapted from [44], panels (**c** and **d**) from [37], and panels (**e** and **f**) from [46]. See corresponding papers for further details.

the experimental data. Still, several competing models can explain specific data sets equally well, and further tests of the model are required. Hence, in a second and more difficult stage, the model should be able to predict the outcome of experimental perturbations. Examples from the models we have discussed include the ability to predict an adaptive scaling of the stem cell domain to the size of the meristem, later confirmed in experiments [95], and a weak feedback component in the ATML1 sepal model that could later be experimentally verified [44]. Other examples are the predicted growth patterns combining cell size and molecular input in the epidermal root tip (Fig. 6.4c, d) [37], and the prediction of auxin levels given the distributions of auxin transporters extracted from experiments (Fig. 6.4e, f) [46]. Note that sometimes the theoretical parameter explorations can provide predictions of key experimental perturbations. For instance, in a recent study, a mathematical model predicted that the modulation of auxin influx transport would drive a change of vascular spacing in the *Arabidopsis* shoot, and this was experimentally corroborated in influx mutants [93].

As described, several examples of models of tissue morphogenesis exist where a robust simulated behaviour can explain data, and where novel predictions have come from the models. Still, the use of single parameter value explorations is dominating, in part due to the complexity of running models combining molecular regulation with growth. Hence, more efforts will have to be made in this direction. Also, comparisons between data and experiments can be further developed, and probably can be improved with the use of machine learning techniques.

6.5 Conclusion and Outlook

In recent years, several efforts have been made for understanding morphogenesis in plants in a more quantitative manner. Many advances have been produced in both experimental and theoretical methods and applications, and these advances are enabling us to see the potential of cell-centred approaches to better understand tissue morphogenesis. Still, we need to further enhance and promote multidisciplinary efforts to get tangible and significant advances in the field of Computational Morphodynamics in plants. It will be essential to improve the ability to build on and directly compare published experiments and models from several groups to generate a coherent quantitative understanding of plant development.

Alternative experimental approaches, such as the use of plant cell strains, might still have an additional value to shed light into different puzzling phenomena that cannot be disentangled *in planta*. Indeed, single cell approaches, through the use of plant cell strains, have already brought very valuable knowledge into fundamental questions, including how auxin is transported in plants, and how cell shape and cell polarity are generated [98].

A great challenge is to further automate experimental pipelines that can offer quantitative information in a systematic and high-throughput manner. The development of open source and user-friendly platforms for quantitative image analysis,

and the creation of standards between such platforms for elaborating more complex pipelines will be pivotal for the establishment and consolidation of the Computational Morphodynamics field. Furthermore, the development of models with predictive power, close to the experimental data, will be key for a deeper understanding of plant morphogenesis. Such models might need to incorporate and integrate several mechanisms that can be present in different development scenarios, such as integration of mechanics and molecular regulation, dynamic stochasticity, alternative cell division rules, or even time-dependent parameters [68, 99–101], to emulate the evolving and the adaptable nature of regulatory networks in plants.

Finally, the success of the Computational Morphodynamics field will rely on keeping on bridging the gaps between experimental and theoretical approaches (Figs. 6.1 and 6.4) through strengthening the communication between experimentalists and theorists, and through the development of new interdisciplinary scientists and laboratories.

Acknowledgments We apologise to all authors whose valuable contributions are not mentioned in this review due to space limitation. The work in the authors' group was supported by the Gatsby Charitable Foundation (GAT3395/PR4). P. F.-J. and J. T. acknowledge postdoctoral fellowships provided by the Herchel Smith Foundation.

References

1. Gilbert SF, Singer SR (2006) Developmental biology, 8th edn. Sinauer Associates, Inc, Sunderland
2. Martins BM, Das AK, Antunes L, Locke JCW (2016) Frequency doubling in the cyanobacterial circadian clock. Mol Syst Biol 12:896. https://doi.org/10.15252/msb.20167087
3. Cooper S, Bakal C (2017) Accelerating live single-cell Signalling studies. Trends Biotechnol 35:422. https://doi.org/10.1016/j.tibtech.2017.01.002
4. Handly LN, Yao J, Wollman R (2016) Signal transduction at the single-cell level: approaches to study the dynamic nature of signaling networks. J Mol Biol 428:3669–3682. https://doi.org/10.1016/j.jmb.2016.07.009
5. Martins BM, Locke JC (2015) Microbial individuality: how single-cell heterogeneity enables population level strategies. Curr Opin Microbiol 24:104–112. https://doi.org/10.1016/j.mib.2015.01.003
6. Okumoto S, Jones A, Frommer WB (2012) Quantitative imaging with fluorescent biosensors. Annu Rev Plant Biol 63:663–706. https://doi.org/10.1146/annurev-arplant-042110-103745
7. Locke JC, Elowitz M (2009) Using movies to analyse gene circuit dynamics in single cells. Nat Rev Microbiol 7:383–392
8. Schmoller K, Turner JJ, Kõivomägi M, Skotheim JM (2015) Dilution of the cell cycle inhibitor Whi5 controls budding yeast cell size. Nature 526:268. https://doi.org/10.1038/nature14908
9. Sampathkumar A, Yan A, Krupinski P, Meyerowitz EM (2014) Physical forces regulate plant development and morphogenesis. Curr Biol 24:R475–R483. https://doi.org/10.1016/j.cub.2014.03.014
10. Sparks E, Wachsman G, Benfey PN (2013) Spatiotemporal signalling in plant development. Nat Rev Genet 14:631–644. https://doi.org/10.1038/nrg3541
11. Heidstra R, Sabatini S (2014) Plant and animal stem cells: similar yet different. Nat Rev Mol Cell Biol 15:301–312. https://doi.org/10.1038/nrm3790

12. Roeder A, Tarr P, Tobin C, Zhang X (2011) Computational morphodynamics of plants: integrating development over space and time. Nat Rev Mol Cell Biol 12:265
13. Chickarmane V, Roeder AH, Tarr PT et al (2010) Computational Morphodynamics: a modeling framework to understand plant growth. Annu Rev Plant Biol 61:65–87
14. Jönsson H, Gruel J, Krupinski P, Troein C (2012) On evaluating models in computational Morphodynamics. Curr Opin Plant Biol 15:103–110. https://doi.org/10.1016/j.pbi.2011.09.007
15. Mao Y, Green JBA (2017) Systems morphodynamics: understanding the development of tissue hardware. Philos Trans R Soc B 372:20160505. https://doi.org/10.1098/rstb.2016.0505
16. Sahlin P, Jönsson H (2010) A modeling study on how cell division affects properties of epithelial tissues under isotropic growth. PLoS One 5:e11750. https://doi.org/10.1371/journal.pone.0011750
17. Alim K, Hamant O, Boudaoud A (2012) Regulatory role of cell division rules on tissue growth heterogeneity. Front Plant Sci 3:174. https://doi.org/10.3389/fpls.2012.00174
18. Laufs P, Grandjean O, Jonak C et al (1998) Cellular parameters of the shoot apical meristem in Arabidopsis. Plant Cell 10:1375–1390. https://doi.org/10.1105/tpc.10.8.1375
19. Reddy GV, Heisler MG, Ehrhardt DW, Meyerowitz EM (2004) Real-time lineage analysis reveals oriented cell divisions associated with morphogenesis at the shoot apex of Arabidopsis thaliana. Development 131:4225–4237. https://doi.org/10.1242/dev.01261
20. Willis L, Refahi Y, Wightman R et al (2016) Cell size and growth regulation in the Arabidopsis thaliana apical stem cell niche. Proc Natl Acad Sci U S A 113:E8238–E8246. https://doi.org/10.1073/pnas.1616768113
21. Serrano-Mislata A, Schiessl K, Sablowski R (2015) Active control of cell size generates spatial detail during plant organogenesis. Curr Biol 25:2991–2996. https://doi.org/10.1016/j.cub.2015.10.008
22. Jones R, Forero-vargas M, Withers SP et al (2017) Cell-size dependent progression of the cell cycle creates both homeostasis and flexibility of plant cell size. Nat Commun 8:15060. https://doi.org/10.1038/ncomms15060
23. Besson S, Dumais J (2011) Universal rule for the symmetric division of plant cells. Proc Natl Acad Sci U S A 108:6294–6299. https://doi.org/10.1073/pnas.1011866108
24. Louveaux M, Julien J-D, Mirabet V et al (2016) Cell division plane orientation based on tensile stress in Arabidopsis thaliana. Proc Natl Acad Sci U S A 113:E4294–E4303. https://doi.org/10.1073/pnas.1600677113
25. Shapiro BE, Tobin C, Mjolsness E, Meyerowitz EM (2015) Analysis of cell division patterns in the Arabidopsis shoot apical meristem. Proc Natl Acad Sci U S A 112:4815–4820. https://doi.org/10.1073/pnas.1502588112
26. Jun S, Taheri-Araghi S (2015) Cell-size maintenance: universal strategy revealed. Trends Microbiol 23:4–6. https://doi.org/10.1016/j.tim.2014.12.001
27. North AJ (2006) Seeing is believing? A beginners' guide to practical pitfalls in image acquisition. J Cell Biol 172:9–18. https://doi.org/10.1083/jcb.200507103
28. von Wangenheim D, Hauschild R, Fendrych M et al (2017) Live tracking of moving samples in confocal microscopy for vertically grown roots. elife 6:e26792. https://doi.org/10.7554/eLife.26792
29. Sanati Nezhad A (2014) Microfluidic platforms for plant cells studies. Lab Chip 14:3262–3274. https://doi.org/10.1039/C4LC00495G
30. Grossmann G, Guo W-J, Ehrhardt DW et al (2011) The RootChip: an integrated microfluidic chip for plant science. Plant Cell 23:4234–4240. https://doi.org/10.1105/tpc.111.092577
31. Rizza A, Walia A, Lanquar V et al (2017) In vivo gibberellin gradients visualized in rapidly elongating tissues. Nat Plants 3:803–813. https://doi.org/10.1038/s41477-017-0021-9
32. Massalha H, Korenblum E, Malitsky S et al (2017) Live imaging of root–bacteria interactions in a microfluidics setup. Proc Natl Acad Sci United States Am 114:4549–4554. https://doi.org/10.1073/pnas.1618584114

33. Barbez E, Laňková M, Pařezová M et al (2013) Single-cell-based system to monitor carrier driven cellular auxin homeostasis. BMC Plant Biol 13:20. https://doi.org/10.1186/1471-2229-13-20
34. Yoo S-D, Cho Y-H, Sheen J (2007) Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. Nat Protoc 2:1565–1572. https://doi.org/10.1038/nprot.2007.199
35. Zaban B, Liu W, Jiang X et al (2015) Plant cells use Auxin efflux to explore geometry. Sci Rep 4:5852. https://doi.org/10.1038/srep05852
36. Van Der Honing HS, De Ruijter NCA, Emons AMC, Ketelaar T (2010) Actin and myosin regulate cytoplasm stiffness in plant cells: a study using optical tweezers. New Phytol 185:90–102. https://doi.org/10.1111/j.1469-8137.2009.03017.x
37. Bassel GW, Stamm P, Mosca G et al (2014) Mechanical constraints imposed by 3D cellular geometry and arrangement modulate growth patterns in the Arabidopsis embryo. Proc Natl Acad Sci U S A 111:8685–8690. https://doi.org/10.1073/pnas.1404616111
38. Yoshida S, BarbierdeReuille P, Lane B et al (2014) Genetic control of plant development by overriding a geometric division rule. Dev Cell 29:75–87. https://doi.org/10.1016/j.devcel.2014.02.002
39. Bassel GW, Smith RS (2016) Quantifying morphogenesis in plants in 4D. Curr Opin Plant Biol 29:87–94. https://doi.org/10.1016/j.pbi.2015.11.005
40. Morris EC, Griffiths M, Golebiowska A et al (2017) Shaping 3D root system architecture. Curr Biol 27:R919–R930. https://doi.org/10.1016/j.cub.2017.06.043
41. Lee KJI, Calder GM, Hindle CR et al (2017) Macro optical projection tomography for large scale 3D imaging of plant structures and gene activity. J Exp Bot 68:527–538. https://doi.org/10.1093/jxb/erw452
42. McNally JG, Karpova T, Cooper J, Conchello JA (1999) Three-dimensional imaging by deconvolution microscopy. Methods 19:373–385. https://doi.org/10.1006/meth.1999.0873
43. Shaw PJ, Rawlins DJ (1991) The point spread function of a confocal microscopy: its measurement and use in deconvolution of 3-D data. J Microsc 163:151–165. https://doi.org/10.1111/j.1365-2818.1991.tb03168.x
44. Meyer HM, Teles J, Formosa-Jordan P et al (2017) Fluctuations of the transcription factor ATML1 generate the pattern of giant cells in the Arabidopsis sepal. Elife 6:e19131. https://doi.org/10.7554/eLife.19131
45. Fernandez R, Das P, Mirabet V et al (2010) Imaging plant growth in 4D: robust tissue reconstruction and lineaging at cell resolution. Nat Methods 7:547. https://doi.org/10.1038/nMeth.1472
46. Band LR, Wells DM, Fozard JA et al (2014) Systems analysis of Auxin transport in the Arabidopsis root apex. Plant Cell 26:862–875. https://doi.org/10.1105/tpc.113.119495
47. Jackson MDB, Xu H, Duran-Nebreda S et al (2017) Topological analysis of multicellular complexity in the plant hypocotyl. Elife 6:e26023. https://doi.org/10.7554/eLife.26023
48. Soille P (2004) Morphological image analysis: principles and applications, 2nd edn. Springer Science & Business Media, Berlin. https://doi.org/10.1007/978-3-662-05088-0
49. Barbier de Reuille P, Routier-Kierzkowska A-L, Kierzkowski D et al (2015) MorphoGraphX: a platform for quantifying morphogenesis in 4D. elife 4:e05864. https://doi.org/10.7554/eLife.05864
50. Jönsson H (2005) Modeling the organization of the WUSCHEL expression domain in the shoot apical meristem. Bioinformatics 21:i232–i240
51. Jönsson H, Heisler MG, Shapiro BE et al (2006) An auxin-driven polarized transport model for phyllotaxis. Proc Natl Acad Sci U S A 103:1633–1638
52. Maška M, Ulman V, Svoboda D et al (2014) A benchmark for comparison of cell tracking algorithms. Bioinformatics 30:1609–1617. https://doi.org/10.1093/bioinformatics/btu080
53. Haubold C, Schiegg M, Kreshuk A et al (2016) Segmenting and tracking multiple dividing targets using ilastik. In: Focus bio-image informatics. Springer, Basel, pp 199–229

54. Carpenter AE, Jones TR, Lamprecht MR et al (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7:R100. https://doi.org/10.1186/gb-2006-7-10-r100
55. Cordelières FP, Petit V, Kumasaka M et al (2013) Automated cell tracking and analysis in phase-contrast videos (iTrack4U): development of Java software based on combined mean-shift processes. PLoS One 8:e81266. https://doi.org/10.1371/journal.pone.0081266
56. Tinevez J-Y, Perry N, Schindelin J et al (2017) TrackMate: an open and extensible platform for single-particle tracking. Methods 115:80–90. https://doi.org/10.1016/j.ymeth.2016.09.016
57. Michelin G, Refahi Y, Wightman R et al (2016) Spatio-temporal registration of 3D microscopy image sequences of arabidopsis floral meristems. In: 2016 IEEE 13th international symposium. Biomedical imaging. IEEE, Amsterdam, pp 1127–1130
58. Commowick O, Grégoire V, Malandain G (2008) Atlas-based delineation of lymph node levels in head and neck computed tomography images. Radiother Oncol 87:281–289. https://doi.org/10.1016/j.radonc.2008.01.018
59. Ourselin S, Roche A, Prima S, Ayache N (2000) Block matching: a general framework to improve robustness of rigid registration of medical images. Springer, Heidelberg, pp 557–566
60. Fricker MD, Moger J, Littlejohn GR, Deeks MJ (2016) Making microscopy count: quantitative light microscopy of dynamic processes in living plants. J Microsc 263:181–191. https://doi.org/10.1111/jmi.12403
61. Montenegro-Johnson TD, Stamm P, Strauss S et al (2015) Digital single-cell analysis of plant organ development using 3DCellAtlas. Plant Cell 27:1018–1033. https://doi.org/10.1105/tpc.15.00175
62. Sommer C, Gerlich DW (2013) Machine learning in cell biology – teaching computers to recognize phenotypes. J Cell Sci 126:5529–5539. https://doi.org/10.1242/jcs.123604
63. Kan A (2017) Machine learning applications in cell image analysis. Immunol Cell Biol 95(6):1–6. https://doi.org/10.1038/icb.2017.16
64. Van Valen DA, Kudo T, Lane KM et al (2016) Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. PLoS Comput Biol 12:e1005177. https://doi.org/10.1371/journal.pcbi.1005177
65. Held M, A Schmitz MH, Fischer B et al (2010) CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. Nat Methods 7:747–754. https://doi.org/10.1038/nmeth.1486
66. Yin Z, Sadok A, Sailem H et al (2013) A screen for morphological complexity identifies regulators of switch-like transitions between discrete cell shapes. Nat Cell Biol 15:860–871. https://doi.org/10.1038/ncb2764
67. Teles J, Pina C, Edé NP et al (2013) Transcriptional regulation of lineage commitment - a stochastic model of cell fate decisions. PLoS Comput Biol 9:e1003197. https://doi.org/10.1371/journal.pcbi.1003197
68. Schröter C, Rué P, Mackenzie JP, Martinez Arias A (2015) FGF/MAPK signaling sets the switching threshold of a mutual repressor circuit controlling cell fate decisions in ES cells. Development 142:4205–4216. https://doi.org/10.1242/dev.127530
69. Protonotarios ED, Baum B, Johnston A et al (2014) An absolute interval scale of order for point patterns. J R Soc Interface 11:20140342. https://doi.org/10.1098/rsif.2014.0342
70. Greese B, Hülskamp M, Fleck C et al (2014) Quantification of variability in trichome patterns. Front Plant Sci 5:596. https://doi.org/10.3389/fpls.2014.00596
71. Azpeitia E, Benitez M, Vega I, Villarreal C (2010) Single-cell and coupled GRN models of cell patterning in the Arabidopsis thaliana root stem cell niche. BMC Syst Biol 4:134
72. Digiuni S, Schellmann S, Geier F et al (2008) A competitive complex formation mechanism underlies trichome patterning on Arabidopsis leaves. Mol Syst Biol 4:217. https://doi.org/10.1038/msb.2008.54
73. Jaeger KE, Pullen N, Lamzin S et al (2013) Interlocking feedback loops govern the dynamic behavior of the floral transition in Arabidopsis. Plant Cell 25:820–833. https://doi.org/10.1105/tpc.113.109355

74. Vilarrasa-Blasi J, González-García MP, Frigola D et al (2014) Regulation of plant stem cell quiescence by a brassinosteroid signaling module. Dev Cell 30:36–47. https://doi.org/10.1016/j.devcel.2014.05.020
75. Yadav RK, Perales M, Gruel J et al (2013) Plant stem cell maintenance involves direct transcriptional repression of differentiation program. Mol Syst Biol 9:654. https://doi.org/10.1038/msb.2013.8
76. Meyer HM, Roeder AHK (2014) Stochasticity in plant cellular growth and patterning. Front Plant Sci 5:420. https://doi.org/10.3389/fpls.2014.00420
77. Álvarez-Buylla ER, Chaos Á, Aldana M et al (2008) Floral morphogenesis: stochastic explorations of a gene network epigenetic landscape. PLoS One 3:e3626. https://doi.org/10.1371/journal.pone.0003626
78. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81:2340–2361. https://doi.org/10.1021/j100540a008
79. El Samad H, Khammash M, Petzold L, Gillespie D (2005) Stochastic modelling of gene regulatory networks. Int J Robust Nonlinear Control 15:691–711. https://doi.org/10.1002/rnc.1018
80. Gillespie DT (2000) The chemical Langevin equation. J Chem Phys 297:297–306. https://doi.org/10.1063/1.481811
81. Adalsteinsson D, McMillen D, Elston TC (2004) Biochemical network stochastic simulator (BioNetS): software for stochastic modeling of biochemical networks. BMC Bioinformatics 5:24. https://doi.org/10.1186/1471-2105-5-24
82. Weber M, Buceta J (2013) Stochastic stabilization of phenotypic states: the genetic bistable switch as a case study. PLoS One 8:e73487. https://doi.org/10.1371/journal.pone.0073487
83. Robinson S, Barbier De Reuille P, Chan J et al (2011) Generation of spatial patterns through cell polarity switching. Science 333:1436
84. Boudon F, Chopard J, Ali O et al (2015) A computational framework for 3D mechanical modeling of plant morphogenesis with cellular resolution. PLoS Comput Biol 11:e1003950. https://doi.org/10.1371/journal.pcbi.1003950
85. Hamant O, Heisler MG, Jönsson H et al (2008) Developmental patterning by mechanical signals in Arabidopsis. Science 322:1650–1655
86. Uyttewaal M, Burian A, Alim K et al (2012) Mechanical stress acts via Katanin to amplify differences in growth rate between adjacent cells in Arabidopsis. Cell 149:439–451. https://doi.org/10.1016/j.cell.2012.02.048
87. Bhatia N, Bozorg B, Larsson A et al (2016) Auxin acts through MONOPTEROS to regulate plant cell polarity and pattern Phyllotaxis. Curr Biol 26:1–7. https://doi.org/10.1016/j.cub.2016.09.044
88. Stoma S, Lucas M, Chopard J et al (2008) Flux-based transport enhancement as a plausible unifying mechanism for auxin transport in meristem development. PLoS Comput Biol 4:e1000207. https://doi.org/10.1371/journal.pcbi.1000207
89. Morales-Tapia A, Cruz-Ramírez A (2016) Computational modeling of Auxin: a Foundation for Plant Engineering. Front Plant Sci 7:1881. https://doi.org/10.3389/fpls.2016.01881
90. Abley K, De Reuille PB, Strutt D et al (2013) An intracellular partitioning-based framework for tissue cell polarity in plants and animals. Development 140:2061–2074. https://doi.org/10.1242/dev.062984
91. Runions A, Smith RS, Prusinkiewicz P (2014) Computational models of auxin-driven development. In: Auxin its role plant development. Springer, Vienna, pp 315–357
92. van Berkel K, de Boer RJ, Scheres B, ten Tusscher K (2013) Polar auxin transport: models and mechanisms. Development 140:2253–2268. https://doi.org/10.1242/dev.079111
93. Fàbregas N, Formosa-Jordan P, Confraria A et al (2015) Auxin influx carriers control vascular patterning and xylem differentiation in Arabidopsis thaliana. PLoS Genet 11:e1005183. https://doi.org/10.1371/journal.pgen.1005183
94. Yadav SR, Yan D, Sevilem I, Helariutta Y (2014) Plasmodesmata-mediated intercellular signaling during plant growth and development. Development 5:1–7. https://doi.org/10.3389/fpls.2014.00044

95. Gruel J, Landrein B, Tarr P et al (2016) An epidermis-driven mechanism positions and scales stem cell niches in plants. Sci Adv 2:e1500989. https://doi.org/10.1126/sciadv.1500989
96. Liepe J, Kirk P, Filippi S et al (2014) A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation. Nat Protoc 9:439. https://doi.org/10.1038/nprot.2014.025
97. Pullen N, Morris RJ (2014) Bayesian model comparison and parameter inference in systems biology using nested sampling. PLoS One 9(2):e88419. https://doi.org/10.1371/journal.pone.0088419
98. Opatrný Z, Nick P, Petrášek J (2014) Plant cell strains in fundamental research and applications. In: Nick P, Opatrny Z (eds) Applied plant cell biology. Springer, Heidelberg, pp 455–481
99. Koseska A, Bastiaens PI (2017) Cell signaling as a cognitive process. EMBO J 36:568–582. https://doi.org/10.15252/embj
100. Palau-Ortin D, Formosa-Jordan P, Sancho JM, Ibañes M (2015) Pattern selection by dynamical biochemical signals. Biophys J 108:1555–1565. https://doi.org/10.1016/j.bpj.2014.12.058
101. Verd B, Crombach A, Jaeger J (2014) Classification of transient behaviours in a time-dependent toggle switch model. BMC Syst Biol 8:43. https://doi.org/10.1186/1752-0509-8-43

Chapter 7
Modeling Plant Tissue Growth and Cell Division

Gabriella Mosca, Milad Adibi, Soeren Strauss, Adam Runions, Aleksandra Sapala, and Richard S. Smith

Abstract Morphogenesis is the creation of form, a complex process requiring the integration of genetics, mechanics, and geometry. Patterning processes driven by molecular regulatory and signaling networks interact with growth to create organ shape, often in unintuitive ways. Computer simulation modeling is becoming an increasingly important tool to aid our understanding of these complex interactions. In this chapter we introduce computational approaches for studying these processes on spatial, multicellular domains. For some problems, such as the exploration of many patterning processes, simulation can be done on static (non-growing) templates. These can range from abstract idealized cells, such as rectangular or hex grids, to more realistic shapes such as Voronoi regions, or even shapes extracted from bio-imaging data. More dynamic processes like phyllotaxis involve the interaction of growth and patterning, and require the simulation of growing domains. In the simplest case growth can be modeled descriptively, provided as an input to the model. Growth is specified globally, and must be designed carefully to avoid conflicts (growing cells must fit together). We present several methods for this that can be applied to shoots, roots, leaves, and other plant organs. However when shape is an emergent property of the model, different cells or areas of the tissue need to specify their growth locally, and physically-based methods (mechanics) are required to resolve conflicts. Among these are mass-spring, finite element, and Hamiltonian-based approaches.

G. Mosca · M. Adibi · S. Strauss · A. Runions · A. Sapala · R. S. Smith (✉)
Max Planck Institute for Plant Breeding Research,, Köln, Germany
e-mail: smith@mpipz.mpg.de

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_7

7.1 Modeling Plant Morphogenesis

Plants regulate their form through interactions between genetic processes, cell–cell signalling, and mechanics. Since plant cells are attached to each other by a rigid cell wall, emergent shape is controlled almost entirely by regulating the rates and directions of growth and proliferation of individual cells. Genetic programs operating at the cellular level must be integrated over the tissue via signalling mechanisms, to result in reproducible organs, shapes, and sizes. To explore these interactions, which can be complex and result in unintuitive behavior, computer simulation modeling has emerged as an important tool.

Although much modeling work in biology has focused on genetic networks that operate within single cells [39], many processes in morphogenesis require models with a spatial component. One example is models of pattern formation, exploring the differentiation of cells into different cell types within a plant organ. Typical mechanisms include Turing patterns [79] such as reaction-diffusion [52] and auxin-transport feedback based patterning [76]. Although some aspects of patterning can be analyzed with purely mathematical models [71, 79], this is only possible in cases with very simple geometry and interactions involving only a few interacting components.

Spatial models can be represented at different scales. Entire plants are often modeled at the organ level, with model elements representing individual leaves or sections of stem [62]. Organs may be discretized into pieces that represent groups of many cells [44], whereas portions of organs can often be represented directly at the cellular level [38]. In some cases, such as modeling involving proteins that transport auxin to and from the cell to extracellular space, it is necessary to explicitly include compartments for the wall between cells [1, 9, 42, 71]. Both the cell and cell wall themselves may be discretized into compartments, if it is required to represent gradients of morphogens or signaling molecules at a sub-cellular level in the model [28, 43].

7.2 Fixed Templates

Fixed templates are useful to explore patterning aspects of plant development when the explicit simulation of growth is not required. Patterning mechanisms can depend on mass-action reactions, cell–cell communication via diffusion or active transport, mechanical forces, or any combination of these. Boundary conditions can also be important, as can environmental inputs, as they might form pre-patterns or direct patterning in some other way.

Perhaps the most widely used model of pattern formation in developmental biology is the Turing mechanism [52, 59, 79], which is often referred to as reaction-diffusion or diffusion driven instability. Turing proposed the idea to answer a fundamental question in developmental biology: given a field of identical cells, how can a pattern emerge when all cells follow the same reactions and have the same molecular components? In terms of genetic networks one can ask how different cell types emerge if they all have the same genes.

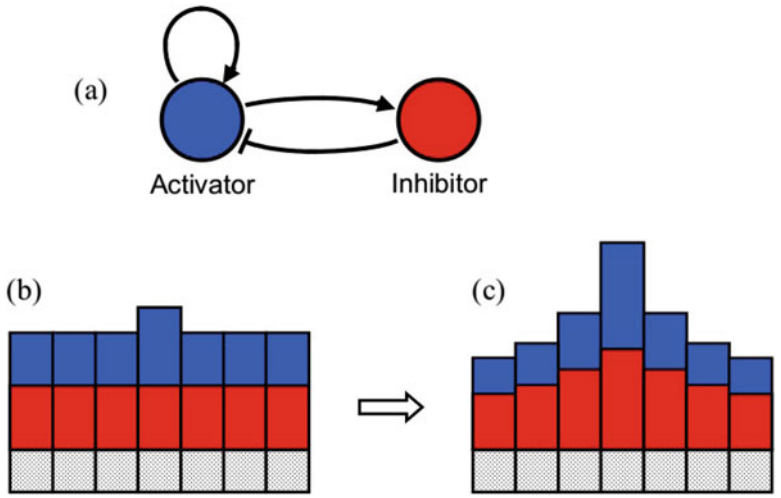


Fig. 7.1 Activator-inhibitor system. The activator is shown in blue, the inhibitor is shown in red (**a**) The activator enhances its own production as well as the production of the inhibitor. The inhibitor inhibits production of the activator. (**b–c**) A line of cells with the height of the blue and red bars indicating activator and inhibitor concentration. (**b**) A slight increase in the concentration of the activator in one cell due to random fluctuations causes an increase in production of both the activator and the inhibitor in that cell. (**c**) The inhibitor diffuses away more quickly than the activator, allowing local activation to escalate, while simultaneously suppressing neighbor cells. Adapted from [76]

Gierer and Meinhardt [25] proposed a Turing system with equations that are easier to relate to molecular regulatory networks than those in Turing's original work. Their formulations have since been used to model many patterning processes in biology [52], mostly on fixed templates. The simplest of these is based on two substances and is called an activator-inhibitor system. The activator a enhances its own production, as well as that of another substance h, termed the inhibitor (Hemmstoff in German), whereas the inhibitor inhibits production of the activator. Figure 7.1a shows the interaction in the format typically used to display genetic or molecular regulatory networks.

To demonstrate the mechanism's pattern forming ability, it is instructive to use the same template as that originally proposed by Turing, a line of identical cells connected at the ends to form a ring. In this case all cells have the same equations and the same boundary conditions. The following equations are used to model the change in concentration of the activator and inhibitor in the cells over time:

$$\frac{da_i}{dt} = \rho_{a_0} + \rho_a \frac{a_i^2}{1 + h_i} - \mu_a a_i - D_a \sum_{j \in N_i} \left(a_i - a_j\right) \tag{7.1}$$

$$\frac{dh_i}{dt} = \rho_{h_0} + \rho_h a_i^2 - \mu_h h_i - D_h \sum_{j \in N_i} \left(h_i - h_j\right) \tag{7.2}$$

where a_i and h_i are the concentrations of activator and inhibitor in cell i, ρ_{a_0} and ρ_{h_0} control a small amount of background production, ρ_a and ρ_h control activator-enhanced production, μ_a and μ_h control decay, D_a and D_h control the rates of cell to cell diffusion. Diffusion is summed over all of the neighboring cells N_i. Each cell is modeled as a single compartment with a single value for the concentration of the activator and inhibitor (i.e. no gradient within the cell). Extracellular space is ignored, with diffusion occurring directly between cells. Meinhardt notes that for stable peaks to form, the rate of activator diffusion must be much smaller than that of the inhibitor.

If the system started with uniform initial conditions, the derivatives will be the same for all cells and no pattern will emerge. The steady state concentrations of the activator and inhibitor will be the same in all cells. If a small amount of noise is added to the system, for example in the initial activator concentration, the system can destabilize, depending on parameter values. Small local maxima in activator concentration lead to a local increase in production of both the activator and the inhibitor (see Fig. 7.1b, c). The inhibitor diffuses away more quickly than the activator, reducing its effect on local activator self-enhancement, while suppressing activator self-enhancement nearby. This can lead to a spatial pattern of peaks in activator concentration (Fig. 7.2), which can trigger selective differentiation and patterning. Parameters in Eqs. (7.1) and (7.2) control the size and spacing between peaks.

Many patterning problems in development require at least a 2D template for simulation. One example is the canalization hypothesis proposed by Sachs for leaf venation patterning [55, 70]. Sachs' canalization mechanism is based on the idea that a cell's ability to transport a morphogen (auxin) increases with the flux of

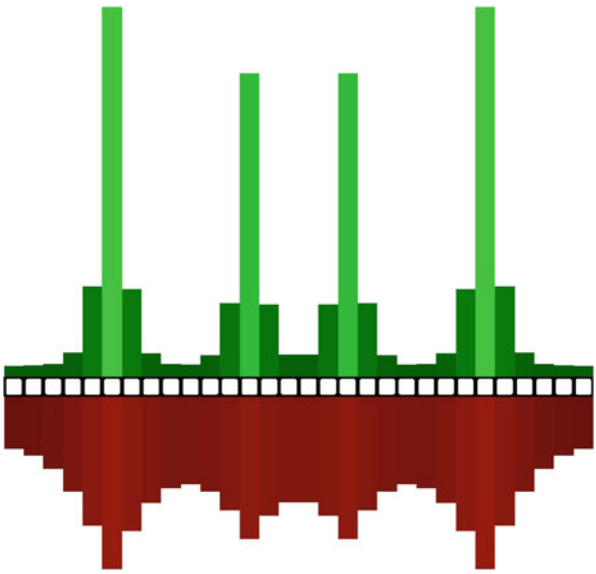


Fig. 7.2 Reaction-diffusion patterning on a line of cells using an activator-inhibitor system. The ends of the line are connected to give the topology of a ring. The activator is in green, and the inhibitor in red, with the height of the bars representing the concentration of the morphogens. Starting from uniform initial conditions with a small amount of noise, peaks in concentration emerge. Note that both the activator and inhibitor have high concentrations in the activated cells. Adapted from [76]

the morphogen through the cell. He proposed that this mechanism is analogous to the process by which water carves preferred routes when running over even terrain. Random variations cause a slight preference for a certain path, which is then enhanced by the flow causing it to attract even more flux. Mitchison showed that this mechanism is capable of pattern formation on a 2D grid of cells [55]. The following equations are from Mitchison's polar transport model, as re-formulated by Rolland-Lagan [66].

The net flux of auxin $\phi_{i\to j}$ from cell i to cell j is given by:

$$\phi_{i\to j} = T\left(a_i c_{i\to j} - a_j c_{j\to i}\right) + D\left(a_i - a_j\right) \tag{7.3}$$

where T is the transport coefficient, a_i is the concentration of auxin in cell i, $c_{i\to j}$ is the amount of carriers allocated to the membrane in cell i facing cell j, and D is the diffusion coefficient.

The change in carriers at the membrane of cell i facing cell j is given by:

$$\frac{dc_{i\to j}}{dt} = \alpha\phi_{i\to j}^2 + \beta - \gamma c_{i\to j} \qquad if\, \phi_{i\to j} > 0 \tag{7.4}$$

$$\frac{dc_{i\to j}}{dt} = \beta - \gamma c_{i\to j} \qquad if\, \phi_{i\to j} \leq 0 \tag{7.5}$$

where α is the carrier allocation coefficient based on flux, β is the flux independent (background) carrier allocation coefficient, and γ is the coefficient representing carrier decay (release) from the cell membrane. If the net flux is less than or equal to zero, only the background carrier allocation and decay are used.

The change in concentration of auxin with respect to time is modeled as:

$$\frac{da_i}{dt} = \rho - \mu a_i - \sum_{j\in N_i} \phi_{i\to j} \tag{7.6}$$

where ρ is the auxin production coefficient in the cells, μ is the auxin decay coefficient, with the summation is taken over all of the neighbors N_i of cell i.

The canalization model is capable of *de novo* pattern formation, and is able to select strands of cells from a tissue of undifferentiated cells. The grid in Fig. 7.3 has a row of source cells at the top (non-zero ρ) and sink cells at the bottom whose auxin concentration a is fixed at zero. The grid is connected at the sides to give the topology of a cylinder. Initially auxin flows via diffusion uniformly from source to sink cells. Like the activator-inhibitor model, without noise no pattern is formed. By adding a small amount of noise to auxin levels, some paths are very slightly favored. The feedback of auxin on its own transport causes the reinforcement of these paths due to the non-linear production of carriers based on auxin flux [55]. This causes the uniform gradient to destabilize into discreet strands of high auxin flux. It is the cells in these strands which are hypothesized to then differentiate into veins.

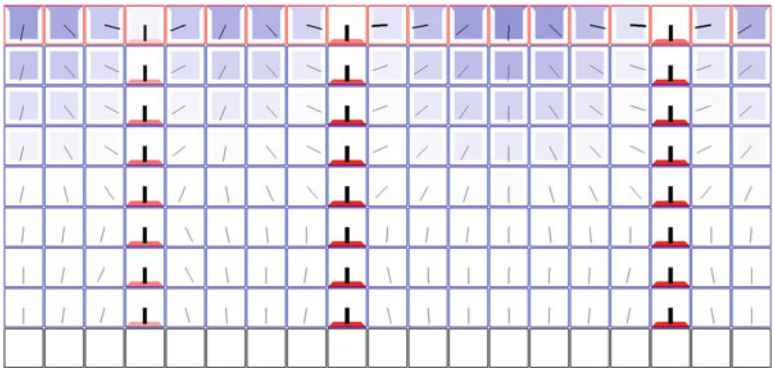


Fig. 7.3 The canalization hypothesis proposed for leaf venation is an example of a patterning mechanism that requires a two-dimensional template. The concentration of the morphogen auxin is shown in blue, with transporters that move auxin from cell to cell localized at the cell boundaries shown in red. The short lines from the center of each cell represent directions of flux through the cell, and their width represents flux strength. The top row of cells are auxin sources, and the bottom row are auxin sinks. The grid is connected at the sides to give the topology of a cylinder. Starting from uniform initial conditions with a small amount of noise to break symmetry, auxin is canalized into discrete strands by the feedback of auxin on its own transport. The exact location of the strands depends on the initial noise and is different for each simulation run, however the spacing is similar. Note that some cells in the top row have low auxin even though they are sources, because of the strong efflux due to transport. Simulations after [66]

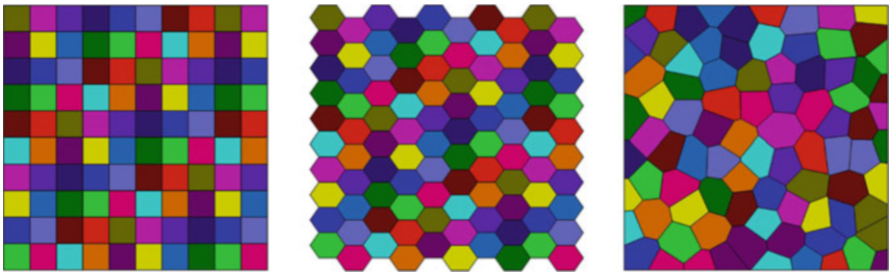


Fig. 7.4 Examples of simple cellular templates made from rectangular, hex and Voronoi regions. Color represents an arbitrary cell labeling

In addition to rectangular grids of cells, other geometric shapes such as hexagonal arrays are also commonly used in plant developmental models [21]. Voronoi regions have also been used for modeling [35, 36] because of their cell-like appearance (see Fig. 7.4). Increasingly more realistic templates are being employed, with geometric idealizations that more accurately capture the cell topology (Fig. 7.5a), or templates based on images taken from biological samples (Fig. 7.5b, c).

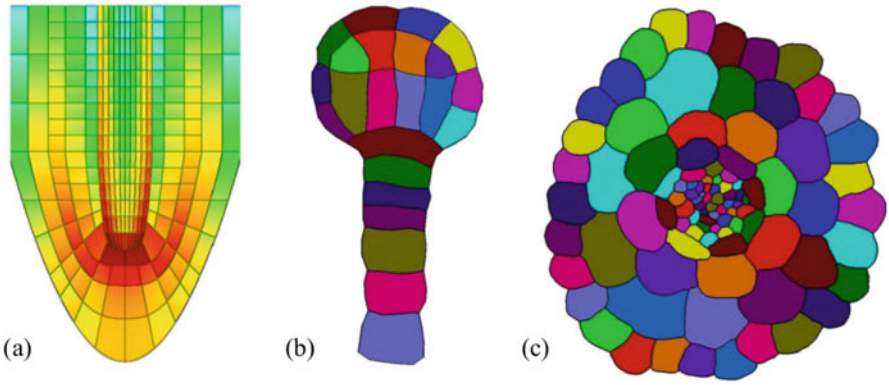


Fig. 7.5 More realistic cellular templates. (**a**) A geometric template model of the root apex of *Arabidopsis thaliana* that more accurately captures root meristem topology. Color shows auxin level, with blue cells having lower concentration and red cells having higher concentration. Adapted from [17]. (**b**) An *Arabidopsis* embryo template extracted from a segmented section taken from a confocal microscopy image stack. Color represents cell labeling. Adapted from [80]. (**c**) An *Arabidopsis* hypocotyl template extracted from a physical cross-section of the hypocotyl. Color represents cell labeling. Adapted from [82]

7.3 Growth and Cell Division

Although fixed domains can be useful to explore patterning mechanisms, many questions in development require the simulation of growing domains [5, 14, 16, 37, 77]. In the one-dimensional case, L-systems are a particularly convenient formalism. Originally developed to describe the development of filamentous organisms [47, 48], they have been widely used to model the branching structures of whole plants [6, 62]. L-systems are rewriting systems based on strings made up of modules or letters from a given alphabet. Modules can have parameters to hold data values such as the size of a cell or stem segment, or the concentration of a morphogen. Beginning with an initial string called the axiom, rules are defined that map each module to one or more new modules. A module may be replaced by the same module with a different state, for example a different size or morphogen concentration, or by two or more new modules, when cells divide. Context sensitive rules can be used to simulate inter-module communication [62].

Figure 7.6 shows an example of an L-system model of the cyanobacterium Anabaena, that contains two cell types. The heterocyst cells fix nitrogen, which requires an anaerobic environment that must be isolated from the vegetative cells that perform photosynthesis. Between the heterocysts there are approximately ten vegetative cells [29]. The L-system rules define growth of the vegetative cells, which

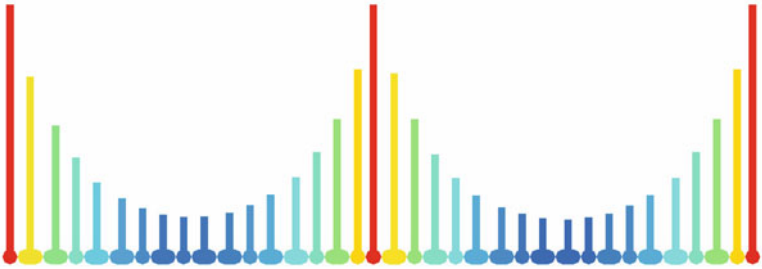


Fig. 7.6 Model of development of the filamentous blue-green algae Anabaena. Specialized heterocyst cells (red) are capable of fixing nitrogen, and are separated by vegetative cells. The heterocysts emit a diffusible signal, and when it drops below a threshold in a vegetative, they differentiate into heterocysts. As the filament grows, new heterocysts form as the space between them increases. Both the height of the bars and the color of the cells show the amount of diffusible inhibitor. The model is simulated with L-systems, and was adapted from [60]

divide when they reach a threshold size. Superimposed on this growing template is a morphogen simulation that controls cell identity. The heterocysts produce a morphogen that diffuses out to the vegetative cells and decays. As it diffuses down the line of cells, the concentration falls, and when it gets below a threshold concentration the vegetative cells differentiate into heterocysts.

The simulation is a simple example of a dynamic system with dynamic structure [24]. In models like this, where the underlying topology is one dimensional, it is straightforward to specify growth at the cellular level. Each cell is free to grow at its own rate without causing conflicts. This is not true in two or more dimensions [10]. For higher dimensions, growth must be specified globally to explicitly avoid spatial conflicts, or if defined locally at the cellular level, conflicts must be reconciled through a physically-based simulation [40, 44]. The important distinction between these two possibilities is that in the former, growth and shape is an input into the model, whereas in the latter shape becomes an emergent property of the model.

A simple method to model cell division in a growing tissue is to model only cell centers and use Voronoi regions to represent the cell boundaries. This is not ideal for modeling symplastic growth in plants since the rearrangement of regions that occurs during cell division creates unrealistic motions of cells (Fig. 7.7). Nonetheless, this can be an efficient method to create a growing cellular template when only the cell centers are needed to be stored [37], and the rearrangements are not likely to affect the outcome.

Models that have an explicit representation of the cell wall are better suited for modeling symplastic growth. In the simplest case, cells are defined as polygons with the vertex positions scaled to simulate uniform growth. The cells expand until a threshold area is reached and the cells divide. There are several possibilities for cell division rules. The most common is based on the division rule by Errera [18], who proposed that cells divide like soap bubbles, to find the surface of minimal energy. In 2D this is often simplified to the “shortest wall” rule (Fig. 7.8), although

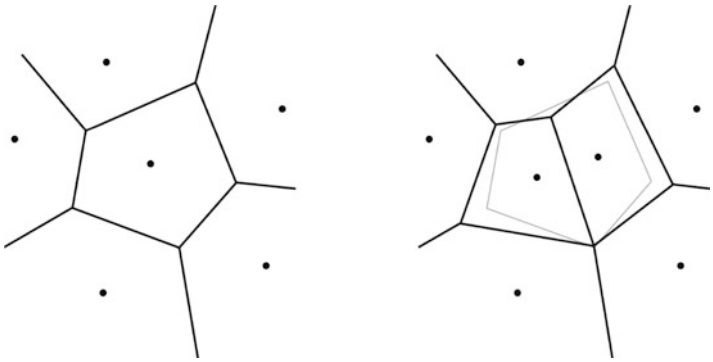


Fig. 7.7 Cell division in a tissue with cells represented by Voronoi regions. Unrealistic changes in the cells' geometry are introduced during cell division. Cell before division (left), and the new Voronoi regions after division (right). Note the shortening and stretching of some of the neighbor cell walls, and the occupation of some of the space of the neighbor cells by the newly created daughter cells

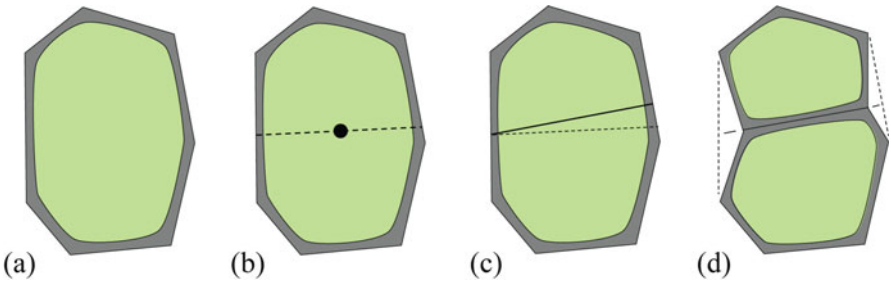


Fig. 7.8 Model of cell division. Dividing a cell through the shortest wall. (**a**) The cell before division. (**b**) Dotted line shows the shortest line passing through the center of the cell. (**c**) If an endpoint of the new wall is too close to an existing junction, it is displaced in order to avoid 4-way junctions. (**d**) The cell is "pinched" by shortening the dividing wall. This gives the daughter cells a more realistic shape. Adapted from [78]

a better approximation is a curved wall that joins existing walls at 90° [4]. Other rules include dividing the cell along the principal directions of growth [32] or along the directions of maximal stress [49, 51]. Nakielski [56] proposed that cells could be "pinched" after division, shortening the dividing wall to give the cells more realistic shapes. When pinching is included, the shapes of cells become very close to those of meristematic cells (Fig. 7.9).

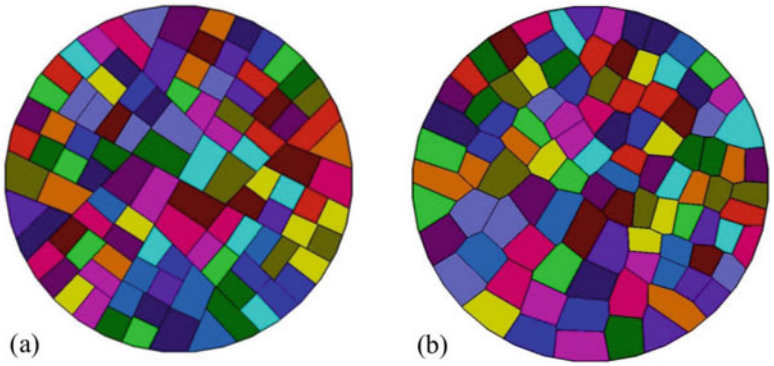


Fig. 7.9 Growing cell disk with dividing cells. Uniform growth displaces cell junctions outward, and the cells divide when they reach a threshold area. The shortest wall through the center of the cell is chosen for division. Color represent the cell label (unique identifier). (**a**) The algorithm without pinching. (**b**) The algorithm when the new cell wall is pinched (shortened up to 80% of its original length)

7.4 Polar Coordinates for Radially Symmetric Organs

For complicated shapes, modeling growth in more than one dimension is more difficult. In general, locally specified growth causes spatial tissue conflicts that require physically-based simulations to reconcile them [10]. Nevertheless, in some cases it is still possible to model growth descriptively. A common technique involves creating organ-centric coordinate systems, that are designed to simplify the specification of growth [32, 56, 77]. In this section the technique is used to model the plant apical meristem at the tip of growing shoots.

The plant shoot apex has a dome-like shape and consists of a group of stem cells at the tip, surrounded by rapidly growing and dividing cells in the peripheral zone where new organs form. The study of the shoot meristem has a long history, as the placement of organs initiating there determines the phyllotaxis of the plant, giving rise to striking geometrical patterns [45]. Since the molecular processes leading to the positioning of new organs are thought to occur largely in the surface layer of cells [64, 77], the simulation of patterning in the shoot apex can be simplified to a cellular surface.

An organ-centric coordinate system is defined for the shoot apex as a surface of revolution, generated by rotating a profile curve around a longitudinal axis (Fig. 7.10a, b). In addition to its three-dimensional coordinates, a point on the apex surface is defined by two local coordinates (θ, a) in this polar coordinate system. θ is the angle of rotation about the longitudinal axis of the apex, and a is the distance from the apex tip along the apex surface. A mapping function is defined to map the polar coordinates $S(\theta, a)$ to Cartesian coordinates (x, y, z):

$$S(\theta, a) = (x(a)\cos(\theta),\ \ y(a),\ \ x(a)\sin(\theta)). \tag{7.7}$$

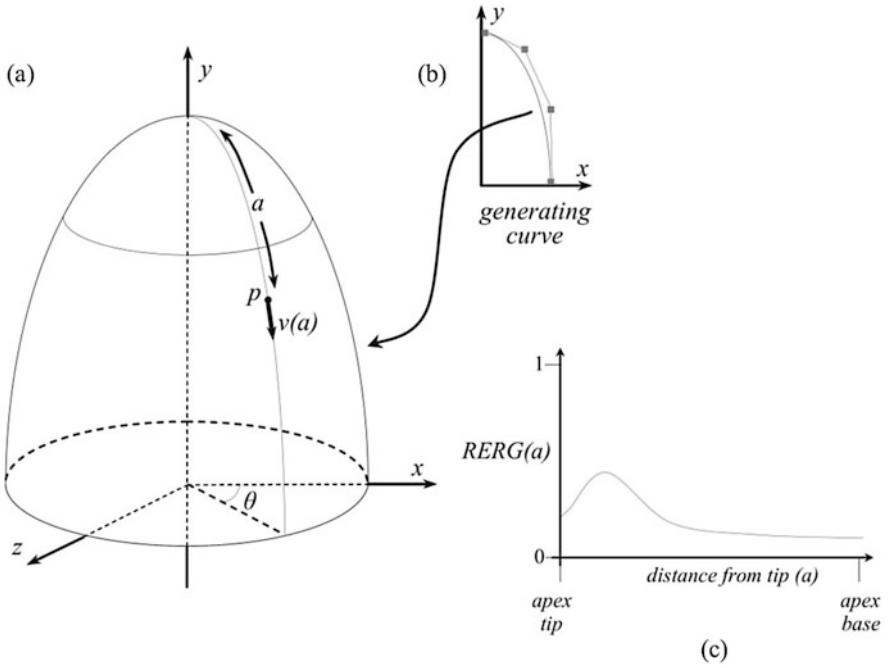


Fig. 7.10 A descriptive model of a growing shoot apex based on polar coordinates. (**a**) The shoot apex surface is defined as a surface of revolution from the generating curve (**b**). A material point P embedded in the surface with local coordinates (θ, a) moves away from the apex tip with the velocity $v(a)$. (**c**) The Relative Elemental Rate of Growth (RERG) is specified as a function of the distance from the tip. Adapted from [78]

Growth is simulated by moving points down the surface away from the tip by "growing" the local coordinate a. The growth function, RERG(a), defines the relative elemental rate of growth [32, 56] on the surface away from the tip as a function of the distance from the tip (Fig. 7.10c). The velocity $v(a)$ with which a point $P = S(\theta, a)$ moves away from the apex tip is used to displace points during a simulation step of growth. This velocity along the generating curve is given by:

$$v(a) = \int_0^a \text{RERG}(a)da. \tag{7.8}$$

In this coordinate system, growth occurs in only one dimension, the local coordinate a. The shape of the generating curve and the angle θ determines the growth in Cartesian coordinates. Note there is no change in angular position due to growth. The coordinate system provides a means to model growth that allows for its local specification, although it has the restriction that the growth rate must be radially symmetric. In the *Arabidopsis* shoot apex the growth rate has the shape

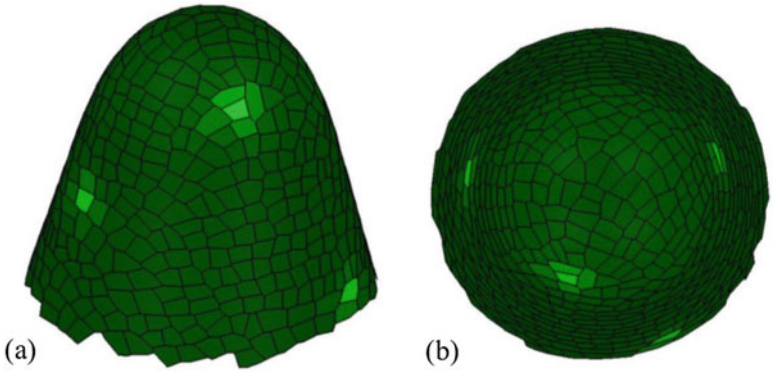


Fig. 7.11 A reaction-diffusion model of phyllotaxis on a growing shoot apex. As the surface grows, cells are displaced away from the tip, opening up space for new organs to form. (**a**) Side view with peaks in activator concentration (bright green) marking the locations where organ primordia would initiate. (**b**) Top view showing spiral pattern that appears when more than one previous primordium affects the location of the newly forming one

shown in Fig. 7.10c, with reduced growth in the central zone at the tip, and increased growth in surrounding peripheral zone.

Phyllotactic patterning is thought to rely on existing organs preventing the formation of new ones nearby [34]. Although molecular work points to a transport-feedback patterning mechanism related to that used for vein formation [37, 77], reaction-diffusion models have also been proposed [53]. In principle, any spacing mechanism on a growing shoot tip would suffice [30], with spiral phyllotaxis appearing when more than one previous primordium influences the position of the newly forming one. Figure 7.11 shows a simulation using the activator-inhibitor system defined in Sect. 7.2 with the same cell division model as described in Sect. 7.3. Activator peak formation is suppressed in the central zone by eliminating background production there. As the apex grows and cells move down the surface structure, space opens up allowing new peaks to form.

7.5 The Hejnowicz Coordinate System for the Root Apex

Another plant organ for which an organ-centric coordinate system has been proposed is the apical meristem in the root tip [31]. The root apex has a highly organized structure in many plant species. In *Arabidopsis thaliana* it is composed of a quiescent center near the root tip, with the adjacent stem cells that divide slowly and create a layered structure. Near the root tip the cells in the different layers divide, before switching to an endoreduplication cycle in the elongation zone where the bulk of cell expansion driving root growth occurs. In a longitudinal section of the root apex, cells are arranged in files that converge towards the quiescent center [31, 57].

This property is conserved as the root matures, although there are changes in root width and length of the meristematic area as the root matures during development.

The organization of the cell files results from the cells growing along the cell file direction and dividing parallel or orthogonal to it [58], with the principal directions of growth (PDGs) aligning with the cell files. This led Hejnowicz [31] to propose an organ centric system that he called a Natural Coordinate System for the root. Similar to the polar coordinate system for the shoot (Sect. 7.4), the system maps the parametric coordinates (u, v) to Cartesian coordinates (x, y). The function for this mapping is as follows:

$$x = 2\arctan(\tanh(u)\tan(v)) \tag{7.9}$$

$$y = \ln(\cosh^2(u) - \sin^2(v)) \tag{7.10}$$

Figure 7.12a shows the parametric coordinates plotted in the (x, y) plane, with the anticlines and periclines representing lines with fixed u or v in this mapping. The anticlines and periclines approximately align with the PDGs and the directions of cell division. Equations (7.9) and (7.10) can be scaled with a suitable multiplicative factor in order to adapt the coordinate system to the physical dimensions of a given root apex.

Growth is implemented by specifying the growth rates in the parametric coordinates (u, v). Hejnowicz [31] proposed dividing the root apex into 4 zones that have different growth rates although care must be taken that the growth functions make a continuous displacement field, preferably with continuous derivatives, at zone boundaries. In the simulation model shown in Fig. 7.12b, growth only occurs along the longitudinal direction (u coordinate) in the area above the green line in Fig. 7.12a. Note that this still results in displacement of points in both the x and y directions, due to the curvilinear coordinate system.

As with the previous templates, cells are represented as polygons, enlarge due to growth, and divide when they reach a threshold area. Hejnowicz proposed that cells in the root divide along one of the principal directions of growth [31], which can be determined by taking the derivatives of Eqs. (7.9) and (7.10). In the simulation shown in Fig. 7.12b the shortest of the two directions was chosen for the dividing wall, although this choice in the *Arabidopsis thaliana* root apex is thought to largely depend on cell type [8]. Once the new intersection points are found in Cartesian coordinates, a numerical inversion is required to evaluate their parametric counterparts. In the simulation in Fig. 7.12b a multidimensional root finding algorithm from the GNU Scientific Library [23] was used to obtain values for the parametric coordinates for newly inserted vertices. The Jacobian of the coordinates transformation is singular at the origin and along the line $u = 0$, so special care is required during the root search. To address this point, the symmetry about the y-axis of the parametric coordinates was exploited so that only positive solutions for v are computed.

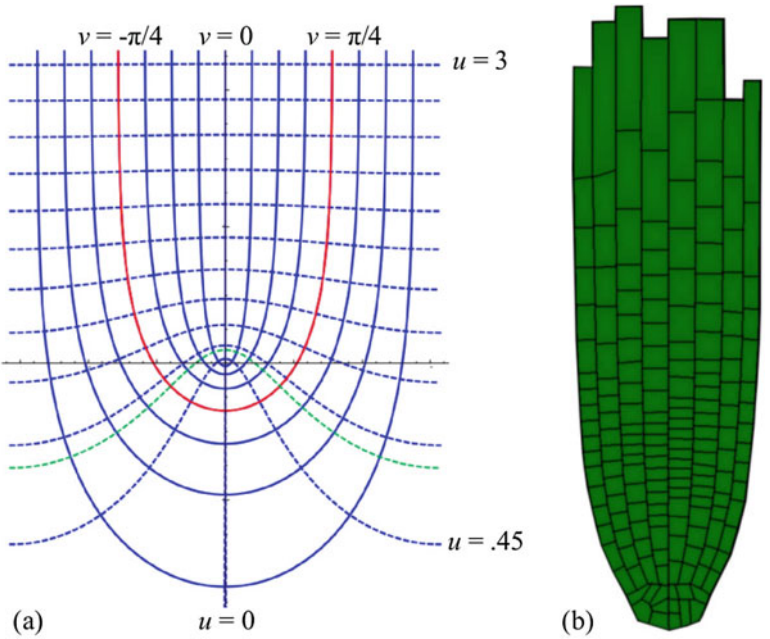


Fig. 7.12 A Natural Coordinate System for the *Arabidopsis thaliana* root apex. (**a**) The mapping of the parametric coordinates (u,v) plotted in the (x, y) plane. The lines are the anticlines and periclines that represent fixed values of the parametric coordinates u and v. The red line ($v = \pi/4$ and $v = -\pi/4$) defines the area of the system used for the simulation. The quiescent center is located approximately at the origin. After Hejnowicz [31]. (**b**) A simulation model of the root meristem with growth and cell division. Growth is implemented by growing the u coordinate in the area above the green line in (**a**). Cells divide along the shortest wall of the two principal directions of growth (PDG). At a specified distance from the root tip, cells stop dividing, separating the meristem from the elongation zone

Similar to the polar apex surface described in Sect. 7.4, the parametric coordinate system makes it straightforward to define a velocity field to model descriptive growth of the root apex. The model in Fig. 7.12 uses an explicit growth function for the u coordinate, but it is possible to determine the growth from a morphogen concentration determined by a molecular process in the cells in the same simulation. This provides a framework for the local specification of growth in a 2D tissue without conflicts, because the growth becomes essentially 1D, although it would require a single longitudinal growth rate for all cell files for a given u position. Possibilities include using a single cell file to control growth, such as the epidermis, or using an average value from multiple cell files.

7.6 A Key-Frame Approach Using B-Splines

The previous sections described organ-centric coordinate systems that simplify growth modeling by transforming Cartesian coordinates so that the principal directions of growth are axis aligned in the parametric coordinates. Another approach using coordinate system transformations is to leave the parametric coordinates fixed, and change the coordinate system itself over time [76]. The mapping from parametric coordinates to Cartesian changes over time causing growth, rather than changes in the parametric coordinates themselves (see Fig. 7.13b). Input for the method is provided as exemplar shapes of the organ at important time points, or key frames, in development. Intermediate time points are determined by interpolation. From this perspective, these methods can be considered as a form of key-frame animation of the evolving tissue geometry.

The leaf structure in Fig. 7.13 uses B-splines to represent the leaf shape at different time points (key frames) during the simulation[68]. The surfaces were defined interactively using the splineEdit surface editor provided with the Virtual Laboratory (vlab) simulation environment [19]. B-spline surfaces were specified by a set of control points describing the overall shape of the surface. Linear interpolation of the control points was used to produce the leaf shape at any desired point between the key frames. By advancing time in small steps, a smoothly growing leaf surface is produced. Key-frame surfaces for leaf one of *Arabidopsis thaliana* are shown in Fig. 7.13c. These key-frames were constructed based on the leaf-shapes and data reported in [15, 73, 74]. Internal growth was interpreted by manually tracking landmarks, such as veins and convergence points, throughout development.

The B-spline surfaces used to model the leaf are two-dimensional parametric surfaces embedded in three-dimensional space (Fig. 7.13b). Points on the surface are represented by parametric coordinates (u, v) which are mapped by B-spline surface evaluators to positions in Cartesian space (x, y, z). As the simulation progresses and the leaf changes shape, a (u, v) coordinate at the tip of the leaf will remain at the tip even though its actual position in space may change considerably.

The same cellular structure described in Sect. 7.3 was implemented on the growing leaf in Fig. 7.13a, along with the shortest-wall rule for cell division. Beginning with a single cell that covers the entire initial shape, cell division proceeds until there are no cells left above the threshold area. As the simulation proceeds, the size and shape of the leaf changes, causing the polygons that define the cells to enlarge. When cells reach the threshold area, they are divided. The gradual division of cells creates more space in which the reaction-diffusion system can organize additional activator peaks.

The library that implements the B-spline surface has two main functions: one to compute the (x, y, z) coordinates of a point on the surface given its (u, v) coordinates, and the other to perform the inverse operation. As with the shoot and root apex models in Sects. 7.4 and 7.5, there is in general no closed formula to compute the (u, v) coordinate given a particular (x, y, z) position that may or may not be on the surface. This inverse operation is required whenever new points

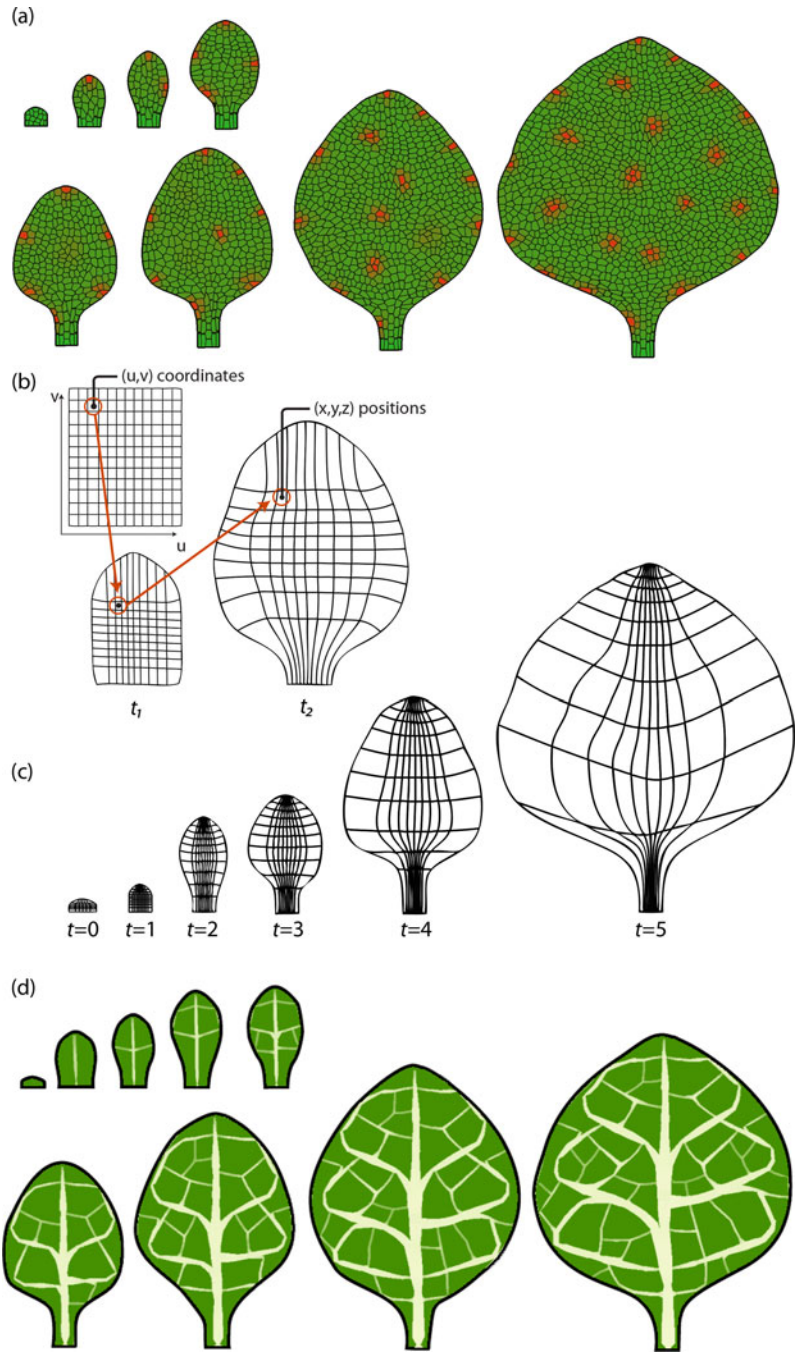


Fig. 7.13 Simulating growth using key framing with B-spline surfaces. (**a**) A reaction-diffusion model on a growing-leaf. Activator concentration is shown in red and is set to zero in cells

are added to the model (e.g., when cells divide) and is implemented by using a steepest decent search algorithm on the evaluator function. Steepest decent requires a set of parameter values, an initial guess from which to begin the search. For points introduced during cell-division, the (u, v) positions of neighboring points provide a good initial value for the descent. Although B-splines are employed here, the method is compatible with most control point based surfaces, such as Bezier surfaces[76].

This method has also been used for the geometric vascular patterning algorithm described in Runions et al. [68, 69]. This model nicely illustrates how growth can feed back on the emergent pattern formation of the vascular network (see Fig. 7.13d). Growth causes space to appear, that results in new sources of a morphogen that attracts the growing veins, reproducing the events seen *in vivo* during initiation of the midvein and primary loops [74]. The iterative formation of prominent secondary loops at the base of the leaf blade results from increased growth of this region relative to the leaf tip. Although the model requires a relatively precise spatial representation of growth in the leaf to reproduce the correct patterns of venation, the growth itself is not an emergent property of the model.

7.7 Data Driven Growth

Another approach to descriptively model tissue growth is to extract growth rates directly from experimental data. By imaging the same plant organ or tissue at consecutive time points, time-lapse data can be obtained and used to calculate growth rates [2, 41, 46, 67], which are then used to drive the tissue growth in a simulation model. As with the B-spline methods a parametric coordinate system is created, which then evolves to simulate growth, rather than changing the parametric coordinates themselves.

The model input begins with a time-lapse data series, in this case from a section on *Arabidopsis thaliana* cotyledon, the initial leaf-like organ that is present in the embryo. Figure 7.14a shows the segmentation of the cells on a 2D surface of the cotyledon made using the MorphoGraphX [13] software from 3D confocal image stacks. After this process is performed for 2 time points and the lineage tracking of the cells is created, the growth rates of the cells are computed (Fig. 7.14b).

Fig. 7.13 (continued) at the leaf base. As space opens up in the leaf blade due to growth, new peaks in activator concentration appear. (**b**) Growth occurs due to the mapping of coordinates in the parameter-space (u, v) to three-dimensional space (x, y, z) as shown. Positions at time steps between key frames are found by linear interpolation. (**c**) Multiple key-frames can be used for simulations depicting more complicated growth sequences. (**d**) A geometric model of leaf venation. The pattern of emerging veins depends on how the growth is distributed in the leaf blade, with secondary loops forming at the base of the growing-leaf where the growth rates are higher

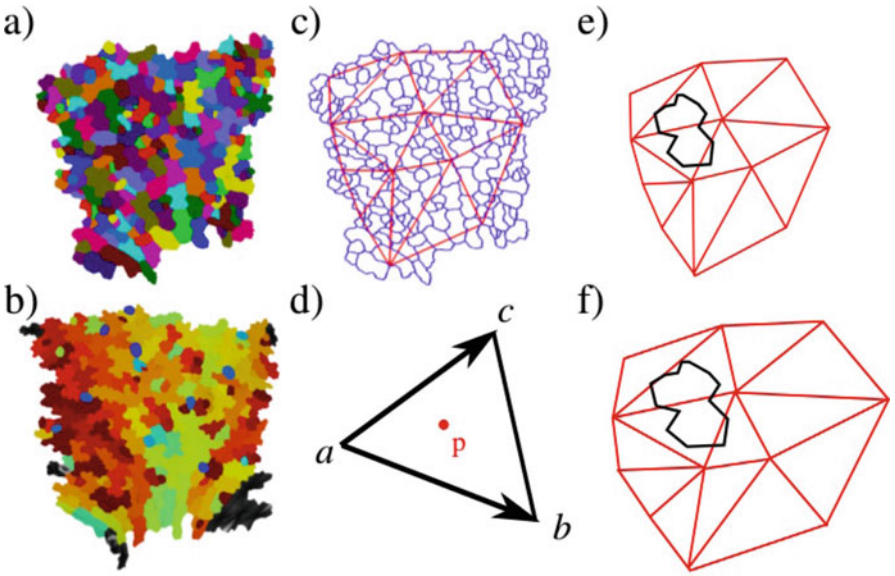


Fig. 7.14 Modeling data-driven tissue growth. (**a**) Surface mesh of the initial timepoint of experimental data of an *Arabidopsis thaliana* cotyledon segmented into cells. (**b**) Later time point with heat map according to the area increase of individual cells. (**c**) Surface mesh with cell contours (blue) of (**a**) with material points and their Delaunay triangulation (red). Growth parameters are assigned to the material points and the growth of the remaining tissue is interpolated using barycentric coordinates. (**d**) The principle of barycentric coordinates: Given a triangle and one of its corners as origin, a coordinate system can be created using the adjacent edges as axis vectors. Any point p then has unique coordinates within this system. (**e**) The Delaunay triangulation of the modeled tissue in the initial time step with an example cell in black. (**f**) This tissue and all its cells are grown to the final time step (only one example cell shown)

The next step is to create a coordinate system to map model vertex positions from one time point to another. This is done by selecting material points in the surface that mark identical positions on the plant tissue at different time points. Between these time points the values are interpolated, as models typically require a finer time resolution than is practical experimentally. Often one point is fixed and the remaining points rotated to remove rigid body translations and rotations.

The material points can be taken from any landmark on the organ surface, however if the surface is already segmented and lineage tracked, then the correspondence between cell centers or cell junctions is known. Since plants have symplastic growth, cell movement and sliding cannot occur, and the cell junctions represent material points. Here a few cell junction points have been selected in the first time point as material points. The next step is to create a triangulation of the material points. If the sample is reasonably flat, it can be projected into 2D and the Delaunay triangulation can be used (Fig. 7.14c). This triangulation is then used for both time points.

As with the previous models of descriptive growth defined in Sects. 7.4–7.6, a coordinate system is defined for each point on the surface. This coordinate system

is defined for each triangle, using barycentric coordinates. One of the points of the triangle is selected as the origin and the other two points define the axis of the coordinate system. Each point inside (or outside) the triangle can be defined by a unique coordinate (u, v). For the triangle (a, b, c), let a be the origin of the barycentric coordinate system, with $b - a$ and $c - a$ representing the axes (Fig. 7.14d). Any point p can be represented as:

$$p = a + u(b - a) + v(c - a). \tag{7.11}$$

To use the coordinate system in a simulation, each model point in Cartesian coordinates is mapped to the barycentric coordinate of the triangle t to which it belongs (t, u, v). If the point lies outside the domain, the closest triangle is used. Note the representation is not unique, since the coordinates are defined for all points for each triangle. However since the growth rates and direction are determined at the triangle level, it is important that each point be assigned to the triangle that contains it or the closest one possible. As the simulation proceeds, the triangle positions (a, b, c) are interpolated linearly between the time points. As with the B-spline simulation in Sect. 7.6 the parametric coordinates (t, u, v) remain the same, and the mapping to Cartesian coordinates evolves, simulating growth.

The triangulation of the surface represents a piecewise linear mapping of the space from the first time point to the second, with a single growth rate for all the points in each triangle. Therefore, the number and density of material points in relation to the homogeneity of growth is important. Areas with a high variance in growth rates will require more material points. Resolution in time is a consideration as well, with more time points better capturing the temporal variability in growth. In addition to linear mappings, higher order methods such as splines can be used, which can provide a smooth transition in growth rates over the tissue.

7.8 Mass Spring Models

Descriptive growth models are useful to explore patterning mechanisms on growing domains, where the growth and changing shape is provided as an input to the model. However many questions in morphogenesis aim to understand how genes control shape, with the shape of organ becoming an emergent property of the model [33, 44, 72]. Gradients of morphogens created by patterning processes control the expression of growth factors that drive the expansion of tissue locally, resulting in the final shape of the organ. The requirement to specify growth locally presents problems when simulating growth in more than one dimension, as conflicts will develop if the growth is not carefully coordinated across the tissue [40, 63]. To reconcile these conflicts, a mechanical model of the tissue is required [65].

There are several methods to simulate physically-based growth, the simplest of which are mass-spring systems [12]. In a mass-spring model the tissue is discretized into vertices that represent mass particles, often assumed to have uniform

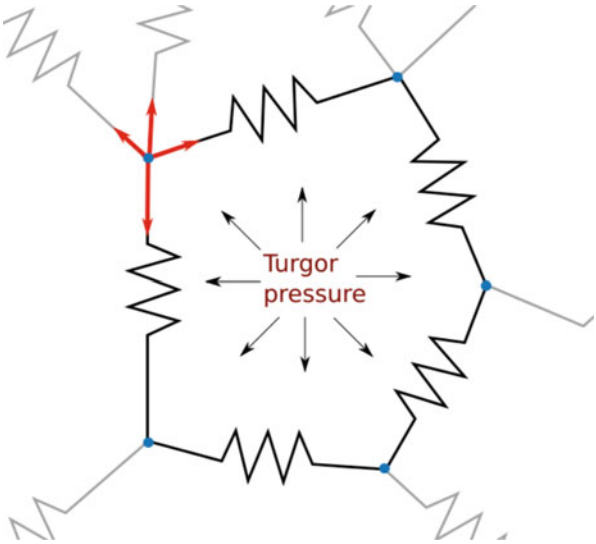


Fig. 7.15 Schematic representation of a mass-spring system approximating a tissue of cells. The system is comprised of vertices that represent mass particles coinciding with cell wall junctions (blue circles) connected to each other by springs that represent cell wall segments. Cell turgor pressure within the cells is applied to the cell walls generating force that is transferred to the vertices. This causes the springs to stretch, exerting force on the vertices (red arrows). An equilibrium is reached when the forces from the springs balance the forces from turgor pressure

mass. These particles are connected by springs, creating a mass-spring network. The discretization can be an arbitrary triangulation of the plant organ, or may explicitly represent the structure of cells. Since the internal structure of plant cells (cytoskeleton) is very weak compared to the cell wall, a cellular discretization typically only includes springs along the walls, and not the interior (see Fig. 7.15). The state of the system is defined by the position of each vertex and the force acting on it. The force F_v acting on a vertex v at position p_v due to springs is calculated by adding forces exerted by the springs connected to it:

$$F_v = \sum_{n \in N_v} k \left(\frac{\| p_n - p_v \|}{l_{v \to n}} - 1 \right) \frac{p_n - p_v}{\| p_n - p_v \|} \tag{7.12}$$

where N_v is the neighbors of vertex v, p_n is the position of a neighbor vertex n, k is the spring constant per unit length, and $l_{v \to n}$ is the rest length of spring joining vertex v to neighboring vertex n. The expression within the brackets is positive when the distance between v and its neighbor n is greater than the rest length, and negative if it is smaller. Note that for a given difference from the rest length, the magnitude of the force is reduced as the rest length of the spring increases. This magnitude is multiplied by the normalized vector $(p_n - p_v)/\| p_n - p_v \|$ and by k, which represents the spring stiffness, to give the forces.

In addition to the forces on a vertex due to springs, other forces can be added to the mass-spring model, such as boundary forces or gravity. In plant cells internal turgor pressure generated internally by the cells is usually included, and acts as a uniform force perpendicular to the cell walls. Once all the forces on the vertices are calculated, the steady-state of the system is determined. There are several methods to do this. One option is to calculate the velocity c and position p of vertex v based on the total force F_v acting upon it:

$$\frac{dc_v}{dt} = (F_v - \varsigma c_v)/m \tag{7.13}$$

$$\frac{dp_v}{dt} = c_v \tag{7.14}$$

where ς is the damping constant and m is the mass of vertex v. The damping constant is required or the system will oscillate indefinitely and never reach equilibrium.

Equations (7.13) and (7.14) are solved for the steady state, which is indicated when the maximum total force acting on any of the vertices is zero. Note that the steady state does not depend on the damping coefficient ς, or the mass m if there is no gravity. In this case a system of equations can be constructed directly from Eq. (7.12) and solved for $F_v = 0$ for all vertices. Since Eq. (7.12) is non-linear this typically requires an iterative approach, such as Newton-Raphson.

After finding the steady-state of the system, a step of growth is performed. One approach is to directly increase the rest lengths of the springs, depending on how much the cell walls should grow. However this method can become problematic, as there is no way to release residual stresses that tend to accumulate over time. A better approach comes from the theory of plant cell wall growth proposed by Lockhart [50]. In Lockhart's model growth depends both on growth factors and turgor pressure that induces strain in the cell wall. Thus the increase in rest length of the springs depends both on the amount of growth factor, combined with the strain in the spring representing the cell wall due to turgor:

$$l_{v\to n}^{t+\Delta t} = l_{v\to n}^{t} + G\Delta t \left(\left\| p_n^t - p_v^t \right\| - l_{v\to n}^{t} \right) \tag{7.15}$$

where $l_{v\to n}^{t+\Delta t}$ denotes the new rest length of the spring connecting vertices v and n, $l_{v\to n}^{t}$ is the current rest length, G is the growth factor, and Δt is the time step. Since plant cell walls normally do not shrink, growth is usually set to zero if the second term on the right-hand side of Eq. (7.15) is negative. Sometimes a yield threshold is also used as in the original Lockhart formulation. After growth is applied a new steady state is found and the simulation proceeds. The simulation loop may include patterning processes to direct the growth, or these may be specified as morphogen gradients at the start of the simulation.

Figure 7.16 shows a mass-spring system starting from the disk of cells presented in Sect. 7.2. The same Meinhardt activator-inhibitor system is used to drive growth, with the activator concentration in each cell used as the growth factor G in

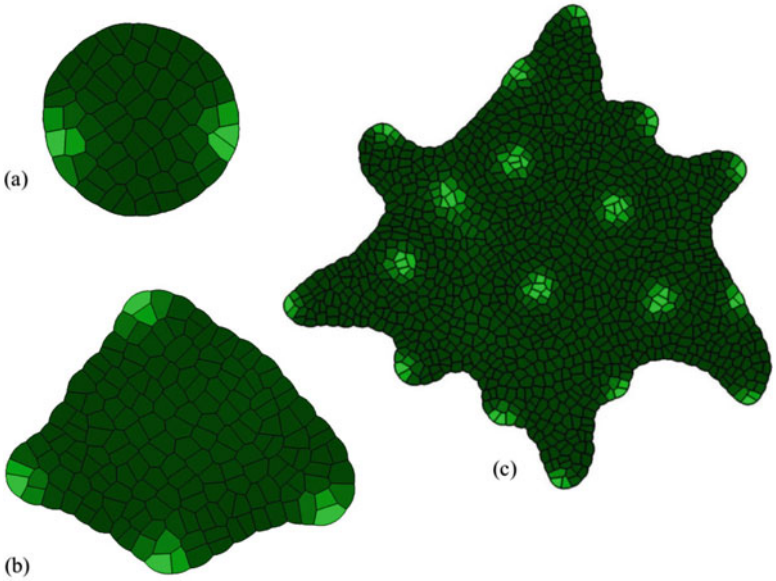


Fig. 7.16 Reaction-diffusion on a growing domain simulated with a mass-spring system. (**a**) In the initial configuration there is enough space for the activator-inhibitor system to create two activator peaks (shown in bright green). (**b**) Model after a period of growth. Growth is implemented by relaxing springs based on activator concentration. This causes increased growth in the areas of high activator, causing them to grow out from the structure. As more space is produced by growth, new peaks form. (**c**) When the structure becomes large enough, peaks form in the interior as well

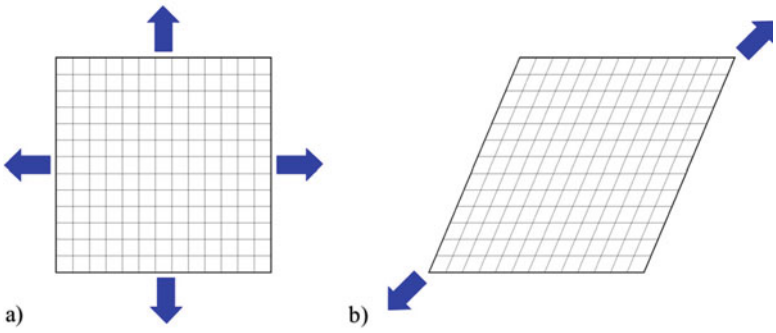


Fig. 7.17 The effect of spring orientation on tissue stiffness. (**a**) In a grid of springs, the stiffness in the directions parallel to the springs is the same. (**b**) In an oblique direction, the stiffness can be much lower

Eq. (7.15). Walls between cells use the average of the concentrations from their adjacent cells. The activator-inhibitor system initially makes two peaks on the domain, which drive increased growth in these areas. As the structure expands, more peaks form in the available space. Note that shape is an emergent property of the model, and it depends on where the peaks form.

Mass-spring systems suffer from an intrinsic short-coming that limits their ability to model growing plant tissues. In general the properties of the material modeled depend on the discretization. For example, consider an array of squares with uniform spring length and stiffness. When stretched in either direction parallel to the springs, the tissue has uniform stiffness. However, stretched in an oblique direction the stiffness is much lower (see Fig. 7.17). This dependence on spring geometry also appears when modeling growth, especially if it is anisotropic.

7.9 Modeling Growth with the Finite Element Method

A more general and mechanically accurate approach to physically-based plant tissue modeling is based on continuum mechanics solved numerically with the Finite Element Method (FEM) [83]. With FEM modeling the solution does not depend on the discretization of the domain, giving an accurate representation of both isotropic and anisotropic materials. It is also much more suitable for modeling continuum formulations of growth [27, 65], which are required for whole organ modeling.

Given a body with a known reference configuration and an infinitesimal specified growth amount, it is possible to decompose the total deformation gradient $\mathbf{F}$ into a pure growth tensor $\mathbf{G}$ and an elastic tensor $\mathbf{A}$, that may be required to ensure that the body grows in a symplastic and continuous manner. In general it is necessary to describe the finite growth process as the composition of many smaller growth steps:

$$\mathbf{F} = (\mathbf{A}_n \cdot \mathbf{G}_n) \cdot (\mathbf{A}_{n-1} \cdot \mathbf{G}_{n-1}) \cdots (\mathbf{A}_1 \cdot \mathbf{G}_1) \tag{7.16}$$

The initial configuration is an input to the mechanical model, whereas the growth tensor $\mathbf{G}_i$ may be specified upfront, or be calculated from a morphogen gradient or patterning simulation [40]. $\mathbf{A}_i$ is computed for each step of the simulation.

It is possible to implement growth by applying the growth tensor directly to the reference configuration of the elements. This "grows" the individual elements as if they were not connected to surrounding tissue. Kennaway et al. [40] refer to this as specified growth, the amount a small portion of tissue would grow if not constrained by neighbor tissue. After the specified growth has been applied, the equilibrium for the tissue is computed using the newly grown reference configuration. This resultant growth [40] in general does not match the specified growth, but includes residual stresses within the tissue to account for spatial differences in growth rates. In simulations where mechanical feedback is not required, these stresses are often released, reflecting a remodeling process that relaxes the cell wall. This is done by simply updating the reference configuration to the newly deformed configuration.

The growth tensor $\mathbf{G}$ can specify different growth amounts for different directions in order to model growth anisotropy. It can also vary depending on the position in the tissue (growth inhomogeneity), integrating morphogen gradients that specify the growth rates and directions of anisotropy [33, 40, 44]. These gradients may be fixed at the start of the simulation and follow the material points during their

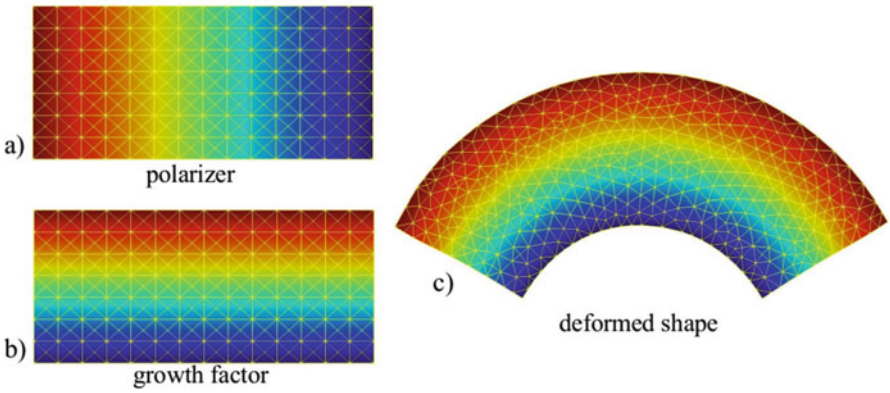


Fig. 7.18 A simple growing canvas controlled by gradients of morphogens. (**a**) The gradient of a morphogen determining the tissue polarity. (red high, blue low). (**b**) The gradient of a morphogen that promotes growth in the direction parallel to the polarizer gradient. (**c**) The simulation after many steps of growth. Since the morphogens are fixed to material points at the beginning of the simulation, the growth factor gradients deform with the tissue as it grows

deformation or they may be computed at each time point, allowing the shape of the tissue to feed back on the growth [40]. This approach has been used to model the development of whole plant organs, such as the leaf or petal of *Arabidopsis thaliana* [44, 72].

In order for a plant tissue to grow anisotropically, there must be a mechanism to determine a reference direction or polarity to orient preferred direction of growth [11]. Kennaway et al. [40] propose the existence of morphogens that act as tissue polarizers, to determine this direction. Other morphogens then act as growth factors that enhance growth either parallel or perpendicular to this direction. Figure 7.18 demonstrates this idea on a simple rectangular template. The morphogens that determine the tissue polarity and growth rates are assigned at the beginning of the simulation and deform with the template as it grows. The polarizer is assigned as a linear gradient from left to right (Fig. 7.18a). The growth factor drives anisotropic growth in the orientation of the polarizer, with no growth in the perpendicular direction. A gradient of growth factor is specified from top to bottom (Fig. 7.18b). As the simulation proceeds, the increase growth at the top causes the rectangle to grow into a curved arch.

Figure 7.19 shows a simulation of a finite element model of growing leaf similar to that proposed by Kuchen et al. [44] with a few simplifications. At the beginning of the simulation, a polarizer morphogen diffuses from a source at the base of the leaf primordium to a sink at the tip (Fig. 7.19a). The gradient of the concentration of the polarizer determines the polarity of the leaf. Two more gradients determine how much the leaf grows parallel to the polarizer (Fig. 7.19b) and perpendicular to it (Fig. 7.19c). The growth simulation is then performed in a series of small growth steps. Since the gradients are set at the start of the simulation, the gradient follows the material points as the tissue deforms. Over time the shape of the initial primordium develops into the shape of a mature leaf (Fig. 7.19d).

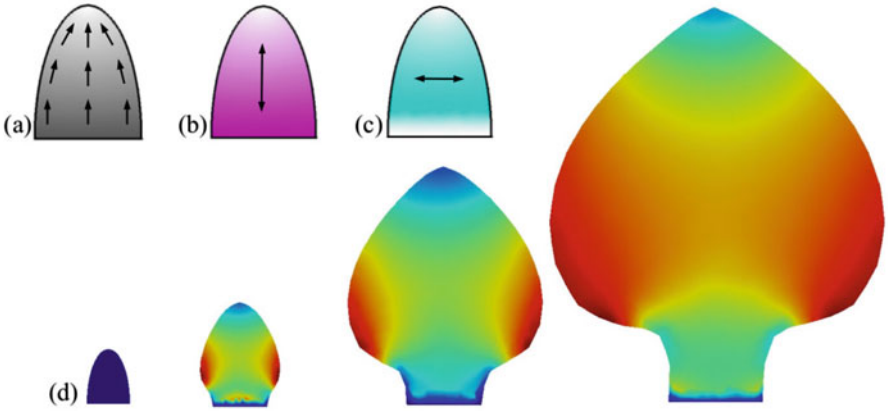


Fig. 7.19 A simple finite element model of a growing leaf controlled by morphogens. (**a**) Visualization of the polarizer gradient on the initial leaf primordium shape (white low, gray high). The arrows indicate the direction of the gradient of the polarization field. (**b**) Gradient of the growth factor that determines growth parallel to the polarizer field (white low, magenta high). (**c**) Gradient of the growth factor that determines growth perpendicular to the polarizer field (white low, cyan high). (**d**) Simulation of the growing leaf over time. The color represents the areal strain from the previous time point (blue low, red high)

7.10 Hamiltonian-Based Methods

Hamiltonian based methods implement cellular behaviors based on a generalized energy function. The most prominent among these is the Cellular Potts Model (CPM) developed for the simulation of animal cells where cell movement and adhesion are important factors. CPM is a grid-based method that differs from all the previous models in that space itself is discretized (Eulerian approach) rather than the model elements (Lagrangian approach). In the CPM each grid point has a label and cells are defined as regions on the grid with the same label (see Fig. 7.20). Changes in cell geometry that occur through growth or cell movement are implemented with a Metropolis algorithm that attempts to copy the label of a given grid point to a randomly chosen adjacent grid point [26]. This change occurs with some probability that depends on the value of the global Hamiltonian energy function. This function contains terms that control individual aspects of the cells' geometry, state, and interdependence with neighbors. In the basic CPM formulation the Hamiltonian is a function of cell area and the boundary length for a 2D model, corresponding to volume and surface area in a 3D model:

$$H = \lambda_V \sum_{\sigma} (v_\sigma - V_\sigma)^2 + \lambda_B \sum_{i} \sum_{j \in N_i} \tau\,(i, j) \tag{7.17}$$

where H is the total energy of the system, σ is a biological cell (set of grid points with the same label), v_σ is the volume of cell σ, V_σ is the target volume, λ_V is

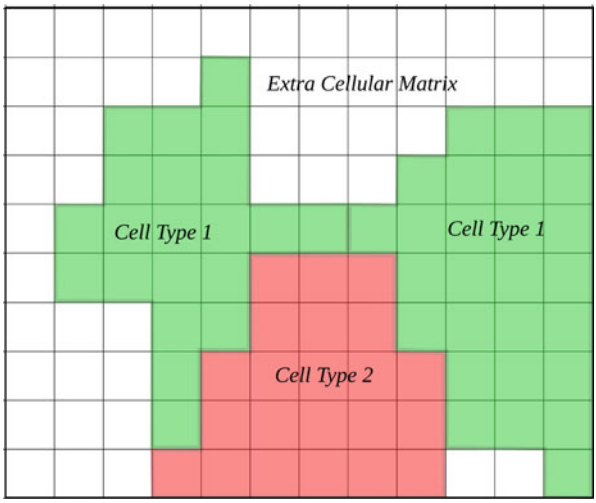


Fig. 7.20 Overview of typical interactions in a Cellular Potts Model. Grid squares at the border of cells interact with different neighboring cell types as well as the extra cellular matrix

a coefficient determining the strength of the volume constraint, i is a grid point, $j \in N_i$ are the neighboring grid points of i, $\tau(i, j)$ determines the boundary cost (reduced for adhesion, zero if $i = j$), and λ_B is a coefficient determining the strength of the boundary constraint. Note that the volume constraint is determined at the cell level, represented by all the grid points that share a label, whereas the boundary constraint is specified on pairs of adjacent grid points belonging to different cells. The boundary energy $\tau(i, j)$ may depend on the cell type, as adhesion can be different between different cell types.

When the target volume is constant, cell growth is driven by cell division, as daughter cells try to reach their target volume. Alternatively, growth can be implemented by incrementally increasing the target volume, with the cells dividing at a threshold volume. Various other cell behaviors can be implemented in a CPM model by the addition of extra terms to the Hamiltonian. For example, chemotaxis can be modeled by adding a terms in that couples the system to an external concentration field.

The simulation loop proceeds by randomly selecting a grid point and attempting to change its label to that of one of its neighbors, again randomly selected. The label transition is accepted if it reduces the total energy, or if it increases the energy, with a probability that depends on a Boltzmann distribution on the change in energy of the Hamiltonian. The exponential nature of the distribution makes large increases in energy very unlikely to be accepted. Nevertheless, this behavior helps to prevent simulations from becoming trapped in local minima.

The probability of a change being accepted P is defined as:

$$P = \begin{cases} 1 & \Delta H \leq 0 \\ e^{-\Delta H/T} & \Delta H > 0 \end{cases} \tag{7.18}$$

where ΔH is the change in value of Hamiltonian resulting from the label transition and T specifies the intrinsic cell motility. If more than one cell type is implemented, T can depend on cell type as $T(\sigma_i)$. The resulting CPM simulations are inherently stochastic and characterized by fluctuating cell borders, which is often the case in vivo for animal systems.

While the Hamiltonian-based CPM methods have been extensively applied in modeling animal development, the method has also been adapted to vertex-based modeling for plant development [14, 54]. A vertex model is better suited for the symplastic growth of plant cells, since it enables a more accurate description of cell walls and their physical properties. Cell wall yielding is captured in the model by insertion of new nodes once a cell wall element is stretched beyond a threshold. The notion of a target area is preserved, along with the idea of a cost for the boundary. Points are randomly chosen and displaced in random direction, and the update is accepted with the same probabilistic function as used for CPM models. The Hamiltonian for the system is defined as:

$$H = \lambda_A \sum_i (a_i - A_i)^2 + \lambda_M \sum_j \left(l_j - L\right)^2 \tag{7.19}$$

where indices i and j are summed over all cells and walls, respectively, a_i is area of cell i, A_i is the target area of cell i, λ_A is a coefficient determining the strength of the area constraint. l_i is the length of cell wall segment i, L is the target length of cell wall segments, and λ_L is a coefficient determining the strength of the wall length constraint. Growth is achieved through iterative updating of the target area of each cell based on its growth rate (Fig. 7.21).

The Hamiltonian-based approach is very general. For example, a mass-spring system can be formulated as an energy minimization problem and solved with similar methods. The approach also simplifies the modeling of diverse cell behaviors, that can be implemented by modifying terms in the energy function. With a fixed target area, cell division becomes the driver of growth, and the orientation of the

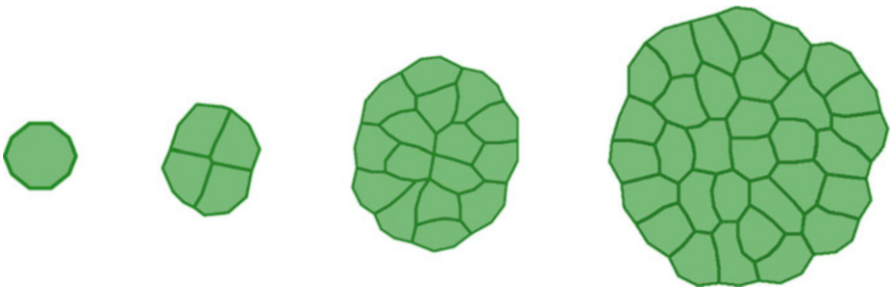


Fig. 7.21 A cellular plant tissue model using a Hamiltonian-based vertex model in the VirtualLeaf simulation environment [54]. Cells grow uniformly and divide according to the shortest wall rule once they double their area following a division

division wall will then control growth anisotropy [22]. Growth anisotropy could also be controlled by favoring cell elongation in a preferred direction, which could be specified by a morphogen gradient. A potential drawback, however, is that it is not difficult to create an energy function that does not relate directly to the underlying cellular biophysics. Nevertheless, Hamiltonian based methods do allow for the local specification of growth, with tissue shape an emergent property of the model.

7.11 Conclusion

Simulation modeling has become an increasingly important tool to understand questions in plant morphogenesis. At the heart of these processes are patterning mechanisms, that direct the differentiation of cells and the local control of growth and growth anisotropy. Often there is a tight connection between growth and the emerging patterns, as in the case of phyllotaxis [37, 77] or leaf vein formation [69]. Because of this, considerable work has been put into modeling patterning mechanisms on growing structures, which present challenges over static representations of plant tissues.

For many questions, the interaction of growth and patterning can be explored by using descriptive models of growth, which are easy to implement and computationally efficient. However the study of morphogenesis is ultimately about the creation of form, requiring shape to be an emergent property of the model. This requires the use of physically-based models such as mass-spring systems or finite element methods. Models for several plant organs have been developed by using FEM models operating on sheets of continuous tissue [33, 40, 44, 72], however these models do not include cells. This makes it difficult to integrate the models with patterning mechanisms that are specified at the cellular level, such as auxin transport-feedback processes [30, 76].

Improvements in microscopy methods are pushing towards more realistic models, with full-3D time-lapse data now possible in some tissues [20]. Although initial progress has been made on physically-based simulation of full-3D plant tissues with finite element models [3, 7], cell division remains a problem. Recent work on cell complexes [61, 81] looks like a promising framework for dynamic systems with dynamic structure in full 3D.

Acknowledgements Funding is gratefully acknowledged from the Bundesministerium für Bildung und Forschung grants 031A492 and 031A494, the Swiss National Science Foundation SystemsX.ch Plant Growth RTD, Human Frontiers Science Program grant RGP0008/2013 to R.S.S., Marie Skodowska-Curie individual fellowship (Horizon 2020, 703886) to A.R., and the Max Planck Institute for Plant Breeding Research, Cologne, Germany. Some sections of this chapter were adapted from the lecture notes of the Les Houches summer school of 2009 [75]. We would also like to acknowledge Przemyslaw Prusinkiewicz and the members of his lab for helping to formulate many of the ideas appearing in this chapter.

References

1. Abley K, De Reuille PB, Strutt D, Bangham A, Prusinkiewicz P, Marée AF, Grieneisen VA, Coen E (2013) An intracellular partitioning-based framework for tissue cell polarity in plants and animals. Development 140(10):2061–2074
2. Armour WJ, Barton DA, Law AM, Overall RL (2015) Differential growth in periclinal and anticlinal walls during lobe formation in arabidopsis cotyledon pavement cells. Plant Cell 27(9):2484–2500
3. Bassel GW, Stamm P, Mosca G, de Reuille PB, Gibbs DJ, Winter R, Janka A, Holdsworth MJ, Smith RS (2014) Mechanical constraints imposed by 3d cellular geometry and arrangement modulate growth patterns in the arabidopsis embryo. Proc Natl Acad Sci 111(23):8685–8690
4. Besson S, Dumais J (2011) Universal rule for the symmetric division of plant cells. Proc Natl Acad Sci 108(15):6294–6299
5. Bilsborough GD, Runions A, Barkoulas M, Jenkins HW, Hasson A, Galinha C, Laufs P, Hay A, Prusinkiewicz P, Tsiantis M (2011) Model for the regulation of arabidopsis thaliana leaf margin development. Proc Natl Acad Sci 108(8):3424–3429
6. Boudon F, Pradal C, Cokelaer T, Prusinkiewicz P, Godin C (2012) L-py: an l-system simulation framework for modeling plant architecture development based on a dynamic language. Front. Plant Sci. 3:76
7. Boudon F, Chopard J, Ali O, Gilles B, Hamant O, Boudaoud A, Traas J, Godin C (2015) A computational framework for 3d mechanical modeling of plant morphogenesis with cellular resolution. PLoS Comput Biol 11(1):e1003950
8. Campilho A, Garcia B, Wijk HV, Campilho A, Scheres B et al (2006) Time-lapse analysis of stem-cell divisions in the arabidopsis thaliana root meristem. Plant J 48(4):619–627
9. Cieslak M, Runions A, Prusinkiewicz P (2015) Auxin-driven patterning with unidirectional fluxes. J Exp Bot 66:5083–5102. https://doi.org/10.1093/jxb/erv262
10. Coen E, Rebocho AB (2016) Resolving conflicts: modeling genetic control of plant morphogenesis. Dev Cell 38(6):579–583
11. Coen E, Rolland-Lagan AG, Matthews M, Bangham JA, Prusinkiewicz P (2004) The genetics of geometry. Proc Natl Acad Sci USA 101(14):4728–4735
12. de Boer MJ, Fracchia FD, Prusinkiewicz P (1992) A model for cellular development in morphogenetic fields. In: Lindenmayer systems. Springer, Berlin, pp 351–370
13. de Reuille PB, Routier-Kierzkowska AL, Kierzkowski D, Bassel GW, Schüpbach T, Tauriello G, Bajpai N, Strauss S, Weber A, Kiss A et al (2015) Morphographx: a platform for quantifying morphogenesis in 4d. Elife 4:e05864
14. De Rybel B, Adibi M, Breda AS, Wendrich JR, Smit ME, Novák O, Yamaguchi N, Yoshida S, Van Isterdael G, Palovaara J et al (2014) Integration of growth and patterning during vascular tissue formation in arabidopsis. Science 345(6197):1255215
15. Donnelly P, Bonetta D, Tsukaya H, Dengler R, Dengler N (1999) Cell cycling and cell enlargement in developing leaves of *Arabidopsis*. Dev Biol 215(2):407–419
16. Dupuy L, Mackenzie J, Haseloff J (2010) Coordination of plant cell division and expansion in a simple morphogenetic system. Proc Natl Acad Sci 107(6):2711–2716
17. el Showk S, Blomster T, Siligato R, Marée AF, Mähönen AP, Grieneisen VA et al (2015) Parsimonious model of vascular patterning links transverse hormone fluxes to lateral root initiation: auxin leads the way, while cytokinin levels out. PLoS Comput Biol 11(10):e1004450
18. Errera L (1888) Über zellfromen und seifenblasen. Botanisches Centralblatt 34:395–398
19. Federl P, Prusinkiewicz P (1999) Virtual laboratory: an interactive software environment for computer graphics. In: Computer graphics international, vol 242, pp 93–100
20. Fernandez R, Das P, Mirabet V, Moscardi E, Traas J, Verdeil JL, Malandain G, Godin C (2010) Imaging plant growth in 4d: robust tissue reconstruction and lineaging at cell resolution. Nat Methods 7(7):547–553

21. Feugier FG, Mochizuki A, Iwasa Y. (2005) Self-organization of the vascular system in plant leaves: inter-dependent dynamics of auxin flux and carrier proteins. J Theor Biol 236(4):366–375
22. Fukushima K, Fujita H, Yamaguchi T, Kawaguchi M, Tsukaya H, Hasebe M (2015) Oriented cell division shapes carnivorous pitcher leaves of sarracenia purpurea. Nat Commun 6:6450
23. Galassi M, Davies J, Theiler J, Gough B, Jungman G, Alken P, Booth M, Rossi F (2002) Gnu Scientific Library. Network Theory Ltd 3
24. Giavitto JL, Michel O (2001) MGS: a ruled-based language for complex objects and collections. Electron Notes Theor Comput Sci 59(4):1–19
25. Gierer A, Meinhardt H (1972) A theory of biological pattern formation. Kybernetik 12(1):30–39
26. Glazier JA, Graner F (1993) Simulation of the differential adhesion driven rearrangement of biological cells. Phys Rev E 47(3):2128
27. Goriely A, Robertson-Tessi M, Tabor M, Vandiver R (2008) Elastic growth models. In: Mathematical modelling of biosystems. Springer, Berlin, pp 1–44
28. Grieneisen VA, Xu J, Marée AFM, Hogeweg P, Scheres B (2007) Auxin transport is sufficient to generate a maximum and gradient guiding root growth. Nature 449(7165):1008–1013. https://doi.org/10.1038/nature06215. http://dx.doi.org/10.1038/nature06215
29. Haselkorn R (1998) How cyanobacteria count to 10. Science 282(5390):891–892
30. Heisler MG, Jönsson H (2006) Modeling auxin transport and plant development. J Plant Growth Regul 25:302–312. https://doi.org/10.1007/s00344-006-0066-x
31. Hejnowicz Z, Karczewski J (1993) Modeling of meristematic growth of root apices in a natural coordinate system. Am J Bot 80:309–315
32. Hejnowicz Z, Nakielski J, Hejnowicz K (1984) Modeling of spatial variations of growth within apical domes by means of the growth tensor. ii. Growth specified on dome surface. Acta Soc Bot Pol 53:301–316.
33. Hervieux N, Dumond M, Sapala A, Routier-Kierzkowska AL, Kierzkowski D, Roeder AH, Smith RS, Boudaoud A, Hamant O (2016) A mechanical feedback restricts sepal growth and shape in arabidopsis. Curr Biol 26(8):1019–1028
34. Hofmeister W (1868) Handbuch der physiologishen botanik. Engelmann, Leipzig
35. Honda H (1978) Description of cellular patterns by Dirichlet domains: the two-dimensional case. J Theor Biol 72:523–543
36. Honda H (1983) Geometrical models for cells in tissues. Int Rev Cytol 81:191–248
37. Jönsson H, Heisler MG, Shapiro BE, Meyerowitz EM, Mjolsness E (2006) An auxin-driven polarized transport model for phyllotaxis. Proc Natl Acad Sci U S A 103(5):1633–1638. https://doi.org/10.1073/pnas.0509839103. http://dx.doi.org/10.1073/pnas.0509839103
38. Jönsson H, Gruel J, Krupinski P, Troein C (2012) On evaluating models in computational morphodynamics. Curr Opin Plant Biol 15(1):103–110
39. Karlebach G, Shamir R (2008) Modelling and analysis of gene regulatory networks. Nat Rev Mol Cell Biol 9(10):770–780
40. Kennaway R, Coen E, Green A, Bangham A (2011) Generation of diverse biological forms through combinatorial interactions between tissue polarity and growth. PLoS Comput Biol 7(6):e1002071
41. Kierzkowski D, Nakayama N, Routier-Kierzkowska AL, Weber A, Bayer E, Schorderet M, Reinhardt D, Kuhlemeier C, Smith RS (2012) Elastic domains regulate growth and organogenesis in the plant shoot apical meristem. Science 335(6072):1096–1099
42. Kramer EM (2008) Computer models of auxin transport: a review and commentary. J Exp Bot 59(1):45–53
43. Kramer EM (2009) Auxin-regulated cell polarity: an inside job? Trends Plant Sci 14(5):242–247
44. Kuchen EE, Fox S, de Reuille PB, Kennaway R, Bensmihen S, Avondo J, Calder GM, Southam P, Robinson S, Bangham A et al (2012) Generation of leaf shape through early patterns of growth and tissue polarity. Science 335(6072):1092–1096
45. Kuhlemeier C (2007) Phyllotaxis. Trends Plant Sci 12(4):143–150

46. Kwiatkowska D (2006) Flower primordium formation at the arabidopsis shoot apex: quantitative analysis of surface geometry and growth. J Exp Bot 57(3):571–580
47. Lindenmayer A (1968) Mathematical models for cellular interactions in development. I. Filaments with one-sided inputs. J Theor Biol 18(3):280–299
48. Lindenmayer A (1968) Mathematical models for cellular interactions in development. II. Simple and branching filaments with two-sided inputs. J Theor Biol 18(3):300–315
49. Lintilhac PM, Vesecky TB (1984) Stress-induced alignment of division plane in plant tissues grown in vitro. Nature 307(5949):363–364
50. Lockhart JA (1965) An analysis of irreversible plant cell elongation. J Theor Biol 8(2):264–275
51. Louveaux M, Julien JD, Mirabet V, Boudaoud A, Hamant O (2016) Cell division plane orientation based on tensile stress in arabidopsis thaliana. Proc Natl Acad Sci 113(30):E4294–303. https://doi.org/10.1073/pnas.1600677113
52. Meinhardt H (1982) Models of biological pattern formation. Academic Press, London
53. Meinhardt H (2003) Complex pattern formation by a self-destabilization of established patterns: chemotactic orientation and phyllotaxis as examples. C R Biol 326(2):223–237
54. Merks RM, Guravage M, Inzé D, Beemster GT (2011) Virtualleaf: an open-source framework for cell-based modeling of plant tissue growth and development. Plant Physiol 155(2):656–666
55. Mitchison GJ (1980) A model for vein formation in higher plants. Philos Trans R Soc Lond B Biol Sci 207:79–109
56. Nakielski J (2000) Pattern formation in biology, vision and dynamics, chap. Tensorial model for growth and cell division in the shoot apex. World Scientific, pp. 252–286
57. Nakielski J, Barlow P (1995) Principal directions of growth and the generation of cell patterns in wild-type and gib-1 mutant roots of tomato (lycopersicon esculentum mill.) grown in vitro. Planta 196(1):30–39
58. Nakielski J, Lipowczan M (2013) Spatial and directional variation of growth rates in arabidopsis root apex: A modelling study. PLOS ONE 8(12). https://doi.org/10.1371/journal.pone.0084337. https://doi.org/10.1371/journal.pone.0084337
59. Neubert MG, Caswell H, Murray J (2002) Transient dynamics and pattern formation: reactivity is necessary for turing instabilities. Math Biosci 175(1):1–11
60. Prusinkiewicz P, Lane B (2012) Pattern formation in morphogenesis. Springer, Berlin
61. Prusinkiewicz P, Lane B (2013) Modeling morphogenesis in multicellular structures with cell complexes and l-systems. In: Pattern formation in morphogenesis. Springer, Berlin, pp 137–151
62. Prusinkiewicz P, Lindenmayer A (1990) Algorithmic beauty of plants. Springer, Berlin
63. Rebocho AB, Southam P, Kennaway JR, Bangham JA, Coen E (2017) Generation of shape complexity through tissue conflict resolution. eLife 6:e20156
64. Reinhardt D, Pesce ER, Stieger P, Mandel T, Baltensperger K, Bennett M, Traas J, Friml J, Kuhlemeier C (2003) Regulation of phyllotaxis by polar auxin transport. Nature 426(6964):255–260. https://doi.org/10.1038/nature02081. http://dx.doi.org/10.1038/nature02081
65. Rodriguez EK, Hoger A, McCulloch AD (1994) Stress-dependent finite growth in soft elastic tissues. J Biomech 27(4):455–467
66. Rolland-Lagan AG, Prusinkiewicz P (2005) Reviewing models of auxin canalization in the context of leaf vein pattern formation in *Arabidopsis*. Plant J 44(5):854–865. https://doi.org/10.1111/j.1365-313X.2005.02581.x
67. Rolland-Lagan AG, Remmler L, Girard-Bock C (2014) Quantifying shape changes and tissue deformation in leaf development. Plant Physiol 165(2):496–505
68. Runions A (2008) Modeling biological patterns using the space colonization algorithm. M.Sc. Thesis, University of Calgary
69. Runions A, Fuhrer M, Lane B, Federl P, Rolland-Lagan AG, Prusinkiewicz P (2005) Modeling and visualization of leaf venation patterns. ACM Trans Graph 24:702–711
70. Sachs T (1981) The control of patterned differentiation of vascular tissues. Adv Bot Res 9:151–262
71. Sahlin P, Söderberg B, Jönsson H (2009) Regulated transport as a mechanism for pattern generation: capabilities for phyllotaxis and beyond. J Theor Biol 258(1):60–70

72. Sauret-Güeto S, Schiessl K, Bangham A, Sablowski R, Coen E (2013) Jagged controls arabidopsis petal growth and shape by interacting with a divergent polarity field. PLoS Biol 11(4):e1001550
73. Scarpella E, Francis P, Berleth T (2004) Stage-specific markers define early steps of procambium development in *Arabidopsis* leaves and correlate termination of vein formation with mesophyll differentiation. Development 131(14):3445–3455
74. Scarpella E, Marcos D, Friml J, Berleth T (2006) Control of leaf vascular patterning by polar auxin transport. Genes Dev 20(8):1015–1027. https://doi.org/10.1101/gad.1402406
75. Smith R (2011) Modeling plant morphogenesis and growth. New Trends Phys Mech Biol Syst 92:301–336
76. Smith RS, Bayer EM (2009) Auxin transport-feedback models of patterning in plants. Plant Cell Environ 32(9): 1258–1271. https://doi.org/10.1111/j.1365-3040.2009.01997.x. http://dx.doi.org/10.1111/j.1365-3040.2009.01997.x
77. Smith RS, Guyomarc'h S, Mandel T, Reinhardt D, Kuhlemeier C, Prusinkiewicz P (2006) A plausible model of phyllotaxis. Proc Natl Acad Sci U S A 103(5):1301–1306. https://doi.org/10.1073/pnas.0510457103
78. Smith RS, Kuhlemeier C, Prusinkiewicz P (2006) Inhibition fields for phyllot actic pattern formation: a simulation study. Can J Bot 84(11):1635–1649
79. Turing A (1952) The chemical basis of morphogenesis. Philos Trans R Soc Lond B Biol Sci 237:37–52
80. Wabnik K, Robert HS, Smith RS, Friml J (2013) Modeling framework for the establishment of the apical-basal embryonic axis in plants. Curr Biol 23(24):2513–2518
81. Yoshida S, de Reuille PB, Lane B, Bassel GW, Prusinkiewicz P, Smith RS, Weijers D (2014) Genetic control of plant development by overriding a geometric division rule. Dev cell 29(1):75–87
82. Žádníková P, Wabnik K, Abuzeineh A, Gallemi M, Van Der Straeten D, Smith RS, Inzé D, Friml J, Prusinkiewicz P, Benková E (2016) A model of differential growth-guided apical hook formation in plants. Plant Cell 28(10):2464–2477
83. Zienkiewicz OC, Taylor RL (2005) The finite element method for solid and structural mechanics. Butterworth-Heinemann, Boston

Chapter 8
Modeling Plant Development with L-Systems

Przemyslaw Prusinkiewicz, Mikolaj Cieslak, Pascal Ferraro, and Jim Hanan

Abstract Since their inception in 1968, L-systems have become a key conceptual, mathematical and software tool for modeling plant development at different levels of plant organization spanning molecular genetics, plant physiology, whole plant architecture and plant communities. The models can be descriptive, directly recapitulating observations and measurements of plants; mechanistic, explaining higher-level processes in terms of lower-level ones; or they may combine features of both classes. We present the basic idea of L-systems, motivate and outline some of their most useful extensions, and give a taste of current techniques for modeling with L-systems. The sample models progress in the scale of organization from a bacterium to a herbaceous plant to a tree, and simulate different forms of information transfer during the development, from communication between adjacent cells to bidirectional information exchange with the environment.

8.1 Genesis of the Idea

The discovery of the structure and functioning of DNA placed molecular genetics in the centre of modern biology, and opened the door for reducing diverse biological processes to their biochemical (and, ultimately, physical) basis. However, the sequencing of the genome of numerous organisms, including humans, has also highlighted the gap between knowing the genome and understanding how it regulates the development and form of an organism. This regulation is not direct; instead, genes and molecular-level processes establish local rules for the spatially distributed dynamic processes of morphogenesis from which the developing forms arise. The emergence of global patterns, forms or behaviours through the local interaction

P. Prusinkiewicz (✉) · M. Cieslak · P. Ferraro
Department of Computer Science, University of Calgary, Calgary, AB, Canada
e-mail: pwp@ucalgary.ca; msciesla@ucalgary.ca; pferraro@ucalgary.ca

J. Hanan
Centre for Horticultural Science, University of Queensland, Brisbane, QLD, Australia
e-mail: j.hanan@uq.edu.au

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_8

of their components is the defining feature of self-organization and highlights the inherent difficulty in studying and understanding development: properties of self-organizing systems are often non-intuitive and difficult to analyse. Computational models and simulations facilitate studies of self-organizing phenomena by revealing the consequences of different hypothetical rules, explicitly specified by the modeler. A series of computational experiments, in which the model is modified and refined until the emerging developmental dynamics approximates reality with sufficient accuracy, leads to insights into the plausible processes underlying the analysed processes, patterns and forms.

The use of computational models raises the question of how they should be conceptualized, specified and executed to effectively support the process of scientific discovery. Similar questions appear in any application of computing, but in the case of morphogenesis the answers are particularly elusive. This is due to the discrepancy between the standard view of computation that underlies commonly used programming languages on the one hand, and the nature of the problems of morphogenesis on the other.

The standard view of computation is related to the classical (von Neumann) model of computer architecture. According to this model, computation is organized around a central processing unit (CPU), which executes a sequence of operations on a set of numbers in a well-defined order. The geometric concept of space is not part of this model. In contrast, morphogenesis involves multiple processes taking place in parallel, and it is an inherently geometric phenomenon. Spatial relations between components of the developing organism play a key role in defining which specific processes take place in each component at each point of time. Further complicating the matter, the number of components and processes may change as the organism grows. Because of these changes, the standard mathematical framework of dynamical systems, in which temporal evolution of a system is described using a predefined set of variables and equations, becomes too limiting [20].

In view of these discrepancies, nonstandard models of computation, programming languages and computational devices have been proposed to specify morphogenetic processes and implement simulations effectively. A well-known example is that of *cellular automata* (also called cellular spaces), invented in the 1950s by von Neumann and Ulam. Their early applications included a biologically-inspired model of self-replication [70], and—of particular importance to the history of computational studies of morphogenesis—the first simulations of growing branching structures [68, 69].

With cellular automata, space is represented as an array of cells, each of which houses an automaton that indicates, depending on its state, whether the cell is empty or occupied by a component of the modeled structure. The automata change their states in parallel (synchronously) to simulate a uniform progress of time within the entire structure. Transition functions governing state changes can be conveniently specified using declarative statements (rules), which characterize the next state of each automaton/cell as a function of its current state and the state of its neighbours. An example of such a rule for a one-dimensional cellular automaton defined over the set of states {0,1} might be “a cell in state 1, situated between cells in state 0 to the left and state 1 to the right, will change its state to 0."

L-systems [35] were inspired by cellular automata, and in some simple cases are indistinguishable from them. Using the terminology of L-systems, the above rule would be termed a *production* that replaces the strict *predecessor*, symbol 1, occurring in the *context* of symbol 0 to the left and symbol 1 to the right, with the *successor* 0. Such a production can be written as $0 < 1 > 1 \rightarrow 0$ (here symbols $<$ and $>$ do not denote inequalities, but simply separate the strict predecessor from the left and right contexts). Four sample L-systems inspired by cellular automata are specified in Table 8.1, and the first 32 steps of their operation are shown in Fig. 8.1.

The L-system in Fig. 8.1a most closely mimics the operation of a cellular automaton, in the sense that the space in which it operates—a row of 65 cells—is given in advance. In the course of the simulation, the states of individual cells may change, but their number and spatial configuration are fixed. In contrast, the L-systems in Fig. 8.1b–d employ productions $A \rightarrow A0$ and $B \rightarrow 0B$ operating on additional symbols A and B (colours used for emphasis) to gradually expand the strings of symbols that grow from the short L-system *axiom* (initial string) $A00100B$. The ability to expand (or contract) the modeled structure over time is a distinctive feature of L-systems.

There is also a more subtle distinction between cellular automata and L-systems. Cellular automata describe a quantity (state) associated with (discrete) points of space. This space-centred perspective is known as Eulerian. In contrast, L-systems focus on the developing structure itself, a perspective known as Lagrangian. The difference between these perspectives comes to light when the modeled structure grows over time, and is particularly evident when new components are added not only at the structure boundary, but in its interior as well. Some components then change their position and "flow" through space as they are pushed by other

Table 8.1 Sample L-systems inspired by cellular automata. L-system **a** rewrites symbols of the initial string (axiom) without changing its length, as typical of cellular automata (the power notation 0^{32} denotes a string of 32 zeroes)

Figure	**a**	**b**	**c**	**d**
Axiom	$0^{32}10^{32}$	$A00100B$		
Predecessor	Successor			
$0 < 0 > 0$	0	0	0	0
$0 < 0 > 1$	1	1	1	1
$0 < 1 > 0$	0	0	1	1
$0 < 1 > 1$	0	0	0	1
$1 < 0 > 0$	1	1	1	1
$1 < 0 > 1$	0	0	1	0
$1 < 1 > 0$	0	0	0	0
$1 < 1 > 1$	0	0	0	0
A	–	$A0$	$A0$	$A0$
B	–	$0B$	$0B$	$0B$

In L-systems **b–d**, productions operating on symbols A and B gradually extend the generated strings.

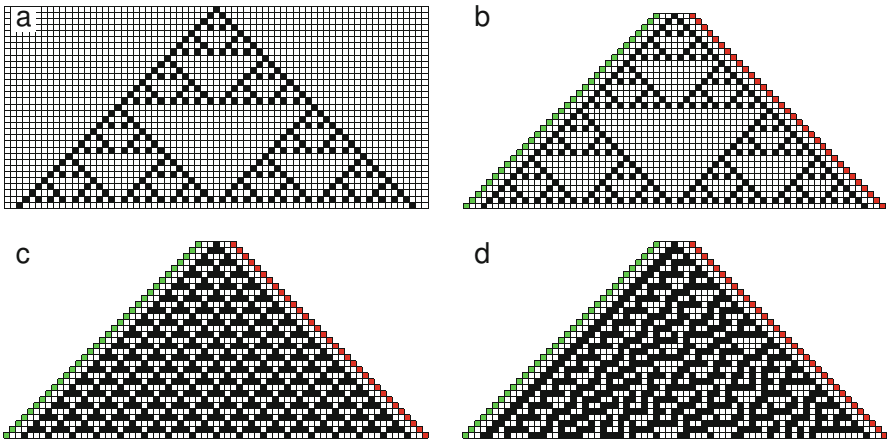


Fig. 8.1 Operation of L-systems from Table 8.1. White cells represent the symbol 0, black cells the symbol 1. In each simulation, the top row represents the initial state, and the subsequent rows represent the results of each simulation step. The space in which simulations operate is fixed in model (**a**) and gradually expands in models (**b**)–(**d**), in which coloured cells produce new white cells. These examples also show that small differences in productions may result in qualitatively different patterns, including self-similar (**a**), (**b**), repetitive (**c**) and chaotic (**d**) patterns [73, 74]. These differences highlight the need for the modeling and simulations in the studies of self-organization. The rules used in the above examples are commonly referred to as Rule 18 (**a**), (**b**), Rule 54 (**c**) and Rule 30 (**d**). These names result from interpreting the list of successors in the respective columns of Table 8.1, read from the bottom up, as binary numbers, e.g. binary 00010010 equals 30 in common decimal notation

components that need room to fit. It makes a difference whether this process is described from the perspective of points of space or individual components [56]. The latter perspective, afforded by L-systems, is often natural and convenient in the description of development.

8.2 Modeling Cell Division Patterns

A model of vegetative segments of the filamentous blue-green bacterium *Anabaena catenula* provides a simple example illustrating the capability of L-systems to simulate growth and cell division patterns in one dimension. A vegetative *Anabaena* segment consists of sequences of longer and shorter cells that appear to be arranged at random, but in fact divide deterministically [42]. During the development, short cells elongate, and long cells divide asymmetrically into a short and long cell. A short daughter cell is always produced on the side of the older wall separating the mother cell from its neighbours. The following L-system captures these principles, with symbols S and L denoting a short and a long cell, respectively, and symbols W denoting cell walls that delimit cells within the filament:

Table 8.2 L-system modeling the development of a vegetative segment of *Anabaena*

Axiom	$W(0)LW(1)$
Productions:	
p_1	$W(a) \rightarrow W(a+1)$
p_2	$W(a_l) < L > W(a_r) : a_l \geq a_r \rightarrow SW(0)L$
p_3	$W(a_l) < L > W(a_r) : a_l < a_r \rightarrow LW(0)S$
p_4	$S \rightarrow L$

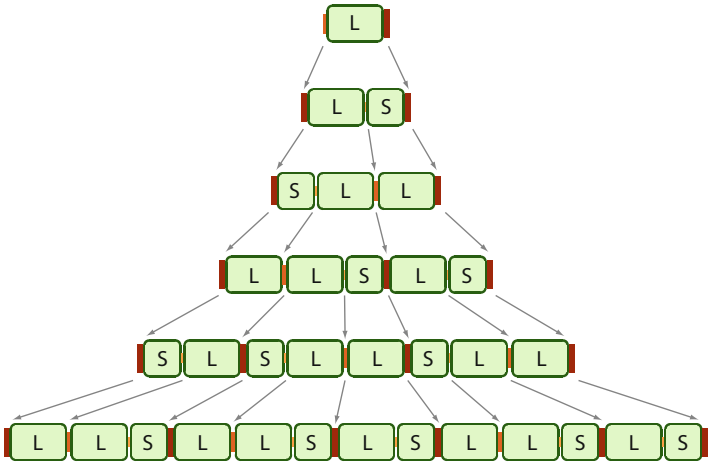


Fig. 8.2 The first five steps of the simulation of the development of a vegetative filament of *Anabaena*, simulated using the L-system in Table 8.2. The size and colour of bars between the cells and on the outside of the filament indicates the age of the walls (0, 1, or ≥ 2). Arrows indicate how the walls propagate from one simulation step to the next

Compared to the L-systems in Table 8.1, this L-system incorporates two extensions. The first one is the introduction of a numerical *parameter a* representing wall age. It is incremented in every derivation step by production p_1. The second extension is the introduction of *conditions* that guard production application according to the parameter values. Specifically, conditions $a_l \geq a_r$ and $a_l < a_r$ determine whether a long cell L will divide into a short cell S followed by a long cell L (production p_2) or into a long cell L followed by a short cell S (production p_3). The choice depends on the age of the walls surrounding L. The remaining production, p_4, represents the growth of a short cell S into a long cell L. The first five simulation steps, beginning with a single long cell bound by walls with the age set arbitrarily to 0 and 1 in the axiom, are shown in Fig. 8.2. As this example illustrates, numerical parameters and conditional application of productions are very useful in modeling practice. They first appeared as a programming construct in the L-system-based simulator CELIA [4, 5, 37] and were subsequently formalized by Prusinkiewicz and Hanan [53], Prusinkiewicz and Lindenmayer [55], Hanan [23].

Experimenting with the L-system in Table 8.2 provides insights beyond the original objective of simulating the development of the vegetative segment of

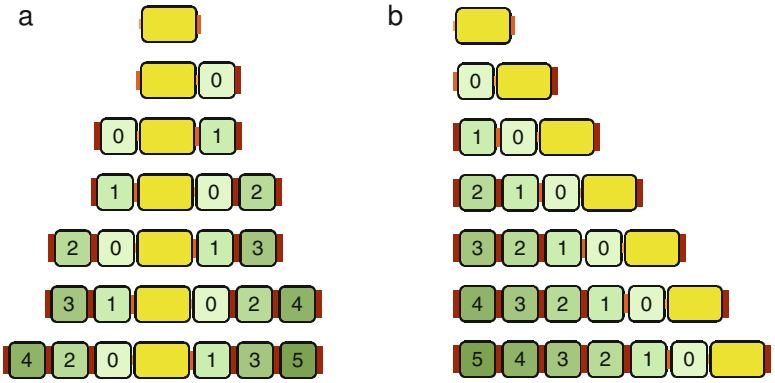


Fig. 8.3 Two developmental sequences generated by modifying the L-system for the vegetative segment of *Anabaena*. The size and colour of bars between the cells and on the outside of the filament indicates the age of the walls (0, 1, or $\geq$ 2), as in Fig. 8.2. As the long cell divides, the daughter short cell is positioned on the side of the older (**a**) or younger (**b**) wall of the mother cell. Numbers and colours of the short cell indicate their age

Anabaena. For instance, Fig. 8.3a shows the developmental sequence obtained by removing production p_4. Cell L, present in the axiom, is then the only cell that divides. Cells S, created by L in consecutive derivation steps, appear on both of sides of L and "internalize" it near the filament centre. A similar mechanism, operating in two dimensions, explains the distribution of stomata in the leaf epidermis. As in the linear case, a dividing cell—precursor of a stoma—produces new cells towards its older walls. This process results in the internalization of the precursor cell and, consequently, the separation of one stoma from another by non-stomatal cells: a fundamental feature of stomata distribution [62].

A further modification reverts conditions in productions p_2 and p_3 as follows (colour added for emphasis):

$$p_2' : \quad W(a_l) < L > W(a_r) : a_l \leq a_r \rightarrow SW(0)L$$
$$p_3' : \quad W(a_l) < L > W(a_r) : a_l > a_r \rightarrow LW(0)S$$

The resulting L-system creates a drastically different developmental sequence, in which the dividing cell L is positioned at the end of a growing structure, extending it in one direction only (Fig. 8.3b). A similar process is commonly observed in plants, from filamentous algae to mosses, ferns and higher plants, where an apical cell (or multicellular apical meristem) creates an axis of plant development. The contrast between the development of a filament with multiple dividing cells, the division of a single cell leading to its internalization, and the maintenance of an axis with the dividing cell situated at an apex shows that even very simple L-systems can shed light on nontrivial relations between the local rules of development and the global characteristics of the resulting structures.

8.3 Programming with L-Systems

The L-system in Table 8.2 has been specified using a syntax inspired by the notion of rewriting, as studied in mathematics and theoretical computer science (in particular, formal language theory, where it underpins the concept of formal grammars). The relation of L-systems to rewriting systems was first recognized by Lindenmayer [36]. With the subsequent extensions to parametric L-systems, productions such as those shown in Table 8.2 are written as:

$$\textit{left context} < \textit{strict predecessor} > \textit{right context} : \textit{condition} \rightarrow \textit{successor}$$

for instance

$$A(u) < B(x, y) > C(v)D : u < v \rightarrow E(x + \sin(u), y + \cos(v))$$

For simple L-systems, this mathematically-inspired syntax works well; in particular, its conciseness makes it easy to specify and modify productions, an inherent component of developing models in an interactive computing setting. Unfortunately, it does not easily scale up to more complex L-systems [51]. For instance, sequences of single-letter symbols associated with multiple parameters look cryptic, making model specifications difficult to read and maintain, and there are no constructs for evaluating conditions and parameter values programmatically. Extensions to the mathematical notation addressing some of these limitations have been proposed and are useful [55], but a fundamental solution is to combine L-system constructs with a programming language. Existing systems include binding L-systems with C/C++ [29, 61], Java [30, 31] and Python [10]. For instance, the basic format of productions in the L+C language [29, 61] is:

$$\textit{left context} < \textit{strict predecessor} > \textit{right context} : \{\textit{statement block}\}.$$

The statement block is (relatively unrestricted) C++ code, allowing for the specification of arbitrarily complex computations involving module parameters, and extended with `produce` statements, which precede the specification of a successor. A single production may include several `produce` statements, which provides a means of specifying alternative outcomes to production execution, depending on the evaluation of conditions. Furthermore, modules may have parameters of different types, including entire data structures. This use of structures improves the clarity of L-system specification if numerous parameters are involved, and implies that in L+C modules are declared.

8.4 Incorporating a Genetic Regulatory Network into an L-System Model

The next model illustrates L+C language constructs and provides an example of an L-system incorporating a genetic regulatory network (GRN). This example returns to *Anabaena*, but focuses on an additional element of its development: the differentiation of a special type of cell, the heterocyst.

In an *Anabaena* filament grown in an environment without nitrogenous compounds there is a division of functions: vegetative cells assimilate carbon from the atmosphere in the process of photosynthesis, while heterocysts fix nitrogen [2, 26]. These functions are performed by different cells because the enzyme involved in nitrogen fixation, the nitrogenase, disintegrates in the presence of the oxygen produced during photosynthesis. On the average, heterocysts are separated by approximately ten vegetative cells [25]. New heterocysts differentiate as the distance between existing heterocysts increases due to the division of the vegetative cells in-between. The mechanism controlling this differentiation has been of interest for a long time, and has included the construction of L-system models both to address the biological problem itself and to illustrate model construction with L-systems (e.g., [4, 5, 13, 14, 20, 37, 54, 55]).

It was initially thought that heterocyst differentiation is triggered by low concentration of nitrogenous compounds [16], which are synthesized in the heterocyst and diffuse into vegetative segments, where they are used. The concentration of these compounds would decrease as the length of a vegetative segments increases, and eventually would fall below a threshold near the centre of the vegetative segment, triggering the differentiation of a new heterocyst [18]. Advances in molecular biology have shown that, while the general idea of a diffusing substance controlling the differentiation of heterocysts is correct, the morphogenetically active diffusing substance is a small protein, PatS, acting as a proxy for the nitrogenous compounds [75]. The production of PatS is regulated by another protein, HetR, which is present in high concentration in heterocysts and defines their identity [11]. The interaction between PatS and HetR is qualitatively depicted in Fig. 8.4.

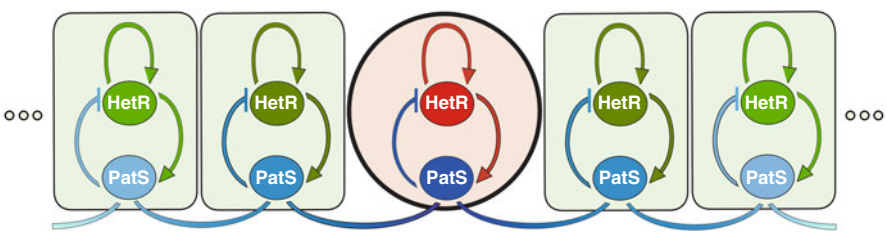


Fig. 8.4 Interaction between HetR and PatS in an *Anabaena* filament. Arcs with arrows denote upregulation, arcs with bars downregulation, and horizontal arcs diffusion of the respective proteins. Colour intensity indicates concentration of HetR and PatS in the heterocyst (circular cell) and the neighbouring vegetative cells. Adapted from [13]

To incorporate these interactions into an L-system model it is necessary to give them a mathematical form. Following [20, 22, 72], we cast it as an activator-inhibitor model [21, 41]: a type of reaction-diffusion model in which the interaction between the intervening substances is qualitatively consistent with the interaction between PatS and HetR [2]. The equations for an arbitrary cell in the filament then have the form:

$$\frac{d[\text{HetR}]}{dt} = \rho \frac{[\text{HetR}]^2}{[\text{PatS}]} + \rho_0 - \mu[\text{HetR}], \tag{8.1}$$

$$\frac{d[\text{PatS}]}{dt} = \rho[\text{HetR}]^2 + \rho_0 - \nu[\text{PatS}] + \Phi. \tag{8.2}$$

Symbols [HetR] and [PatS] denote the concentration of the corresponding proteins, μ and ν control their turnover (use and decay), ρ and ρ_0 control production rates, and Φ represents the flux of PatS into the cell under consideration from its neighbouring cells (for simplicity, we assume that all cells have a unit volume, and the walls between them have a unit area). Given the diffusive character of PatS transport, this flux is proportional to the difference in concentration of PatS in the cell under consideration and its left and right neighbours, here identified by subscripts l and r (Fick's law):

$$\Phi = D\left(([\text{PatS}]_l - [\text{PatS}]) + ([\text{PatS}]_r - [\text{PatS}])\right). \tag{8.3}$$

The coefficient of proportionality D controls the rate of diffusion through cell walls.

The specification of an L+C model corresponding to these equations begins with definitions of the simulation parameter values. These definitions are given in the standard C/C++ format:

```
/* Definition of constants used in simulation */
#define rho    1.0      // controls rate of the production of HetR and PatS
#define rho0   0.01     // controls basic production of HetR and PatS
#define mu     0.1      // turnover rate of HetR
#define nu     1.0      // turnover rate of PatS
#define D      2.0      // diffusion constant for PatS
#define thr    20       // threshold HetR concentration defining a heterocyst
#define dt     0.05     // time step
```

In addition to parameters controlling the biochemical aspects of the simulation, the parameter list includes geometric parameters that control the dimensions and growth rates of the cells:

```
#define gr1    0.004    // relative elementary growth rate of vegetative cells
#define MAX    1.0      // threshold length at which a vegetative cell divides
#define L      0.61804  // length of the longer daughter cell after division
#define S      0.38196  // length of the shorter daughter cell after division
#define TW     0.5      // target width of vegetative cells
#define TD     0.7      // target diameter of heterocysts
#define gr2    0.015    // rate of cell dimension adjustment towards target
```

The genetic regulatory network is then specified as a system of two functions that determine the rates of expression (production) of proteins HetR and PatS, given their concentrations in the cell:

```
20 /* Functions defining the genetic regulatory network */
21 float hetR(float HetR, float PatS)  // HetR production rate
22 {
23    return rho*HetR*HetR/PatS + rho0;
24 }
25
26 float patS(float HetR)              // PatS production rate
27 {
28    return rho*HetR*HetR + rho0;
29 }
```

The data structures involved in the simulation are specified next, following standard C/C++ syntax:

```
31 /* Declaration of data structures characterizing cells and walls */
32 struct CellData
33 {
34    float HetR, PatS; // concentration of HetR and PatS
35    float l, w;       // cell length and width
36 };
37
38 struct WallData
39 {
40    float flux;       // PatS flux through the well
41    float age;        // wall age
42 };
```

The L+C-specific code begins with the declaration of module types and the initial state (axiom) of the simulation:

```
44 /* Initial values */
45 CellData icd1 = {0.02, 100, S, TW};    // HetR, PatS, length, width
46 CellData icd2 = {0.01, 100, L, TW};    // HetR, PatS, length, width
47 WallData iwd  = {0, 0};                // flux, age
48
49 /* Declaration of module types used in the model */
50 module Cell(CellData);
51 module Wall(WallData);
52
53 /* Defintion of the initial structure and state of the simulation */
54 axiom: Right(90) Wall(iwd) Cell(icd1) Wall(iwd) Cell(icd2) Wall (iwd);
```

Module `Right` orients the filament horizontally on the screen, cf. Sect. 8.5. The key part of the model is specified using two type of rules. Ordinary productions are associated with the progress of time by interval `dt`. The production for walls simply advances the age of walls, as in the L-system in Table 8.2:

```
58 production:
59 Wall(wd) :  // advance wall's age in each simulation step
60 {
61    wd.age += dt;
62    produce Wall(wd);
63 }
```

The production for cells has two components. The first component (lines 67–76) represents a numerical solution to the initial value problem given by Eqs. (8.1) and (8.2) using the forward Euler method. For clarity, different components of this solution (protein diffusion, production, and turnover of proteins) are split [39] into separate statements. The second component (lines 78–86) uses ad-hoc rules

to simulate cell growth. A vegetative cells elongates with the relative elementary growth rate `gr1`, while its width tends to target `TW`. A heterocyst asymptotically tends to a rounded shape with diameter `TD`.

```
65 Wall(wdl) < Cell(cd) > Wall(wdr) :     // update cell state
66 {
67 // Diffusion of patS
68    cd.PatS += (wdl.flux - wdr.flux) * dt;
69
70 // Gene expression
71    cd.HetR += hetR(cd.HetR,cd.PatS) * dt;
72    cd.PatS += patS(cd.HetR) * dt;
73
74 // Decay of proteins
75    cd.HetR -= mu*cd.HetR * dt;
76    cd.PatS -= nu*cd.PatS * dt;
77
78 // Growth
79    if (cd.HetR < thr) {                // vegetative cell...
80       cd.l += cd.l * gr1 * dt;         // grows in length;
81       cd.w += gr2 * (TW - cd.w) * dt;  // width is adjusted towards target
82    }
83    else {                              // heterocyst...
84       cd.l += gr2 * (TD - cd.l) * dt;  // grows towards target in length
85       cd.w += gr2 * (TD - cd.w) * dt;  // and in width
86    }
87    produce Cell(cd);
88 }
```

These productions are followed by decomposition rules [29, 60], which are applied immediately after the productions (in the same derivation step) and perform computations that do not involve advancing time. The decomposition rule for walls computes the flux of PatS through each wall according to Eq. (8.3):

```
93 /* Determine the flux of PatS through the wall */
94 Cell(cdl) < Wall(wd) > Cell(cdr) :
95 {
96    wd.flux = D * (cdl.PatS - cdr.PatS);   // Fick's law
97    produce Wall(wd);
98 }
```

The decomposition rule for cells specifies asymmetric division of vegetative cells according to the age of the incident walls in a manner similar to the L-system in Table 8.2, except that the rule is applied when the mother cell reaches the threshold length rather than threshold age. Cell division is assumed to be instantaneous, because in the context of a continuous-time simulation it represents a somewhat arbitrarily chosen moment of the completion of the division process. Note that both patterns of division, with the shorter cell located on the left or on the right side of the mother cell, are specified within the same production, using alternative `produce` statements (lines 107 and 109).

```
/* Cell division (assumed to be instantaneous in the model) */
Wall(wdl) < Cell(cd) > Wall(wdr) :
{
   if (cd.l > MAX) {                        // if cell length exceeds limit
      CellData shorter = cd, longer = cd; // daughters inherit mother's state
      shorter.l = S; longer.l = L;         // ... except for length
      if (wdl.age > wdr.age)               // wall age sets division polarity
         produce Cell(shorter) Wall(iwd) Cell(longer);
      else
         produce Cell(longer) Wall(iwd) Cell(shorter);
   }
}
```

The model is completed by interpretation rules that specify the visual output of the simulations. Snapshots of the simulation are shown in Fig. 8.5.

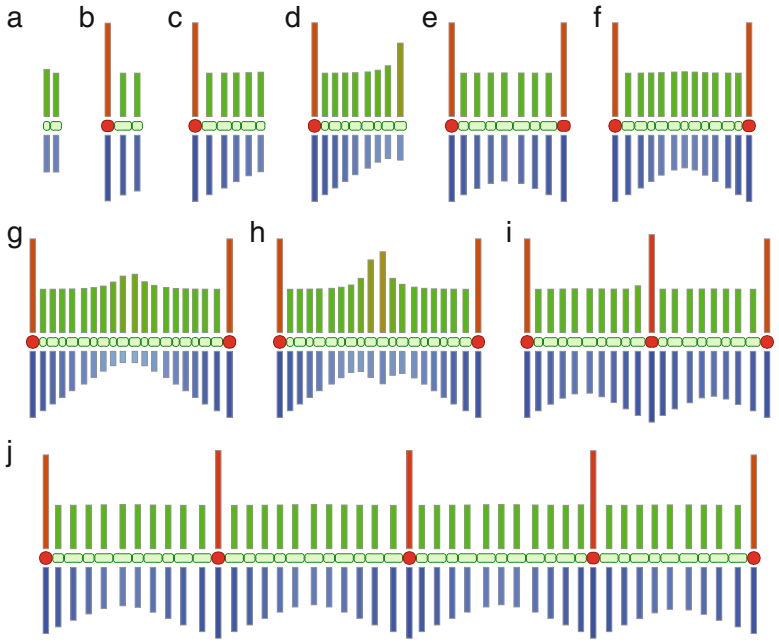


Fig. 8.5 Simulation of the development of an *Anabaena* filament with vegetative cells (green) and heterocysts (red). Bars above cells indicate the concentrations of HetR, and bars below indicate the concentrations of PatS, both on a logarithmic scale. The simulation begins with a short assembly of vegetative cells (**a**), in which a heterocyst quickly differentiates (**b**). As the filament grows and vegetative cells further divide (**c**), the concentration of PatS decreases away from this heterocyst (**d**), which eventually triggers the differentiation of the second heterocyst (**e**). This process repeats, producing a growing filament in which the average distance between heterocysts is approximately ten cells (**f**)–(**j**). Note that vegetative cells distant to previously formed heterocysts may compete to become a heterocyst (**g**), (**h**), until one of them eventually wins (**i**). This competition increases the robustness of heterocyst patterning by reducing the likelihood of heterocysts differentiating next to each other [72]

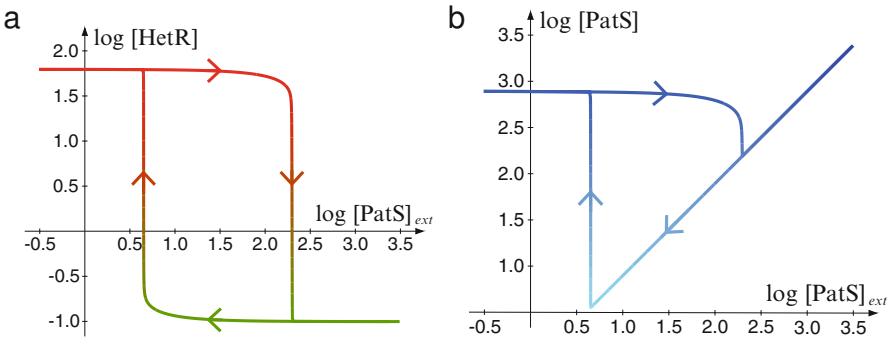


Fig. 8.6 Response of a single *Anabaena* cell to changing concentration of external PatS. Arrows indicate directions of changes. As the concentrations of external PatS decrease, the cell switches from the vegetative state characterized by low concentration of HetR (**a**) and PatS (**b**) to the heterocyst state characterized by high concentration of both proteins. It then remains in the heterocyst state in spite of the subsequent increases in $[\text{PatS}]_{ext}$

An intriguing aspect of the development of an *Anabaena* filament, highlighted by the simulation, is that the concentration of PatS—an inhibitor of heterocyst differentiation—is highest in the heterocysts themselves. This observation leads to the question of why heterocysts do not inhibit themselves from being heterocysts: a problem known as the refractory behaviour of heterocysts [2, 19]. An explanation is given in Fig. 8.6, which shows how concentrations of HetR and PatS change within a single cell when the concentrations of PatS external to it ($[\text{PatS}]_{ext}$) decrease or increase. The plot reveals a bistable region in which the concentrations of HetR and PatS can both be low or high for the same value of $[\text{PatS}]_{ext}$, depending on the history of the system. This form of bistable behaviour (hysteresis) is the key to maintaining the heterocyst state: once the switch from the vegetative to the heterocyst state has occurred, the cell remains in the heterocyst state in spite of subsequent increases in $[\text{PatS}]_{ext}$.

8.5 Geometric Interpretation of L-Systems

In the examples considered so far we have not discussed the geometric aspects of generated structures. This follows the spirit of the original definition of L-systems, which was focused on the topology of organisms—the pattern of connections between cells or larger modules—rather than their size and shape. Nevertheless, the incorporation of geometry greatly extends the modeling power of L-systems. Different geometric interpretations of L-system-generated strings have been pro-

posed. Among them, an interpretation based on the notion of turtle geometry [1, 49] turned out to be particularly useful in diverse implications, including developmental modeling of plants [50, 55]. In this interpretation, a number of predefined L-system symbols act as commands that control a three-dimensional cursor—conceptualized as a "turtle"—that moves in space and traces the skeleton of a branching structure. The key commands to which the turtle responds are listed in Table 8.3 and further illustrated in Fig. 8.7. Many other operations are also useful in plant modeling [3].

For example, the following L-system with turtle interpretation generates the developmental sequence of branching structures shown in Fig. 8.8.

Table 8.3 Basic turtle commands in a symbolic notation and in the L+C language

Turtle command	Symbol		L+C keyword	
Draw line segment	F		`F`	
Move without drawing a line	f		`f`	
Turn left | right	+	−	`Left`	`Right`
Bend up | down	∧	&	`Up`	`Down`
Roll to the left | right	\	/	`RollL`	`RollR`
Start | end branch	[	]	`SB`	`EB`
Set line width	#		`SetWidth`	
Set line colour	,		`SetColor`	

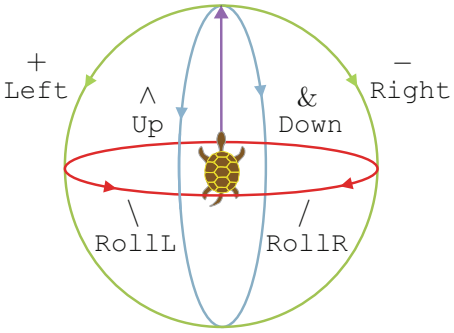


Fig. 8.7 Specification of turtle rotations in three dimensions

Table 8.4 L-system modeling the development of the branching structure in Fig. 8.8

Axiom	$\#(0.4)A$
Productions:	
p_1	$A \rightarrow I(1)[+A][-A]I(1)A$
p_2	$I(s) \rightarrow I(2 * s)$
Interpretation rules:	
h_1	$A \rightarrow ,(1)f(0.2)F(0.8)$
h_2	$I(s) \rightarrow ,(2)f(0.2)F(s - 0.2)$

In this example, all branching angles are assumed to have a globally specified magnitude of 45°. In contrast, the length of internode segments I is specified as a numerical parameter s. Production p_1 sets its initial value to 1, whereas production p_2 multiplies it in each simulation step by 2, thus simulating the growth (exponential elongation) of the internodes.

The L-system in Table 8.4 also includes two *interpretation* rules, h_1 and h_2. Interpretation rules [29, 32, 60] are similar to decomposition rules, discussed previously, in that they are applied immediately after ordinary productions, in the same simulation step. In contrast to decomposition, however, the successor of an interpretation rule is substituted for the predecessor only temporarily, for the purpose of calculating the geometric, graphical representation of the predecessor symbol. In Table 8.4, the interpretation rules state that both the apices A and internodes I are visualized as line segments F preceded by short invisible segments f. The invisible segments highlight the composition of the branching structure by separating its individual components. The modules A and I continue to exist as the predecessor of productions p_1 and p_2 in the subsequent simulation steps.
An L+C program equivalent to the symbolic notation in Table 8.4 is shown below:

```
module A;          // apex
module I(float);  // internode (length)

Axiom: SetWidth(0.4) A;

A :    produce I(1) SB Left(45) A EB SB Right(45) A EB I(1) A;

I(s) : produce I(2*s);

interpretation:
A    : produce SetColor(1) f(0.2) F(0.8);
I(s) : produce SetColor(2) f(0.2) F(s-0.2);
```

8.6 Descriptive Modeling of Plant Architecture

Bracketed parametric L-systems with turtle interpretation are well suited to model plant development at the architectural level. In this case, L-system modules do not correspond to individual cells, but represent higher-level plant components, such as apices (apical meristems), internodes, leaves, flowers and fruits [63]. Architectural models include both descriptive and mechanistic models. Descriptive models recreate plant development according to direct observations and measurements of growth, whereas mechanistic models aim at simulating and explaining development in terms of the underlying lower-level biological processes. In both cases, hypothetical parameter values can be used to test predictions, fill gaps in data, or explore the range of forms that the model can theoretically generate. Below we present a model of the perennial herbaceous plant *Lychnis coronaria* as an

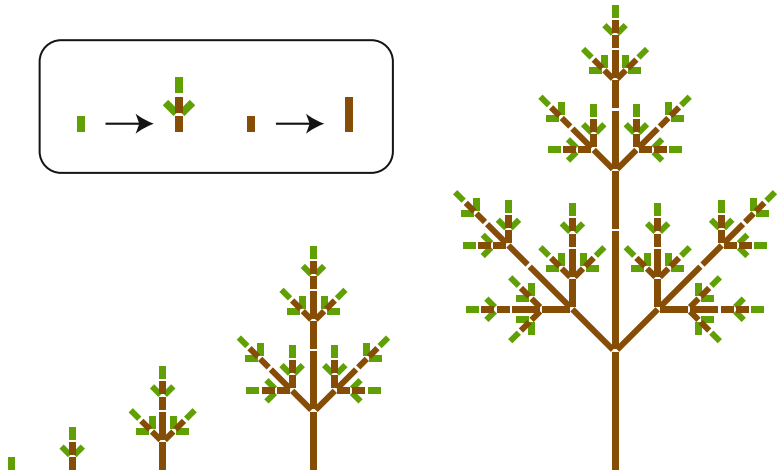


Fig. 8.8 Five stages of the developmental of a simple branching structure modeled using the L-system in Table 8.4. The inset shows a graphical representation of the productions

Fig. 8.9 A photograph of a mature *Lychnis coronaria* plant. The inset shows the characteristic sympodial branching pattern. The architecture of the whole plant results from iterating this motif

example of a descriptive model. The appeal of *Lychnis* stems from the contrast between the apparent complexity of the mature plant (Fig. 8.9) and the simplicity of its architectural development. Its key feature is repetitive sympodial branching,

in which each axis is terminated by a flower, and the main thrust of development is transferred to a pair of lateral branches (Fig. 8.9, inset).

The entire model has approximately 160 lines of L+C code, thus we limit the model description to the key constructs. The main modules are declared as follows:

```
module A(float);                // apex (age)
module I(float,float,float);    // internode (target length, width, age)
module L(float);                // leaf (age)
module K(float);                // bud/flower/fruit (age)
module B(float);                // branching point (age)
```

The first four modules represent components of the plant. The fifth one, `B`, represents branching points, and is introduced to specify time-varying branching angles. All modules are parametrized by age `t`, measured with respect to the time of their creation.

In each simulation step the age of each module is advanced by increment `dt`, which can be made arbitrarily small (as in the model of *Anabaena* with heterocysts) to simulate development in continuous time. The age of apices, leaves, flowers and branching points is advanced by simple productions:

```
A(t) : produce A(t+dt);
L(t) : produce L(t+dt);
K(t) : produce K(t+dt);
B(t) : produce B(t+dt);
```

The age of internodes is advanced by more complex productions that also update the width of internodes and are discussed later.

The sympodial branching structure results from the activity of shoot apices, which give rise to a flower in the terminal position and create a pair of new lateral apices. This process is modeled by the following decomposition rule:

```
decomposition:
A(t):
{
   if (t > 0 && T < BRANCHING_ENDS) produce
      I(LEN1,0,t)
      SB RollR(PHI) B(t) SB L(t) EB A(D1+t) EB
      SB RollL(PHI) B(t) SB L(t) EB A(D2+t) EB
      I(LEN2,0,t) K(t);
}
```

According to this production, an apex reaching the threshold age of zero creates an internode `I`, a pair of lateral apices `A` in the axils of leaves `L`, and a flower bud `K` subtended by another internode `I`. The age values t associated with the active apices are always negative: they indicate the time leading to the production of lateral branches. The branches are initiated after unequal delays `D2 < D1 < 0`, resulting in an asymmetric branching structure (cf. Fig. 8.9, inset). The pattern of repetitive production of lateral apices and branches repeats until the overall plant age represented by a global variable `T` (also incremented by `dt`, statement not shown) reaches the threshold age `BRANCHING_ENDS`.

The apex also creates two modules `B`, which specify the magnitude of the branching angles via the interpretation rule:

```
105 interpretation:
106 B(t) : produce Down(br_angle(t));   // branching angle changes over time
```

According to this rule, in each simulation step module `B` will turn a lateral branch down by the modeler-specified function of time `br_angle(t)`. Meanwhile, module `B` will continue to exist as the object of the time-advancing production in line 85. The development of leaves, flowers and fruits (pods) is modeled in a similar way. For example, leaves are modeled using the production:

```
120 L(t) :
121 {
122    produce Down(leaf_angle(t)) RollToVert() CurrentTexture(LEAF_TEXTURE)
123            SetColor(leaf_color(t)) Surface(LEAF_SURFACE, leaf_length(t));
124 }
```

The turtle command `Surface(LEAF_SURFACE, leaf_length(t))` visualizes a leaf as the surface identified by its name. The shape of this surface is predefined by the modeler using an interactive editor, and the size is determined by the scaling function `leaf_length(t)`, which allows for a coarse approximation of growth over time. (More powerful techniques for modeling organ growth include anisotropic scaling capable of capturing allometric dependencies of organ proportions on size [27, 28, 45, 46], and interpolation between shapes modeled at key developmental stages [23, 47, 57]). The remaining modules specify the orientation, texture and colour of the leaf.

The above description highlights the need for defining numerous constants and functions as a part of the model. In principle, the constants could be defined directly in the L-system code, as in lines 1–18 of the *Anabaena* example, and the functions could be given by mathematical expressions, for example resulting from statistical data analysis. For the sake of model conceptualization, presentation and manipulation it is useful, however, to organize and describe its input graphically. A timeline editor—a concept borrowed from computer animation—can be applied for this purpose. A screenshot of a timeline editor open on the *Lychnis* model is shown in Fig. 8.10. A modeler can use it to interactively modify all time points and function domains, and invoke the graphical editor of specific functions.

The last element of the model is a set of the productions characterizing the internodes:

```
89 I(l,w,t) >> SB I(l1,w1,t1) EB SB I(l2,w2,t2) EB I(l3,w3,t3) :
90        produce I(l, w1+w2+w3, t+dt);
91 I(l,w,t) >> SB I(l1,w1,t1) EB I(l2,w2,t2) : produce I(l, w1+w2, t+dt);
92 I(l,w,t) >> I(l1,w1,t1) : produce I(l, w1, t+dt);
93 I(l,w,t) >> K(t1) : produce I(l, width(t), t+dt);
```

The complicated form of these productions, compared to those describing the aging of the other modules (lines 82–85), stems from the need to determine the internode diameter, which is not simply a function of age. We follow [38] by assuming that, at every branching point, the diameter d_0 of the parent internode is related to the diameters d_1, d_2 and d_3 of the child internodes by equation

$$d_0^\eta = d_1^\eta + d_2^\eta + d_3^\eta. \tag{8.4}$$

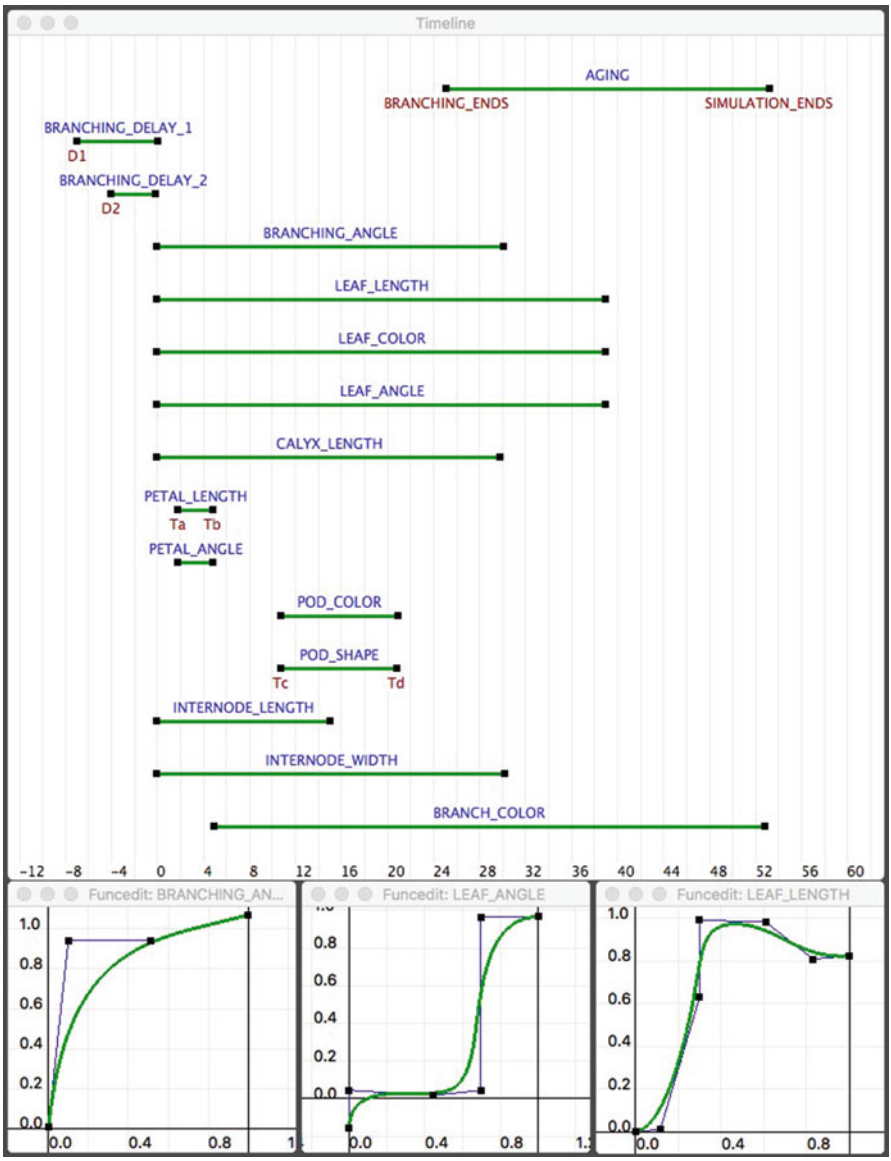


Fig. 8.10 A snapshot of the timeline editor controlling the *Lychnis* model. In addition to the function domains (green lines with blue labels above) and time points (black squares with red labels below) related to branching and leaf development discussed in the text, the editor controls the development of the flowers, fruits (pods) and internodes. The timeline editor also makes it possible to invoke a function editor, here open on three sample functions shown at the bottom. The default domain [0, 1] of each function is mapped onto the domain indicated by its timeline, e.g. [0, 30] in the case of `BRANCHING_ANGLE`. Outside their default domains the functions are assumed to be constant: $f(x) = f(0)$ for $x < 0$, and $f(x) = f(1)$ for $x > 1$. Note that, within the *Lychnis* L-system, the `AGING` and `BRANCH_COLOR` timelines and their associated time points are defined over the whole lifetime `T` of the plant, whereas the remaining timelines and points are defined relative to the age `t` of the modules they describe. This difference is not visualized by the editor

To satisfy this equation, we assume that the diameter of the terminal internodes (supporting a bud, flower or fruit K) is determined directly as a function of their age (line 93), whereas the diameter of the remaining internodes is calculated by setting the width measure $w = d^{\eta}$ in the parent internode to the sum of measures w_i in the supported internodes. The productions in lines 89–92 capture this propagation in the case of three, two and one supported internodes. The symbol >> in the production predecessors indicates fast information transfer, a variant of context sensitivity that makes it possible to propagate information in one direction (from the extremities of a branching structure towards the base or vice versa) in a single derivation step [29]. The basipetal propagation of width information is a version of the pipe model [65], which relates the diameter of each branch in a tree to the number of leaves it supports. The specification of internodes is complemented by the interpretation rule:

```
111 I(l,w,t) :
112 {
113    nproduce SetColor(br_color(T)) SetWidth(pow(w, 1/eta));
114    for (int i=0; i < N_SEG; i++)
115       nproduce F(l*int_length(t) / N_SEG);
116    produce;
117 }
```

The loop in lines 114 and 115 divides the internode into N_SEG segments of length l*int_length(t). This division makes it possible to simulate gravitropism by slightly reorienting the turtle towards the vertical at each junction between consecutive segments [55]. (A more advanced model of tropisms is presented by Bastien et al. [6], commented on by Dumais [15] and expanded by Bastien et al. [7], and Chelakkot and Mahadevan [12]). The nproduce statement used in lines 113 and 115 differs from the produce statement introduced previously in that it does not terminate the production application, making it possible to successively append modules to the successor [29]. The productions in line 113 sets the internode colour according to plant age T (a global variable) and determines the internode diameter as a function of the width measure w using the equation $d = w^{1/\eta}$.

The operation of the complete *Lychnis* model is illustrated in Fig. 8.11. A comparison with Fig. 8.9 shows that the model correctly captures the architecture of *Lychnis coronaria* plants, although it does not reach branch density of a plant grown over many seasons. Modeling of dense plant structures requires incorporating collisions between plant organs and is a topic of current research [47].

Descriptive models have many applications. For example, they can be used as a synthetic representation of our knowledge of plant form and development [17, 43, 59], as a vehicle for exploring the range of forms that plants in a given class can potentially attain [40] (such explorations may find practical applications in plant breeding), or as the source of ground truth for training artificial intelligence programs intended to automatically recognize plant traits from images [67]. A distinctive feature of plant development is, however, their phenotypic plasticity: the ability of plants with the same genotype to assume different forms depending on the environment in which they grow. We describe the incorporation of environment into L-system plant models next.

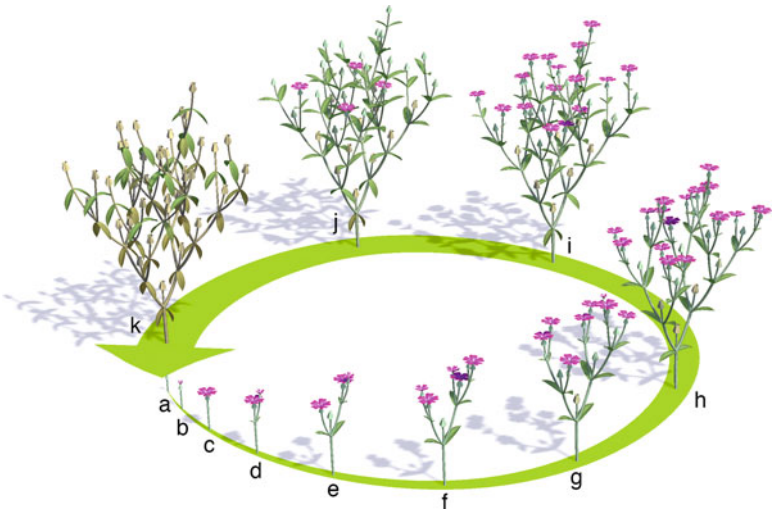


Fig. 8.11 Selected steps in the development of a *Lychnis coronaria* shoot. In the initial shoot axis (**a**), (**b**), (**c**), lateral apices supported by leaves initiate first one (**d**), then another (**e**) lateral branch. In the meantime, the apex terminating the main axis undergoes a transition from the vegetative to the flowering stage, producing a flowering bud (**b**) that gives rise to an open flower (**c**)–(**e**) and, eventually, matures into a seed pod (**f**)–(**k**). This branching pattern repeats, producing branches of increasingly high order (**e**)–(**h**). Upon reaching a threshold age (simulating the end of the growing season), further branching stops (**i**), and the plant gradually reaches maturity (**j**), (**k**). Its seeds give rise to the next generation of plants (**a**)

8.7 Modeling Phenotypic Plasticity

Irrespective of the diversity of environmental factors that affect plant development, including space, light, water, nutrients and pathogens, a unified method for modeling plants in the context of their environment is available. The idea is to model the plant and its environment as separate programs that run concurrently and communicate using standardized programming constructs [44] (Fig. 8.12). In general, this communication is bilateral—the environment affects the plant while the plant reciprocally affects the environment—but simple models may focus on the unilateral influence of the environment on the plant.

An example of such influence is the dependence of the growth rate on temperature. Within some range, this rate is known to be proportional to the difference between the actual temperature τ and a base temperature τ_{base} characteristic of a given plant species, below which growth does not occur [24, 71]. For instance, Fig. 8.13 shows forms created by the *Lychnis* model extended to include the influence of temperature, under the assumption that the temperature of each organ is determined by its position within the canopy. The plant shape is affected by temperature distribution in the environment in which it grows.

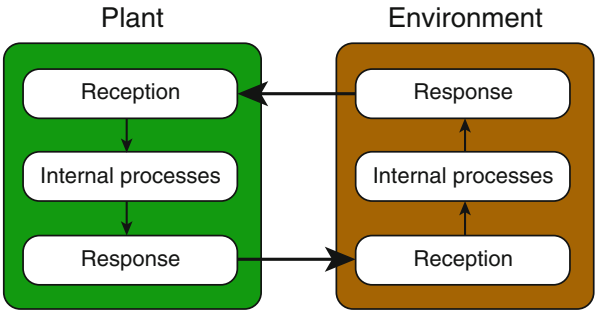


Fig. 8.12 Information flow in a simulation of a plant interacting with its environment. The plant and the environment are modeled by separate communicating programs

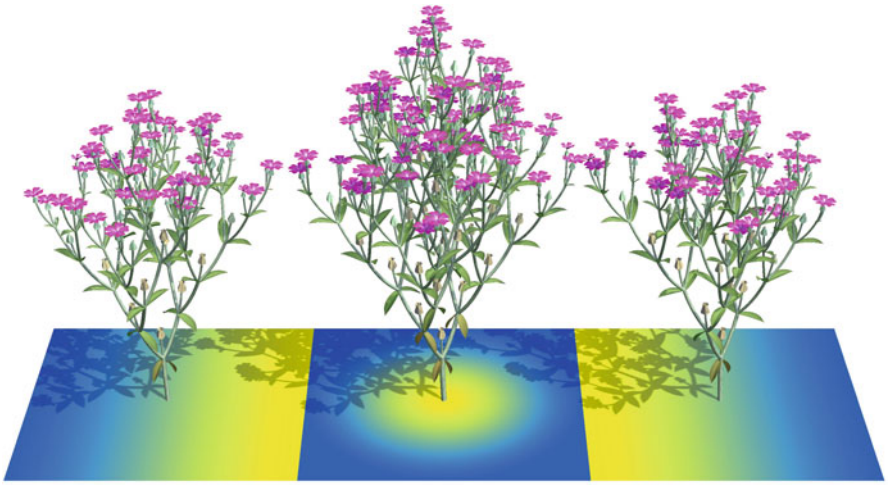


Fig. 8.13 A model of phenotypic plasticity. The simulated *Lychnis* plant grows relatively faster in regions with higher temperatures. Temperature distribution in the plant canopy is defined by the heat map laid on the ground, where yellow indicates warmer regions and blue indicates colder

Specifically, the environmental program returns temperature as defined by an image (texture map): points above yellow regions are warmer than points above blue regions (Fig. 8.13). Information between the plant and its environment is exchanged using predefined communication modules [44, 58]. In the L+C language, these modules are named `E1`, `E2`, `E3`, etc., depending on the number of parameters they carry. In the *Lychnis* model only the `E1` module is used. First, it is inserted into the L-system string at the locations at which the temperature will be queried. This is accomplished by modifying the decomposition rule that controls *Lychnis* branching (lines 95–103 in the original *Lychnis* model) as follows:

```
95  decomposition:
96  A(t):
97  {
98     if (t > 0 && T < BRANCHING_ENDS) produce
99        E1(0) I(LEN1,0,t)
100       SB RollR(PHI) E1(0) B(t) SB E1(0) L(t) EB E1(0) A(D1+t) EB
101       SB RollL(PHI) E1(0) B(t) SB E1(0) L(t) EB E1(0) A(D2+t) EB
102       E1(0) I(LEN2,0,t) E1(0) K(t);
103 }
```

The value (`0`) of the parameter passed to the communication modules `E1` is here irrelevant: there is no information transferred from the plant to the environment other than the location of each module `E1`, which is passed on automatically. This parameter is needed, however, as a placeholder for the temperature information that the environmental program returns to the plant model. With the local temperatures known, the dependence of plant growth on temperature is captured by replacing constant time increments `dt` with the increments `DD(temp)*dt`, a product of the temperature above the threshold measured in Degree Days, and the time increment. For instance, for apices and leaves this dependence is simulated by replacing lines 82 and 83 in the original *Lychnis* model with productions:

```
82  E1(temp) < A(t) : produce A(t+DD(temp)*dt);
83  E1(temp) < L(t) : produce L(t+DD(temp)*dt);
```

The advancement of time in the remaining modules is controlled in a similar manner, resulting in the simulation of plastic behaviour shown in Fig. 8.13.

8.8 Modeling the Development of Trees

In a recursive branching pattern with all internodes of approximately the same length, the number of branches increases exponentially, while the canopy radius—the reach of branches—grows only linearly with time. Eventually, there is not enough room in the canopy for all these branches [9, 52]. The shortage of space is a critical factor in the development of trees due to the high number of branches they could potentially produce during their long life. Borchert and Honda [8] and Sachs and Novoplansky [64] proposed that the necessary limitation of branch proliferation in trees results from their competition for space or light. This competition is not merely a manifestation of tree plasticity, but an essential component of the development in most trees (we exclude here non-branching palms and tree ferns, for example) [48]. Below we show how such competition can be implemented with L-systems, by simulating bilateral communication between the plant and its environment.

The key components of the model and the initial state of the simulation are specified as follows:

```
25 struct internode_data {
26    float l;     // internode length
27    float w;     // internode width
28    int age;     // needed to know when to shed
29 };
30
31 module A(int);              // apex (0: dead, 1: alive)
32 module I(internode_data);   // internode
33
34 internode_data trunk_init = {1.0, w_init, 0};
35
36 Axiom: I(trunk_init) A(1) E1(1) GetHead(0,0,0);
```

The axiom describes the initial configuration of the simulated structure as an apex `A` supported by an internode `I`. The apex is followed by two modules, `E1` and `GetHead`. The predefined module `GetHead`, inserted with arbitrary parameter values—here $(0, 0, 0)$—returns the current direction of the turtle. Its role in the model is described further down. The communication module `E1` provides an interface between the plant model and the environmental program, which in this case tests whether the apex is in proximity of other apices. (Obviously, there is no other apex at the beginning of simulation, but the same sequence of modules is produced as the plant develops and new apices are formed.) The parameter passed to module `E1`—in the axiom, it is a 1—represents apex vigor, and is an input to the environmental program. If the distance of a given apex to all other apices (or environmental modules indicating the presence of branches) is greater than a predefined threshold d, the same module `E1` will return 1, indicating that this apex is not dominated by any other tree component. On the other hand, if this distance is less than d, the environmental program will return 1 or 0, depending on whether the given apex has the highest vigour among all the nearby apices or not. A given apex is thus not dominated if there is enough free space around it, or it has higher vigor than all the modules with which it competes for space (if two or more colliding apices have the same high vigor, all of them are considered dominated). This environmental information affects the development of the simulated tree via the following production:

```
57 I(sl) < A(alive) E1(not_colliding) GetHead(x,y,z) : {
58    internode_data s1 = {sl.l*r1, w_init, 0};
59    internode_data s2 = {sl.l*r2, w_init, 0};
60    if (alive && not_colliding && y > GrowDown) {
61       produce E1(1)
62       SB Left(a1)  RollL(phi)      I(s1) A(1) E1(v1) GetHead(0,0,0) EB
63       SB Right(a2) RollL(phi-180) I(s2) A(1) E1(v2) GetHead(0,0,0) EB;
64    }
65    else
66       produce A(0);
67 }
```

According to Line 57, the production predecessor consists of three modules, `A`, `E1` and `GetHead`, in the context of a supporting internode `I`. Line 60 states that the fate of an apex depends qualitatively on three factors: whether it is “alive" (i.e., was never dominated), whether it is dominated now, and what orientation it has. An apex that has never been dominated and is not oriented too steeply down gives rise to two branches (lines 62 and 63). Their length is reduced with respect to that

of the supporting internode (`sl.l`) by predefined factors `r1` and `r2` (lines 58 and 59). The apices of the new branches are followed by modules `E1` and `GetHead`, allowing for the application of the same production and repetition of the same decision process in the next simulation step. The vigor values $0 < \texttt{v1} < \texttt{v2} < 1$ are predefined constants. The additional environmental module with argument 1 (maximum vigor), introduced in line 61, marks the branching point itself to assure that no apex will ever grow into a previously formed branch. The modules specified in lines 61–63 are not produced, however, if the condition in line 60 is not satisfied. In that case, the apex changes its state to "dead" (`A(0)` in line 66), and it will produce no further branches.

The effect of the detection and response to overcrowding is illustrated by Fig. 8.14. The branching structure in Fig. 8.14a was generated by a model in which detection of overcrowding was disabled (the condition in line 60 was replaced by the always-true statement `if(1)`). The produced structure is then exceedingly dense, with apices and branches running into each other and overlapping. In contrast, the structure in Fig. 8.14b was generated with the condition in line 60 present. Predictably, it is much more sparse: overcrowding has been prevented.

In nature, branches that have ceased growing are often shed by the tree. A simple instance of this process is illustrated in Fig. 8.14c. The shedding criterion used here is the age of a non-growing branch: the time since the last apex supported by it has become dominated. Shedding is implemented by the following production and decomposition rules:

```
77 production:
78 I(s) >> A(alive) : // advance branch age if apex is dead
79 {
80    if (!alive)
81       s.age += 1 ;
82    produce I(s);
83 }
84
85 I(s) >> SB I(sr) EB SB I(srr) EB :  // propagate girth and age
86 {
87    s.w = sr.w + srr.w;
88    s.age = min(sr.age, srr.age);
89    produce I(s);
90 }
91
92 decomposition:
93 I(s) :   // tag branch for shedding if too old
94 {
95    if(s.age >= ShedAge)
96       produce I(s) Cut();
97 }
```

The first production (lines 78–83) increments the age of an internode followed by a dominated apex. The second production (lines 85–90) propagates the age information basipetally. When branches that meet at a given branching point have different age values, the supporting internode is assigned the smaller value of the two (line 88); consequently, each internode "knows" the age of the youngest internode it supports. The actual shedding is implemented by the decomposition rule in lines 93–97. When the age of an internode exceeds threshold `ShedAge`, this rule inserts a predefined module `Cut`, which (in the next simulation step) removes the entire branch that follows it.

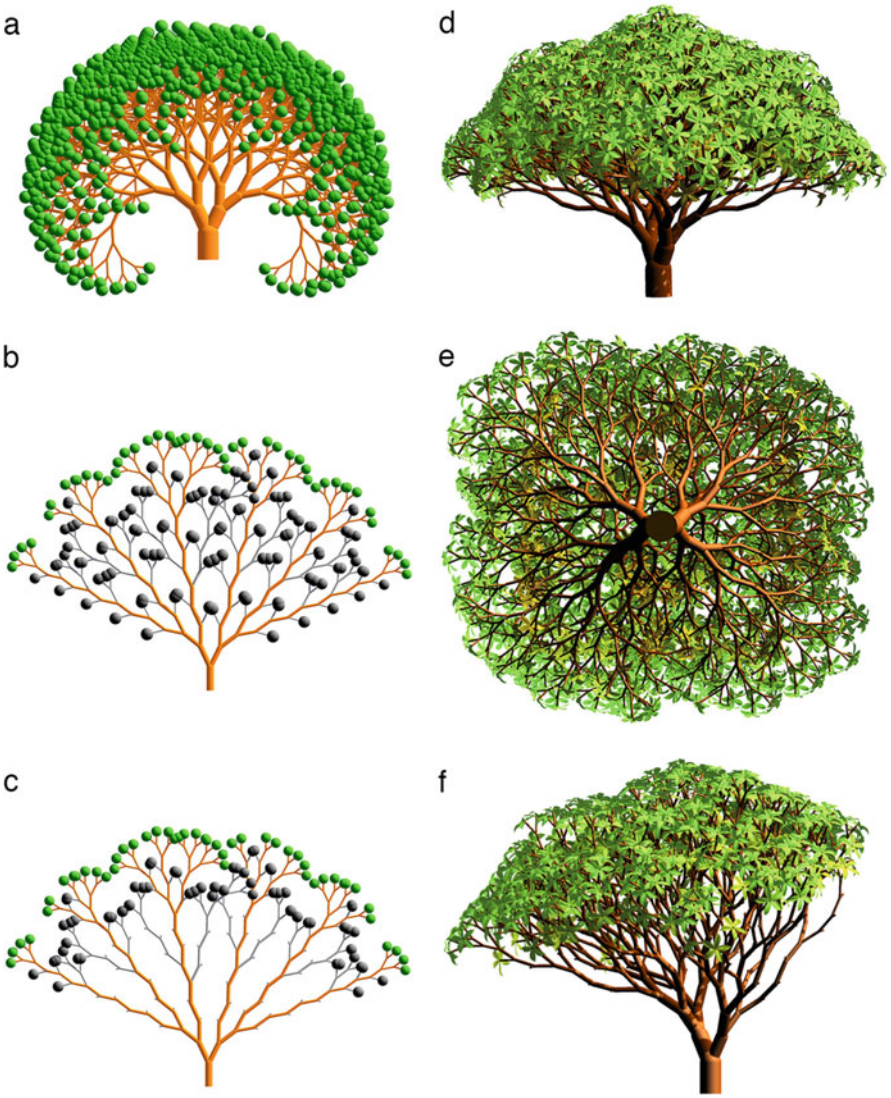


Fig. 8.14 Elements of tree modeling. (**a**) A planar structure developing with branch proliferation unchecked. The “tree" crown is clearly overcrowded. (**b**) A modification of the tree from Figure (**a**), in which dominated branches die. Dead apices and internodes are shown in grey. The structure is no longer overcrowded. (**c**) A modification of the model from Figure (**b**), in which older dead branches are shed. (**d**) A three-dimensional modification of the model from Figure (**c**), in which phyllotaxis has been changed from distichous ($\varphi = 0$) to decussate ($\varphi = 90°$), and leaves grow at the tips of branches. The difference in the diameter of branches results from the application of the pipe model. (**e**) The tree from panel (**d**) seen from the bottom. Note the plausible architecture resulting from the competition of branches for space. (**f**) A variant of the model from Figure (**d**), in which the environmental program simulates competition for light rather than for space. A light source is positioned above and to the left of the tree, resulting in an asymmetric crown, leaning to the left

The production in lines 85–90 also includes the assignment in line 87 that propagates a measure of the internode diameter according to the pipe model as discussed for *Lychnis*. Proper modeling of the girth of branches is an important element of the visually realistic modeling of trees. Remarkably, the inclusion of these two principles, the competition for space combined with shedding (extended from two to three dimensions by changing the phyllotaxis-defining parameter `phi` from 0 to 90° in lines 62 and 63) and the assignment of branch width according to the pipe model, suffice to generate a visually plausible tree architecture (Fig. 8.14d, e) resembling *Dracaena draco* (the dragon tree).

The use of separate programs to model the plant and its environment (Fig. 8.12) facilitates simulation of different plants in the same environment or, conversely, the same plant in different environments. For instance, Fig. 8.14f shows the result of substituting competition for light for competition for space in the simulation of tree development. The required modification of the tree model itself was limited to a single production in lines 57–67 (compare the original listing with the code below):

```
57 I(sl) < A(alive) E1(light) GetHead(x,y,z) : {
58    internode_data s1 = {sl.l*r1, w_init, 0};
59    internode_data s2 = {sl.l*r2, w_init, 0};
60    if (alive && light>Th && y > GrowDown) {
61       produce
62       SB Left(a1)  RollL(phi)     I(s1) A(1) E1(R) GetHead(0,0,0) EB
63       SB Right(a2) RollL(phi-180) I(s2) A(1) E1(R) GetHead(0,0,0) EB;
64    }
65    else
66       produce A(0);
67 }
```

The module `E1` passes the size of leaf clusters to the environment (predefined constant `R` in lines 63 and 64) and receives information about the intensity of light reaching each cluster in return (variable `light` in line 57). An apex is not dominated and can produce new branches (provided that the remaining conditions in line 60 are satisfied) if this intensity is greater than a predefined threshold `Th` (condition `light>Th`). Note the absence of module `E1` at the branching point of the light-driven model (compare lines 61 in both listings): this module is now not needed, because there are no leaf clusters at the branching points. A comparison of Fig. 8.14d, f shows that the tree forms generated assuming competition for space or competition for light may substantially differ, even though the underlying tree model is basically the same. Not surprisingly, the environment has a significant impact on tree development and must be modeled carefully when a faithful simulation of developmental processes in nature is sought.

8.9 Conclusion

In the 50 years since its inception, the formalism of L-systems has become a powerful research tool in developmental biology. Remarkably, it was not imported from another discipline, but created specifically for developmental biology. With

their many extensions, L-systems can be and have been applied to an ever-increasing variety of problems spanning a wide range of scales, from sub-cellular to whole plants and plant ecosystems [34], as well as modeling styles, from descriptive to mechanistic. The key limitation of L-systems is their restriction to linear (filamentous) and branching structures. Nevertheless, the topological approach to modeling development, inherent in L-systems has also inspired research and the establishment of practical methods for modeling growing cellular layers and volumetric structures [33, 55, 66]. L-systems, their extensions and applications continue to be an active and fascinating area of research.

Acknowledgements We thank John Hall for the prototype version of the timeline editor featured in Fig. 8.10, Andrew Owens for help with Fig. 8.11, and Lynn Mercer for insightful discussions. The authors' research on the L-system-based modeling methods, specific models, and the Virtual Laboratory (`vlab`) software used in the preparation of this paper was supported by the Natural Sciences and Engineering Research Council and the Plant Phenotyping Imaging and Research Centre/Canada First Research Excellence Fund. This support is gratefully acknowledged.

References

1. Abelson H, diSessa AA (1982) Turtle geometry. MIT Press, Cambridge
2. Adams D, Duggan P (1999) Heterocyst and akinete differentiation in cyanobacteria. New Phytol 144(1):3–33
3. Algorithmic botany (2018) The Virtual Laboratory/L-studio software distribution. http://algorithmicbotany.org/virtual_laboratory
4. Baker R, Herman GT (1970) CELIA — a cellular linear iterative array simulator. In: Proceedings of the fourth conference on applications of simulation, 9–11 December 1970, pp 64–73
5. Baker R, Herman GT (1972) Simulation of organisms using a developmental model, Parts I and II. Int J Bio-Med Comput 3:201–215, 251–267
6. Bastien R, Bohr T, Moulia B, Douady S (2013) Unifying model of shoot gravitropism reveals proprioception as a central feature of posture control in plants. Proc Natl Acad Sci 110(2):755–760
7. Bastien R, Douady S, Moulia B (2015) A unified model of shoot tropism in plants: photo-, gravi-and propio-ception. PLoS Comput Biol 11(2):e1004037
8. Borchert R, Honda H (1984) Control of development in the bifurcating branch system of *Tabebuia rosea*: a computer simulation. Bot Gaz 145:184–195
9. Borchert R, Slade N (1981) Bifurcation ratios and the adaptive geometry of trees. Bot Gaz 142(3):394–401
10. Boudon F, Pradal C, Cokelaer T, Prusinkiewicz P, Godin C (2012) L-Py: an L-system simulation framework for modeling plant architecture development based on a dynamic language. Front Plant Sci 3:76
11. Buikema W, Haselkorn R (1991) Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. Genes Dev 5(2):321–330
12. Chelakkot R, Mahadevan L (2017) On the growth and form of shoots. J R Soc Interface 14(128):1–6
13. Coen E, Rolland-Lagan AG, Matthews M, Bangham A, Prusinkiewicz P (2004) The genetics of geometry. Proc Natl Acad Sci USA 101:4728–4735
14. de Koster CG, Lindenmayer A (1987) Discrete and continuous models for heterocyst differentiation in growing filaments of blue-green bacteria. Acta Biotheor 36:249–273

15. Dumais J (2013) Beyond the sine law of plant gravitropism. Proc Natl Acad Sci 110(2):391–392
16. Fogg G (1949) Growth and heterocyst production in *Anabaena cylindrica* Lemm. in relation to carbon and nitrogen metabolism. Ann Bot 13(51):241–259
17. Fournier C, Andrieu B (1998) A 3D architectural and process-based model of maize development. Ann Bot 81:233–250
18. Fritsch F (1951) The heterocyst: a botanical enigma. Proc Linnean Soc Lond 162(2):194–211
19. Gerdtzen Z, Salgado J, Osses A, Asenjo J, Rapaport I, Andrews B (2009) Modeling heterocyst pattern formation in cyanobacteria. BMC Bioinf 10(6):S16
20. Giavitto JL, Godin C, Michel O, Prusinkiewicz P (2002) Computational models for integrative and developmental biology. LaMI Rapport de Recherche 72–2002, CNRS — Université d'Evry val d'Essonne
21. Gierer A, Meinhardt H (1972) A theory of biological pattern formation. Kybernetik 12:30–39
22. Hammel M, Prusinkiewicz P (1996) Visualization of developmental processes by extrusion in space-time. In: Proceedings of graphics interface '96, pp 246–258
23. Hanan JS (1992) Parametric L-systems and their application to the modelling and visualization of plants. PhD thesis, University of Regina
24. Hanan JS (1997) Virtual plants — integrating architectural and physiological models. Environ Model Softw 12:35–42
25. Haselkorn R (1998) How cyanobacteria count to 10. Science 282:891–892
26. Herrero A, Stavans J, Flores E (2016) The multicellular nature of filamentous heterocyst-forming cyanobacteria. FEMS Microbiol Rev 40(6):831–854
27. Huxley JS (1924) Constant differential growth ratios and their significance. Nature 114:895–896
28. Huxley JS (1932) Problems of relative growth. MacVeagh, London
29. Karwowski R, Prusinkiewicz P (2003) Design and implementation of the L+C modeling language. Electron Notes Theor Comput Sci 86(2):134–152
30. Kniemeyer O (2004) Rule-based modelling with the XL/GroIMP software. In: The logic of artificial life: abstracting and synthesizing the principles of living systems; Proceedings of the 6th German workshop on artificial life, April 14–16, 2004, Bamberg, AKA Akademische Verlagsgesellschaft, Berlin, pp 56–65
31. Kniemeyer O, Buck-Sorlin G, Kurth W (2007) GroIMP as a platform for functional-structural modelling of plants. In: Vos J et al (eds) Functional-structural modeling in crop production. Springer, Dordrecht, pp 43–52
32. Kurth W (1994) Growth grammar interpreter GROGRA 2.4: a software tool for the 3-dimensional interpretation of stochastic, sensitive growth grammars in the context of plant modeling. Introduction and reference manual. Forschungszentrum Waldökosysteme der Universität Göttingen, Göttingen
33. Lane B (2015) Cell complexes: the structure of space and the mathematics of modularity. PhD thesis, University of Calgary
34. Lane B, Prusinkiewicz P (2002) Generating spatial distributions for multilevel models of plant communities. In: Proceedings of graphics interface 2002, pp 69–80
35. Lindenmayer A (1968) Mathematical models for cellular interaction in development, Parts I and II. J Theor Biol 18:280–315
36. Lindenmayer A (1971) Developmental systems without cellular interaction, their languages and grammars. J Theor Biol 30:455–484
37. Lindenmayer A (1974) Adding continuous components to L-systems. In: Rozenberg G, Salomaa A (eds) L Systems. Lecture notes in computer science, vol 15. Springer, Berlin, pp 53–68
38. MacDonald N (1983) Trees and networks in biological models. Wiley, New York
39. MacNamara S, Strang G (2016) Operator splitting. In: Glowinski R, Osher S, Yin W (eds) Splitting methods in communication, imaging, science, and engineering. Springer, Berlin, pp 95–114

40. McGhee G (1999) Theoretical morphology: the concept and its applications. Columbia University Press, New York
41. Meinhardt H (1982) Models of biological pattern formation. Academic, London
42. Mitchison G, Wilcox M (1972) Rules governing cell division in *Anabaena*. Nature 239:110–111
43. Mündermann L, Erasmus Y, Lane B, Coen E, Prusinkiewicz P (2005) Quantitative modeling of Arabidopsis development. Plant Physiol 139:960–968
44. Měch R, Prusinkiewicz P (1996) Visual models of plants interacting with their environment. In: Proceedings of SIGGRAPH 1996, pp 397–410
45. Niklas KJ (1994) Plant allometry: the scaling of form and process. The University of Chicago Press, Chicago
46. Niklas KJ (2004) Plant allometry: is there a grand unifying theory? Biol Rev 79(4):871–889
47. Owens A, Cieslak M, Hart J, Classen-Bockhoff R, Prusinkiewicz P (2016) Modeling dense inflorescences. ACM Trans Graph 35(4):136
48. Palubicki W, Horel K, Longay S, Runions A, Lane B, Měch R, Prusinkiewicz P (2009) Self-organizing tree models for image synthesis. ACM Trans Graph 28:58
49. Papert S (1980) Mindstorms: children, computers and powerful ideas. Basic Books, New York
50. Prusinkiewicz P (1986) Graphical applications of L-systems. In: Proceedings of graphics interface '86 — vision Interface '86, pp 247–253
51. Prusinkiewicz P (2004) Art and science for life: designing and growing virtual plants with L-systems. Acta Horti 630:15–28
52. Prusinkiewicz P, de Reuille PB (2010) Constraints of space in plant development. J Exp Bot 61:2117–2129
53. Prusinkiewicz P, Hanan J (1990) Visualization of botanical structures and processes using parametric L-systems. In: Thalmann D (ed) Scientific visualization and graphics simulation. Wiley, Chichester, pp 183–201
54. Prusinkiewicz P, Lane B (2013) Modeling morphogenesis in multicellular structures with cell complexes and L-systems. In: Capasso V et al (eds) Pattern Formation in Morphogenesis, Springer, Berlin, pp 137–151
55. Prusinkiewicz P, Lindenmayer A (1990) Hanan JS, Fracchia FD, Fowler DR, de Boer MJM, Mercer L (eds) The algorithmic beauty of plants. Springer, New York
56. Prusinkiewicz P, Runions A (2012) Computational models of plant development and form. New Phytol 193:549–569
57. Prusinkiewicz P, Hammel M, Mjolsness E (1993) Animation of plant development. In: Proceedings of SIGGRAPH 1993, pp 351–360
58. Prusinkiewicz P, James M, Měch R (1994) Synthetic topiary. In: Proceedings of SIGGRAPH 1994, pp 351–358
59. Prusinkiewicz P, Remphrey W, Davidson C, Hammel M (1994) Modeling the architecture of expanding *Fraxinus pennsylvanica* shoots using L-systems. Can J Bot 72:701–714
60. Prusinkiewicz P, Hanan J, Měch R (2000) An L-system-based plant modeling language. In: Nagl M, Schürr A, Münch M (eds) Applications of graph transformations with industrial relevance. Lecture notes in computer science, vol 1779. Springer, Berlin, pp 395–410
61. Prusinkiewicz P, Karwowski R, Lane B (2007) The L+C plant-modeling language. In: Vos J et al (eds) Functional-structural modeling in crop production. Springer, Dordrecht, pp 27–42
62. Robinson S, de Reuille PB, Chan J, Bergmann D, Prusinkiewicz P, Coen E (2011) Generation of spatial patterns through cell polarity switching. Science 333:1436–1440
63. Room PM, Maillette L, Hanan J (1994) Module and metamer dynamics and virtual plants. Adv Ecol Res 25:105–157
64. Sachs T, Novoplansky A (1995) Tree form: architectural models do not suffice. Isr J Plant Sci 43:203–212
65. Shinozaki K, Yoda K, Hozumi K, Kira T (1964) A quantitative analysis of plant form — the pipe model theory. I. Basic analyses. Jpn J Ecol 14(3):97–104
66. Smith C (2006) On vertex-vertex systems and their use in geometric and biological modeling. PhD thesis, University of Calgary

67. Ubbens J, Cieslak M, Prusinkiewicz P, Stavness I (2018) The use of plant models in deep learning: an application to leaf counting in rosette plants. Plant Methods 14:6:1–10
68. Ulam S (1962) On some mathematical properties connected with patterns of growth of figures. In: Proceedings of symposia on applied mathematics, vol 14. American Mathematical Society, Providence, RI, pp 215–224
69. Ulam S (1966) Patterns of growth of figures: mathematical aspects. In: Kepes G (ed) Module, Proportion, Symmetry, Rhythm. Braziller, New York, pp 64–74
70. von Neumann J (1966) Burks AW (ed) Theory of self-reproducing automata. University of Illinois Press, Urbana
71. Wardlaw I (1999) Thermal time. In: Atwell B, Kriedemann P, Turnbull (eds) Plants in action: adaptation in nature, performance in cultivation, Macmillan Education Australia, Melbourne
72. Wilcox M, Mitchison GJ, Smith RJ (1973) Pattern formation in the blue-green alga, *Anabaena*. I. Basic mechanisms. J Cell Sci 12:707–723
73. Wolfram S (1984) Universality and complexity in cellular automata. Physica D 10(1–2):1–35
74. Wolfram S (2002) A new kind of science. Wolfram Media/Cambridge University Press, Champaign, IL
75. Yoon HS, Golden JW (1998) Heterocyst pattern formation controlled by a diffusible peptide. Science 282:935–938

Chapter 9
Flowering Time as a Model Trait to Bridge Proximate and Evolutionary Questions

Akiko Satake

Abstract The transition from vegetative to reproductive growth is a key developmental event that is directly linked to reproductive success in plants. To understand this developmental trait in the context of evolution, the synthesis of molecular/genetic studies and ecology/evolutionary studies is necessary. In this chapter, we present a mathematical approach for synthesizing this body of knowledge to consider different flowering strategies. We model plant life history based on molecular mechanisms of the underlying gene regulatory network of flowering. We then extend the model to predict future flowering phenology under a changing climate and end with conclusions and perspectives.

9.1 Introduction

Almost 50 years ago, Niko Tinbergen, one of the founders of ethology, suggested that four fundamentally different types of problems need to be answered to fully understand the behavior of animals [1]. This is now well known as "Tinbergen's four questions." Four questions about any feature of an organism can be expressed as: what is it for (adaptive significance)? How did it develop during the lifetime of the individual (ontogeny)? How did it evolve over the history of the species (evolution)? And, how does it work (causation)? Two of Tinbergen's four questions (ontogeny and causation) address the *proximate causes* of behavior and emphasize mechanisms within an organism that are responsible for the behavior. The other two questions (adaptive significance and evolution) address the *ultimate causes* of behavior and focus on why a living organism has evolved the aforementioned mechanisms to produce behavior.

Tinbergen's four questions apply broadly to any characteristic of living organisms including plants. In plant biology, the proximate mechanisms underlying

A. Satake (✉)
Department of Biology, Faculty of Science, Kyushu University, Fukuoka, Japan
e-mail: akiko.satake@kyudai.jp

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_9

plant development and environmental responses have been unraveled by remarkable progress in molecular and genetic studies over the past decade. A number of genes and their regulatory networks involved in stress response, metabolism, growth, and reproduction have been identified given the genetic blueprints of two model organisms *Arabidopsis thaliana* [2] and rice *Oryza Sativa* [3, 4]. Ultimate causes of plant behavior in diverse species have been studied in ecology and evolutionary biology considering the environmental response of plants as an adaptive trait that facilitates efficient survival and reproduction in a given environment [5].

After 50 years of Tinbergen's seminal paper, the real question now is how to best integrate the four questions. The integration of the two different aspects of biology would offer great potential to understand the molecular mechanisms by which plants respond to environmental signals but also to address the evolutionary significance of the underlying molecular basis that should have benefit plants to cope with environmental fluctuation and heterogeneity across the globe [6]. However, relatively few cases have had all four questions answered for a single characteristic [7], probably because researchers are focused on a particular aspect of biological problem rather than integration. This chapter aims to show that studies about flowering time, one of critical reproductive traits in angiosperm, is a model trait to show the usefulness of this integrated approach.

9.2 Flowering Time as a Model Trait for Integration of Proximate and Ultimate Causes

The transition from vegetative to reproductive growth is a key developmental event that is directly linked to reproductive success in plants. Flowering behaviors vary widely among taxa. Some plants are annuals that flower only once and complete their life cycle after reproduction in a year. Other plants are perennials that live for many years and flower repeatedly. Many woody plants exhibit a highly variable flowering effort between years. Understanding why and how such variations in flowering behavior arise is a major challenge in ecology and evolutionary biology. In ecology and evolutionary biology, since the 1970s, the growth and reproductive schedule of terrestrial plants has been studied as an adaptive strategy that maximizes lifetime reproductive success by applying optimal control theory or game theory that has been exploited in the fields of engineering and economics. For example, the evolution of the reproductive schedule of annual and perennial plants has been discussed theoretically based on the optimal resource allocation between vegetative and reproductive tissues throughout the plant lifecycle [8–12], and these mathematical models had been extended to explain how plant–animal interaction drives the evolution of reproductive timing [13, 14]. Moreover, how plant size and age influence flowering time have been studied in the framework of adaptive dynamics or integral projection model [15, 16].

In the 1970s, little was known about molecular mechanisms that underlie the diversification of flowering behavior. Meanwhile, in the 1980s, rapid progress in genetic and molecular analyses has been made in model plants, providing the genetic description of well-established flowering pathways. Studies of floral homeotic mutants of *Arabidopsis* and *Antirrhinum* led to the proposal of the ABC model of flower development, in which the spatially localized activity of homeotic gene products specifies the identities of floral organs such as sepal, petal stamen, and carpel [17–19]. Genes involved in controlling the timing of the floral transition have been identified through mutagenesis and analysis of natural variation ([20, 21]), uncovering the genetic pathways that integrate environmental cues and developing the state of plants particularly well in *A. thaliana* [22–25].

Now it is time to synthesize the body of knowledge from molecular/genetic studies and ecology/evolutionary studies to draw the history of evolutionary alterations of the underlying molecular bases of plant's flowering behavior in geographically diverse environments [6, 26]. Plants' environments have changed continuously in the past, and will change more rapidly under the influence of climate change in the future. Adaptive responses to local climate in the past may no longer be optimal in the future [27]. Forecasting future changes in plant flowering behavior and understanding how natural selection will lead to the further evolution of the gene regulatory mechanisms of flowering time in future environments would offer advantages for systematic conservation planning of natural ecosystems and also will benefit the vulnerability assessment of an important ecosystem service, food production, on which human societies depend. Based on these perspectives, in this chapter, two topics, a genetic-physiological model for plant life history and a predictive model of flowering phenology, will be introduced.

9.3 Modeling Life History Evolution Based on Molecular Mechanisms

9.3.1 Natural Variations in the Vernalization Responses

Many plants that inhabit seasonally changing environments synchronize their flowering time to coincide with appropriate seasons by monitoring and responding to environmental cues. Photoperiod and temperature are two of the main environmental signals used as cues to interpret seasons. Despite an accurate and predictable change in photoperiod, temperature fluctuates day-by-day in a very noisy manner, and the short-term trend in temperature is not foreseeable. To cope with unpredictable fluctuations of temperature, the plant has evolved to utilize a long-term trend in temperature, especially a prolonged period of cold. The ability to respond to a prolonged winter season of cold, but not to a short spell of cold, prevents flowering prior to (or during) winter and permits flowering in the favorable conditions in spring because it effectively filters out daily variations and extract a reliable long-term

trend in seasonal changes of temperature. The process that promotes flowering after exposure to prolonged period of winter cold is known as vernalization (reviewed in [28, 29]).

Arabidopsis thaliana, an annual herb, provides an opportunity to determine the genetic basis that contributes to natural variations in the vernalization response [30–34]. *A. thaliana* consists of winter annuals, in which flowering is strongly delayed unless plants experience a prolonged period of cold (i.e., become vernalized), as well as summer annuals that complete flowering and seed production before winter comes [35, 36]. Analysis of a large number of accessions across multiple environments showed that two genes, *FRIGIDA* (*FRI*) and *FLOWERING LOCUS C* (*FLC*), are key factors in the vernalization pathway [37–39].

FRI delays flowering by maintaining high expression levels of *FLC*, which encodes a MADS domain protein that represses flowering [40–42]. Vernalization represses this repressor to accelerate flowering time by activating floral pathway integrators, which include FLOWERING LOCUS T (FT; [43, 44]), known as a universal factor of floral transition that functions as a long-distance signal of floral transition and is transported from leaves to the shoot meristem where flowers are formed. Floral pathway integrators activate target genes promoting floral organ identity, including *LEAFY* (*LFY*; [45, 46]), *FRUITFULL* (*FUL*; [45, 47]), and *APETALA 1*(*AP1*; [45]), which results in the initiation of flowering. Molecular analysis of *FRI* revealed that loss-of-function mutations at this locus promote flowering without vernalization by reducing expression levels of *FLC*, which allows the evolution of summer annuals from winter annuals. *FLC* has been shown to interact epistatically with *FRI* [48], and summer annuals can occur when a nonfunctional *FRI* is present and *FLC* is weak or defective [49, 50].

In addition to natural variations in the timing of flowering in annuals, the vernalization response in perennials, such as *Arabis alpina* and *Arabidopsis halleri* subsp. *Gemmifera*—a close relative of *Arabidopsis*—has been studied [51, 52]. In *A. thaliana*, the expression of *FLC* is reduced during cold and epigenetic modifications are acquired at the *FLC* locus, which maintain low levels of expression even after a return to warm conditions. Stable maintenance of *FLC* silencing even after the arrival of spring is called winter memory [29]. In contrast, the ortholog of *FLC* in *A. alpina* and *A. halleri* is not stably repressed after winter cold. It is instead reactivated once the temperature increases both in laboratory [52] and in natural conditions [51]. This reversible expression of *FLC* facilitates a reversion from reproductive to vegetative development and promotes the perennial life cycle [51, 52]. Thus, regulation of *FLC* is considered one of the key factors explaining natural variation in flowering timing both in annuals and perennials, and variations in *FLC* regulation result in different fitness consequences, providing a basis for adaptation to local temperature environments. Mathematical models that describe genetic-physiological processes of the plant life cycle will be useful to understand how plants adjust their vernalization pathways to adapt to local environmental conditions.

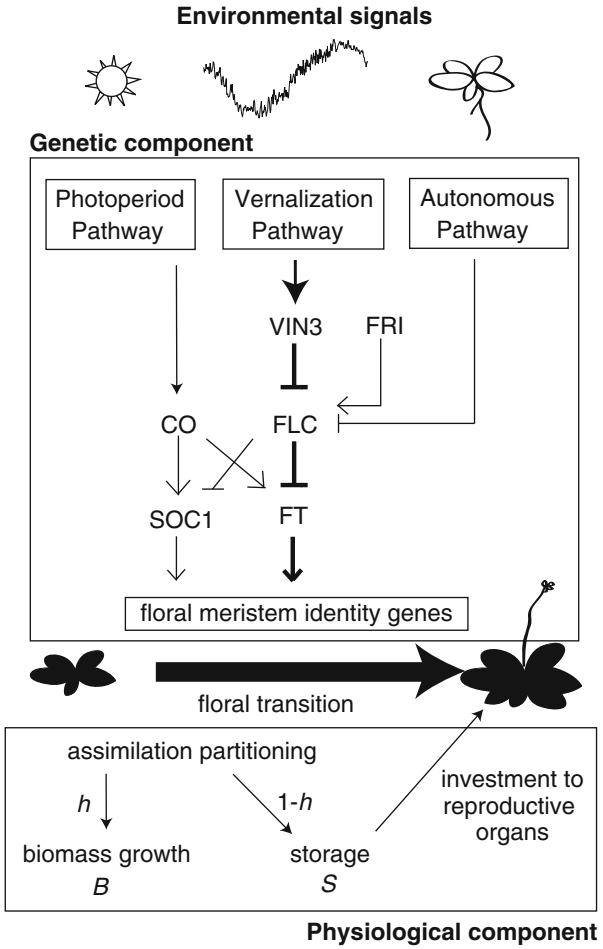


Fig. 9.1 Schematic diagram of the genetic-physiological model for flowering. In the genetic component, different environmental signals transmitted in different pathways are integrated by floral pathway integrators such as *FT* and *SOC1*. In the physiological component, plants are assumed to allocate the *h* fraction of photoassimilates to vegetative growth, and rest of photoassimilates (1-*h*) are stored for reproduction

9.3.2 Genetic-Physiological Model for Flowering

Signal transmission in the vernalization pathway and the physiological process regarding growth and resource partitioning between vegetative and reproductive growth have been modeled previously ([53]; Fig. 9.1). In the signal transmission process, the model focused on the key temperature-dependent pathway to flowering—the transcriptional regulation of two vernalization-specific genes (*VERNALIZATION-INSENSITIVE 3*: *VIN3* and *FLC*) and one floral integrator gene

(*FT*). *VIN3* is one of the initial repressors of *FLC* during cold treatment. *VIN3* expression is induced after several weeks of cold treatment [36]. Dynamics of *VIN3* transcript level per unit leaf biomass can be formalized as

$$\frac{dx_{VIN3}}{dt} = \beta_V \theta\left(T < K_T\right) - \alpha_V x_{VIN3}, \tag{9.1a}$$

where

$$\theta\left(T < K_T\right) = \begin{cases} 1 & \text{if } T < K_T \\ 0 & \text{otherwise} \end{cases}. \tag{9.1b}$$

β_V and α_V in Eq. (9.1a) describe the production and degradation rates of the *VIN3* transcript, and K_T is the threshold temperature below which transcription of *VIN3* is activated. θ is the step function that satisfies Eq. (9.1b)—if the temperature (T) is below the threshold temperature (K_T), θ equals 1; otherwise, it is 0.

Activation of *VIN3* triggers chromatin remodeling and leads to an increase in the levels of repressive histone modifications at *FLC* chromatin, such as histone H3 lysine 27 di- and trimethylation (H3K27me2, H3K27me3), as well as the loss of histone modifications associated with active transcription, such as histone H3 acetylation and histone H3 lysine 4 di- and trimethylation (H3K4me2, H3K4me3) [36, 54–57]. This chromatin remodeling results in a stable silencing of *FLC*, which functions as a stable epigenetic memory of winter. Although the quantitative relationship between *FLC* repression and the chromatin modifications at the cellular level is still unclear because of the difficulty to measure *FLC* histone modification levels at the cellular level, mathematical models suggest that *FLC* expression level at the cellular level should take binary states either in fully active or fully repressed to realize resistance to high noise due to random gain or loss of histone modifications and to random partitioning upon DNA replication [58, 59]. *FLC* expression patterns under continuous and interrupted cold exposure has been best explained by the model assuming digital response of cells to cold signal [60]. Moreover, the importance of bistability of cellular systems for robust response to noisy signals has also been demonstrated in other eukaryotic system [61]. Thus, in a model, it is reasonable to assume that a single cell can be in two states, repressed or active. Let R_{FLC} be the fraction of cells that are in the repressed state. We assume that R_{FLC} changes according to

$$\frac{dR_{FLC}}{dt} = q\theta\left(x_{VIN3} > K_{VIN3}\right)\left(1 - R_{FLC}\right) - pR_{FLC}. \tag{9.2}$$

q represents the transition rate from the active to the repressive state. I call it the "repression rate" of *FLC* activity. When *VIN3* transcripts are accumulated and the transcript level exceeds the threshold K_{VIN3}, repressive histone modification is initiated, and active cells change to the repressed state at a rate q. Repressed cells change to the active state at a rate p. The second term on the right-hand side

of Eq. (9.2) describes the fraction of cells that lose repressive chromatin marks. The difference in q represents the differential duration of winter memory between annuals and perennials [59]. *FLC* transcript level per unit leaf biomass (x_{FLC}) is given by

$$x_{FLC}(t) = r\,(1 - R_{FLC}(t))\,, \tag{9.3}$$

where r is the maximum transcript level. A large r represents a genetic background with functional *FRI*, while a small r corresponds to that of nonfunctional *FRI*. *FLC* transcript levels linearly decrease as the fraction of repressed cells increases.

The repression of *FLC* in turn activates transcription of *FT* and production of *FT* protein is initiated in the vascular tissue of leaves. The concentration of *FT* protein per unit leaf biomass (x_{FT}) changes according to

$$\frac{dx_{\mathrm{FT}}}{dt} = \beta_{FT}\theta\,(x_{FLC} < K_{FLC}) - \alpha_{FT}x_{\mathrm{FT}}, \tag{9.4}$$

in which β_{FT} and α_{FT} describe the production and degradation rates of *FT* protein, and K_{FLC} represents the threshold *FLC* transcript level below which production of *FT* protein is initiated. Because strong correlation between *FT* expression level and the onset of flowering has been reported in many plant species [62–65], we assumed that plants start bolting when the total amount of FT protein is above the threshold K_{FT}:

$$x_{\mathrm{FT}}B > K_{FT}\,, \tag{9.5}$$

where x represents the size of total leaf biomass.

The growth of total leaf biomass (B) was formalized based on carbon partitioning between vegetative to reproductive growth (Fig. 9.1). Plants gain carbohydrates by photosynthesis and allocate some fraction of carbon gain to vegetative growth. B changes according to

$$\frac{dB}{dt} = h\,[1 - \theta\,(x_{\mathrm{FT}}B > K_{FT})]\,\frac{aB}{1+bB} - d_B B, \tag{9.6}$$

where h describes the fraction of carbon gain allocated to vegetative growth. $aB/(1 + bB)$ indicates the gain of carbohydrates after deducing maintenance costs of plant organs, and it saturates with an increase in leaf biomass B because of self-shading. a and b are parameters that determine the function of carbon gain. This formulation is built on the classic optimization model of plant life history studied by Iwasa and Cohen [11]. d_B represents the loss of leaf biomass due to natural damages such as herbivory. When floral transition occurs, resource allocation switches from vegetative to reproductive growth, and entire carbon gains (remained after maintenance costs of plant organs are deduced) are used to produce flowers and seeds.

Before the floral transition, the residual of the resource gain, denoted as $(1 - h)aB/(1 + bB)$, is allocated to storage that will be used to produce flowers, fruits, and seeds. The amount of stored resource (S) changes as follows:

$$\frac{dS}{dt} = (1 - h\,[1 - \theta\,(x_{\mathrm{FT}}B > K_{FT})])\,\frac{aB}{1 + bB} - c_f\theta\,(x_{\mathrm{FT}}B > K_{FT}) - c_S y - d_S S. \tag{9.7}$$

The first term on the right-hand side describes a resource increment by photosynthesis, and the second and the third terms represent the cost to produce reproductive organs. c_f indicates the rate of resource investment for flowers. Flowers are fertilized and fruits and seeds are developed. The rate of resource investment to produce a set of fruit and seeds per fertilized flower is represented as c_S. When the number of fertilized flowers is y, the total rate of resource investment for fruits and seeds is $c_S y$. Stored resources are consumed for respiration within the storage compartment itself at a constant rate d_S. y changes according to

$$\frac{dy}{dt} = \frac{c_f\theta\,(xFT > K_{FT})}{u_f} - d_y y, \tag{9.8}$$

where u_f describes the unit cost to produce a single flower. Produced flowers are degraded and transfer to mature fruit and seeds at a constant rate d_y. Thus resource investment for fruit and seeds stops when development of fruit and seeds is completed. The plant dies when all of stored resource is consumed.

Similarly, the level of resource investment to develop fruits and seeds changes according to

$$\frac{dz}{dt} = c_S y. \tag{9.9}$$

The number of fruits and seeds produced is assumed to be in proportion to the amount of resources invested in fruits and seeds. Given this assumption, lifetime reproductive success of the plant is given by

$$\varphi(m) = c_S \int e^{-mt} y(t)\mathrm{d}t, \tag{9.10}$$

where m indicates the instantaneous mortality rate. The optimal flowering strategy can be calculated by maximizing Eq. (9.10). Maximizing lifetime reproductive success is equivalent to maximization of intrinsic rate of population growth (r) when a population is at a steady state [66]. Allelic variations among genotypes and inter-specific genetic variations associated with environmental sensitivities are represented by different parameter values of the model. Which genotypes or species will be selected under given environments is analyzed by searching for a parameter set that maximizes the lifetime reproductive success.

9.3.3 *Differential Regulation of FLC Generates Diversity in Flowering Behaviors*

Based on the genetic-physiological model, Satake [53] numerically calculated the transcript and resource dynamics associated with flowering using temperature profiles monitored in central Japan from 1998 to 2008. For this calculation, germination date is fixed to July 1st. The model predicted four flowering behaviors depending on the repression and activation rate of *FLC* activity (q and p in Eq. (9.2), respectively). Detailed parameter values are given in Satake [53].

1. Monocarpic annuals (Fig. 9.2a, b): When repressive histone modification is stable (i.e., $p \cong 0$), *FLC* is stably repressed even after low temperature signals disappear in spring. This allows a prolonged production of the flowering signal, FT protein. The plant produces flowers and fruits continuously by allocating stored

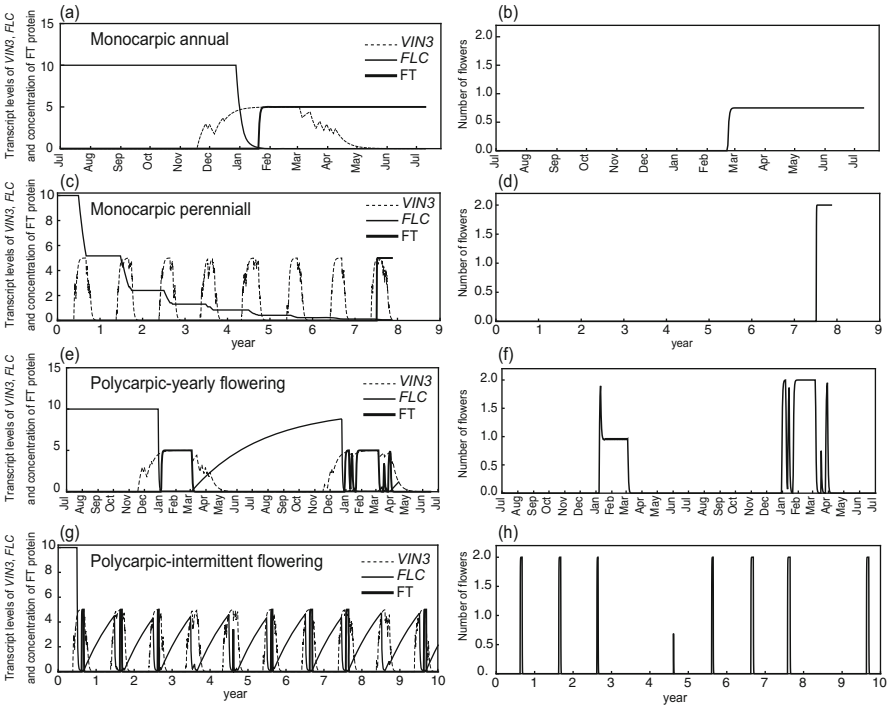


Fig. 9.2 Graphic representation of four flowering behaviors. Monocarpic annual flowering ($p = 0.0$, $q = 0.2$): (**a**) and (**b**), monocarpic perennial flowering ($p = 0.0$, $q = 0.01$): (**c**) and (**d**), polycarpic-yearly flowering ($p = 0.0072$, $q = 0.98$): (**e**) and (**f**), and polycarpic-intermittent flowering ($p = 0.00089$, $q = 0.11$): (**g**) and (**h**). (**a**), (**c**), and (**e**) Temporal change in transcript levels of *VIN3* (dashed line) and *FLC* (solid thin line), and concentration of FT protein (solid thick line). (**b**), (**d**), and (**f**) Plot of number of flowers produced. $h = 0.1$, and $K_T = 8.0$, $K_{FT} = 120$

resources. Finally, stored resources are completely depleted, and eventually, the plant dies. This is the monocarpic annual life cycle.

2. Monocarpic perennials (Fig. 9.2c, d): When repressive histone modification is stable and when the repression rate of *FLC* activity is small (i.e., $p \cong 0$ and $q \cong 0$), *FLC* is repressed very slowly. Experience of winter cold in a year is not sufficient to suppress *FLC* transcript levels below the threshold, and the plant waits for years to accumulate sufficiently high levels of repressive histone modification marks to effectively repress *FLC* to induce production of FT protein. This results in a monocarpic perennial life cycle.
3. Polycarpic-yearly flowering (Fig. 9.2e, f): When repressive histone modification is unstable (i.e., activation rate of *FLC* activity (p) is not negligibly small), *FLC* transcript levels recover once low temperature signals disappear in spring. When *FLC* transcript levels recover and exceed the threshold, *FT* expression is suppressed, and production of FT protein stops. Concentration of *FT* protein declines to below the level required for flowering, which causes a reversion from reproductive to vegetative growth. The reversion occurs before stored resources are completely depleted. This enables the plant to grow and further accumulate photosynthates and flower again in the next year. This is polycarpic-yearly flowering.
4. Polycarpic-intermittent flowering (Fig. 9.2g, h): When the activation rate of *FLC* activity (p) is close to the boundary separating flowering and no-flowering conditions (Appendix B in [53]), plants flower once in several years. If the low temperature signal is weak in a warm winter, *FLC* is not effectively repressed to initiate flowering, while *FLC* is fully repressed in a cold winter. This causes intermittent flowering.

In summary, the genetic–physiological model predicts a diverse flowering pattern ranging from monocarpic annuals to polycarpic-intermittent flowering, and the activation rate (p) and repression rate (q) of *FLC* activity is a critical factor that separate monocarpic and polycarpic life cycles.

9.3.4 Evolution of Flowering Strategies

The optimal flowering strategy can be analyzed by searching for a parameter set that maximizes the lifetime reproductive success defined in Eq. (9.10). Important parameters that influence the plant response to low temperature signals in a vernalization pathway are the temperature threshold (K_T), which triggers expression of *VIN3*, and the repression and activation rate of *FLC* activity (q and p). Thus, it was assumed that the plant adapts to given temperature environments by adjusting the magnitude of these three parameters. Because the fraction of carbon gain allocated to vegetative growth (h) and mortality (m) are also important parameters, a search for parameter values (K_T, q, and p) that maximizes lifetime reproductive success was performed under various magnitude of h and m.

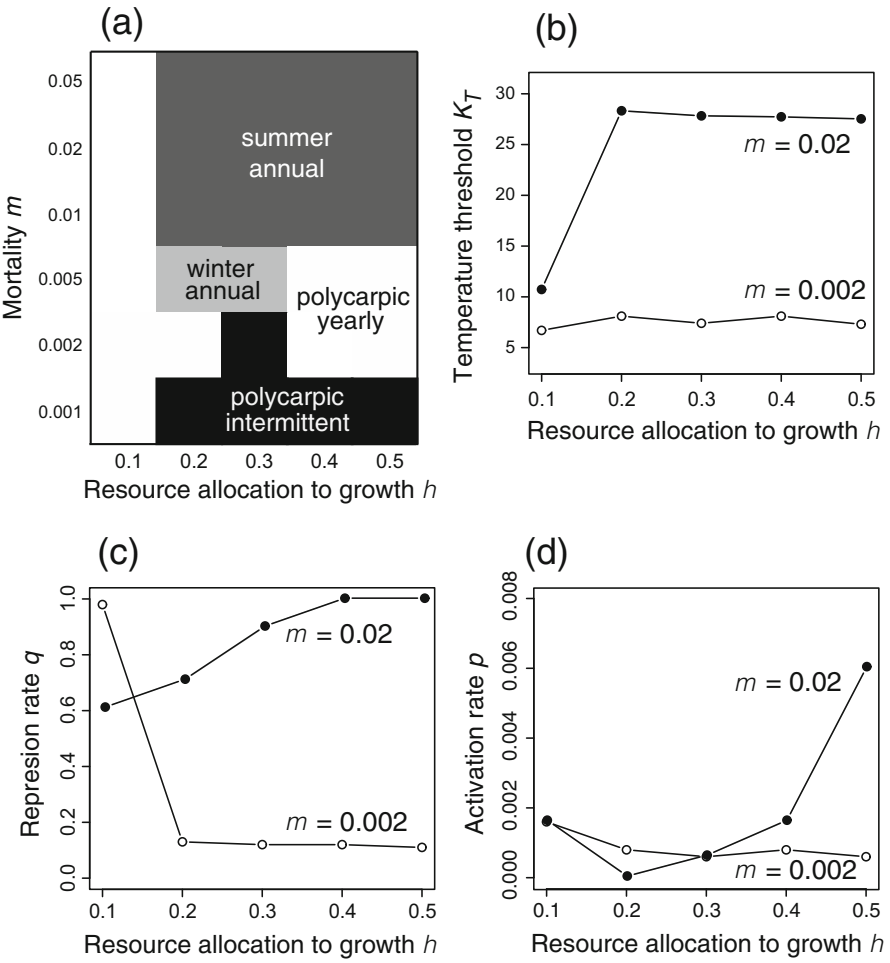


Fig. 9.3 (**a**) Classification of evolutionarily favored phenotypes under different levels of mortality and resource allocation to growth. (**b**) Evolutionary favored temperature threshold (K_T) along the level of resource allocation to growth (h). (**c**) Repression rate of *FLC* activity (q). (**d**) Activation rate of *FLC* activity (p)

When 90% of resources were allocated to storage (i.e., $h = 0.1$), polycarpic-yearly flowering always attained the highest lifetime reproductive success (Fig. 9.3a). When the fraction of resource allocation to growth was larger than or equal to 20% (i.e., $h \geq 0.2$), the mortality rate was the major determinant of the selected phenotype. Polycarpic-intermittent or polycarpic-yearly flowering was evolutionarily favored under environments with low mortality rates. When the mortality rate was intermediate (i.e., $m = 0.005$, $h = 0.2$ or 0.3), evolution favored a winter annual habit in which the plant flowered after winter and then died. As the mortality rate increased further, the winter annual was replaced by the summer

annual, whereby the plant finished flowering before winter. Flowering early is beneficial to minimize the risk of dying before seed production, but flowering with a small size is disadvantageous. In contrast, flowering later faces a greater risk to die, but the plant is large enough to produce a large number of seeds. Regardless of temperature profiles, a slow growth strategy, in which most of the resources were allocated for storage, was evolutionarily favored under environments with low mortality, while a strategy of relatively fast growth is adaptive in environments with high mortality [53].

The plant-tuned temperature threshold (K_T) is high when resource allocation to growth and mortality are large (Fig. 9.3b). When mortality (m) is 0.02, the monocarpic phenotype that employs a rapid cycling habit uses a temperature threshold close to the daily maximum, which enables the plant to initiate flowering without cold signals. In contrast, when $m = 0.002$, the temperature threshold was low (e.g., 6 °C) generating polycarpic phenotypes (Fig. 9.3a, b). Except when $h = 0.1$, the polycarpic phenotype employed a small repression rate of *FLC* activity (q) than the monocarpic phenotype (Fig. 9.3c). No clear difference was observed in the activation rate of *FLC* activity (p), except for very large h (Fig. 9.3d).

Analysis of lifetime reproductive success estimated by the genetic-physiological model for flowering permits investigation of how parameters that describe environmental sensitivity evolve under given environments to produce adaptive phenotypes. This modeling approach can be extended to understand the evolution of other signaling pathways associated with flowering. In the next section, the model that integrates the vernalization and photoperiod pathways will be presented. The model will be parameterized using gene expression data monitored in the controlled laboratory conditions and will be used to predict changes in seasonal responses to future climate change in natural conditions.

9.4 Forecasting Flowering Phenology in Natural Environments Based on Genetic Information

9.4.1 Flowering Phenology and Climate Change

The timing of recurring seasonal activities of animals and plants, or phenology, has been shifting under the influence of climate change [67–70, 81]. In contemporary environments, flowering time and seed maturation timing are becoming earlier than before, owing to warming temperature and elevated atmospheric CO_2 concentrations [71–75]. The trend for advancing flowering in contemporary environments may drive directional selection for earlier flowering, which could potentially deplete populations of the genetic variation needed for continued adaptation [71]. Yet, little is known about the molecular genetic basis for the shifts occurring in nature, which hampers the genetically informed prediction of future phenology changes. Satake et al. [65] developed a model that predicts flowering phenology based on seasonal

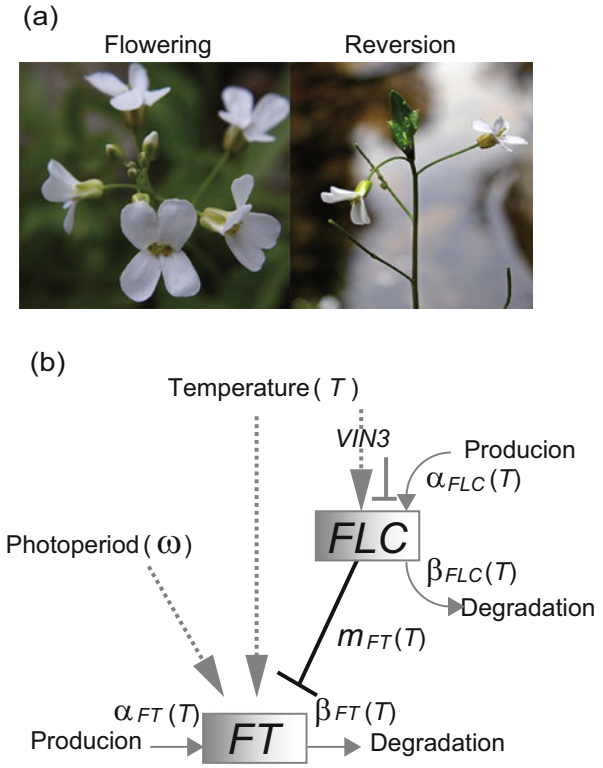


Fig. 9.4 (**a**) Flowering and reversion of *A. halleri.* (**b**) A kinetic model for *FLC* and *FT* regulatory dynamics proposed for *A. halleri*. The model integrates temperature and photoperiod signals into the network of floral regulatory genes. Regulatory dynamics $\overline{\omega}$ represents the critical photoperiod above which *FT* is activated

expression dynamics of *FLC* and *FT* in *A. halleri*, a perennial relative of the annual *A. thaliana*. *A. halleri* is characterized by the formation of aerial rosettes on apical and lateral meristems following flowering (Fig. 9.4a). This perennial life cycle adds interesting complexity to the study of flowering phenology. In monocarpic plants, flowering is irreversible, leading to fruiting and plant death. However, perennials are characterized by a reversion to vegetative growth (hereafter, "reversion") after a period of flower development. Thus, studying perennials permit investigation of when to return to vegetative state in addition to when to start flowering. This enables the vulnerability assessment of flowering period in response to rising temperature.

9.4.2 Model Construction and Parameterization

A kinetic model for *FLC* and *FT* regulatory dynamics has been proposed for *A. halleri* ([65]; Fig. 9.4b). The dynamics of *FLC* and *FT* expression levels were formalized as follows:

$$\frac{dx_{FLC}}{dt} = \alpha_{FLC}(T) f\left(x_{FLC}\right) - \beta_{FLC}(T) x_{FLC}, \tag{9.11}$$

$$\frac{dx_{FT}}{dt} = \alpha_{FT}(T) g\left(x_{FLC}\right) h\left(\omega\right) - \beta_{FT}(T) x_{FT}. \tag{9.12}$$

$\alpha_{FLC}(T)$ and $\beta_{FLC}(T)$ in Eq. (9.11) describe the production and degradation rates of *FLC* transcripts at the cell population level in leaves that are functions of temperature (T). $f(x_{FLC})$ represent the suppression of *FLC* activation at an initial stage after vernalization:

$$f\left(x_{FLC}\right) = \begin{cases} 1 & \text{before and during vernalization} \\ x_{FLC} / \left(m_{FLC} + x_{FLC}\right) & \text{after vernalization} \end{cases}, \tag{9.13}$$

where m_{FLC} is the constant that determines the irreversibility of *FLC* activation. $\alpha_{FLC}(T)$ and $\beta_{FLC}(T)$ represent the combined effects of *VIN3* and other components involved in chromatin modifications at the *FLC* locus. $\alpha_{FT}(T)$ and $\beta_{FT}(T)$ in Eq. (9.12) indicate the production and degradation rates of *FT* transcript and they are assumed to be functions of temperature. The regulation of *FT* by *FLC* was formalized using the Hill function $g(x_{FLC})$, which is defined as

$$g\left(x_{FLC}\right) = \frac{1}{1 + \left(x_{FLC} / m_{FT}\left(T\right)\right)^{H}}, \tag{9.14}$$

where $m_{FT}(T)$ is the dissociation constant of *FT* suppression by *FLC* (which is temperature dependent) and H is the Hill constant. The photoperiod effect on *FT* activation was incorporated by $h(\omega)$, defined as

$$h\left(\omega\right) = \begin{cases} 1 \text{ if } \omega > \overline{\omega} \\ 0 \text{ otherwise} \end{cases}, \tag{9.15}$$

where ω (min) is the photoperiod and $\overline{\omega}$ is the critical photoperiod above which *FT* is activated.

To parameterize the model, Satake et al. [65] obtained long-term gene expression data by using plants from two natural populations (Hyogo (HG) and Hakodate (HK) in Japan) for controlled laboratory experiments. In the experiments, two seasonal transitions in temperature—autumn to winter and winter to spring—were imitated under long-day conditions and the transcriptional levels of two

vernalization-specific genes (*AhgVIN3* and *AhgFLC*) and one floral integrator gene (*AhgFT*) were monitored in *A. halleri* in each experimental condition. In the autumn to winter experiment, plants were exposed to temperature changes from warm (22 °C) to three different levels of cold (5 °C, 10 °C, and 12.5 °C) and control (20 °C) conditions. In the winter to spring experiment, plants were exposed to temperature changes from prolonged cold (10 weeks of 5 °C) to two warmer (10 °C and 20 °C) and one control (5 °C) conditions.

The gene expression data obtained in the controlled laboratory conditions were used to parameterize the model [65]. The model successfully explained the dynamics of *AhgFLC* and *AhgFT* expression under various temperature conditions manipulated in the laboratory (Fig. 9.5a). The *AhgFLC* expression level declined gradually, with a greater rate at lower temperatures (Fig. 9.5a). *AhgFT* was not activated in both populations irrespective of temperature conditions (Fig. 9.5a). In the winter to spring experiment, *AhgFLC* expression recovered gradually with a greater rate at warmer temperature (10 °C and 20 °C; Fig. 9.1b). *AhgFT* expression was induced immediately after prolonged cold, but it was suppressed again earlier at warmer temperatures, resulting in a transient expression peak (10 °C and 20 °C; Fig. 9.5b). At 5 °C, *AhgFLC* expression levels remained low, and *AhgFT* expression levels increased continuously (Fig. 9.5b). Gene specific functions of temperature response were also estimated from the data. Production rate of *AhgFLC* increases, whereas degradation rate decreases as temperature rises (Fig. 9.6a), which allows elevated accumulation of *AhgFLC* transcript as temperature increases. By contrast both production and degradation rates of *AhgFT* increase with rising temperature (Fig. 9.6b), maintaining the levels of *AhgFT* transcript almost constant under most of temperature range. This difference could play a critical role in regulating flowering phenology under climate warming.

9.4.3 *Predicting Flowering Phenology in Natural Environments*

The model parameterized by the data in laboratory conditions facilitates the prediction of expression dynamics of *FLC* and *FT* under arbitrary temperature conditions in natural conditions. When the model was applied to the plants under natural temperature and photoperiod conditions planted in the common garden in central Japan, *AhgFLC* expression levels were predicted to decrease gradually from December in response to temperature decline and start recovering in spring as the temperature increased in both populations ([65]; Fig. 9.7a, b). These predictions accurately accounted for observations from common garden experiments ($R^2 = 0.87$ for HG and 0.90 for HK plants). The model also reproduced the observed dynamics of *AhgFT* expression, showing a transient peak in April, with exceptional accuracy (Fig. 9.7a, b; $R^2 = 0.76$ for HG and 0.92 for HK plants). Moreover the timings of bolting and reversion predicted based on the *AhgFT* expression level were in

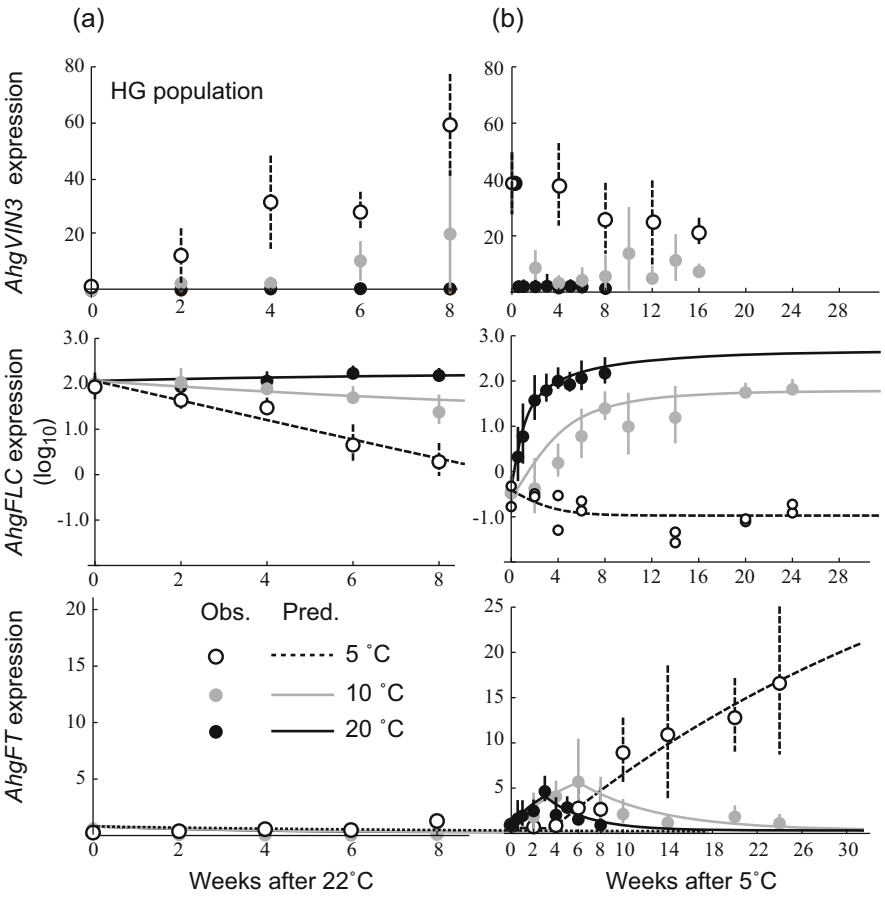


Fig. 9.5 Gene expression analysis in the laboratory-based system using *A. halleri*. Observed (mean ± s.d. of three replicates) and predicted (lines) expression levels of *AhgVIN3, AhgFLC,* and *AhgFT* during the temperature change from autumn to winter (**a**) and from winter to spring in the Hyogo population (**b**). At 5 °C, the raw data were plotted as small dots after removing unmeasurable data points

good agreement with the observed phenology (Fig. 9.7a, b). This study shows that a seemingly complex trait, flowering phenology, can be modeled successfully based on the regulatory dynamics of a surprisingly small number of genes.

The model can be used to assess the impact of climate change on the flowering phenology. Under warming conditions, the level of *AhgFLC* expression declined slower in winter, and recovery occurred earlier in spring in both populations (Fig. 9.7c, d). In response to the early recovery of *AhgFLC*, the timing of *AhgFT* suppression was significantly advanced, while the timing of *AhgFT* activation was less affected by warming (Fig. 9.7c, d). Warming of 4 °C advanced the timing of *AhgFT* suppression by 31 days, while it advanced the activation timing of *AhgFT*

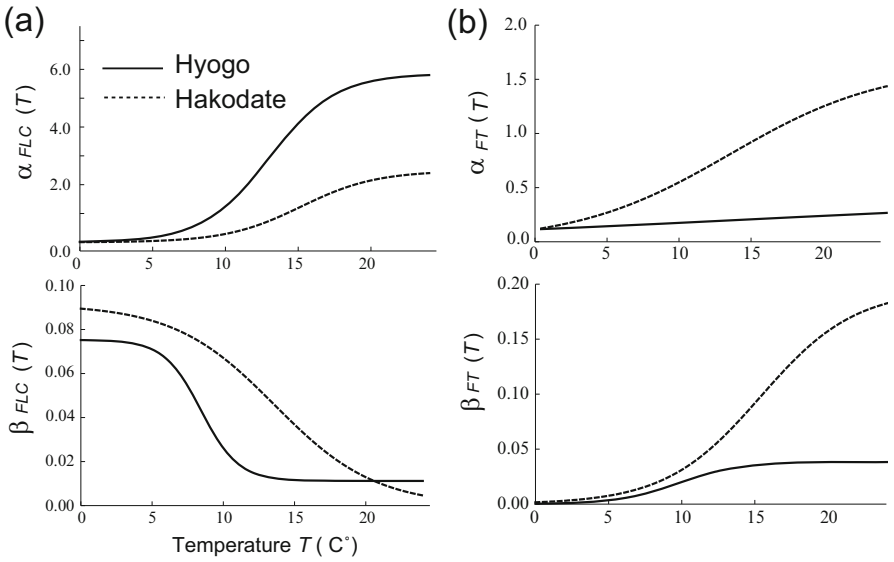


Fig. 9.6 (**a**) Temperature response functions for production and degradation rates of the *AhgFLC* transcript at the tissue level ($\alpha_{\mathrm{FLC}}(T)$ and $\beta_{\mathrm{FLC}}(T)$). (**b**) Temperature response functions for the production and degradation rates of the *AhgFT* transcript at the tissue level ($\alpha_{\mathrm{FT}}(T)$ and $\beta_{\mathrm{FT}}(T)$)

by only 7 days. Given the simulated dynamics of *AhgFT* expression, flowering phenology under wide range of temperatures can be forecasted. As the temperature increased, both bolting and reversion time advanced, but the shift in reversion time was significantly greater (Fig. 9.8a, b). Because of the asymmetric response, the duration of flowering decreased gradually with warming. Besides, the HG plants were expected to lose the opportunity to flower when the temperature increased by 4.5 °C (Fig. 9.8a). The HK plants were more robust, but they were also expected to be flowerless at 5.3 °C warming (Fig. 9.8b). A recent study confirmed the model predictions by demonstrating that increased temperature decreased the flowering duration in *A. halleri* [83]. This striking result implies the potentially high vulnerability of natural populations of *A. halleri* under climate warming.

The model to describe *FLC* and *FT* expression dynamics for perennial herb highlights that the timing to return to vegetative growth is more sensitive to changing temperature than the timing of floral initiation. This asymmetric temperature response in bolting and reversion was caused by the different temperature sensitivity between *AhgFLC* and *AhgFT* (Fig. 9.6). Despite its simplicity, the model accurately predicted complex phenological responses to climate warming in a dynamically changing natural environment, which would offer advantages for systematic conservation planning.

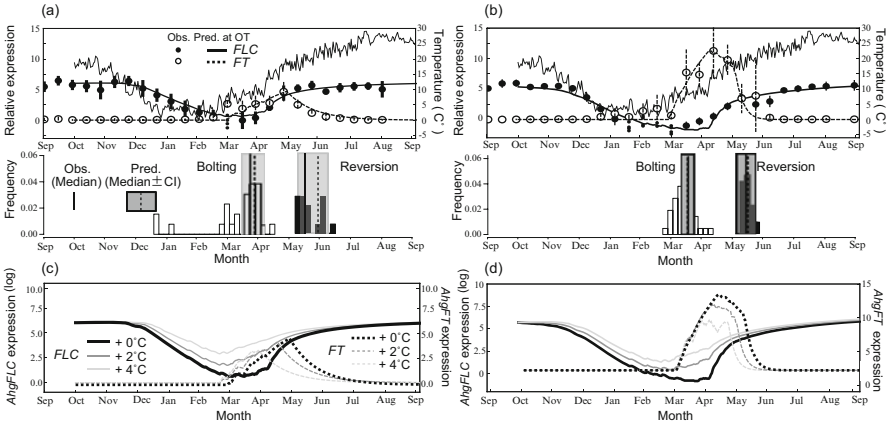


Fig. 9.7 Observed (mean ± s.d. of five replicates) and predicted (lines) *AhgFLC* (red; natural logarithmic scale) and *AhgFT* (blue) expression and the timing of bolting and reversion in the Hyogo population (**a**) and in the Hakodate population (**b**) transplanted in central Japan (Otsu; OT). Black lines represent the mean daily surface soil temperature. Black vertical lines indicate the medians of observed (solid) and predicted (dashed) timings of bolting and reversion. Gray squares represent the predicted range (95% CI of median). Given the uncertainty in the relationship between *AhgFT* expression and phenotypic output, the 95% CI of median was calculated by bootstrapping based on 1000 replications. The predicted transcript levels under warming in the Hyogo population (**c**) and in the Hakodate population (**d**)

9.5 Conclusion and Perspectives

Synthesizing the body of knowledge from molecular/genetic and ecology/evolutionary studies will be increasingly required to understand the molecular bases for the evolution of flowering behaviors and to draw genetically informed predictions for alteration of future phenology under climate change. The use of mathematical models that can connect the genetic base for flowering and its evolutionary consequence is an effective tool for this synthesis, contributing to advance our understanding of adaptive significance of genetic variations associated with flowering responses to environmental cues. The physiological mechanism of floral initiation for diverse plant species can be more accurately modeled by combining molecular and ecological data. Given increased molecular data in non-model plant species, comparison of genetic pathways for flowering between diverse plant species will become possible which provides the opportunity to identify the evolutionary footprint in the regulatory pathways controlling flowering responses to environmental cues.

A molecular phenology approach that monitors seasonal recurring gene expression patterns *in natura*, reviewed by Kudoh [76], has been used for a wide range of species including herbs [51, 65, 82] and trees [63, 77, 84]. This approach was successful in identifying differential environmental cues required for flowering in different species. For example, Miyazaki et al. [63] monitored expression levels of

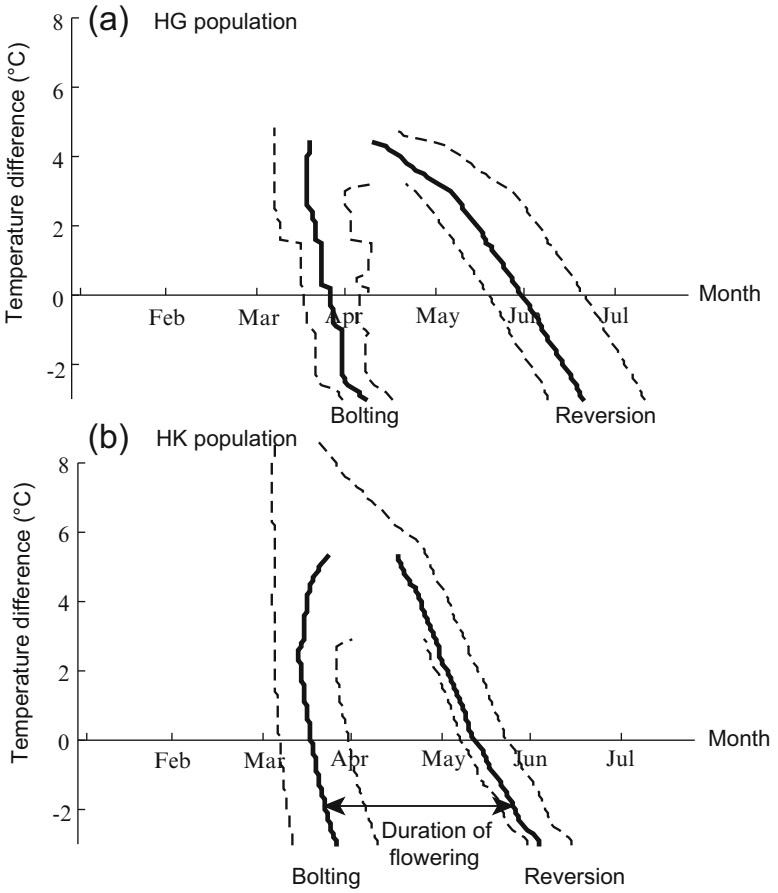


Fig. 9.8 Forecasted timing of bolting and reversion in the Hyogo (**a**) and Hakodate populations (**b**). 0 °C corresponds to the daily surface soil temperature monitored at Otsu from Oct 2011 to Sep 2012

key flowering genes including *FT* for 5 years in the field using *Fagus crenata*, a deciduous tree that shows intermittent and synchronized flowering at the population level. They demonstrated that these flowering genes were activated every other year, leading to an alternate flowering pattern. They showed that the between-year fluctuation in gene expression levels was coincided with the change in nitrogen content in current-year shoots, and identified that nitrogen is a key regulator for the floral transition in this species by nitrogen fertilization experiments [63].

The molecular phenology approach has shown to be useful to unravel the mechanism of general flowering, a spectacular phenomenon in tropical rain forests in Southeast Asia. In general flowering, hundreds of plant species reproduce synchronously and their synchronous reproduction occurs at irregular, multi-year intervals [78, 79]. General flowering is limited to the least seasonal lowland tropical

rainforests of Southeast Asia. Trees of the family Dipterocarpaceae dominate these forests, and most dipterocarps join species from at least 41 other families in general flowering events [80]. Yeho et al. [84] applied a molecular phenology approach to tropical trees in the genus *Shorea* section Mutica to identify proximate environmental cues for flowering. The activation of flowering genes including *FT* was observed twice over 4 years, which was followed by anthesis in these species. The observed flowering gene expression pattern in *Shorea* was precisely explained by the model that assumed the interplay between drought and cool temperature signals to activate the flowering gene, suggesting the synergistic effect of drought and cool temperature is the major driving factor of flowering in these tropical trees.

The molecular studies for flowering time have been advanced using plants inhabiting seasonal environments. Seasonal variation in day length and temperature decline towards the equator. Even in aseasonal environments, most tropical plants can sense the subtle climatic fluctuation as a cue for flowering and still produce flowers and fruits at specific times of year. As explained in this chapter, an integrated approach of ecology, molecular biology, and mathematical modeling is useful to dissect proximate and ultimate factors of flowering traits of plants in different climate zones. We will be able to draw the history of evolutionary alterations of the underlying molecular bases of plant's flowering behavior in geographically diverse environments in near future, and global ecosystem response to climate change.

References

1. Tinbergen N (1963) On aims and methods of ethology. Ethology 20(4):410–433
2. Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408(6814):796
3. Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296(5565):92–100
4. Yu J, Hu S, Wang J, Wong GKS, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science 296(5565):79–92
5. Smith JM (1978) Optimization theory in evolution. Annu Rev Ecol Syst 9(1):31–56
6. Satake A, Sakurai G, Kinoshita T (2015) Modeling strategies for plant survival, growth and reproduction. Plant Cell Physiol 56:583–585
7. Bateson P, Laland KN (2013) Tinbergen's four questions: an appreciation and an update. Trends Ecol Evol 28(12):712–718
8. Cohen D (1971) Maximizing final yield when growth is limited by time or by limiting resources. J Theor Biol 33(2):299–307
9. Cohen D (1976) The optimal timing of reproduction. Am Nat 110(975):801–807
10. Iwasa Y (2000) Dynamic optimization of plant growth. Evol Ecol Res 2(4):437–455
11. Iwasa Y, Cohen D (1989) Optimal growth schedule of a perennial plant. Am Nat 133(4):480–505
12. King D, Roughgarden J (1982) Multiple switches between vegetative and reproductive growth in annual plants. Theor Popul Biol 21(2):194–204
13. Yamamura N, Tsuji N (1995) Optimal strategy of plant antiherbivore defense: implications for apparency and resource-availability theories. Ecol Res 10(1):19–30

14. Yamauchi A, Yamamura N (2004) Herbivory promotes plant production and reproduction in nutrient-poor conditions: effects of plant adaptive phenology. Am Nat 163(1):138–153
15. Metcalf CJE, Rose KE, Childs DZ, Sheppard AW, Grubb PJ, Rees M (2008) Evolution of flowering decisions in a stochastic, density-dependent environment. Proc Natl Acad Sci 105(30):10466–10470
16. Rees M, Rose KE (2002) Evolution of flowering strategies in Oenothera glazioviana: an integral projection model approach. Proc R Soc Lond B Biol Sci 269(1499):1509–1515
17. Bowman JL, Smyth DR, Meyerowitz EM (1991) Genetic interactions among floral homeotic genes of Arabidopsis. Development 112(1):1–20
18. Carpenter R, Coen ES (1990) Floral homeotic mutations produced by transposon-mutagenesis in Antirrhinum majus. Genes Dev 4(9):1483–1493
19. Schwarz-Sommer Z, Huijser P, Nacken W, Saedler H, Sommer H (1990) Genetic control of flower development by homeotic genes in Antirrhinum majus. Science 250(4983):931
20. Koornneef M, Alonso-Blanco C, Peeters AJ, Soppe W (1998) Genetic control of flowering time in Arabidopsis. Annu Rev Plant Biol 49(1):345–370
21. Koornneef M, Hanhart CJ, Veen J (1991) A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. Mol Gen Genet MGG 229(1):57–66
22. Amasino R (2010) Seasonal and developmental timing of flowering. Plant J 61(6):1001–1013
23. Andrés F, Coupland G (2012) The genetic basis of flowering responses to seasonal cues. Nat Rev Genet 13(9):627–639
24. Bäurle I, Dean C (2006) The timing of developmental transitions in plants. Cell 125(4):655–664
25. Simpson GG, Dean C (2002) Arabidopsis, the Rosetta stone of flowering time? Science 296(5566):285–289
26. Blackman BK (2017) Changing responses to changing seasons: natural variation in the plasticity of flowering time. Plant Physiol 173(1):16–26
27. Fournier-Level A, Perry EO, Wang JA, Braun PT, Migneault A, Cooper MD, Metcalf CJE, Schmitt J (2016) Predicting the evolutionary dynamics of seasonal adaptation to novel climates in Arabidopsis thaliana. Proc Natl Acad Sci 113(20):E2812–E2821 p.201517456
28. Dennis ES, Peacock WJ (2007) Epigenetic regulation of flowering. Curr Opin Plant Biol 10(5):520–527
29. Sung S, Amasino RM (2005) Remembering winter: toward a molecular understanding of vernalization. Annu Rev Plant Biol 56:491–508
30. Burn JE, Bagnall DJ, Metzger JD, Dennis ES, Peacock WJ (1993) DNA methylation, vernalization, and the initiation of flowering. Proc Natl Acad Sci 90(1):287–291
31. Clarke JH, Dean C (1994) Mapping *FRI*, a locus controlling flowering time and vernalization response in Arabidopsis thaliana. Mol Gen Genet MGG 242(1):81–89
32. Koornneef M, Vries H, Hanhart C, Soppe W, Peeters T (1994) The phenotype of some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg erecta wild-type. Plant J 6(6):911–919
33. Lee I, Bleecker A, Amasino R (1993) Analysis of naturally occurring late flowering in Arabidopsis thaliana. Mol Gen Genet MGG 237(1):171–176
34. Napp-Zinn K (1987) Vernalization-environmental and genetic regulation. In: Atherton JG (ed) Manipulation of flowering. Butterworths, London
35. Koornneef M, Alonso-Blanco C, Vreugdenhil D (2004) Naturally occurring genetic variation in Arabidopsis thaliana. Annu Rev Plant Biol 55:141–172
36. Sung S, Amasino RM (2004) Vernalization and epigenetics: how plants remember winter. Curr Opin Plant Biol 7(1):4–10
37. Lempe J, Balasubramanian S, Sureshkumar S, Singh A, Schmid M, Weigel D (2005) Diversity of flowering responses in wild Arabidopsis thaliana strains. PLoS Genet 1(1):e6
38. Shindo C, Aranzana MJ, Lister C, Baxter C, Nicholls C, Nordborg M, Dean C (2005) Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of Arabidopsis. Plant Physiol 138(2):1163–1173

39. Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS, Purugganan MD, Schmitt J (2004) A latitudinal cline in flowering time in Arabidopsis thaliana modulated by the flowering time gene *FRIGIDA*. Proc Natl Acad Sci U S A 101(13):4712–4717
40. Michaels SD, Amasino RM (1999) *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11(5):949–956
41. Michaels SD, Amasino RM (2001) Loss of *FLOWERING LOCUS C* activity eliminates the late-flowering phenotype of *FRIGIDA* and autonomous pathway mutations but not responsiveness to vernalization. Plant Cell 13(4):935–941
42. Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, Dennis ES (1999) The *FLF* MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. Plant Cell 11(3):445–458
43. Kardailsky I, Shukla VK, Ahn JH, Dagenais N, Christensen SK, Nguyen JT, Chory J, Harrison MJ, Weigel D (1999) Activation tagging of the floral inducer FT. Science 286(5446):1962–1965
44. Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T (1999) A pair of related genes with antagonistic roles in mediating flowering signals. Science 286(5446):1960–1962
45. Mandel MA, Yanofsky MF (1995) A gene triggering flower formation in Arabidopsis. Nature 377(6549):522
46. Weigel D, Alvarez J, Smyth DR, Yanofsky MF, Meyerowitz EM (1992) *LEAFY* controls floral meristem identity in Arabidopsis. Cell 69(5):843–859
47. Ferrándiz C, Gu Q, Martienssen R, Yanofsky MF (2000) Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. Development 127(4):725–734
48. Scarcelli N, Cheverud JM, Schaal BA, Kover PX (2007) Antagonistic pleiotropic effects reduce the potential adaptive value of the *FRIGIDA* locus. Proc Natl Acad Sci 104(43):16986–16991
49. Michaels SD, He Y, Scortecci KC, Amasino RM (2003) Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in Arabidopsis. Proc Natl Acad Sci 100(17):10102–10107
50. Werner JD, Borevitz JO, Uhlenhaut NH, Ecker JR, Chory J, Weigel D (2005) *FRIGIDA*-independent variation in flowering time of natural Arabidopsis thaliana accessions. Genetics 170(3):1197–1207
51. Aikawa S, Kobayashi MJ, Satake A, Shimizu KK, Kudoh H (2010) Robust control of the seasonal expression of the Arabidopsis *FLC* gene in a fluctuating environment. Proc Natl Acad Sci 107(25):11632–11637
52. Wang R, Farrona S, Vincent C, Joecker A, Schoof H, Turck F, Alonso-Blanco C, Coupland G, Albani MC (2009) *PEP1* regulates perennial flowering in Arabis alpina. Nature 459(7245):423–427
53. Satake A (2010) Diversity of plant life cycles is generated by dynamic epigenetic regulation in response to vernalization. J Theor Biol 266(4):595–605
54. Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C (2004) Vernalization requires epigenetic silencing of FLC by histone methylation. Nature 427(6970):164–167
55. Finnegan JE, Kovac KA, Jaligot E, Sheldon CC, James Peacock W, Dennis ES (2005) The downregulation of *FLOWERING LOCUS C* (*FLC*) expression in plants with low levels of DNA methylation and by vernalization occurs by distinct mechanisms. Plant J 44(3):420–432
56. Schmitz RJ, Sung S, Amasino RM (2008) Histone arginine methylation is required for vernalization-induced epigenetic silencing of *FLC* in winter-annual Arabidopsis thaliana. Proc Natl Acad Sci 105(2):411–416
57. Sung S, He Y, Eshoo TW, Tamada Y, Johnson L, Nakahigashi K, Goto K, Jacobsen SE, Amasino RM (2006) Epigenetic maintenance of the vernalized state in Arabidopsis thaliana requires LIKE HETEROCHROMATIN PROTEIN 1. Nat Genet 38(6):706–710
58. Angel A, Song J, Dean C, Howard M (2011) A Polycomb-based switch underlying quantitative epigenetic memory. Nature 476(7358):105–108
59. Satake A, Iwasa Y (2012) A stochastic model of chromatin modification: cell population coding of winter memory in plants. J Theor Biol 302:6–17

60. Angel A, Song J, Yang H, Questa JI, Dean C, Howard M (2015) Vernalizing cold is registered digitally at FLC. Proc Natl Acad Sci 112(13):4146–4151
61. Dodd IB, Micheelsen MA, Sneppen K, Thon G (2007) Theoretical analysis of epigenetic cell memory by nucleosome modification. Cell 129(4):813–822
62. Krzymuski M, Andrés F, Cagnola JI, Jang S, Yanovsky MJ, Coupland G, Casal JJ (2015) The dynamics of *FLOWERING LOCUS T* expression encodes long-day information. Plant J 83(6):952–961
63. Miyazaki Y, Maruyama Y, Chiba Y, Kobayashi MJ, Joseph B, Shimizu KK, Mochida K, Hiura T, Kon H, Satake A (2014) Nitrogen as a key regulator of flowering in Fagus crenata: understanding the physiological mechanism of masting by gene expression analysis. Ecol Lett 17(10):1299–1309
64. Miyazaki Y, Satake A (2017) Relationship between seasonal progression of floral meristem development and *FLOWERING LOCUS T* expression in the deciduous tree *Fagus crenata*. Ecol Res 32(4):627–631
65. Satake A, Kawagoe T, Saburi Y, Chiba Y, Sakurai G, Kudoh H (2013) Forecasting flowering phenology under climate warming by modelling the regulatory dynamics of flowering-time genes. Nat Commun 4:2303
66. Charlesworth B (1980) Evolution in age-structured populations. Cambridge University Press, Cambridge
67. Chen IC, Hill JK, Ohlemüller R, Roy DB, Thomas CD (2011) Rapid range shifts of species associated with high levels of climate warming. Science 333(6045):1024–1026
68. Root TL, MacMynowski DP, Mastrandrea MD, Schneider SH (2005) Human-modified temperatures induce species changes: joint attribution. Proc Natl Acad Sci U S A 102(21):7465–7469
69. Rosenzweig C, Karoly D, Vicarelli M, Neofotis P, Wu Q, Casassa G, Menzel A, Root TL, Estrella N, Seguin B, Tryjanowski P (2008) Attributing physical and biological impacts to anthropogenic climate change. Nature 453(7193):353–357
70. Thackeray SJ, Sparks TH, Frederiksen M, Burthe S, Bacon PJ, Bell JR, Botham MS, Brereton TM, Bright PW, Carvalho L, Clutton-Brock TIM (2010) Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. Glob Chang Biol 16(12):3304–3313
71. Anderson JT, Inouye DW, McKinney AM, Colautti RI, Mitchell-Olds T (2012) Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. Proc R Soc Lond B Biol Sci 279(1743):3843–3852
72. Franks SJ, Sim S, Weis AE (2007) Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. Proc Natl Acad Sci 104(4):1278–1282
73. Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. Nature 421(6918):37–42
74. Post ES, Pedersen C, Wilmers CC, Forchhammer MC (2008) Phenological sequences reveal aggregate life history response to climatic warming. Ecology 89(2):363–370
75. Springer CJ, Ward JK (2007) Flowering time and elevated atmospheric CO2. New Phytol 176(2):243–255
76. Kudoh H (2015) Molecular phenology in plants: in natura systems biology for the comprehensive understanding of seasonal responses under natural environments. New Phytol 210(2):399–412
77. Kobayashi MJ, Takeuchi Y, Kenta T, Kume T, Diway B, Shimizu KK (2013) Mass flowering of the tropical tree Shorea beccariana was preceded by expression changes in flowering and drought-responsive genes. Mol Ecol 22(18):4767–4782
78. Chan HT (1981) Reproductive biology of some Malaysian dipterocarps, 3: breeding systems. Malaysian Forester, Kepong
79. Janzen DH (1974) Tropical Blackwater rivers, animals, and mast fruiting by the Dipterocarpaceae. Biotropica 6(2):69–103
80. Appanah S (1985) General flowering in the climax rain forests of south-East Asia. J Trop Ecol 1(03):225–240

81. Fitter AH, Fitter RSR (2002) Rapid change in flowering time in British plants. Science 296:1689–1691
82. Nagano AJ, Sato Y, Mihara M, Antonio BA, Motoyama R, Itoh H, Nagamura Y, Izawa T (2012) Deciphering and prediction of transcriptome dynamics under fluctuating field conditions. Cell 151(6):1358–1369
83. Nagahama A, Kubota Y, Satake (2018) A Climate warming shortens flowering duration: a comprehensive assessment of plant phenological responses based on gene expression analyses and mathematical modeling. Ecol Res 5:1059–1068
84. Yeoh SH, Satake A, Numata S, Ichie T, Lee SL, Basherudin N, Muhammad N, Kondo T, Otani T, Hashim M, Tani N (2017) Unraveling proximate cues of mass flowering in the tropical forests of Southeast Asia from gene expression analyses. Mol Ecol 26:5074–5085

Chapter 10
All But Sleeping? Consequences of Soil Seed Banks on Neutral and Selective Diversity in Plant Species

Daniel Živković and Aurélien Tellier

Abstract An axiom of modern evolutionary theory is that intra-species genetic diversity determines the adaptive potential of any species. This diversity results from the interaction between three factors: the effective population size, natural selection and the rate of recombination. All three factors are influenced by the occurrence of long dormant stages (seed banking or seed persistence), which is an evolutionary bet hedging strategy and a key characteristic of many angiosperms, but also bacteria, fungi or invertebrates. Perhaps surprisingly, this ecological trait has so far been almost ignored in evolutionary genomics. Seed banking is expected to have a fundamental influence on neutral and selective evolutionary processes, and is therefore a key factor to comprehend angiosperm genomic evolution. Theoretical modeling aims to predict the effect of seed banking on patterns of nucleotide diversity. We first adapt for seed banks the two classical mathematical frameworks of population genetics: (1) the backward in time process of the Kingman *n*-coalescent, and (2) the forward in time diffusion approach. This allows us to derive population genetics quantities and statistics that can be obtained from DNA sequence data. Second, we generate new predictions on neutral diversity and past demographic inference for single and multiple populations under seed banks. Third, we compute the expected effect of seed banks on unlinked (genome wide selection) and linked (gene level selection) sites. Finally, we conclude by suggesting three hypotheses, which can be tested by contrasting polymorphism data in seed banking and non-seed banking species.

10.1 Introduction

The amount of genetic diversity and polymorphism at the molecular level (in DNA) between individuals is a cornerstone of modern evolutionary theory because it determines the phenotypic diversity on which selection acts, and it is thus key

D. Živković · A. Tellier (✉)
Section of Population Genetics, Technical University of Munich, Freising, Germany
e-mail: zivkovic@bio.lmu.de; tellier@wzw.tum.de

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5_10

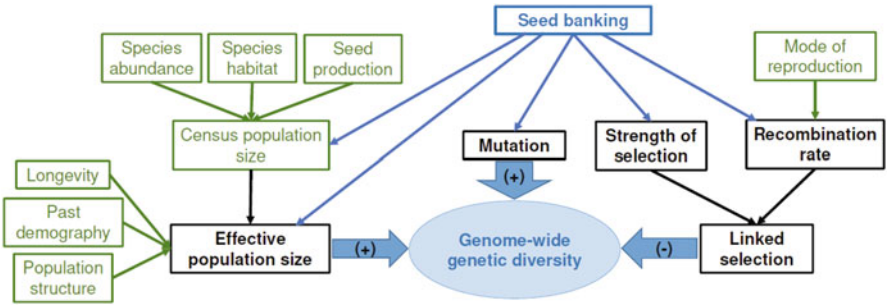


Fig. 10.1 Overview of the determinants of genetic diversity in plant species (adapted from [13]). In black the key evolutionary factors, in green the key ecological factors, and in blue the effects of seed banking. The sign on the arrows indicates the direction of correlation to genetic diversity

for predicting the adaptive potential of any species [7, 13, 40]. Our ability to predict evolutionary changes has implications for conservation biology, disease management in medicine and agriculture, plant breeding and future ecosystems management in response to global change [1, 35, 40]. Theoretical and empirical evolutionary genetics describe three main factors controlling the amount of neutral and selective genetic diversity: the species effective population size, mutation and linked selection (Fig. 10.1).

The effective population size is determined directly by species characteristics such as spatial structure, past demographic events (population increase or decrease, [35, 40]) and life span (longevity), as well as the census size, i.e. the number of observed individuals (Fig. 10.1). Several life history traits affect directly the census size, such as the size of propagules in animals [45], or in plants crucial ecological factors such as the geographic range, ecological habitat, density and abundance [47] and seed production [17].

The mutation rate is a characteristic of the species which may vary with latitude and longitude, e.g., due to exposure to UV light. It is also important to quantify if mutations occurs in the resting or dormant stages of some species: seeds for plants, eggs for insects and crustaceans (e.g. Daphnia), spores in bacteria and fungi [12, 26, 37, 38, 42].

Linked selection is a function of the strength of selection and the recombination rate (Fig. 10.1, [59, 60]). The strength of selection depends on the trait affected by selection and the advantage (or disadvantage) in a given environment compared to other genotypes. By linkage disequilibrium (LD, e.g. [8]) between genomic sites, natural selection affects (1) neutral genomic diversity around the targets of selection, and (2) the efficiency of selection at other neighbouring sites (or loci), the latter is termed the Hill-Robertson effect [25]. The extent of LD depends on the recombination rate with higher rates decreasing the effect of linked selection on diversity (Fig. 10.1, [11, 13]). In flowering plants recombination rate has been primarily studied as determined by the mode of reproduction (selfing, outcrossing) and the reproductive characteristics (Hermaphrodite, Dioecy or Monoecy) [9, 19, 20, 46].

Higher mutation rates and larger effective population size increase genetic diversity, while increased linked selection has the opposite effect (Fig. 10.1). The joint action of these three factors explains the occurrence of the so-called Lewontin paradox [39], stating that the census size of a species, N, defined as the number of individuals observed, does counterintuitively not correlate with genetic diversity. Census sizes can vary over several orders of magnitude between species, e.g. between elephants, humans and crustaceans, while genetic diversity varies only by two to three orders of magnitude in animals [11, 13, 36, 39, 45]. Recent studies show that such a discrepancy between expected and observed genetic diversity arises because selection counteracts the effect of a large population size on diversity: in large populations, positive and purifying selection are more efficient thus increasing the effect of linked selection and the Hill-Robertson effect [11, 13]. Resolving this paradox is instrumental to understand the adaptability of species in the current age of global climatic and anthropogenic changes [1, 36].

In plants, in particular in angiosperms (flowering plants), despite the recent progress in sequencing and the large amount of available genome data, we still do not know the relative importance of the several ecological and life history traits highlighted in Fig. 10.1 in shaping genetic diversity [11, 13, 36]. A key characteristic of many plant species is generally ignored: the ability of forming persistent seed banks with seeds remaining in the soil for many years. It is classically observed as plant adaptation to unpredictable semi-arid to desertic habitats, but is found also in many species in temperate climate in Eurasia and North-America [17, 52]. As an ecological adaptation, seed persistence has major consequences on the genomic evolution by increasing genetic diversity and also by buffering against fast ecological changes in population sizes and preventing population extinction. Seed persistence is therefore a key life-history trait linking the ecology of a species with its abundance, the production and survival of seeds over years, and consequently defining the census size above ground and the available genetic diversity (effective population size). To give a first general idea of its potential importance (see details below), let us define the germination rate, b, as the probability ($0 < b \leq 1$) that a seed will germinate on average in the next year. Theory predicts that persistent seed banks increase genetic diversity by a factor $1/b^2$ and the recombination rate by a factor $1/b$. If the germination rate is realistically small ($b < 0.5$), the effect of seed banking becomes prominent on genetic diversity and recombination rates [10, 14, 34, 62].

In this chapter we describe the effect of seed banking on shaping genetic diversity in plant species. We first build here backward and forward population genetics models which allow us to derive statistics as the expected time of coalescence, the time of fixation and the site frequency spectrum (SFS) of a sample (Sect. 10.2). We then derive a set of predictions that should be observable in polymorphism data from diverse plant species for neutral diversity (Sect. 10.3) and for natural selection (including footprints of selection in the genome, Sect. 10.4).

10.2 Model Description

Seed banking is an evolutionary strategy termed as bet hedging because it consists in reducing short-term reproductive success in favour of longer-term risk reduction to maximize individual fitness over time in variable and unpredictable environments [10, 14, 15, 54, 55]. It is also commonly found in bacteria, fungi, protozoans (including human parasites) [27, 37] and invertebrates (e.g. Daphnia, [42]), thus our predictions can be extended to these species. This ecological trait is classically observed as plant adaptation to unpredictable semi-arid to desertic habitats [55], but is found also in many species in temperate climate in Eurasia and North-America [2, 17]. A non-persistent seed bank is defined as seeds remaining in the soil for up to 5 years (hereafter defining non-seed banking, non-SB species), while persistent seed bank is longer (defining seed banking, SB species, [2, 17]). Note that seed bank persistence is independent of dormancy [17]. While dormancy is defined as the best timing for germination within a year, seed banking occurs over many years and seeds may alternate between periods of dormancy and non-dormant stages when in the soil.

In population genetics, it is common practice to either model an entire population forwards in time within a diffusion framework [16] or to trace a sample of a population backwards in time by applying coalescent theory [30]. In both cases a population of large and finite size N is initially assumed that evolves on a discrete generation-wise time scale and whose reproduction mechanism will basically follow a Wright–Fisher model throughout this chapter. In the backwards formulation of the usual non-SB version of this model, the entire population is replaced after reproduction and descendants pick their ancestors by random sampling from the foregoing generation, while in its SB version, ancestors can be chosen from the previous and up to m generations according to the probabilities $b_1, \ldots, b_m$. In the dual prospective perception of time, these probabilities determine when ancestors give rise to their descendants within the subsequent m generations. We focus throughout on the average germination rate b, which is the inverse of the mean time a seed will spend in the bank, i.e. $b = 1/\sum_{i=1}^{m} i\, b_i$. In practice, the seed bank age distribution is often assumed to follow a geometric distribution, so that seeds are more likely to germinate at earlier than later ages.

Since it is mathematically favourable to treat the ancestral (or the equivalent reproductive) process on a continuous time-scale, time is scaled in units of N generations (in a haploid model as assumed throughout) while $N \to \infty$. In the retrospective setting, for instance, the waiting times until two sampled individuals coalesce into a common ancestor can be geometrically distributed in the discrete, and exponentially distributed in the continuous model. Mutations are the main source of variation among the individuals of a population. As rare events these are assumed throughout to follow a Poisson distribution with a scaled mutation rate, $\theta = \lim_{N\to\infty} 2N\mu/b^2$, and that each of them arises on a previously unmutated (or monomorphic) site according to the infinitely-many sites model of Kimura [33]. In other words, the effective population size, N_e, consists of all plants and seeds,

while the census size is only consisting of living plants, so that we obtain the relation $N_e = N/b^2$ [28, 43]. In other words, mutations are assumed to arise in seeds and living plants [12, 58]. Alternatively, mutations may not arise in seeds, so that the scaled mutation rate reduces by a factor b [52, 62]. Except in the case of linked selection, all polymorphic sites will be assumed to be unlinked (due to recombination in between them) and therefore independent of each other.

10.2.1 The Retrospective Coalescent View for Neutrally Evolving Sites

Kaj et al. [28] introduced an urn model to describe neutral seed bank dynamics for a population of a constant size, where the population of a new generation is formed via multinomial sampling from the m previous generations in this initially discrete setting. A sample represented as balls and consisting of n balls at present is repeatedly relocated across the previous generations by sliding a window comprising the m consecutive generations as cells in a stepwise manner. Sliding the window one generation backwards, all balls from the first cell of the previous window are relocated into one of the m cells of the actual window according to the probabilities $b_1, \ldots, b_m$. More precisely, each ball is relocated into one of the N slots of a given cell, each representing an individual of the population in the respective generation. Thereby, two types of coalescent events may occur in the sample's history: either two balls are placed into the same slot of the same cell or a ball is placed into a previously occupied slot. The probability of one coalescent event at a time is $O(1/N)$, while more than one coalescent event at a time occur with the negligible probability of $O(1/N^2)$. Therefore, coalescences occur on a 'slow' timescale of $O(N)$ steps, while the relocation process runs on a 'fast' time scale making the separation of time scales possible.

The ancestral process of the discrete seed bank model is denoted by $(A_n^N(k))_{k\geq 0}$, where $A_n^N(k)$ is the number of ancestors at step k with population size N and initial sample size n. It has been shown [28] that the continuous seed bank model is the n-coalescent [30] run on a slower time-scale by proofing that the time-rescaled ancestral process $(A_n^N([Nt]))_{t\geq 0}$ converges as $N \to \infty$ to the continuous-time Markov chain $(A_n(t))_{t\geq 0}$ with infinitesimal generator matrix $\mathbf{Q} = (q_{ij})_{i,j\in\{1,\ldots,n\}}$ defined by

$$\begin{aligned} q_{ii} &= -b^2\tbinom{i}{2}, && 2 \leq i \leq n, \\ q_{ii-1} &= b^2\tbinom{i}{2}, && 2 \leq i \leq n, \\ q_{ij} &= 0, && \text{otherwise.} \end{aligned} \tag{10.1}$$

From Eq. (10.1) one can derive the probability that the process $A_n(t)$ is in the state of $j = n, \ldots, 2$ ancestors at time t via a matrix decomposition (e.g. [61]) to obtain

$$P(A_n(t) = j) = \sum_{k=j}^{n} c_{nk} r_{kj} \exp\left(-b^2 \binom{k}{2} t\right), \tag{10.2}$$

where $c_{nk} = \binom{n}{k} k_{(k)}/n_{(k)}$ and $r_{kj} = (-1)^{k-j}\binom{k}{j} j_{(k-1)}/k_{(k-1)}$ are the elements of the matrices of column and row eigenvectors of $\mathbf{Q}$, respectively, and $a_{(b)} = a(a+1)\cdots(a+b-1)$, $a_{(0)} = 1$. The mean waiting times between coalescent events are given by

$$E(T_j) = \int_0^\infty P(A_n(t) = j)dt. \tag{10.3}$$

as the inverse of the coalescent rate.

10.2.2 The Prospective Diffusion Framework for Neutral and Selected Sites

Following allele frequencies in the limit $N \to \infty$ forward in time leads to diffusion approximations with time and allele frequencies being again measured on a continuous scale (i.e. in units of N generations). We assume two allelic types A and a with frequencies x and $1-x$. The advantage of the diffusion framework over the coalescent setting is that selection can be more straightforwardly taken into account. The effect of weak selection on the fertility of plants and on the fraction of surviving seeds can be summarized into a coefficient s, so that the scaled coefficient in the haploid model is given by $\sigma = N s$. By making use of a perturbation approach, it has been shown [34] that in the diffusion limit, as $N \to \infty$, the probability $f(y,t)dy$ that the type-A genotype has a frequency in $(y, y+dy)$ is determined by the following forward equation (see [31] for the non-SB model):

$$\frac{\partial}{\partial t} f(y,t) = -\frac{\partial}{\partial y}\{\mu(y)\, f(y,t)\} + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left\{\sigma^2(y)\, f(y,t)\right\}, \tag{10.4}$$

where the drift and the diffusion terms are, respectively, given by $\mu(y) = \sigma\ b\ y(1-y)$ and $\sigma^2(y) = b^2 y(1-y)$. For neutrality, $\mu(y) = 0$ so that the exclusive consequence of genetic drift is characterized by the diffusion term. For the derivation of the time to fixation and the SFS we require the following measures. The scale density of the diffusion process is given by

$$\xi(y) = \exp\left(-\int_0^y \frac{2\mu(z)}{\sigma^2(z)} dz\right).$$

The speed density is obtained (up to a constant) as $\pi(y) = [\sigma^2(y)\xi(y)]^{-1}$ and the probability of absorption at $y = 0$ is given by

$$u_0(x) = \frac{\int_x^1 \xi(z)dz}{\int_0^1 \xi(z)dz}$$

and $u_1(x) = 1 - u_0(x)$ gives the probability of absorption at $y = 1$.

Assuming that both $y = 0$ and $y = 1$ are absorbing states the mean time $\bar{t}$ until one of these states is reached is given by [16]

$$\bar{t}(x) = \int_0^1 t(x, y)dy, \tag{10.5}$$

where

$$t(x, y) = 2\,u_0(x)[\sigma^2(y)\xi(y)]^{-1} \int_0^y \xi(z)dz, \quad 0 \le y \le x,$$

$$t(x, y) = 2\,u_1(x)[\sigma^2(y)\xi(y)]^{-1} \int_y^1 \xi(z)dz, \quad x \le y \le 1.$$

The time until a mutant allele is fixed conditional on fixation can be evaluated as

$$\bar{t^*}(x) = \int_0^1 t^*(x, y)dy, \tag{10.6}$$

where $t^*(x, y) = t(x, y)u_1(y)/u_1(x)$.

10.2.3 Statistical Measures for the Analysis of Genomic Data

The SFS is one of the most commonly used statistics for the analysis of genomewide distributed SNPs. It is defined as the distribution of the number of times i a mutation is observed in a population or a sample of n sequences conditional on segregation. In the coalescent setting the SFS is (either theoretically or empirically) evaluated for a sample at the present time. In the diffusion framework, the SFS of a population or a sample can often even be derived in dependence of a time variable t and an equilibrium solution can either be implied by letting $t \to \infty$ or by making use of the measures from the foregoing section. Although the forward version of the SFS has interesting applications on time-series data (i.e. DNA sequences sampled over time), we will, for convenience, focus throughout on coalescent formulations and results and only mention equilibrium solutions with reference to diffusions.

Let the site frequencies be denoted as $f_{n,i}$ $(1 \leq i \leq n-1)$, where the index n shall indicate that the SFS depends on the sample size (except in the case of neutrality and a constant population size). It has been shown [22, 63] that the following equation holds for general binary coalescent trees, i.e. ancestral trees with pairwise coalescences at a time with arbitrary continuous waiting times and following the mutation model given above:

$$f_{n,i} = \frac{\theta}{2} \sum_{k=2}^{n-i+1} k \frac{\binom{n-i-1}{k-2}}{\binom{n-1}{k-1}} E(T_k). \tag{10.7}$$

Two related measures of the SFS are the expected number of segregating sites S_n equalling the total number of mutations in the infinitely-many sites model and the average number of pairwise differences Π_n. The relationships are

$$S_n = \sum_{j=1}^{n-1} f_{n,j},$$

$$\Pi_n = \frac{1}{\binom{n}{2}} \sum_{j=1}^{n-1} j(n-j) f_{n,j},$$

and the expectations of both quantities can be easily obtained via Eq. (10.7). One can also derive the normalized version of the SFS as $r_{n,i} = f_{n,i}/S_n$.

In the diffusion framework the equilibrium version of the SFS as the proportion of sites where the mutant frequency is in $(y, y + dy)$ is given by [21]

$$\hat{f}(y) = \theta \, \pi(y) \, u_0(y), \tag{10.8}$$

and the finite version can be immediately obtained via binomial sampling as

$$\hat{f}_{n,i} = \int_0^1 \binom{n}{i} y^i (1-y)^{n-1} \hat{f}(y) dy. \tag{10.9}$$

10.3 Effect of Seed Banking on Neutral Evolutionary Processes

The major effect of seed persistence is to increase genetic diversity by a storage effect of seeds in the soil. In coalescent theory terms, seed banks increase the time for two lineages to coalesce by a factor $1/b^2$ under any neutral model.

10.3.1 Constant Population Size

The expected time to coalescence is simply given by

$$E(T_j) = \left(b^2 \binom{j}{2}\right)^{-1},$$

which can, e.g., be derived by applying Eq. (10.2) to (10.3). The consideration of the average germination rate b diminishes the rate of genetic drift (Fig. 10.3, [28]) and is bounded as $1/m \leq b \leq 1$. The lower and upper bounds result from the scenarios, where all seeds, respectively, rest m and one generation in the bank. So the expected coalescent tree can be up to m^2 generations longer in the SB model compared to the usual non-SB Wright–Fisher model. Persistent seed banks, acting as a genetic storage, increase diversity by a factor $1/b^2$. For a neutral model of a constant population size, Eq. (10.7) simply yields $f_{n,i} = \theta/(b^2\, i)$.

Applying diffusion theory, the population and the sample SFS can be, respectively, obtained as $\hat{f}(y) = \theta/(b^2\ y)$ and $\hat{f}_{n,i} = \theta/(b^2\ i)$ by applying Eqs. (10.8) and (10.9), which are equivalent to the coalescent results. The mean time to absorption and the time to fixation can be, respectively, derived via Eqs. (10.5) and (10.6) as $\bar{t}(x) = -2/b^2\,(x\,\log(x) + (1-x)\,\log(1-x))$ and $\bar{t}^*(x) = -2/b^2(1-x)/x\,\log(1-x)$. Therefore, the mean time to absorption/loss and the time to fixation are both slowed down in the SB model by a factor of $1/b^2$ compared to the non-SB model.

10.3.2 Estimation of Past Demographic Events

The change in coalescent rate due to SB decreases the strength of genetic drift, lengthens the time of fixation of neutral alleles and thus diminishes the genetic differentiation between populations [56]. This consequently affects the inference of past demography of a population or a species [62]. In the following, we present a simplified demographic setting that allows us to make the results of Sects. 10.2.1 and 10.2.3 applicable via a simple time rescaling argument. We assume that plants and seeds of all age classes are equivalently affected by changes in the population size such that the relative proportions of all type of seeds, and therefore the seed bank age distribution $b_1, \ldots, b_m$ and its average germination rate b remain constant over time. Furthermore, the population size changes are assumed to occur on a coalescent time scale so that the relocation process can reach an equilibrium between coalescent events as in the case of a constant population size. This particularly implies that the population size is approximately constant over any given time window, i.e. for a given m-window k_0 the relative population size function (as scaled by the population size N at present time) at the i-th cell $\rho_N(i + k_0 - 1) = N(i + k_0 - 1)/N \approx \rho_N(k_0)$. This simplification holds for a

geometrically growing population, if the growth rate and m are chosen realistically small. In the case of an instantaneous population decline, this relationship is violated, but only for $m-1$ generations, so that for small m instantaneous changes within a window can be neglected due to the small corresponding coalescence probability for large population sizes.

In continuous time, let $\rho(t)$, which arises from $\rho_N([Nt])$ as $N \to \infty$ and time being measured in units of N generations, be piecewise continuous, bounded and follow the conventions in discrete time. The time-rescaling argument (in terms of the harmonic mean of the relative population sizes) for the coalescent approximation of the non-SB Wright–Fisher model [32], $t \to \int_0^t \rho(s)^{-1} ds$, can then be applied to the ancestral process $(A_n(t))_{t\geq 0}$ to obtain the process with time-varying population size $(A_n^{\rho}(t))_{t\geq 0}$. Therefore, the corresponding results to Eqs. (10.2) and (10.3) are, respectively, given by Živković and Tellier [62]

$$P(A_n^{\rho}(t) = j) = \sum_{k=j}^{n} c_{nk} r_{kj} \exp\left(-b^2 \binom{k}{2} \int_0^t \rho(s)^{-1} ds\right) \tag{10.10}$$

and

$$E(T_j^{\rho}) = \int_0^{\infty} P(A_n^{\rho}(t) = j) dt. \tag{10.11}$$

The SFS can be simply obtained by applying Eq. (10.11) to (10.7) and applied for the estimation of demographic parameters, which is only feasible to some extent with prior knowledge on the average germination rate b [62]. The SFS can also be obtained within the diffusion framework (e.g. via a moment based approach) to solve a time inhomogeneous-version of Eq. (10.4) with the diffusion term given by $\sigma^2(y,t) = b^2 y(1-y)/\rho(t)$ and the drift term $\mu(y) = 0$ (see [61] for non-SB models). As an application of these results, it has been shown in a seminal study [52] that b can be estimated by applying an Approximate Bayesian Computation (ABC) method to polymorphism data with prior knowledge on metapopulation structure and N using ecological data.

10.3.3 Seed Banks Decrease the Rate of Divergence Between Populations/Species

As the coalescent tree is longer by a factor $1/b^2$, we also predict that the inference of recent splits between populations (or species) is affected (Fig. 10.2). Seed banking decreases the genetic differentiation, measured as F_{st}, in spatially structured populations [49]. Note that the population migration rate should be multiplied by b if pollen only disperse, and not if pollen and seeds disperse. Ignoring seed banking thus influences the estimation of migration rates if the appropriate dispersal scaling

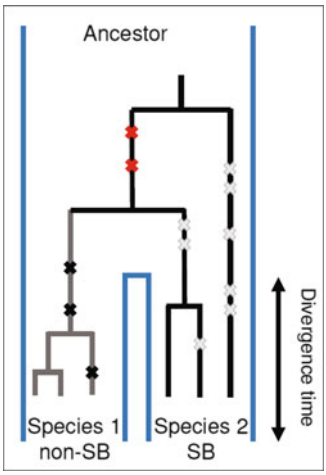


Fig. 10.2 In a model of speciation (without gene flow) from an ancestor into two incipient species, persistent seed banks increase the amount of incomplete lineage sorting due to the lower rate of genetic drift in species 2 (SB) compared to species 1 (non-SB). The number of shared alleles (red mutations) between species with and without seed bank is increased by seed banking compared to private alleles (black and grey crosses)

is not taken into account. On a longer time scale, species with persistent seed banks may exhibit higher rates of shared alleles between divergent species, the so-called Incomplete Lineage Sorting (ILS), which is obscuring the phylogenetic signal of speciation (red mutations in Fig. 10.2). For example, several species in the tomato clade show evidence for seed banking [43], and high rates of ILS are found [44].

10.3.4 Seed Banks Increase the Rate of Recombination

Seed banking decreases the recombination rate per year because only seeds germinating with probability b can undergo recombination upon gamete production. However, as the overall coalescent tree is longer by a factor $1/b^2$, the net effect of longer seed persistence is to increase the amount of recombination events per locus by a factor $1/b$ (Fig. 10.3). We can thus predict that at equal census size SB species compared to non-SB species should exhibit (1) a higher genetic diversity (number of SNPs), and (2) a higher recombination rate per locus (Fig. 10.3). Using prior information on N, nucleotide diversity and recombination rate per gene, it should thus be possible to estimate b for any given species with persistent seed bank.

10.4 Effect of Seed Banking on Selective Processes

Seed banks slow down natural selection, and decrease the Hill-Robertson effect. Persistent seed banking slows down the action of positive (and purifying) selection, so that it takes longer for a selected allele to get fixed in a population [23, 34, 49]. However, we have recently demonstrated that even though selection is slower under

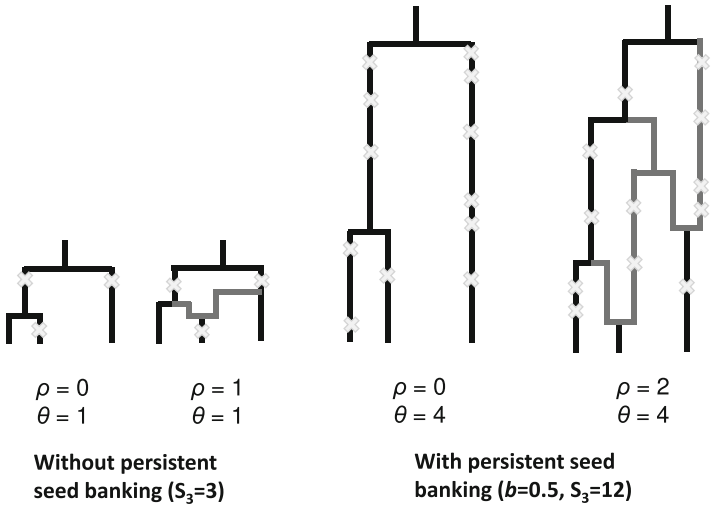


Fig. 10.3 Effect of a persistent seed bank on the length of the coalescent tree, diversity and recombination rate per gene. The comparison is made for two species of equal census size: one without ($b = 1$, left panel) and one with seed banking ($b = 0.5$, right panel). The number of segregating sites S_3 is increased by the seed bank (grey crosses). The per gene scaled recombination (ρ) and mutation (θ) rates are exemplarily given. The branches in grey are generated by recombination events

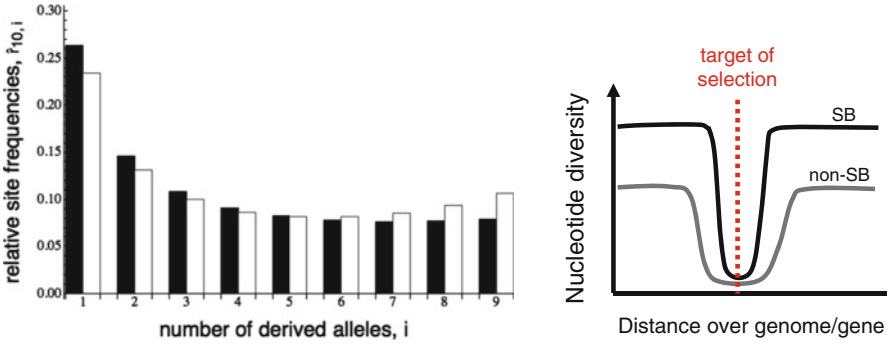


Fig. 10.4 Effect of a persistent seed bank on signatures of selection. (left panel) An increase of the strength of positive selection ($\sigma = 2$) among independent SNPs is illustrated in terms of the normalized SFS, $\hat{r}_{n,i} = f_{n,i}/S_n$; for species without (black) and with (white, $b = 0.5$) seed bank. (right panel) Persistent seed banks increase the strength of selection and recombination rates, so that narrower but deeper selective sweep signatures are expected around the target of selection in SB versus non-SB species

more persistent seed banks, the strength of selection is enhanced when observed at equilibrium (Fig. 10.4 for positive selection, [34]). This is explained by selection acting on every lineage germinating from the seed bank at rate b, whereas genetic drift acts on the time scale of b^2.

10.4.1 Selection on Unlinked Sites

As already mentioned above, it is more straightforward to model selection within the diffusion framework than via coalescent theory. For non-SB models, a partly numerical and partly explicit solution of Eq. (10.4) can be obtained by finding a spectral representation of $f(y,t)$ for a constant population size [48] and even for piecewise population size changes [64]. Since the average germination rate b is included as a linear factor in Eq. (10.4), the underlying method can even be extended to SB models. For convenience, we will only present the equilibrium solution of the SFS for a constant population size, which anyway demonstrates the main affect of selection on seed banks.

The population and the sample SFS are, respectively, given via Eqs. (10.8) and (10.9) as

$$\hat{f}(y) = \frac{\theta}{b^2 y(1-y)} \frac{1-\exp(-2\sigma(1-y)/b)}{1-\exp(-2\sigma/b)}$$

and

$$\hat{f}_{n,i} = \frac{\theta n}{b^2 i(n-i)} \frac{1-{}_1F_1(i;n;2\sigma/b)e^{-2\sigma/b}}{1-e^{-2\sigma/b}}.$$

For the mean time to absorption and the time to fixation one can only obtain onerous equations [34], which as well as the results for the SFS show that seed banks enhance the effect of selection (Fig. 10.4, left panel) while the considered allele takes longer to reach that equilibrium state.

10.4.2 Selection on Linked Sites

Here we propose four novel theoretical predictions which have not yet been tested with experimental data.

1. As selection is strong and recombination rates are higher in SB species, classic selective sweeps[8, 29] will exhibit a strong but reduced extent hitchhiking signature [41] around the target site (Fig. 10.4, right panel).
2. As seed banking generates higher nucleotide diversity and slows down positive selection, it is also expected that positive selection acts on standing genetic variation and incomplete sweeps (in which the selected allele has not reached fixation) should be observed (the so-called soft sweeps, [24, 57]).
3. Seed banks also favour the maintenance of polymorphism and enhance the strength of balancing selection [49, 51, 53], which should thus be more observable in SB species.

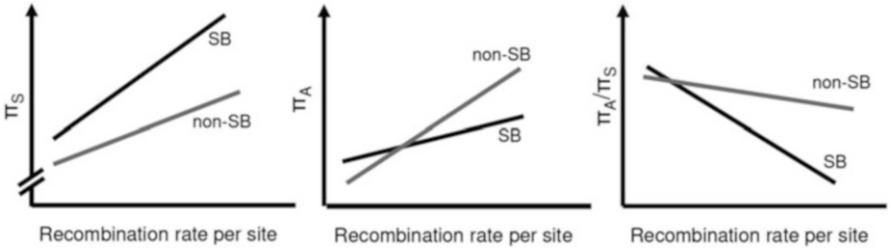


Fig. 10.5 Effect of a persistent seed bank on increasing pervasive purifying selection and decreasing the effect of background selection in the genome when comparing SB (black line) and non-SB (grey line) species. Genes with similar recombination rates are binned (see [4]). Nucleotide diversity at synonymous sites increases faster for increasing recombination rates in SB than non-SB (left panel), diversity at non-synonymous sites is driven by more efficient selection in SB versus non-SB (middle panel) predicting the ratio Π_A/Π_S to differ across recombination rates between SB and non-SB species (right panel)

4. The Hill-Robertson effect [25] should be reduced in outcrossing SB compared to outcrossing non-SB species. Purifying selection is also predicted to be efficient and narrow around the target sites (decreasing the extent of background selection, [5, 6]) in SB species (Fig. 10.5). This yields from the enhanced recombination rate compared to the coalescent rate and the increased strength of selection under seed banking.

Positive and purifying selection are thus predicted to be more efficient and show a reduced LD signature in SB than in non-SB species. In SB species, positive selection should be pervasive across the genome, signatures of classic selective sweeps would be narrow around the target site (Fig. 10.4), and the extent of purifying linked (background) selection should be limited (Fig. 10.5). More selective sweeps are also expected to arise from standing genetic variation and/or be incomplete in SB compared to non-SB species.

10.5 Conclusions

We have built here a model of a so-called weak seed bank for plants species, but also applicable to fungi or invertebrate species producing dormant resting stages. An assumption that we have used in both of our models is namely that the maximum time seeds can spend in the bank, m, is bounded and small compared to the time of coalescence of lineages. This condition ensures that our coalescent model converges to the n-Kingman coalescent [28], and that the separation of time scale and the Markovian property can be used to analyse the diffusion model. Seed banking is thus a unique life-history trait which links ecology and genetic diversity defining the Lewontin paradox in angiosperms. Following the hypotheses and predictions above (summarized in Table 10.1), we predict that SB species exhibit smaller census sizes,

Table 10.1 Summary of hypotheses and predictions

Hypothesis	Predictions
Seed banking increases neutral diversity and the rate of recombination	Seed banking species exhibit more diversity and higher recombination rates per locus
Seed banking enhances the efficiency of selection and decreases the Hill-Robertson effect	Signatures of selection are enhanced, selection is more pervasive, and footprints of selection narrower in genomes of SB species
Seed banking enhances the time of allele fixation and decreases genetic differentiation	Seed banking decreases the accuracy of past demographic inference and increases the amount of Incomplete Lineage Sorting and incongruence in phylogenies

i.e. species abundance, but much higher genetic diversity than non-SB species. This could be termed as an inverse Lewontin paradox for plant species when comparing SB and non-SB species, which generates the opposite pattern of diversity versus census size than found in animals. As a corollary, we predict that selfing non-SB species exhibit large census sizes but show the lowest genetic diversity compared to SB species.

An alternative model has been proposed, the so-called strong seed bank, where the time m can be infinite or at least very large compared to coalescent times [3]. These models produce very different predictions regarding the SFS and the effect of recombination, because they do not consist in a rescaled version of the Kingman n-coalescent. Strong seed banks can generate so-called multiple merger coalescents [3, 50] with signatures differing from our classic results for neutral sites. It is suggested that these models may be better adapted to bacteria species with resting stages that can survive for many years (especially for soil bacteria [37]), which is a much larger amount of time compared to their generation time [18]. As more and more full genome intra-species polymorphism data are becoming available in bacteria, fungi, invertebrates (Daphnia, Mosquito) and plants, it is a fascinating prospect to test the derived predictions and to assess which of the weak or strong seed bank models applies to a given species.

Acknowledgements This contribution is supported in part by Deutsche Forschungsgemeinschaft grants TE 809/1 (AT) and STE 325/14 from the Priority Program 1590 (DZ).

References

1. Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11:697–709
2. Baskin CC, Baskin JM (2014) Seeds: ecology, biogeography, and evolution of dormancy and germination, 2nd edn. Elsevier, Amsterdam
3. Blath J, González-Casanova A, Kurt N, Spano D (2013) On the ancestral process of long-range seedbank models. J Appl Probab 50:741–759

4. Campos JL, Halligan DL, Haddrill PR, Charlesworth B (2014) The relation between recombination rate and patterns of molecular evolution and variation in Drosophila melanogaster. Mol Biol Evol 31:1010–1028
5. Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289–1303
6. Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet Res 70:155–174
7. Charlesworth B (2009) Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. Nat Rev Genet. 10:195–205
8. Charlesworth B, Charlesworth D (2010) Elements of evolutionary genetics. Roberts & Company Publishers, Greenwood Village
9. Chen J, Glémin S, Lascoux M (2017) Genetic diversity and the efficacy of purifying selection across plant and animal species. Mol Biol Evol 34:1417–1428
10. Cohen D (1966) Optimizing reproduction in a randomly varying environment. J. Theor. Biol. 12:119–129
11. Corbett-Detig RB, Hartl DL, Sackton TB (2015) Natural selection constrains neutral diversity across a wide range of species. PLoS Biol 13:e1002112
12. Dann M, Bellot S, Schepella S, Schaefer H, Tellier A (2017) Mutation rates in seeds and seed-banking influence substitution rates across the angiosperm phylogeny. bioRxiv. https://doi.org/10.1101/156398
13. Ellegren H, Galtier N (2016) Determinants of genetic diversity. Nat Rev Genet 17:422–433
14. Evans MEK, Dennehy JJ (2005) Germ banking: bet-hedging and variable release from egg and seed dormancy. Q Rev Biol 80:431–451
15. Evans MEK, Ferriere R, Kane MJ, Venable DL (2007) Bet hedging via seed banking in desert evening primroses (Oenothera, Onagraceae): demographic evidence from natural populations. Am Nat 169:184–194
16. Ewens WJ (2004) Mathematical population genetics: I. Theoretical introduction. Springer, Berlin
17. Fenner M, Thompson K (2004) The ecology of seeds, Cambridge University Press, Cambridge
18. González-Casanova A, Aguirre-von-Wobeser E, Espín G, Servín-González L, Kurt N, Spanò D et al Strong seedbank effects in bacterial evolution. J Theor Biol 356:62–70 (2014)
19. Gossmann TI, Song BH, Windsor AJ, Mitchell-Olds T, Dixon CJ, Kapralov MV et al Genome wide analyses reveal little evidence for adaptive evolution in many plant species. Mol Biol Evol 27:1822–1832 (2010)
20. Gossmann TI, Keightley PD, Eyre-Walker A (2012) The effect of variation in the effective population size on the rate of adaptive molecular evolution in eukaryotes. Genome Biol Evol 4:658–667
21. Griffiths RC (2003) The frequency spectrum of a mutation, and its age, in a general diffusion model. Theor Popul Biol 64:241–251
22. Griffiths RC, Tavaré S (1998) The age of a mutation in a general coalescent tree. Stoch. Models 14:273–295
23. Hairston Jr NG, De Stasio Jr BT (1988) Rate of evolution slowed by a dormant propagule pool. Nature 336:239–242
24. Hermisson J, Pennings PS (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 169: 2335–2352
25. Hill WG, Robertson A (1966) The effect of linkage on limits to artificial selection. Genet Res 8:269–294
26. Finch-Savage WE, Leubner-Metzger G (2006) Seed dormancy and the control of germination: Tansley review. New Phytol 171:501–523
27. Jones SE, Lennon JT (2010) Dormancy contributes to the maintenance of microbial diversity. Proc Natl Acad Sci 107:5881–5886
28. Kaj I, Krone SM, Lascoux M (2001) Coalescent theory for seed bank models. J Appl Probab 38:285–300

29. Kim Y, Stephan, W (2002) Detecting a local signature of genetic hitchhiking long a recombining chromosome. Genetics 160:765–777
30. Kingman JFC (1982) On the genealogy of large populations. J Appl Probab 19A:27–43
31. Kimura M (1955) Stochastic processes and distribution of gene frequencies under natural selection. In: Cold spring harbor symposia on quantitative biology, vol 20. Cold Spring Harbor Laboratory Press, pp 33–53
32. Kimura M (1955) Random genetic drift in multi-allelic locus. Evolution 9:419–435
33. Kimura M (1969) The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. Genetics 61:893–903
34. Koopmann B, Müller J, Tellier A, Živković D (2017) Fisher-Wright model with deterministic seed bank and selection. Theor Popul Biol 114:29–39
35. Lande R (1988) Genetics and demography in biological conservation. Science 241:1455–1460
36. Leffler EM, Bullaughey K, Matute DR, Meyer WK, Segurel L, Venkat A et al (2012) Revisiting an old riddle: what determines genetic diversity levels within species? PLoS Biol 10:e1001388
37. Lennon JT, Jones SE (2011) Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat Rev Microbiol 9:119
38. Levin DA (1990) The seed bank as a source of genetic novelty in plants. Am. Nat. 135:563–572
39. Lewontin RC The genetic basis of evolutionary change. Columbia University Press (1974)
40. Lynch M, Lande R (1998) The critical effective size for a genetically secure population. Anim. Conserv. 1:70–72
41. Maynard-Smith J, Haigh J (1974) Hitch-hiking effect of a favorable gene. Genet Res 23:23–35
42. Möst M, Oexle S, Markova S, Aidukaite D, Baumgartner L, Stich HB et al (2015) Population genetic dynamics of an invasion reconstructed from the sediment egg bank. Mol Ecol 24:4074–4093
43. Nunney L (2002) The effective size of annual plant populations: the interaction of a seed bank with fluctuating population size in maintaining genetic variation. Am Nat 160:195–204
44. Pease JB, Haak DC, Hahn MW, Moyle LC (2016) Phylogenomics reveals three sources of adaptive variation during a rapid radiation. PLoS Biol 14:e1002379
45. Romiguier J, Gayral P, Ballenghien M, Bernard A, Cahais V, Chenuil A et al (2014) Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature 515:261–263
46. Roselius K, Stephan W, Städler T (2005) The relationship of nucleotide polymorphism, recombination rate and selection in wild tomato species. Genetics 171:753–763
47. Salguero-Gómez R (2017) Applications of the fast-slow continuum and reproductive strategy framework of plant life histories. New Phytol 213:1618–1624
48. Song YSS, Steinrücken M (2012) A simple method for finding explicit analytic transition densities of diffusion processes with general diploid selection. Genetics 190:1117–1129
49. Templeton AR, Levin DA (1979) Evolutionary consequences of seed pools. Am Nat 114:232–249
50. Tellier A, Lemaire C (2014) Coalescence 2.0: a multiple branching of recent theoretical developments and their applications. Mol Ecol 23:2637–2652
51. Tellier A, Brown JKM (2009) The influence of perenniality and seed banks on polymorphism in plant-parasite interactions. Am Nat 174:769–779
52. Tellier A, Laurent SJ, Lainer H, Pavlidis P, Stephan W (2011) Inference of seed bank parameters in two wild tomato species using ecological and genetic data. Proc Natl Acad Sci USA 108:17052–17057
53. Turelli M, Schemske DW, Bierzychudek P (2001) Stable two-allele polymorphisms maintained by fluctuating fitnesses and seed banks: protecting the blues in Linanthus parryae. Evolution 55:1283–1298
54. Venable DL (1989) Modeling the evolutionary ecology of seed banks. In: Leck MA (ed) Ecology of soil seed banks. Elsevier, Amsterdam, pp 67–87
55. Venable DL, Lawlor L (1980) Delayed germination and dispersal in desert annuals: escape in space and time. Oecologia 46:272–282

56. Vitalis R, Glémin S, Olivieri I (2004) When genes go to sleep: the population genetic consequences of seed dormancy and monocarpic perenniality. Am Nat 163:295–311
57. Vy HMT, Kim Y (2015) A Composite-likelihood method for detecting incomplete selective sweep from population genomic data. Genetics 200:633–649
58. Waterworth WM, Footitt S, Bray CM, Finch-Savage WE, West CE (2016) DNA damage checkpoint kinase ATM regulates germination and maintains genome stability in seeds. Proc Natl Acad Sci 113:9647–9652
59. Wright SI, Andolfatto P (2008) The impact of natural selection on the genome: emerging patterns in Drosophila and Arabidopsis. Annu Rev Ecol Evol Syst 39:193–213
60. Wright SI, Gaut BS (2005) Molecular population genetics and the search for adaptive evolution in plants. Mol Biol Evol 22:506–519
61. Živković D, Stephan W (2011) Analytical results on the neutral non-equilibrium allele frequency spectrum based on diffusion theory. Theor Popul Biol 79:184–191
62. Živković D, Tellier A (2012) Germ banks affect the inference of past demographic events. Mol Ecol 21:5434–5446
63. Živković D, Wiehe T (2008) Second-order moments of segregating sites under variable population size. Genetics 180:341–357
64. Živković D, Steinrücken M, Song YSS, Stephan W (2015) Transition densities and sample frequency spectra of diffusion processes with selection and variable population size. Genetics 200:601–617

Index

R. J. Morris (ed.), *Mathematical Modelling in Plant Biology*,
https://doi.org/10.1007/978-3-319-99070-5