METHODS IN PLANT BIOCHEMISTRY

— SERIES EDITORS —

P. M. DEY AND J. B. HARBORNE

VOLUME 10b
Molecular Biology

Edited by J. Bryant

Methods in Plant Biochemistry

Volume 10b
Molecular Biology

METHODS IN PLANT BIOCHEMISTRY

Series Editors
P. M. DEY
Department of Biochemistry, Royal Holloway and Bedford New College, UK
J. B. HARBORNE
Plant Science Laboratories, University of Reading, UK

1 Plant Phenolics: J. B. HARBORNE
2 Carbohydrates: P. M. DEY
3 Enzymes of Primary Metabolism: P. J. LEA
4 Lipids, Membranes and Aspects of Photobiology: J. L. HARWOOD and J. R. BOWYER
5 Amino Acids, Proteins and Nucleic Acids: L. ROGERS
6 Assays for Bioactivity: K. HOSTETTMANN
7 Terpenoids: B. V. CHARLWOOD and D. V. BANTHORPE
8 Alkaloids and Sulphur Compounds: P. G. WATERMAN
9 Enzymes of Secondary Metabolism: P. J. LEA
10 Molecular Biology: J. A. BRYANT
10b Molecular Biology: J. A. BRYANT

Methods in Plant Biochemistry

Series editors
P. M. DEY and J. B. HARBORNE

Volume 10b
Molecular Biology

Edited by

J. A. BRYANT
Department of Biological Sciences, University of Exeter, UK

ACADEMIC PRESS
San Diego London Boston
New York Sydney Tokyo Toronto

This book is printed on acid-free paper.

Academic Press, Inc.
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
http://www.apnet.com

Academic Press Limited
24-28 Oval Road, London NW1 7DX, UK
http://www.hbuk.co.uk/ap/

ISBN 0-12-461021-8

A catalogue record for this book is available from the British Library

Typeset by Saxon Graphics Ltd., Derby, UK
Transferred to digital printing 2006
96 97 98 99 00 01 EB 9 8 7 6 5 4 3 2 1

Contents

Contributors

J. A. Bryant, Department of Biological Sciences, University of Exeter, Exeter, EX4 4QG, UK

R. J. Ferl, Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, USA

R. Goldbach, Department of Virology, Agricultural University, Binnenhaven 11, 67 09 PD, Wageningen, The Netherlands

T. Hohn, Friedrich Miescher Institut, PO BOX 2543, CH-4002, Basel, Switzerland

A.-L. Paul, Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, USA

M. A. Vega-Palas, Instituto de Bioquimica Vegetal y Fotosintesis, Universidad de Sevilla-CSIC, Apdo 1113, 41080-Sevilla, Spain

R. Walden, Max-Planck-Institut für Züchtungforschung, Carl-von-Linné-Weg 10, D-50829 Köln 30, Germany

Z. A. Wilson, Department of Life Science, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

Preface to the Series

Scientific progress hinges on the continual discovery and extension of new laboratory methods and nowhere is this more evident than in the subject of biochemistry. The application in recent decades of novel techniques for fractionating cellular constituents, for isolating enzymes, for electrophoretically separating nucleic acids and proteins and for chromatographically identifying the intermediates and products of cellular metabolism has revolutionised our knowledge of the biochemical processes of life.

While there are many books and series of books on biochemical methods, volumes specifically catering for the plant biochemist have been few and far between. This is particularly unfortunate in that the isolation of DNA, enzymes or metabolites from plant tissues can often pose special problems not encountered by the animal biochemist. For a long time, the Springer series *Modern Methods in Plant Analysis*, which first appeared in the 1950s, provided the only comprehensive guide to experimental techniques for the investigation of plant metabolism and plant enzymology. This series, however, has never been completely updated; a second series has recently appeared but this is organised on a techniques basis and thus does not provide the comprehensive coverage of the first series. One of us (JBH) wrote a short guide to modern techniques of plant analysis *Phytochemical Methods* in 1976 (second edition, 1984) which showed the need for an expanded comprehensive treatment, but which by its very nature could only provide an outline of available methodology.

The time therefore seemed ripe to us to produce an entirely new multi-volume series on methods of plant biochemical analysis, which would be both thoroughly up-to-date and comprehensive. The success of *The Biochemistry of Plants*, edited by P. K. Stumpf and E. E. Conn and published by Academic Press, was an added stimulus to produce a complementary series on the methodology of the subject. With these thoughts in mind, we planned individual volumes covering: phenolics, carbohydrates, amino acids, proteins and nucleic acids, terpenoids, nitrogen and sulphur compounds, lipids, membranes and light receptors, enzymes of primary and secondary metabolism, plant molecular biology and biological techniques in plant biochemistry. Thus we have tried to cover all the major areas of current endeavour in phytochemistry and plant biochemistry.

The main aim of the series is to introduce to the scientist current knowledge of techniques in various fields of biochemically-related topics in plant research. It is also intended to present the historical background to each topic, to give experimental details of methods and analyses and appraisal of them, pointing out those methods that are most

suitable for immediate application. Wherever possible illustrations and structures have been used and one or more case treatments presented. The compilation of known data and properties, where appropriate, is included in many chapters. In addition, the reader is directed to relevant references for further details. However, for the sake of clarity and completeness of individual reviews, some overlap between chapters of volumes has been allowed.

Finally, we extend our warmest thanks to our volume editors for undertaking the important task of organising each volume and cooperating in preparing the contents lists. Our special thanks go to the staff of Academic Press and to the many colleagues who have made this project a success.

P. M. DEY
J. B. HARBORNE

Preface to Volume 10b

In 1976, I published a book entitled *Molecular Aspects of Gene Expression in Plants.*[1] I had wished originally to entitle it Plant Molecular Biology. However, I was advised by the Series Editor for the Academic Press Botanical Monographs Series, the late Professor James Sutcliffe, that the term 'molecular biology' would not necessarily be understood in the way I intended by all members of the plant science community. At that time, his advice was correct, but how different things are today. One of the most spectacular features of research in the plant sciences over the past 20 years has been the very rapid expansion of plant molecular biology, and I am sure that all who read the term 'molecular biology' know that it is about 'genes and their function as seen at the molecular level'.[1]

This expansion has followed the advent in 1973 of recombinant DNA technology – the ability to manipulate DNA *in vitro* – coupled with an amazing range of associated technical advances undreamt of in the late 1970s. When I published *Molecular Aspects of Gene Expression in Plants*,[1] recombinant DNA techniques had barely entered the plant science world, but I predicted that 'During the next decade we shall undoubtedly see . . . direct efforts to modify the genetic content of plant cells and attempts to regulate gene expression artificially'.[1] Both of those predictions have been amply fulfilled, but what I had failed to see was that recombinant DNA techniques would have an enormous impact on our understanding of plant genes themselves and of the ways in which the genes work. It is in this latter area that the expansion of our knowledge has been most spectacular, whilst the applications in agriculture and horticulture, although very important, have proceeded more slowly.

So, it is this flowering of plant molecular biology which has led to the production of volume 10 and of this supplement to Volume 10. In this present publication, I have concentrated on two main themes: methods for working with and analysing DNA, chromatin and genes and methods for genetic manipulation of plants. The first group of chapters in particular draws together material that has been difficult to find between the covers of one book and I hope this will be especially useful to our readers. In the second group of chapters, we concentrate exclusively on the methods of manipulation based on natural vectors since it is these methods, and particularly those based on the Ti plasmid of *Agrobacterium*, which have found the widest use. I am conscious that there are also some excellent "vectorless" methods for genetic manipulation, including biolistics and electroporation, but circumstances prevented their inclusion in this volume.

I should also add that practitioners of plant molecular biology are very busy people, active in their research and fired up by the general enthusiasm (and funding opportunities!)

in this exciting field. It is a significant sacrifice to take time away from the bench to write a chapter for a book such as this. Some have, entirely justifiably, been unwilling to do so. However, those authors who did respond to my invitation to contribute have done so enthusiastically, and I am very grateful to them. I am also grateful to the series editors, Professor J. B. Harborne and Dr P. M. Dey for their help in the early stages of preparing this book, and to the staff of Academic Press, not only for their help and support, but also for their cheerfully administered 'prods' at times when it seemed that things were moving rather slowly.

JOHN A. BRYANT

1. Bryant, J. A. ed. (1976). *Molecular Aspects of Gene Expression in Plants*. Academic Press, London & New York.

1 DNA Extraction

JOHN A. BRYANT

Department of Biological Sciences, University of Exeter, Exeter, EX4 4QG, UK

I. INTRODUCTION

The extraction of DNA from plant tissue lies at the heart of any attempt to analyse plant genomes at the molecular level. The method to be employed will depend on the specific purpose to which the isolated DNA is to be put, the amount and purity of the DNA required for this specific purpose and, as will become very apparent as the chapter proceeds, the nature of the plant material which is being used as the source of the DNA. This chapter thus starts with methods for the preparation of purified nuclear DNA where rapidity of preparation and total yield of DNA are not major criteria. It then proceeds to

METHODS IN PLANT BIOCHEMISTRY Vol. 10b
ISBN 0-12-461021-8

deal with methods for isolating DNA from cells/tissues without prior preparation of nuclei; some of these methods are suitable for relatively rapid processing of large numbers of samples of plant material. Finally there is a discussion of the problems encountered in preparing DNA from plants in which there is an extensive content of phenolic compounds, of polysaccharides or of other material which interferes with the more routine preparation methods. In practice it is found that many plants fall into this category, such that almost anyone who wishes to work with a plant species other than the 'standard' range of 'laboratory plants' is likely to encounter some problems in preparing DNA.

II. PREPARATION OF NUCLEAR DNA

A. Introduction

For some purposes, for example in making a nuclear genomic library, it is important to exclude as far as possible the contamination of nuclear DNA by DNA from the other genomes in plant cells, i.e. the genomes of the chloroplasts and mitochondria. Chloroplast DNA is quantitatively by far the more important of these two and even in non-green tissues can constitute a significant percentage of the total cellular DNA. The obvious way to overcome this problem is to prepare DNA from nuclei rather than from total cell lysates. Preparation of DNA from nuclei has also been recommended as a means of avoiding the problems caused by contamination by polyphenols in tissues which are rich in these compounds (Couch and Fritz, 1990; Collins and Symons, 1992).

B. Isolation of Nuclei

Isolation of nuclei is a topic which almost deserves a chapter in its own right, such is the variety of techniques which have been employed. For this reason, I concentrate here on a relatively simple technique which has been used (sometimes with modifications) with success in several laboratories and which may be readily adapted for a range of the more 'standard' species used in research.

1. Isolation of nuclei from shoot tips of dark-grown pea seedlings

This method is based on that of Dunham and Bryant (1983). Dark-grown seedlings are used firstly in order to prevent development of etioplasts into chloroplasts. This increases the size differential between the plastids and the nuclei and reduces the potential for contamination of nuclear DNA by plastid DNA. Secondly, in the absence of photosynthesis the amount of starch which accumulates is reduced significantly (however, some starch is still made from the the soluble sugars which are transported to the shoot tip during mobilisation of the reserves in the cotyledons). In general, even if plants are not grown completely in the dark, transferring them to darkness for 48 h prior to preparing nuclei (and indeed, prior to the use of other DNA extraction methods) can be very helpful in lessening the problems caused by carbohydrate contamination.

2. Nuclear isolation buffer

Tris–HCl	50 mM, pH 8.0
β-mercaptoethanol	10 mM
magnesium chloride	5 mM
sucrose	600 mM

3. Method

Harvest 1–5 g of the shoot tips (apical 10 mm) and homogenise in four volumes of nuclear isolation buffer with a pestle and mortar at 1–4°C. Add a further two volumes of buffer, and gently disperse the homogenate. Filter the homogenate through two layers of Miracloth into a chilled centrifuge tube. The two layers of Miracloth should be arranged with the mesh of one layer at right angles to the mesh of the other layer. Wash the pestle and mortar twice, each time with two volumes of nuclear isolation buffer, and filter the washings through the same two layers of Miracloth. The relatively large volume of buffer in relation to the amount of plant material used in this phase of the procedure helps to prevent nuclei being trapped in fragments of cell wall or of incompletely homogenised tissue. The filtrate is then centrifuged in the cold for 8 min at 350 × *g*. In this laboratory, we use the swing-out rotor of a rapid-acceleration centrifuge such as the MSE 'Cool-Spin'. The pellet is gently resuspended in 20 ml of nuclear isolation buffer and again centrifuged for 8 min at 350 × *g*. The pellet from this second centrifugation is essentially free from contamination by plastids and mitochondria (the low centrifugation speed is important here). However, the nuclei are contaminated by starch grains. These are not a major problem, particularly if the DNA extraction procedure to be used incorporates a CsCl centrifugation step prior to precipitating the DNA (see Section II.C). This method for making a crude nuclear preparation is ideal for etiolated or non-green tissue. However, if it is not possible to avoid green tissue, then contaminating chloroplasts may be lysed by adding Triton X-100 (to 1%) to the buffer prior to the first centrifugation. The lysed chloroplasts do not pellet at 350 × *g*, and the Triton X-100 may be removed by resuspension of the pellet and subsequent centrifugation in Triton-free buffer.

C. Lysis of Nuclei and Extraction of DNA

The best way of ensuring that DNA is free from contaminants is to refrain, if possible, from precipitating the DNA until all possible contaminants have been removed. This will avoid the co-precipitation of the contaminants with the DNA. One way of doing this is to lyse a nuclear pellet into a solution of CsCl and then to centrifuge the CsCl solution in order to collect the DNA as a buoyant band.

1. Nuclear lysis buffer

Nuclear isolation buffer is made up to 20 mM with respect to Na_2EDTA. One gram of sodium sarkosinate ('Sarkosyl') per 100 ml is then added.

2. Method

To the nuclear pellet obtained after the second centrifugation (as described above) add 1.0 ml of nuclear lysis buffer for every gram of fresh weight of the original plant material used. Disperse the pellet gently into the lysis buffer and then add 1.6 ml of CsCl solution (saturated at 22°C) for every 1.0 ml of nuclear lysis buffer. Mix very thoroughly by gentle stirring and continue to stir at 1–4°C for 4 h. Add 50 μl of ethidium bromide solution (10 mg ml^{-1}) for every 1.0 ml of nuclear lysis buffer, and mix thoroughly by gentle stirring. Check that the refractive index of the solution is between 1.398 and 1.400 (adjust if necessary by adding distilled water or saturated CsCl). Centrifuge the solution for 18 h in an ultracentrifuge. The exact conditions for this will depend on the particular model of ultracentrifuge and the availability of rotors. We carry out the ultracentrifugation at 50 000 rpm in the 70Ti rotor of a Beckman centrifuge at 25°C. Following centrifugation, the DNA, stained with ethidium bromide, should be visible as a deep-pink band against the paler orange of the CsCl–ethidium bromide solution. However, if yields of DNA are low, it may be necessary to visualise the fluorescence of the stained DNA under ultraviolet light, e.g. from a transilluminator. RNA will be visible at the bottom of the tube as a deep-red precipitate while proteins and polysaccharides usually form a 'skin' on top of the CsCl solution (polysaccharides may form a lightly stained buoyant band above the DNA band). The DNA is removed from the centrifuge tube by puncturing the tube with a hypodermic needle and drawing out the DNA *via* the needle, using a syringe. The next step is to remove the ethidium bromide. My own preference here is to gently mix the DNA solution with an equal volume of *iso*propanol which has been previously equilibrated with saturated CsCl. The ethidium bromide partitions into the *iso*propanol, colouring it a vivid pink, and the difference in density between the organic and inorganic phases means that they separate very readily, usually without the need for centrifugation. The extraction with *iso*propanol must be repeated until all the ethidium bromide is removed (this can be verified by checking for fluorescence under UV light). Other members of my research group prefer to use water-saturated butanol as the extractant for removing the ethidium bromide, but the method is identical to that described for *iso*propanol. Although the use of butanol lessens the cost of this procedure (because it is not neccessary to equilibrate the butanol against saturated CsCl), it should be noted that butanol is more toxic than *iso*propanol and that butanol vapour can cause severe headaches in some people. Following the removal of ethidium bromide from the DNA solution, the CsCl is then removed by dialysing the solution against either 100 mM NaCl or 50 mM sodium acetate. Dialysis may be completed within about 4 h, especially if the dialysis sac containing the DNA is removed to a fresh batch of dialysis medium after about 2 h. The DNA may then be precipitated by addition of two volumes of absolute ethanol followed by storage at –20°C. DNA prepared in this way is usually well in excess of 50 kb in size, and is entirely suitable for restriction enzyme digestion, establishment of genomic libraries and many other uses.

III. PREPARATION OF DNA FROM TISSUE HOMOGENATES *VIA* CENTRIFUGATION IN CsCl

Although there are situations where it is specifically necessary to prepare nuclear DNA, there are many more where this is not necessary. In these experiments, the protocol

described above may be readily simplified. One such procedure is outlined in Chapter 3. Here, I briefly describe one which is used routinely in my laboratory.

1. Method

Tissue is homogenised in nuclear isolation buffer and the homogenate is filtered through two layers of Miracloth as described for nuclear DNA. The homogenate is centrifuged at 2000 × *g* for 20 min, and the pellet from this centrifugation is subject to lysis in nuclear lysis buffer as previously described. From here, the protocol is as described for nuclear DNA, but it must be remembered that although yields of DNA may be higher, there will also be more by way of contaminants, such as cell wall polysaccharides, although for most of the 'standard' laboratory species this will not present any new problems. However, one problem which must be mentioned is that of nucleases. In procedures of this type, the DNA comes more readily into contact with nucleases than it does when the emphasis is on the preparation of 'whole' nuclei. If nucleases are indeed a problem, it is advisable to freeze the tissue in liquid nitrogen and then to homogenise the frozen tissue to a powder before adding the nuclear isolation buffer. In this way the tissue is in a very finely divided state but too cold for nucleases to work, and it is this finely divided state which allows the buffer to gain rapid access to all of the homogenate. If this method is used, it may be advisable to increase the velocity at which the Miracloth filtrate is centrifuged in order to increase DNA yields.

IV. 'RAPID' METHODS FOR PREPARATION OF PLANT DNA

A. Introduction

Although the methods just described certainly yield clean DNA which is suitable for further manipulation or analysis, they are neither rapid nor convenient for parallel preparation of more than a few samples. To overcome these disadvantages, a number of methods have been developed which do not necessitate the use of CsCl gradients. This feature cuts down on both expense and time and also makes more feasible the preparation of DNA from larger numbers of samples of plant tissue. In parallel with rapid methods for preparing small amounts of plasmid DNA, these methods are often termed 'mini-preparations'. Interestingly, the basis of these methods goes back to a protocol which was first developed over 30 years ago for the isolation of bacterial DNA (Marmur, 1961) and then modified for plants in the 1970s (e.g. see Bryant *et al.*, 1975). In outline, such procedures make use of detergents to lyse cell homogenates; some use phenol (or phenol–chloroform) to de-proteinise the lysates; some employ ribonuclease to remove RNA. The most widely used modification to this general procedure is the use of cetyltrimethylammonium bromide (CTAB; also known as hexadecyltrimethylammonium bromide) as an aid in the removal of polysaccharides.

B. A general plant DNA 'mini-preparation' technique

The method to be described here forms the basis for plant DNA mini-preparations used in many laboratories. It closely follows the protocol published by Dellaporta *et al.* (1983) and avoids the use of phenol or of phenol–chloroform.

1. Isolation of DNA from shoots of dark-grown pea seedlings

Extraction buffer

Tris–HCl,	100 mM, pH 8.0
Na_2EDTA	50 mM
NaCl	500 mM
β-mercaptoethanol	10 mM

Method

Freeze 0.5 g of tissue rapidly in liquid nitrogen and homogenise to a powder with a pestle and mortar. Transfer the still-frozen powder to a 30 ml Oak Ridge centrifuge tube (if necessary, wash the powder from the mortar to the tube with a small amount of liquid nitrogen). Before the homogenate thaws, add 15 ml of extraction buffer and then allow any remaining liquid nitrogen to evaporate. Add 1.0 ml of 20% (w/v) sodium dodecyl sulfate and mix vigorously; incubate at 65°C for 10 min. This is the lysis step, and the length of time may be increased to up to 1 h if necessary. After lysis, add 5.0 ml of 5 M potassium acetate, mix vigorously and allow to stand at 0°C for 20 min. An exchange occurs between K^+ and Na^+ ions, leading to formation of potassium dodecyl sulfate, which is insoluble. Proteins and polysaccharides form a complex with the potassium dodecyl sulfate, and are thus precipitated. The precipitate is removed by centrifugation at 25 000 × *g* for 20 min, and the supernatant is then filtered through one layer of Miracloth into another Oak Ridge tube containing 10 ml of *iso*propanol. After mixing thoroughly, the tubes are incubated at –20°C for 30 min in order to precipitate the nucleic acids (although no specific steps have been employed to remove RNA, *iso*propanol tends to promote the precipitation of DNA rather than RNA). The (mainly) DNA precipitate is collected by centrifugation at 20 000 × *g* for 15 mins, drained dry and then redissolved in 700 μl of 50 mM Tris–HCl buffer (pH 8.0) containing 10 mM Na_2EDTA. Any material which fails to redissolve may be pelleted at this stage by a 10 min centrifugation in a microcentrifuge. The supernatant from this centrifugation is transferred to a fresh microcentrifuge tube, and the DNA is precipitated by addition of 75 μl of 3 M sodium acetate and 500 μl of *iso*propanol. According to Marmur (1961), as cited by Dellaporta *et al.* (1983), precipitation with *iso*propanol in the presence of sodium acetate at this point leaves any remaining polysaccharides in solution, thus effectively removing them from the DNA. The DNA is collected by a brief centrifugation in the microcentrifuge, washed with 70% ethanol, dried and then dissolved in 100 μl of 10 mM Tris–HCl/1 mM Na_2EDTA, pH 8.0 (although many investigators find it equally satisfactory to dissolve the DNA in sterile double-distilled water).

C. Commonly encountered problems with the mini-preparation procedure

The most frequently encountered problems with the procedure described above all relate to the presence of contaminating compounds in the DNA samples and, in particular, one or more of the following: RNA, proteins, polysaccharides and polyphenols.

1. RNA

Although the protocol described above employs no specific precautions to preserve RNA and at the same time involves the use of two (partially) selective precipitation steps, nearly all DNA preparations made in this way are contaminated with RNA. If the amounts are small, then the RNA may be digested with ribonuclease at the same time as the DNA is being treated with a restriction endonuclease, as described by Dellaporta *et al.* (1983). However, for tissues which are particularly rich in RNA in comparison to the amount of DNA it may be necessary to incorporate a ribonuclease digestion step into the main protocol. In my laboratory, in common with many others, this is done after the first precipitation of nucleic acids. The pellet is dissolved in an acidic buffer (such as MOPS–KOH), deoxyribonuclease-free ribonuclease is added to a final concentration of 100 $\mu g\ \mu l^{-1}$ and the solution is incubated at 37°C for 1 h. (Note: if there is any doubt that the RNAase is DNAase-free, it should be heated at boiling point for 10 min prior to use). Following the RNAase treatment, the enzyme must be removed by treatment of the DNA solution with a 1:1 mixture of phenol and chloroform (phenol – 500 g of phenol, 55 ml of water, 70 ml of *m*-cresol, 500 mg of 8-hydroxyquinoline; chloroform – 24 volumes of chloroform: 1 volume of *iso*amyl alcohol). The phenol–chloroform is mixed thoroughly with the DNA solution, the mixture is centrifuged at about 5000 × *g* for 15 min and the upper phase, containing the DNA, is carefully removed and retained for the next precipitation step.

2. Protein

For tissues which are rich in proteins, the co-precipitation of proteins as an insoluble complex with potassium dodecyl sulfate may prove inadequate for removal of protein from the nucleic solution. In such situations, it is important not to precipitate the nucleic acids until as much as possible of the protein has been removed because nucleic acid precipitates which are contaminated with proteins are difficult to redissolve. To achieve this, nucleic acid solutions are treated either with phenol or with phenol–chloroform (as described above) until no more denatured protein appears at the interface between the aqueous and phenol phases. The nucleic acids may then be precipitated in the usual way (e.g. with *iso*propanol, as described above).

3. Polysaccharides

Plant cells are notoriously rich in polysaccharides, not least because polysaccharides make up the bulk of the plant cell wall (except in lignified cells) and because many plant cells contain starch grains. Further, some plants also accumulate gelling polysaccharides, and these may be especially abundant in particular tissues or in plants which have been subject to environmental stress. As with proteins, the objective must be to avoid as far as is possible the co-precipitation of polysaccharides with DNA. Again in common with protein contamination, contamination by polysaccharides may prevent successful redissolving of the DNA; further, even if the DNA does dissolve, polysaccharides associated with DNA often prevent access of enzymes to DNA, thus inhibiting processes such as cutting with restriction enzymes, PCR or *in vitro* labelling. Some authors advocate the use of hydrolytic enzymes (Rether *et al.*, 1993) or of ion-exchange resins (Guillemaut and Maréchal-Drouard, 1992; Maréchal-Drouard and Guillemaut, 1995) to remove poly-

saccharides from nucleic acid solutions. However, in more generally used methods, the key feature in avoiding co-precipitation of polysaccharides with the DNA is to keep one of them in solution, whilst precipitating the other (although I note that Dellaporta *et al.* (1983) do not attempt to avoid co-precipitation of nucleic acids with polysaccharides at the first precipitation step in their method for preparing DNA from 'difficult' species; I guess it depends on how one defines difficult!). The most widely used method for avoiding co-precipitation of polysaccharides and DNA is to keep the polysaccharides in solution at the DNA precipitation step. CTAB is an ideal reagent here because it precipitates nucleic acids whilst leaving polysaccharides in solution. The following procedure was worked up in my own laboratory (P. Barnwell, A. N. Blanchard, J. A. Bryant and N. Smirnoff, unpublished; based on Rogers *et al.* (1989)) for use with succulent plants which are rich in mucilaginous polysaccharides, such as *Sedum telephium*. The method has been successfully used with as little as 1 g of fresh weight of starting material, but here I describe its use with 24 g of leaf material.

4. Extraction of DNA from leaves of Sedum telephium

Extraction buffer

Tris–HCl	100 mM, pH 8.0
Na_2EDTA	20 mM
CTAB	2% (w/v)
NaCl	1.4 M
polyvinylpyrrolidone	1% (w/v)

Precipitation buffer

Tris–HCl	50 mM, pH 8.0
Na_2EDTA	10 mM
CTAB	1% (w/v)

High-salt TE buffer

Tris–HCl	10 mM, pH 8.0
Na_2EDTA	1 mM
NaCl	1 M

10% CTAB

CTAB	10% (w/v)
NaCl	0.7 M

Method

Freeze 24 g of freshly harvested leaves in liquid nitrogen and grind to a powder with a pestle and mortar. Add 120 ml of extraction buffer and mix until the homogenate has thawed to form a paste. Distribute the paste equally between four 50 ml Oak Ridge tubes and incubate at 65°C with occasional shaking. Centrifuge at 3000 × *g* for 5 min at 22°C. Now distribute the supernatants from the four tubes to eight clean Oak Ridge tubes; to each of these add 3.75 ml of 10% CTAB, mix by vortexing for 10 s and centrifuge at 3000 × g for 5 min at 22°C. Remove 14 ml of supernatant from each tube, combine to a single sample, add 336 ml (i.e. three volumes) of precipitation buffer. Mix thoroughly and divide the sample between two 250 ml centrifuge bottles. Incubate at 22°C for 30 min and then centrifuge at 5000 × *g* for 15 min at the same temperature. Carefully pour off

the supernatant, drain the bottles and redissolve each pellet in 3 ml of high-salt TE buffer. Disperse the solution to 12 microcentrifuge tubes (500 μl to each tube). If necessary, centrifuge for 2 min to precipitate any material which has not redissolved, and transfer the supernatants to clean microcentrifuge tubes. To each, add 1 ml of ethanol (from a stock kept at –20°C) and mix gently by inverting the tubes several times; incubate at –20°C for 1 h. Harvest the nucleic acid precipitates by centrifugation, pour off the supernatants and wash the pellets twice in 70% aqueous ethanol. At this stage, the pellets consist of a mixture of DNA and RNA. The latter may be removed by treatment with RNAase, as described earlier. Alternatively, it is possible to carry out a differential precipitation. To do this, redissolve the nucleic acid pellets in 100 μl of sterile double-distilled water and then add 20 μl of ice-cold 12 M LiCl, and incubate at 4°C for 30 min before centrifuging for 10 min at the same temperature. The pellet contains rRNA and mRNA (which are susceptible to LiCl precipitation by virtue of their secondary structure), and the supernatant contains DNA and tRNA, which may be precipitated by addition of two volumes of cold ethanol followed by incubation at –20°C for 30 min. The presence of tRNA in the DNA sample does not present any problems; its low molecular weight means that it runs very much farther than nearly all DNA fragments on agarose gels, and it certainly does not interfere with processes such as restriction enzyme digestion of the DNA.

S. telephium is undoubtedly the most difficult plant we have worked on, but with this method it has proved possible to prepare DNA which is clean enough to cut with a wide range of restriction enzymes, to use as templates in polymerase chain reaction (PCR) and for *in vitro* labelling and for use in Southern blots. The yields of DNA are certainly low (20 $\mu g.g^{-1}$ fresh weight, compared with 300 $\mu g.g^{-1}$ fresh weight for pea shoot tips subjected to the same method), but the purity of the DNA is adequate for all usual applications.

5. *Polyphenols*

Many plant cells contain a range of phenolic compounds, including polyphenols, which are readily oxidised. In intact cells, the oxidases are separated by subcellular compartmention from their substrates, but when cells are broken open, enzymes and substrates have ready access to each other. Oxidised polyphenols are brown in colour and tend to co-precipitate with proteins and nucleic acids, thereby 'tanning' them. For most proteins, this results in denaturation, whilst DNA is rendered insoluble (or, at best, very difficult to redissolve) and cannot be cut by restriction enzymes nor used as a template for PCR.

As mentioned earlier, some investigators recommend isolation of nuclei as the most effective way of separating DNA from polyphenols and their oxidases (Couch and Fritz, 1990; Collins and Symons, 1992). Modifications to the more generally used extraction techniques include the addition to extraction buffers of polyvinylpyrrolidone, an absorbant of polyphenols (Howland *et al.*, 1991; Guillemaut and Maréchal-Drouard, 1992), especially at acid pH (Kanazawa and Tsutsumi, 1992) and/or of SH group reagents such as cysteine or β-mercaptoethanol as antioxidants (Guillemaut and Maréchal-Drouard, 1992; Chung *et al.*, 1994; Maréchal-Drouard and Guillemaut, 1995). It is also possible to add a more specific inhibitor of polyphenol oxidases, such as diethyldithiocarbamic acid (Couch and Fritz, 1990; Howland *et al.*, 1991).

D. The range of possible extraction techniques

The preceding sections have mainly dealt with methods which are generally applicable, albeit with modifications to cope with problems encountered during DNA extraction. However, it is becoming obvious that, as molecular biological techniques are being

applied to a wider range of problems (and hence to a wider range of plant species), the techniques may need specific modification for the problem in question. Hence, regular readers of a journal such as *Plant Molecular Biology Reporter* are aware that each issue is likely to contain a description of yet another protocol for extracting DNA from 'problem plants'. Space does not permit dealing with all the published methods in detail, but Table 1.1 gives examples of the range of plants and problems dealt with by different investigators.

TABLE 1.1. Protocols for extracting DNA from 'difficult' plant species.

Reference	Species	Problem
Aitchitt *et al.* (1993)	*Phoenix dactylifera* (date palm)	Nucleases, polyphenols, polysaccharides
Chung *et al.* (1994)	*Prunus persica* (peach)	Nucleases, polysaccharides
Collins and Symons (1992)	*Vitis vinifera* (grape vine)	Polyphenols, polysaccharides
Couch and Fritz (1990)	*Theobroma cacoa* (cocoa)	Polyphenols
de Kochko and Hamon (1990)	*Abelmoschus* sp. (okra)	Polysaccharides
Guillemaut and Maréchal-Drouard (1992)	Several	Polyphenols, polysaccharides
Honeycutt *et al.* (1992)	*Saccharum officinarum* (sugar cane)	Polysaccharides
Howland *et al.* (1991)	*Betula* sp. (birch)	Polyphenols
Kanazawa and Tsutsumi (1992)	*Nelumbo* sp. (lotus)	Polyphenols
Li *et al.* (1994)	Several	Polysaccharides
Lodhi *et al.* (1994)	*Vitis vinifera* (grape vine)	Polyphenols, polysaccharides
Maguire *et al.* (1994)	Family Proteaceae	Polyphenols, polysaccharides
Maréchal-Drouard and Guillemaut (1995)	Several	Polysaccharides
Mettler (1987)	Several	Small amounts of tissue
Rether *et al.* (1993)	Several	Polysaccharides
Varadarajan and Prakash (1991)	*Ipomoea batatus* (sweet potato)	Latex, polyphenols, polysaccharides, secondary metabolites
Ziegenhagen *et al.* (1993)	*Abies alba* (silver fir)	Polyphenols, polysaccharides, resin, terpenoids

V. ISOLATION OF VERY HIGH MOLECULAR WEIGHT DNA

The techniques described in the previous sections of this chapter usually yield DNA molecules of about 50–100 kb in size. For most purposes, this is entirely acceptable; DNA

of this size is suitable for the construction of genomic libraries, for Southern blotting (after restriction enzyme digestion), for use as a PCR template and for many other applications. However, large though these DNA molecules may appear to be, they actually represent small fragments of the very long molecules which make up plant chromosomes; it is no exaggeration at all to state that many DNA molecules *in vivo* are hundreds of millions of base pairs in length. This means that there is a big gap between the level of information provided by methods which map chromosomes and the level of information provided by methods which analyse DNA. However, that gap is closing because of the advent of techniques such as pulsed-field gel electrophoresis which permit the separation of very large DNA molecules (e.g. see Schwartz and Cantor, 1984; Carle *et al.*, 1986). In order to take advantage of such techniques, it is of course necessary to be able to isolate large DNA molecules. To this end, methods have been developed which, whilst not yielding molecules of native length, certainly result in obtaining DNA molecules of 2 million base pairs or more (Ganal and Tanksley, 1989; Creusot *et al.*, 1992; Wing *et al.*, 1993). These methods rely on reducing as far as possible the shearing forces which act on DNA during preparation. To achieve this, either protoplasts or isolated nuclei are embedded in plugs of low melting temperature agarose and are then lysed by diffusing a lysing solution (usually buffered Na_2EDTA, sarkosyl and proteinase K) into the agarose (Ganal and Tanksley, 1989). Following lysis, the DNA may be digested, while still remaining in the agarose plug, by restriction enzymes which have rare cutting sites in order to generate very large fragments. Alternatively, the fragment size may be controlled by digesting the DNA with one of the more usually used restriction enzymes such as *Eco*RI, in the presence of the corresponding DNA methylase (Bonnema *et al.*, 1996). By changing the ratio of nuclease to methylase it is possible to regulate the extent to which the DNA is digested and thus to generate fragments in the size range required. The agarose plugs are then loaded into the wells of a suitable agarose gel for pulsed field electrophoresis. This type of technique is greatly aiding our ability to analyse and map plant chromosomes (see Chapter 3).

VI. CRUDE DNA PREPARATIONS FOR PCR

The ability to amplify specific DNA sequences by PCR does not rely on having purified the DNA. The main requirements are that the DNA should be accessible for the processes of primer annealing and template copying and free from contaminants which may introduce artefacts into the system. Further, since some applications of PCR require comparison of several samples, rapidity is at a premium. Therefore, for plant species which do not present any major problems, several rapid partial purification techniques have been developed (Oard and Dronavalli, 1992; Paterson *et al.*, 1993), the simplest of which applies PCR to DNA simply released into the soluble phase of a tissue homogenate by a drastic freeze–thaw procedure followed by a brief centrifugation (Luo *et al.*, 1992).

VII. CONCLUDING REMARKS

This chapter has dealt with different types of DNA extraction procedures developed for different purposes. It has also dealt with the problems which may occur when these tech-

niques are used in the 'real world' rather than with the 'standard' laboratory plants which may serve us as good models but cannot necessarily be used to answer the questions we need to ask in our particular line of work. The main approach to DNA extraction must be first to decide what type of procedure is appropriate, second to ascertain whether the plant species under investigation is likely to present any particular problems, third to see whether previous investigators have dealt with similar problems (and if so to try their methods) and fourth to be prepared to work out variants of published methods to suit the needs of the situation.

REFERENCES

Aitchitt, M., Ainsworth, C. C. and Thangavelu, M. (1993). *Plant Mol. Biol. Rep.* **11**, 317–319.
Bonnema, G., Hontelez, J., Verkerk, R., Zhang, Y. Q., van Daelen, R., van Kammen, A. and Zabel, P. (1996). *Plant J.* **9**, 125–133.
Bryant, J. A., Wildon, D. C. and Wong, D. (1975). *Planta* **127**, 53–68.
Carle, G. F., Frank, M. and Olson, M. V. (1986). *Science*, **232**, 65–68.
Chung, K. H., Korban, S. S. and Buetow, D. E. (1994). *Plant Mol. Biol. Rep.* **12**, 304–309.
Collins, G. G. and Symons, R. H. (1992). *Plant Mol. Biol. Rep.* **10**, 233–235.
Couch, J. A. and Fritz, P. J. (1990). *Plant Mol. Biol. Rep.* **7**, 8–12.
Creusot, F., Sevignac, M. and Dron, M. (1992). *Plant Mol. Biol. Rep.* **10**, 207–218.
de Kochko, A. and Hamon, S. (1990). *Plant Mol. Biol. Rep.* **7**, 3–7.
Dellaporta, S. L., Wood, J. and Hicks, J. B. (1983). *Plant Mol. Biol. Rep.* **1**, 19–21.
Dunham, V. L. and Bryant, J. A. (1983). *In* 'Isolation of Membranes and Organelles from Plant Cells' (J. L. Hall and A. L. Moore, eds) pp. 237–275, Academic Press, London.
Ganal, M. W. and Tanksley, S. D. (1989). *Plant Mol. Biol. Rep.* **7**, 17–27.
Guillemaut, P. and Maréchal-Drouard, L. (1992). *Plant Mol. Biol. Rep.* **10**, 60–65.
Honeycutt, R. J., Sobral, B. W. S., Keim, P. and Irvine, J. E. (1992). *Plant Mol. Biol. Rep.* **10**, 66–72.
Howland, D. E., Olive, R. P. and Davy, A. J. (1991). *Plant Mol. Biol. Rep.* **9**, 340–344.
Kanazawa, A. and Tsutsumi, N. (1992). *Plant Mol. Biol. Rep.* **10**, 319–323.
Li, Q.-B., Cai, Q. and Guy, C. L. (1994). *Plant Mol. Biol. Rep.* **12**, 215–220.
Lodhi, M. A., Ye, G-N., Weeden, N. F. and Reisch, B. I. (1994). *Plant Mol. Biol. Rep.* **12**, 6–13.
Luo, G., Hepburn, A. G. and Widholm, J. M. (1992). *Plant Mol. Biol. Rep.* **10**, 319–323.
Maguire, T. L., Collins, G. G. and Sedgeley, M. (1994). *Plant Mol. Biol. Rep.* **12**, 106–109.
Maréchal-Drouard, L. and Guillemaut, P. (1995). *Plant Mol. Biol. Rep.* **13**, 26–30.
Marmur, J. (1961). *J. Mol. Biol.* **3**, 208–218.
Mettler, I. J. (1987). *Plant Mol. Biol. Rep.* **5**, 346–349.
Oard, J. H. and Dronavalli, S. (1992). *Plant Mol. Biol. Rep.* **10**, 236–241.
Paterson, A. H., Brubaker, C. L. and Wendel, J. F. (1993). *Plant Mol. Biol. Rep.* **11**, 122–127.
Rether, B., Delmas, G. and Laouedj, A. (1993). *Plant Mol. Biol. Rep.* **11**, 333–337.
Rogers, S. O., Rehner, S., Bledsoe, C., Mueller, G. J. and Ammirati, J. F. (1989). *Can. J. Bot.* **67**, 1235–1243.
Schwartz, D. C. and Cantor, C. R. (1984). *Cell*, **37**, 67–69.
Varadarajan, G. S. and Prakash, C. S. (1991). *Plant Mol. Biol. Rep.* **9**, 6–12.
Wing, R. A., Rastogi, V. K., Zhang, H-B., Paterson, A. H. and Tanksley, S. D. (1993). *Plant J.* **4**, 893–898.
Ziegenhagen, B., Guillemaut, P. and Scholz, F. (1993). *Plant Mol. Biol. Rep.* **11**, 117–121.

2 Chromatin

MIGUEL A. VEGA-PALAS[1], ANNA-LISA PAUL and ROBERT J. FERL

Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, USA

[1]*Instituto de Bioquimica Vegetal y Fotosintesis, Universidad de Sevilla-CSIC, Apdo 1113, 41080-Sevilla, Spain.*

METHODS IN PLANT BIOCHEMISTRY Vol. 10b
ISBN 0-12-461021-8

I. INTRODUCTION

Most of the genome of eukaryotes is contained in the nucleus. The nucleus is a very complex organelle that is separated from the cytoplasm by the nuclear membrane. This membrane bears nuclear pores that allow the interchange of macromolecules between the nucleus and the cytoplasm. Inside the nuclear membrane is the lamina, which has been proposed to contribute to the structure of the nucleus and positioning of chromosomes by anchoring them (Galcheva-Gorgova *et al.*, 1988). The genome size in plants is variable, being in some cases up to 10^{10} bp, which would extend over some metres if unravelled. To fit into the nucleus, the genomic DNA needs to undergo a high degree of condensation. The solution to this steric problem has been resolved through the organisation of the nuclear DNA into chromatin.

The smallest level of organisation of the DNA in chromatin is the nucleosome. One nucleosome is composed of 160–200 bp of DNA that is wrapped around a core of basic proteins called histones. The nucleosome is bound by a single molecule of histone H1, which also binds to the internucleosome linker DNA and facilitates the organisation of nucleosomes into higher-order structures (Clark and Kimura, 1990). The nucleosomal array is organised into a helical structure with a diameter of 30 nm that is referred to as the chromatin fibre. Specific regions of this chromatin fibre bind to a proteinaceous structure known as the nuclear scaffold. As a consequence of this attachment to the scaffold, the fibre organises into loops that distribute along the chromosome and range in size from a few to hundreds of kilobases (Mirkovitch *et al.*, 1984). The most abundant protein component of the scaffold is the enzyme topoisomerase II, which is found in metaphase chromosomes along the entire length of the chromatid (Gasser *et al.*, 1986). The DNA regions associated with the nuclear scaffold have been named scaffold attachment regions (SARs), and show sequence similarities to each other (Gasser and Laemmli, 1987).

Many different techniques have been applied to the study of chromatin structure. In this chapter we will concentrate on the utilisation of nucleases and chemical agents to study lower levels of chromatin organisation, such as nucleosome distribution and the interactions between DNA and non-histone proteins. We will also describe some techniques to study SARs.

II. NUCLEASE SENSITIVITY ANALYSES

A correlation between the transcriptional state of plant genes and sensitivity to digestion with nucleases has been described. Thus, when the genes are transcribed they become more sensitive to nucleases (Murray and Kennard, 1984; Ferl, 1985; Steinmuller and Apel, 1986; Conconi and Ryan, 1993). However, not all nucleases digest chromatin in an identical way as they recognise different structural elements. Therefore, the digestion of the chromatin with a variety of nucleases allows different aspects of chromatin structure to be examined.

Although the nucleases cannot enter the cell, they can enter the nucleus through the nuclear pores and digest the chromatin. Thus, all the nuclease treatments described below begin with the isolation of nuclei. The procedures to purify the DNA and visualise the

results are also identical for each of the nuclease procedures. Since the fragments of DNA generated by the nucleases are resolved on agarose gels, we refer to these types of studies as nuclease sensitivity analyses at agarose resolution level.

An alternative to conducting nuclease digestions in nuclei is to make the chromatin accessible to nucleases by permeabilising intact cells with NP-40, lysolecithin or saponin. Chromatin sensitivity analyses of intact animal cells using these agents have recently been described (Stewart *et al.*, 1991). However, chromatin sensitivity analyses of permeabilised plant cells have not as yet been reported.

A. Nuclei Isolation

Nuclei isolation is a critical step as the quality of the results relies on the integrity of the nuclei. We recommend checking nuclear integrity with several procedures such as transcription run-on assays, microscopy and digestion with micrococcal nuclease (see below). Although different protocols have been described to isolate nuclei from plant tissue, they generally have many steps in common. We have had success studying maize and *Arabidopsis* genes, using leaves, roots and cultured cells, with the very basic nuclei isolation procedure detailed below:

(1) Collect 10 g of tissue and grind in a cold mortar in the presence of 5 ml of ice-cold nuclei isolation buffer (50 mM Tris, pH 8.0, 5 mM $MgCl_2$, 15 mM NaCl, 0.5 mM β-mercaptoethanol (β-ME), 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 0.1 mM EGTA, 0.3 M sucrose). Add more cold buffer to a ratio of 4 ml of buffer to 1 g of tissue.
(2) Transfer the resulting slurry to a motor-driven homogeniser and do four passes with the pestle at full speed (Wheaton Safe-Grind Potter–Elvehjer tissue grinder). When homogenising a tissue for the first time, the release of nuclei and their quality should be checked by microscopy. The nuclei can be stained by adding a few drops of ethidium bromide.
(3) Filter the homogenate through four layers of cheese cloth and one layer of Miracloth (Calbiochem) into an ice-cold centrifuge tube.
(4) Pellet the nuclei by centrifugation at 5600 ×*g* (8500 rpm in a Beckman JA-20 rotor) at 4°C for 5 min, resuspend 30 ml of nuclei isolation buffer and centrifuge again at 5600 ×*g* and 4°C for 5 min. Repeat the wash with nuclei isolation buffer once. We recommend using a paint brush to gently resuspend the pellet.
(5) Finally, resuspend the nuclei preparation in the appropriate nuclease sensitivity assay buffer as detailed below.

The resulting preparations can be subjected to digestion with nucleases directly, which enter the nuclei through the nuclear pores. These nuclei preparations are contaminated with mitochondria and chloroplasts, but since the nucleases cannot pass the membranes of either of these organelles, the chloroplast and mitochondrial DNAs will be seen in agarose gels as a high molecular weight band that is unaffected by nuclease concentration. Other procedures to isolate nuclei include a purification step of the nuclei in a discontinuous Percoll gradient (Luthe and Quatrano, 1980). We have observed that this step is not necessary for nuclease digestions and decreases the yield of nuclei.

B. Digestion of Nuclei with DNAase I

DNAase I (bovine pancreatic deoxyribonuclease) is an endonuclease that has been widely used to analyse the chromatin structure of many eukaryotic organisms. This enzyme degrades double-stranded DNA in a non-specific manner, but shows a preference for certain sequences (Ehrlich *et al.*, 1973). For this reason, any study of chromatin structure using DNAase I should include the digestion of naked DNA for reference.

DNAase I is commonly used to characterise the chromatin structure of genes under different transcriptional states. In general, it has been shown that when the genes are transcriptionally active they exhibit a higher sensitivity to nucleases than when they are inactive (Wood and Felsenfeld, 1982; Thomas *et al.*, 1985). In addition, enhanced sensitivities of DNA sequences within the promoter and in other regions of the genes have been proposed to be associated with the presence of *trans*-acting regulatory factors (Elgin, 1981; Ashraf *et al.*, 1987; Ferl and Nick, 1987; Paul *et al.*, 1987). In the following sections, we describe a procedure to digest nuclei with DNAase I. It has been successful in showing chromatin modifications associated with transcriptional activation of several plant genes.

(1) Resuspend the nuclei preparation in an adequate volume of DNAase I buffer (50 mM Tris, pH 8.0, 5 mM $MgCl_2$, 0.05 mM β-ME, 0.01 mM PMSF, 0.1 mM EGTA, 0.3 M sucrose) and separate aliquots of 500 μl.
(2) Preincubate the aliquots of nuclei at 37°C for 1–2 min.
(3) Prepare several concentrations of DNAase I by diluting the enzyme in DNAase I dilution buffer (20 mM sodium acetate, 5 mM $CaCl_2$, 0.1 mM PMSF, 50% glycerol) and add 10 μl of each resulting dilution to a different aliquot of nuclei. We recommend the final concentrations of DNAase I to range between 0.0 and 1.0 μg ml^{-1}. If instead of 10 μl, some other volume of the DNAase I solution is added to the 500 μl aliquots of nuclei, the concentration of $CaCl_2$ in the DNAase I dilution buffer should be modified to keep it at a final concentration of 0.1 mM.
(4) Incubate at 30°C for 9 min. Spin in a microcentrifuge for 15 s and resuspend the nuclei in nuclei extraction buffer (see Section II.E). Resuspending the nuclei in nuclei extraction buffer stops the reaction.

C. Digestion of Nuclei with Micrococcal Nuclease

Micrococcal nuclease has been widely used to analyse the chromatin structure of eukaryotic genes. This enzyme preferentially digests the linker DNA between nucleosomes. Thus, when the DNA from nuclei that have been extensively digested with micrococcal nuclease is resolved on high-percentage agarose gels (e.g. 1.5%), a 'nucleosome ladder' is observed (Fig. 2.1). Hybridising the nucleosome ladder to a precise fragment of DNA can determine whether the specific DNA fragment is associated with nucleosomes in the chromatin. In addition, micrococcal nuclease also can be used to identify phasing of nucleosomes in a specific region of DNA (Bryan *et al.*, 1981).

Like DNAase I, micrococcal nuclease also degrades double-stranded DNA in a non-specific manner, but shows a preference for certain sequences. Therefore, studies of chromatin structure with micrococcal nuclease should also include digestion of naked DNA.

In this section we describe a method to digest nuclei with micrococcal nuclease that has been successful in generating nucleosome ladders from maize and *Arabidopsis* chromatin.

(1) Resuspend the nuclei preparation in an adequate volume of micrococcal nuclease buffer (50 mM Tris, pH 8.0, 5 mM $MgCl_2$, 15 mM NaCl, 1 mM $CaCl_2$, 0.5 mM β-ME, 0.3 M sucrose) and separate into aliquots of 500 μl.
(2) Preincubate the nuclei aliquots at 37°C for 2 min.
(3) Prepare several concentrations of micrococcal nuclease by diluting the enzyme in micrococcal nuclease buffer, and add 10 μl of each resulting dilution to a different aliquot of nuclei. We recommend the final concentrations of micrococcal nuclease to range between 0.0 and 1.0 unit ml^{-1}.
(4) Incubate at 37°C for 1–2 min and stop the reaction by adding 100 μl of 0.1 M EDTA, pH 7.0. Spin in a microcentrifuge for 15 s and resuspend the nuclei in nuclei extraction buffer (see Section II.E).

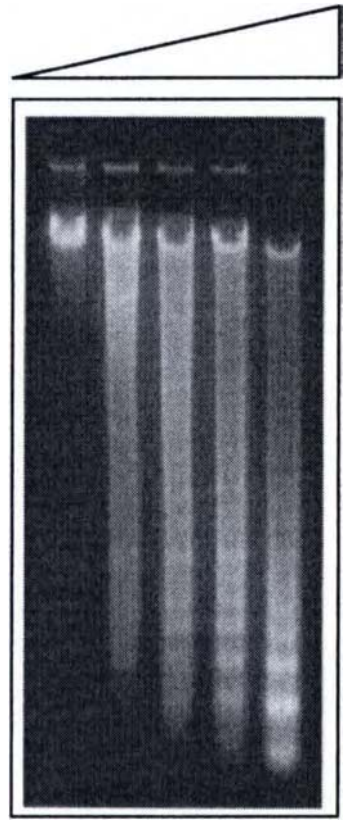


FIG. 2.1. Generation of nucleosome ladders with micrococcal nuclease. Increasing concentrations of micrococcal nuclease are indicated by the triangle over the lanes.

D. Digestion of Nuclei with Restriction Enzymes

The accessibility of a particular region of chromatin to restriction enzymes depends on its structure. Restriction enzymes have been used to characterise changes in the chromatin conformation of plant genes, associated with their transcriptional state (Ferl, 1985). In general, the results obtained with restriction enzymes corroborate equivalent results obtained with other nucleases, but analyses are limited to the specific area of the recognition sequences.

The restriction enzymes cut double-stranded DNA in an absolutely sequence-specific manner. Thus, changes in sensitivity can be precisely located when restriction enzymes are used. This level of precision cannot be obtained with either DNAase I or micrococcal nuclease on agarose gels

To digest nuclei preparations with restriction enzymes we have been successful using the nuclei isolation procedure and the buffers described above for DNAase I and micrococcal nuclease. The final concentrations of the enzymes in the assays should be determined empirically.

E. DNA Purification

In this section, a simple purification procedure that renders high-quality DNA is described (after Dellaporta *et al.*, 1983). This procedure can be used for all the nuclei digestion protocols described in the previous sections.

(1) Once the nuclei preparations have been digested with the nucleases and centrifuged, resuspend the pelleted nuclei in 300 μl of nuclei extraction buffer (100 mM Tris, pH 8.0, 50 mM EDTA, 500 mM NaCl, 10 mM β-ME). Add 5 μl of 20 mg ml^{-1} proteinase K and 50 μl of 10% sodium dodecyl sulfate (SDS). Incubate at 65°C for 45 min.
(2) Add 100 μl of 5 M potassium acetate, pH 4.5, mix by inversion and set on ice for 30 min.
(3) Remove the debris by centrifugation in a microcentrifuge for 10 min. Add 600 μl of isopropanol to the supernatant, mix and incubate at –20°C for 30 min.
(4) Centrifuge at 4°C for 10 min, wash the pelleted DNA twice with 75% ethanol, let the pellet dry and resuspend in 150 μl of TE.
(5) Extract once with phenol:chloroform:isoamyl alcohol (25:24:1) and a second time with chloroform:isoamyl alcohol (24:1). Add 1 μl of 10 mg ml^{-1} RNAase A, and incubate at 37°C for 10 min. Repeat both organic extractions and adjust the total volume of each sample to 150 μl. The first organic extraction is done to avoid possible degradation of the DNA by remaining endogenous nucleases during the following incubation in the presence of RNAase.
(6) Add 50 μl of 7.5 M ammonium acetate and 500 μl of 95% ethanol. Incubate at –20°C for 30 min.
(7) Centrifuge at 4°C for 10 min, wash the pelleted DNA twice with 75% ethanol, let the pellet dry and resuspend in 20–40 μl of TE (10 mM Tris, pH 8.0, 1 mM EDTA).

When the isopropanol and the ethanol are added to the samples, the typical fibrilar structure of precipitated DNA is usually seen in the non-digested samples or samples barely

digested with nucleases. The higher the nuclease concentration is, the lower the amount of fibrilar structure that is observed. However, when the total amount of DNA is high, this fibrilar structure is often seen in all the samples. In addition, when the total amount is low, it may not be seen in any case.

F. Visualisation of the Results

The samples of DNA treated with different concentrations of nucleases can be used to study the sensitivity of any region of the genome, as long as it has been previously cloned and mapped with restriction enzymes. The technique used for this purpose is called indirect end-labelling, and involves the following steps:

(1) Digest the DNA samples with a restriction enzyme to generate a fragment containing the region of interest.
(2) Resolve the resulting restriction fragments on an agarose gel and transfer them to a nylon membrane. GeneScreen (NEN/DuPont) and Hybond-N$^+$ (Amersham) membranes are recommended. The percentage of agarose in the gel will determine the extent and the degree of resolution of the nuclease sensitivity analysis. We have used gels containing 1.5% agarose to analyse fragments of 0.6–2 kb at a high level of resolution. The gel should include a lane of genomic DNA digested with restriction enzymes that cut inside the region of interest. This lane of markers is subsequently used to map the sensitive areas.

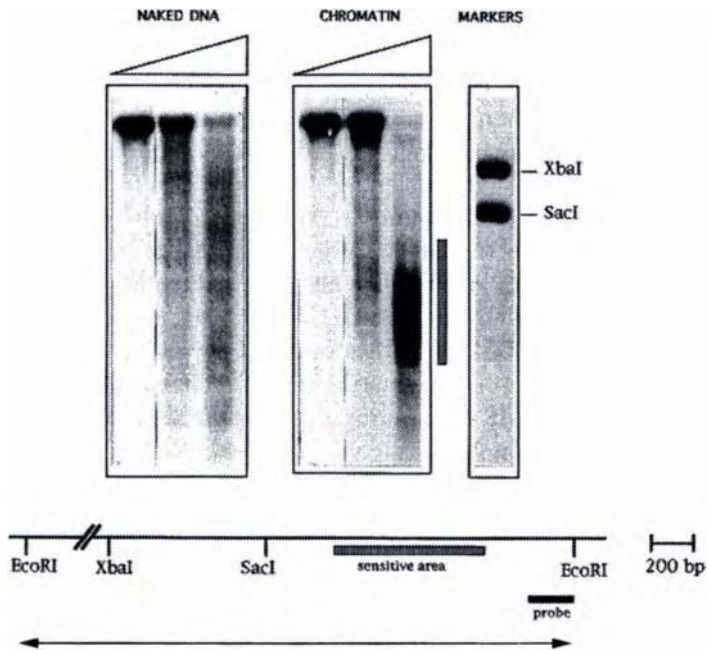


FIG. 2.2. DNAase I digestion patterns of the *Adh* gene from *Arabidopsis thaliana*. Increasing concentrations of DNAase I are indicated by the triangles over the lanes. A restriction map of the region of interest is given below the DNAase I digestion patterns, and indicates the DNAase-sensitive areas of chromatin (grey bars) and the probe homology (black bar).

(3) Clone a small portion of DNA that abuts one of the ends of the restriction fragment containing the region of interest. It will be used as a template to synthesise the probe for the hybridisation of the membrane. We recommend the template fragment to be 100–200 bp long.
(4) Synthesise a radioactive probe of high specific activity (10^8–10^9 cpm μg^{-1}). The more complex the genome to be studied, the more radioactive the probe should be. We have had success with M13 probes and riboprobes. The templates for the riboprobes can be synthesised according to Maniatis *et al.* (1989), and the riboprobe synthesis reactions can be done by using any commercial kit and 800 Ci $mmol^{-1}$ UTP. M13 probes can be synthesised according to Church and Gilbert (1984).
(5) Hybridise the membranes. Hybond membranes are hybridised according to the procedure recommended by the suppliers (Amersham), and GeneScreen membranes are hybridised following the method described by Church and Gilbert (1984). We have used Riboprobes to hybridise Hybond membranes and M13 probes to hybridise GeneScreen membranes.
(6) Autoradiograph the membranes using high sensitivity film and enhancer screens. Figure 2.2 shows an example of DNAase I hypersensitive site in the 5′ region of the *Arabidopsis Adh* gene.

III. *IN VIVO* FOOTPRINTING

There are several types of specific DNA–protein interactions and DNA structures that are often found within gene promoters. The observation of these aspects of chromatin structure *in vivo* requires a method of analysis that will not compromise the native chromatin structure, yet creates a modification in the DNA at the point of interaction that can be detected after the DNA has been purified from the system under analysis. One method that is capable of defining specific binding sites of proteins to DNA as they occur *in vivo* is dimethyl sulfate (DMS) footprinting in conjunction with genomic sequencing. DMS penetrates intact plant cells to modify the guanine residues in the genomic DNA within the nucleus (Ferl and Nick, 1987) by the chemistry of the Maxam and Gilbert sequencing reactions (Maxam and Gilbert, 1980). However, where proteins are in intimate contact with the DNA, the DMS modifications can be either enhanced or suppressed, depending upon the nature of the amino acid residue in contact with the guanine (Nick and Gilbert, 1985). The DMS footprinting and genomic sequencing techniques have been successfully applied to plant systems (e.g. see Ferl and Nick, 1987; Schulze-Lefert *et al.*, 1989; Paul and Ferl, 1991), making possible the *in vivo* detection of protein-binding sites in plant genomic DNA.

There are several points that should be taken under consideration while designing an *in vivo* footprinting experiment. First, the DNA sequence of the gene or region of interest must be known, so that any *in vivo* modifications can be mapped to specific base positions. Second, subclones need to be constructed in a suitable vector to enable the production of single-stranded probes for hybridisation. The cloned DNA should be 100–200 bp long and be defined by a genomic restriction site that is no more than 200 bases away from the centre of the region of study. Third, a homogeneous population of cells (such as

cultured cell suspensions) will yield more consistent results than a mixed population of cells (as in a plant organ).

A. DMS Treatment of Cells

The steps involving the direct use of DMS should be conducted in a chemical fume hood while wearing gloves and protective clothing, as DMS is a potent carcinogen. The steps outlined below are geared for using a 50 ml cell suspension culture of about 5 g that is contained in a 250 ml culture flask.

(1) Add 100 μl of DMS to the culture and swirl the flask vigorously for about 30 s to completely dissolve the DMS in the culture medium, then proceed with more gentle shaking for an additional 30–90 s. One or two minutes of exposure to DMS at 0.2% is usually sufficient to produce single cuts in a 1 kb section of DNA in intact cells.
(2) Recover the cells by vacuum filtration promptly after treatment, and rinse them with 1 l of deionised water. The media and wash should be collected into a vacuum flask containing enough NaOH to keep the solution at a concentration of 0.3 M. This concentration of NaOH is sufficient to inactivate the residual DMS.
(3) Freeze the filtered cells in liquid nitrogen. If the DNA is not prepared immediately, the frozen cells should be stored at –80°C.

B. Isolation of Genomic DNA

The cells recovered from the DMS treatments can be processed by almost any protocol for DNA isolation. Where possible, however, we prefer to recover genomic DNA from CsCl gradients. The protocol presented here has been adapted from several sources (e.g. see Shure *et al.*, 1983) and is also used to prepare untreated genomic DNA for the control reactions. The same method should be used to isolate both the control DNA and the DNA from the *in vivo* DMS treatments. The control DNA will provide a reference pattern of *in vitro* DMS cleavages that occurs when no proteins are associated with the DNA.

(1) Freeze the cells in liquid nitrogen, then grind to a fine powder in a prechilled, small electric coffee grinder.
(2) Drop the powdered cells directly into a 50 ml screw-top centrifuge tube containing 5 ml of TESE buffer (50 mM Tris, pH 8.0, 50 mM EDTA, 50 mM NaCl, 400 μg ml^{-1} of ethidium bromide, 2% *N*-lauroyl sarcosine), mix well with a glass rod and allow the slurry to come to room temperature, mixing occasionally.
(3) Centrifuge for 10 min at 15 000 rpm in a Beckman J-20 or equivalent rotor at 5°C, then remove the supernatant to a fresh 50 ml screw-top centrifuge tube. Add 1 g of CsCl per millilitre of supernatant.
(4) After the CsCl has dissolved completely, centrifuge as above, but at 20°C.
(5) Carefully pour off the liquid, leaving behind the pellet and the pellicle, and prepare for ultracentrifugation (additional ethidium bromide can be added at this point if it looks as if too much of it was lost to the pellet and pellicle). After ultracentrifugation, recover the banded genomic DNA, remove the ethidium bromide by butanol extraction and dialyse the DNA against TE as indicated in Maniatis *et*

al. (1989). Recovering the genomic DNA at concentrations of at least 0.1 mg ml^{-1} will enable the remaining steps to be handled in one microfuge tube per sample.

C. Preparation of the DNA for Genomic Sequencing

From this point, the control naked genomic DNA can be processed concomitantly with the genomic DNA isolated from the DMS treated cells. The control samples require a few extra steps.

(1) Digest 20 μg of both the *in vivo* DMS-treated and control DNA samples with the restriction enzyme of choice in a 200 μl volume.
(2) After digestion is complete, set the tubes containing the DNA from the *in vivo* treatments aside. To the control DNA only, add 1 μl of DMS, vortex to dissolve the DMS, and let it sit at room temperature for 1–2 min in a fume hood. Stop the reaction by extracting with phenol:chloroform:isoamyl alcohol (25:24:1).
(3) Once the *in vitro* DMS reactions have been stopped, the restricted DNAs from the *in vivo* DMS treatments can be phenol:chloroform:isoamyl alcohol extracted as well. Repeat organic extraction at least once for all samples, then precipitate the DNA from the aqueous phase with 80 μl of 7.5 M ammonium acetate and 500 μl of 95% ethanol for 15 min on ice. Centrifuge at top microfuge speed for 10 min and wash the pellets with 70% ethanol and then 95% ethanol.
(4) Resuspend the air-dried pellet in 50 μl of a freshly made 10% solution of piperidine in water. Incubate at 90°C for 20 min in a rack designed to keep the lids of the microfuge tubes tightly closed.
(5) Cool the tubes on ice, then add 300 μl of distilled water. Freeze in liquid N_2 and lyophilise the samples to dryness. Add 50 μl of distilled water and repeat the lyophilisation.
(6) Bring the samples up in 5 μl of sequencing dye, boil for 3 min, and load on a standard sequencing gel. The percentage of acrylamide can be varied (6% is usual), but the thickness of the gel must be in the 0.75 mm range. Thinner gels are easily overloaded by such large amounts of genomic DNA, and do not stand up to the rigours of electrotransfer very well.
(7) After electrophoresis, dismantle the gel plates from the sequencing apparatus and place on a large working surface. Carefully separate the gel plates so that the sequencing gel remains on the bottom plate.
(8) Cover the gel with a piece of plastic wrap and mark the plastic wrap with an outline of the area of the gel to be transferred. Cut through the plastic wrap and the gel along the indicated outline with a sharp scalpel, peel back the plastic, and overlay with a dry piece of Whatman 3MM paper. Smooth the paper over the gel surface and lift the gel–paper piece off of the plate. Lay the gel–paper piece with the gel side up on one surface of the electrotransfer sandwich support.
(9) Cut a piece of nylon membrane slightly larger than the size of the gel piece to be transferred, mark one surface of the membrane with an indelible pen, then wet the membrane in TBE (89 mM Tris, 89 mM boric acid, 2.6 mM EDTA).
(10) Wet the exposed surface of the gel with TBE. Carefully lay the wetted membrane on the gel with the marked surface of the membrane facing, and in contact with, the sequencing gel. Remove any trapped air bubbles and excess TBE with a roller.

(11) Place the other side of the electrotransfer sandwich support on the membrane and electrotransfer the DNA from the gel to the membrane for 1.5–2 h at 1.6–1.8 A.
(12) After transfer, dismantle the electrotransfer sandwich, peel the membrane away from the gel and cross-link the DNA to the membrane. We have used ultraviolet irradiation for GeneScreen membranes and 0.4 N NaOH for Hybond-N^+ membranes.

D. Generation of the Probe and Hybridisation

The only requirements for the probes used in genomic sequencing are that they be single stranded and of a high specific activity. DNA probes produced by synthesis from M13 cloned templates, and RNA probes produced by *in vitro* transcription systems, appear to be equally effective (see Section II.F). The hybridisation protocols are also as indicated in Section II.F.

Figure 2.3 shows an example of *in vivo* footprinting within the promoter region of the *Adh1* gene from maize.

E. Other Applications of Genomic Sequencing

Genomic sequencing is a powerful tool that facilitates strand-specific analyses of DNA at the nucleotide level and in situations where it is impossible, or inconvenient, to label a strand of DNA directly. Additional applications of the technique include detecting cytosine methylation (e.g. see Nick *et al.*, 1986; Saluz *et al.*, 1988), applying a higher degree of resolution to the analysis of nuclease sensitivity (e.g. see Thomas and Elgin, 1988) and the detection of non-B DNA structures. Our laboratory has used the latter application extensively to identify tracts of Z-DNA (Ferl and Paul, 1992), H-DNA (Lu and Ferl, 1992) and an altered DNA structure associated with a scaffold attachment region (Paul and Ferl, 1993).

A modification of genomic sequencing is the ligation-mediated polymerase chain reaction (LMPCR) genomic sequencing technique. This technique uses PCR to amplify the region of genomic DNA to be analysed (e.g. see Mueller and Wold, 1989). One of the drawbacks of LMPCR genomic sequencing is the possibility of introducing artifactual results through the PCR reactions. Recently, Garrity and Wold (1992) have addressed this problem and have introduced some modifications to their original procedure to improve the fidelity of the reactions. Detailed protocols for the application of LMPCR to plant systems can be found in Sorenson (1992) and Hammond-Kosack *et al.* (1993).

IV. IDENTIFICATION OF SARs

The genome of higher eukaryotes is organised into looped domains of chromatin by attachments to a nuclear scaffold. In addition to having an obvious function in the organisation of the genome, the SARs are also operationally important for the proper regulation of gene expression within topological domains. DNA topology itself influences gene expression in eukaryotes (e.g. see Dunaway and Ostrander, 1993; Lee and Garrard, 1991) and there are several examples of cohabitation of SARs and regulatory elements in

animal genes. In each of these cases SARs coincide with the boundaries of general nuclease sensitivity (Gasser and Laemmli, 1986; Steif *et al.*, 1989).

Identification of a SAR is usually approached in two different ways. The first establishes the general area of the SAR in its likely *in vivo* configuration, and is conducted in histone-depleted nuclei. The second approach is conducted *in vitro*, using end-labelled DNA fragments and isolated nuclear scaffolds. Procedures for the scaffold experiments in nuclei are after the methods of Mirkovitch *et al.* (1984), and have been successfully applied to plants (Hall *et al.*, 1991; Slatter *et al.*, 1991; Breyne *et al.*, 1992; Allen *et al.*, 1993; Avramova and Bennetzen, 1993; Paul and Ferl, 1993).

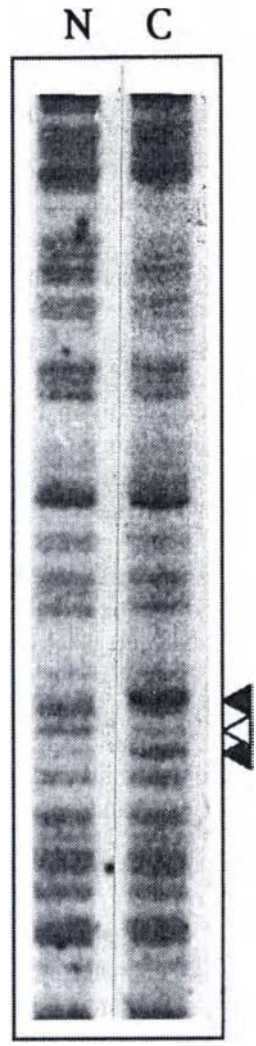


FIG. 2.3. *In vivo* footprinting in the promoter region of the maize *Adh1* gene. Examples of enhancements (solid triangles) and protections (open triangle) of DMS reactivity *in vivo* are indicated to the right of lane C (chromatin), and should be compared to the corresponding positions from the *in vitro* control lane N (naked DNA).

A. *In Vivo* Analysis of SARs: The Endogenous Assay

The endogenous assay is conducted in histone-depleted nuclei, in the hope that since the nuclei are gently lysed under physiological conditions the higher-order chromatin structure closely resembles the configuration of its native state. Thus, the SARs established through the endogenous assays are less likely to be an artefact of non-specific binding than the SARs defined by the *in vitro* analysis using isolated scaffold components (discussed in the next section).

(1) Prepare nuclei as described in Section II.A, resuspend in 2 ml of nuclei isolation buffer and set on ice.

(2) Set up a discontinuous Percoll gradient composed of four layers: a 2 M sucrose cushion in 25 mM Tris–10 mM MgCl, 80% Percoll, 60% Percoll and 40% Percoll. The various Percoll concentrations are made up in a solution of 0.44 M sucrose, 25 mM Tris and 10 mM $MgCl_2$ (after Luthe and Quatrano, 1980). Starting from the bottom, gently load 2 ml of each solution into a 30 ml Corex tube. Precentrifuging the gradient at 4000 ×*g* for 30 min in a swinging bucket rotor and 4°C will help set up the layers. After the pre-run, layer the 2 ml of nuclei reserved from the first step on to the gradient and centrifuge as described for the pre-run.

(3) Recover the nuclei with a glass pipette from the 80% Percoll–2 M sucrose interface, dilute with 10 ml of nuclei isolation buffer and respin at 4000 ×*g* for 10 min. Resuspend the pellet in 5 ml of nuclei resuspension buffer (50 mM Tris, pH 8, 5 mM $MgCl_2$, 10 mM β-ME, 20% glycerol). Divide into 1 ml aliquots for either storage or analysis. The subsequent steps are basically after Hall *et al.* (1991).

(4) Bring a 1 ml aliquot of nuclei in a microfuge tube to 1 mM $CuSO_4$ and let it stand at room temperature for 10 min. Spin down nuclei (4000 ×*g*, 10 min) and resuspend in 1 ml of histone extraction buffer (10 mM lithium diiodosalycilate acid (LIS), 5 mM Hepes, pH 7.4, 2 mM EDTA, 2 mM KCl, 0.1% digitonin, 0.2 mM PMSF, 5 μg ml^{-1} of aprotinin, 10 μM E-64 (*trans*-epoxysuccinyl-L-leucylamido(4-guanidino) butane) and incubate at room temperature for 10 min (*note*: the optimum concentration of LIS may need to be determined empirically for each plant system). Centrifuge at 4000 ×*g* for 10 min to pellet the histone-depleted nuclei (they will be very stringy and delicate at this stage).

(5) Gently resuspend in 1 ml of scaffold restriction–binding buffer (20 mM Hepes, pH 7.4, 20 mM KCl, 70 mM NaCl, 10 mM $MgCl_2$, 0.2 mM PMSF, 5 μg ml^{-1} of aprotinin, 10 μM E-64) and respin. Wash the histone-depleted nuclei three times in this manner to remove all traces of LIS, which will inhibit the subsequent restriction digests.

(6) Resuspend the final pellets in 600–1200 μl of restriction–binding buffer and divide into 200 μl aliquots. Digest with 250 units of restriction enzymes in two sequential 2 h incubation steps at 37°C. The restriction enzymes should be suitable to define large domains and facilitate indirect end labelling.

(7) Separate the DNA in the supernatant from the scaffold by centrifugation as above. Save the supernatant to another tube. Wash the pellets with 100 μl of the restriction–binding buffer and add the wash supernatant to the original supernatant. Restriction fragments containing a scaffold attachment site will sediment

with the scaffold pellet, whereas unattached loop fragments will remain in the supernatant.

(8) Resuspend the scaffold pellets in 300 μl of restriction–binding buffer and digest overnight with 600 μg ml^{-1} of Protease K (Sigma) in the presence of 1% SDS.

(9) The next day, extract both the digested pellets and the supernatants with phenol:chloroform:isoamyl alcohol (25:24:1) and precipitate the DNA with a 0.3 volume of 7.5 M ammonium acetate and two volumes of 95% ethanol.

(10) Resolve the recovered DNAs on 1% agarose and transfer to a nylon membrane.

(11) Probe and hybridise as described for the DNAase I assays. Regions of the gene (defined by restriction fragments) which are attached to the scaffold *in vivo* will appear as a darker band on the autoradiograph in the lanes containing DNA from the pellet fractions as compared to lanes from the supernatant fractions.

B. *In Vitro* Analysis of SARs: The Exogenous Assay

The *in vitro* experiments can be used to complement the endogenous assay. These experiments allow the study of the interaction between scaffolds and SARs from different cells that can belong to the same or to different organisms. Although the possibility exists of observing artefactual binding of DNA to the isolated scaffolds, the experiments can be controlled through the use of non-specific competitors.

(1) Divide histone-depleted nuclei that have been prepared as indicated above (steps (1)–(5)) and resuspend in restriction–binding buffer into 200 μl aliquots. Digest with 250 units of restriction enzymes that will not interfere with the subsequent incubation in the presence of end-labelled exogenous DNA, as the restriction enzymes are not inactivated before the digested scaffolds are incubated with the labelled DNA fragments. Digest in two sequential 2 h incubation steps at 37°C.

(2) Add 0.05 μg of end-labelled DNA fragments containing the region of interest plus 20 μg of non-specific competitor DNA to a 200 μl aliquot, and incubate at 37°C for 2 h.

(3) Centrifuge at 4000 ×*g* for 10 min to pellet the scaffolds and attached DNA fragments, saving the supernatant to another tube. Wash the pellets with 100 μl of the restriction–binding buffer minus the protease inhibitors, and add the wash supernatant to the original supernatant. As with the endogenous assay, fragments containing a scaffold attachment site will sediment with the scaffold pellet, whereas fragments unable to bind to the scaffold will remain in the supernatant.

(4) Resuspend the scaffold pellets in 20 mM Tris, pH 8, 1% SDS, 20 mM EDTA and 100 μg of proteinase K. Incubate overnight at 37°C.

(5) Purify the DNA and analyse as described in the previous section (steps (8)–(11)).

V. SUMMARY

The protocols we have described in this chapter illustrate a variety of approaches that can be used to elucidate plant chromatin structure at different levels of organisation. The lower levels deal with the packaging of DNA in nucleosomes and the interactions between DNA and non-histone proteins. The higher orders of chromatin organisation result from the association of the chromatin fibre with the nuclear scaffold.

Investigations utilising nucleases such as DNAase I, micrococcal nuclease or restriction enzymes are useful in characterising aspects of the lower levels of chromatin organisation. The protocols outlined in Section II can be used to identify the nucleosome repeat length, the presence or absence of nucleosomes within a specific region of DNA, regions of DNA where the nucleosomes are preferentially positioned in specific places (phasing) and regions of DNA that interact with non-histone proteins. The points of interaction between the chromatin and non-histone proteins, such as *trans*-acting transcription factors, can be identified at the nucleotide level with the *in vivo* footprinting and genomic sequencing techniques outlined in Section III.

The endogenous and exogenous assays described in Section IV facilitate the characterisation of higher orders of chromatin structure. These assays define scaffold attachment sites within a DNA region, giving a researcher insights into how that particular region is organised with regard to the loops of chromatin.

Each level of chromatin structure is interconnected to package the genome in such a way that any given portion can be replicated or transcribed in a highly regulated way. Thus, an understanding of chromatin structure is an integral part of understanding gene regulation and genome organisation.

REFERENCES

Allen, G. C., Hall, G. E., Childs, L. C., Weissinger, A. K., Spiker, S. and Thompson, W. F. (1993). *Plant Cell* **5**, 603–613.

Ashraf, M., Vasil, V., Vasil, I. K. and Ferl, R. J. (1987). *Mol. Gen. Genet.* **208**, 185–190.

Avramova, Z. and Bennetzen, J. L. (1993). *Plant Mol. Biol.* **22**, 1135–1143.

Breyne, P., Van Montagu, M., Depicker, A. and Gheysen, G. (1992). *Plant Cell* **4**, 463–471.

Bryan, P. N., Hofstetter, H. and Birnstiel, M. L. (1981). *Cell* **27**, 459–466.

Church, G. M. and Gilbert, W. (1984). *Proc. Natl Acad. Sci. USA* **81**, 1991–1995.

Clark, D. J. and Kimura, T. (1990). *J. Mol. Biol.* **211**, 883–896.

Conconi, A. and Ryan, C. A. (1993). *J. Biol. Chem.* **268**, 430–435.

Dellaporta, S. L., Wood, J. and Hicks, J. B. (1983). *Plant Mol. Biol. Rep.* **1**, 19–23.

Dunaway, M. and Ostrander, E. A. (1993). *Nature* **361**, 746–748.

Ehrlich, S. D., Bertazzoni, U. and Bernardi, G. (1973). *Eur. J. Biochem.* **40**, 143–147.

Elgin, S. C. (1981). *Cell* **27**, 413–415.

Ferl, R. J. (1985). *Mol. Gen. Genet.* **200**, 207–210.

Ferl, R. J. and Nick, H. N. (1987). *J. Biol. Chem.* **262**, 7947–7950.

Ferl, R. J. and Paul, A.-L. (1992). *Plant Mol. Biol.* **18**, 1181–1184.

Galcheva-Gorgova, Z. I., Marinova, E. I. and Koleva, S. T. (1988). *Plant Cell Environ.* **11**, 819–825.

Garrity, P. A. and Wold, B. (1992). *Proc. Natl Acad. Sci. USA* **89**, 1021–1025.

Gasser, S. M. and Laemmli, U. K. (1986). *Cell* **46**, 521–530.

Gasser, S. M. and Laemmli, U. K. (1987). *TIG* **3**, 16–22.

Gasser, S. M., Laroches, T., Falquet, J., Boy de la Tour, E. and Laemmli, U.K. (1986). *J. Mol. Biol.* **188**, 613–629.

Hall, G., Allen, G. C., Loer, D. S., Thompson, W. F. and Spiker, S. (1991). *Proc. Natl Acad. Sci. USA* **88**, 9320–9324.

Hammond-Kosack, M. C. U., Holdsworth, M. J. and Bevan, M. W. (1993). *EMBO J.* **12**, 545–554.

Lee, M. S. and Garrard, W. T. (1991). *Proc. Natl Acad. Sci. USA* **88**, 9675–9679.

Lu, G.-H and Ferl, R. J. (1992). *Plant Mol. Biol.* **19**, 715–723.

Luthe, D. S. and Quatrano, R. S. (1980). *Plant Phys.* **65**, 305–308.

Maniatis, T., Fritsch, E. F. and Sambrok, J. (1989) 'Molecular Cloning: A Laboratory Manual'. Cold Spring Harbor, New York.

Maxam, A. M. and Gilbert, W. (1980). *Methods Enzymol.* **65**, 499–560.
Mirkovitch, J., Mirault, M.-E. and Laemmli, U. K. (1984). *Cell* **39**, 223–232.
Mueller, P. R. and Wold, B. (1989). *Science* **246**, 780–786.
Murray, M. G. and Kennard W. G. (1984). *Biochem.* **23**, 4225–4232.
Nick, H. and Gilbert, W. (1985). *Nature* **313**, 795–798.
Nick, H., Bowen. B., Ferl, R. J. and Gilbert, W. (1986). *Nature* **319**, 243–246.
Paul, A.-L. and Ferl, R. J. (1991). *Plant Cell* **3**, 159–168.
Paul, A.-L. and Ferl, R. J. (1993). *Plant Mol. Biol.* **22**, 1145–1151.
Paul, A.-L., Vasil, V., Vasil, I. K. and Ferl, R. J. (1987). *Proc. Natl Acad. Sci. USA* **84**, 799–803.
Saluz, H. P., Feavers, I. M., Jiricny, J. and Jost, J. P. (1988). *Proc. Natl Acad. Sci. USA* **85**, 6697–6700.
Schulze-Lefert, P., Dangl, J. L., Becker-André, M., Hahlbrock, K. and Schulz, W. (1989). *EMBO J.* **8**, 651–656.
Shure, M., Wessler, S. and Fedoroff, N. (1983). *Cell* **35**, 225–233.
Slatter, R. E., Dupree, P. and Gray, J. C. (1991). *Plant Cell* **3**, 1239–1250.
Sorenson, M. B. (1992). *Proc. Natl Acad. Sci. USA* **89**, 4119–4123.
Steif, A., Winter, D. M., Statling, W. H. and Sippel, A. E. (1989). *Nature* **341**, 343–345.
Steinmuller, K. and Apel, K. (1986). *Plant Mol. Biol.* **7**, 87–94.
Stewart, A. F., Reik, A. and Schutz, G. (1991). *Nucl. Acids Res.* **19**, 3157.
Thomas, G. H. and Elgin S. C. R. (1988). *EMBO J.* 7, 2191–2201.
Thomas, G. H., Siegfried, E. and Elgin, S. C. R. (1985). *In* 'Chromosomal Proteins and Gene Expression' (Reeck, G., Goodwin, G. and Puigdomenech, P., eds). 77–101. Plenum Press, New York.
Wood, W. I. and Felsenfeld, G. (1982). *J. Biol. Chem.* **257**, 7730–7736.

3 Molecular Genetic Analysis: Chromosome Walking, Gene Tagging and RFLP Analysis

ZOE A. WILSON

Department of Life Science, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

METHODS IN PLANT BIOCHEMISTRY Vol. 10b
ISBN 0-12-461021-8

I. INTRODUCTION

The fields of plant genetics and molecular biology have been rapidly developing over recent years, to the extent that genes can now be identified based on mutagenesis, mapped by the use of genetic, molecular and physical positioning and then cloned based on this map position. This means that the study of complex developmental pathways and the analysis of regulatory genes is now feasible. For some plant systems, such as *Arabidopsis thaliana*, the task of positionally based cloning is made easier by the small size of its genome (~100 Mb) and low levels of repetitive DNA (Pruitt and Meyerowitz, 1986). However, similar cloning exercises are now being attempted for plants with large genomes and high levels of repetitive DNA, such as maize (Edwards *et al.*, 1992), sorghum (Woo *et al.*, 1995) and barley (Kleine *et al.*, 1993). This chapter will describe various techniques and methods for the cloning of genes based on mutant phenotype, concentrating primarily on genome mapping and positional cloning by chromosome walking and T-DNA tagging using *A. thaliana* as a model system.

A. Approaches for Gene Cloning

Two main strategies have been used to clone genes based on phenotype:

(1) *Chromosome walking*, whereby a map position is established for a gene and overlapping clones are linked together until the region containing the gene has been identified in contiguous clones; the region is then delimited and the gene identified, usually by complementation analysis.
(2) *DNA tagging*, where function is disrupted by the incorporation of a 'DNA tag' into the coding or promoter sequence of a gene. This approach can be subdivided into T-DNA tagging using tags generated from the Ti plasmid of *Agrobacterium tumefaciens* and into tagging using transposons.

B. Chromosome Walking

Chromosome walking involves the successive linking of a series of cloned DNA fragments. The walk is started from a predetermined probe (e.g. a restriction fragment length polymorphism, random amplified polymorphic DNA (RFLP, RAPD) clone), which from previous mapping experiments is known to be in the vicinity of the gene of interest (see Section III). The starting probe is used to screen a library of overlapping clones. Positive clones which hybridise to it can then be used as probes to progressively link contiguous clones in the region. By repeating this process one can sequentially screen for sequences distal to the original probe and effectively 'walk' along the chromosome until the gene of interest is reached (Fig. 3.1). Such map-based cloning has been pioneered in a number of animal systems, namely *Drosophila* (Bender *et al.*, 1983), *Caenorhabditis elegans* (Papp *et al.*, 1991) and the human genome project (Green and Olson, 1990; Butler *et al.*, 1992;

Ballabio, 1993). There have, however, been only a limited number of reports of positional cloning in higher plants, and those predominantly from *A. thaliana* (Arondel *et al.*, 1992; Giraudat *et al.*, 1992; Martin *et al.*, 1993; Bent *et al.*, 1994; Mindinos *et al.*, 1994; Pepper *et al.*, 1994). Although a number of laboratories are well advanced with chromosome walks and linking contiguous regions in the *Arabidopsis* genome (Hwang *et al.*, 1991; Putterill *et al.*, 1993; Vijayraghavan *et al.*, 1995), the lengthy time taken to clone genes by chromosome walking is a reflection of the fact that walks tend to be extremely labour-intensive. This is partly due to the unknown orientation of each step, meaning that steps have to be carried out in both directions to effectively walk one way and that many of the techniques for walking are still in the developmental stages. In addition, difficulties have frequently been encountered with individual YAC clones. In the case of existing *Arabidopsis* yeast artificial chromosome (YAC) libraries, Putterill *et al.* (1993) detailed that approximately 40% of 67 YAC clones analysed exhibited problems; these were either deleted in part of the plant DNA, chimeric or contained repetitive DNA at one of their termini. These problems have been identified in most of the *Arabidopsis* YAC libraries analysed to date (Schmidt *et al.*, 1994). Levels of chimeric clones vary between different YAC libraries, and it is thought that the incidence of chimerism may be in part a consequence of the number of size selection steps prior to library construction, since libraries with two selection stages appear to have a lower incidence of chimeric clones (Grill and Sommerville, 1991; Creusot *et al.*, 1994). Many of the technical problems first encountered in chromosome walks have now been circumvented, and, in time, positional cloning, at least in *Arabidopsis*, will become easier still, as the need to identify clones is diminished with the completion of linking and ordering genomic libraries.

C. Requirements for Commencing a Chromosome Walk

1. Mapping

An essential prerequisite for chromosome walking is the accurate mapping of the gene in question and the identification of flanking markers. In the case of *A. thaliana* a well-developed linkage map (Koornneef, 1986), two extensive RFLP maps (Chang *et al.*, 1988; Nam *et al.*, 1989), which have been combined to produce an integrated map (Hauge *et al.*, 1993), and an RAPD (Reiter *et al.*, 1992) and recombinant inbred maps (Lister and Dean, 1993) have made this a relatively easy task. This means that 'starting probes' can be readily obtained in the vicinity of any gene of interest regardless of map position.

2. Genetic markers/recombinants

Recombinants between two different populations, which give easily definable RFLPs, are essential to aid in preliminary marker positioning and in walk orientation. Phenotypic markers are needed in the vicinity of the gene of interest to aid in the screening and identification of recombinants that flank the gene. The recombinants can be identified by the transfer and segregation of mutant phenotypes to the resultant progeny. This allows RFLPs to be fine-mapped to the gene of interest and for end-probes and markers to be closely positioned to identify the direction and extent of the walk. The greater the number of recombinants the closer and more accurate the mapping.

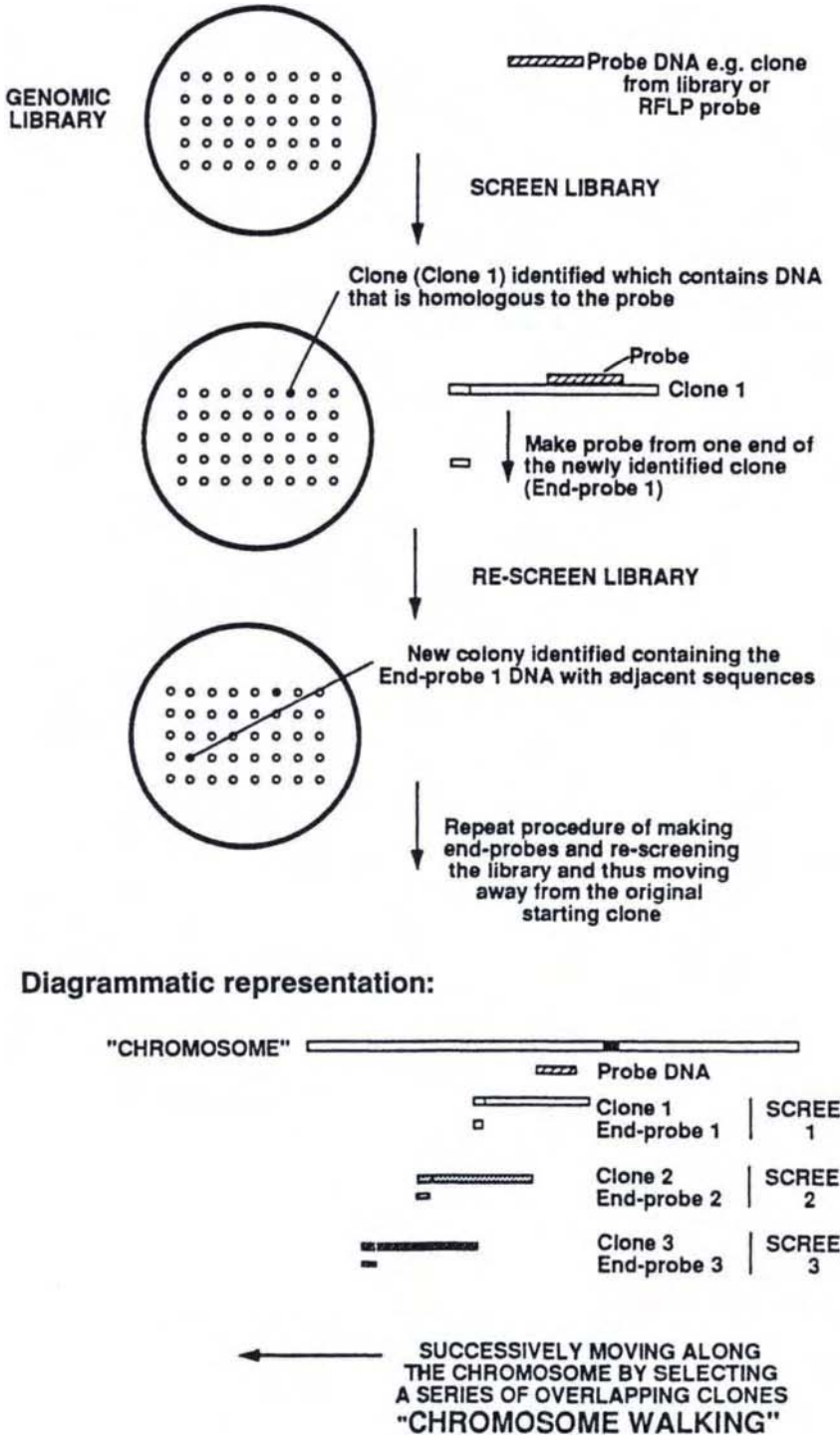


FIG. 3.1. Schematic representation of chromosome walking.

3. Closely linked molecular markers

It is very important that the starting point of any walk be extremely close (if possible < 1 cM) to the gene of interest, to minimise the region that needs to be linked; current estimations suggest that each step in a walk is very small, covering approximately 50 kb (R. Schmidt, personal communication). It is also important to have other markers which are known to be on the alternative side of the gene to the starting probe. This allows identification of when the end of the walk has been reached. It is necessary before embarking on a walk to fully establish the closest markers and the distances to be covered. Fine-scale RFLP mapping using recombinants in this region can determine the closest markers and the relative genetic distances to start a walk. Many of the RFLP markers that have already been mapped are not sufficiently accurately positioned to commence a walk and therefore should not be used without prior detailed mapping. It is also advisable to establish the actual physical distances to be linked; these can be determined by restriction mapping using rare-cutting restriction enzymes (Ganal *et al.*, 1989; Bancroft *et al.*, 1992; Segal *et al.*, 1992; Schumacher *et al.*, 1995).

4. Overlapping genomic library

A good-quality overlapping genomic library, which contains at least three genome equivalents, is required for the walk. The libraries that have been predominantly used for chromosome walking are YAC libraries. Frequently, more than one library has been used in screening to ensure that a sufficient number of clones are identified at each stage. This permits greater assurance of the walk and linking integrity.

II. MAPPING STRATEGIES

Prior to cloning, the map position of a mutation must be identified. This can be conducted by classical genetic approaches or molecular genetic analysis. Mapping strategies are well advanced for *A. thaliana*; for classical linkage analysis a range of multiple marker lines have been constructed (Koornneef and Stam, 1987). The mutants can be crossed to the multimarker lines for linkage analysis, and the resultant population can be scored for all the markers to identify recombinants and locate approximate map positions to the top or bottom arms of each chromosome. The mutants can then be crossed to multimarker lines from the appropriate chromosome to establish detailed map positions (Koornneef and Stam, 1992). Mapping can also be facilitated by the use of recombinant inbred lines (Reiter *et al.*, 1992; Lister and Dean, 1993), RFLPs (Tanksley *et al.*, 1989; Fabri and Schäffner, 1994), RAPD markers (Giovannoni *et al.*, 1991; Klein-Lankhorst *et al.*, 1991; Devos and Gale, 1992), polymerase chain reaction (PCR) primers to specific genes or intron–exon junctions (Weining and Langridge, 1991) and by PCR using co-dominant cleaved amplified polymorphic sequences (CAPS)–PCR (Konieczy and Ausubel, 1993).

III. RFLP ANALYSIS

Genetic and RFLP maps are based upon recombination frequency in large data sets to determine linkage by two-point, three-point or *n*-point analysis. The order and distances

between markers will depend on the size of the population used for the analysis. Other factors that will affect this are the genetic backgrounds and the types of markers used (Helentjaris *et al.*, 1986; Chang and Meyerowitz, 1991; Koornneeff and Stam, 1992).

A. Generation of Recombinants for RFLP Mapping

The mutant that is being studied needs to be crossed to alternative ecotypes to allow the identification of RFLPs flanking the gene. Crossing to lines which carry, preferably reliable and easy to score, phenotypic mutations in this region will assist in the identification of recombinants. This can be done in a number of ways, which include:

(1) The construction of a multimarker line which carries mutations on either side of the gene in question. This line can then be crossed to alternative ecotypes.
(2) The crossing of the mutant in question with an alternative ecotype that carries markers which will flank the gene. In this case recombinants will be selected that carry the flanking markers.

These crosses are most effective when conducted in a 'test cross'. This involves the crossing of a homozygous mutant to a heterozygous 'mutant/alternative ecotype', since this means that recombinants can be identified in the F_1 generation and are not masked by wild-type alleles.

Test crosses can be difficult due to the requirements for double recessive genotypes and specific crossings. In the case of *Arabidopsis* large numbers of selfed seed can be obtained easily; therefore, most recombinant selection occurs in the F_2 and F_3 generations. After the initial cross, the F_1 is allowed to self, and the F_2 screened for recombinants. Since limited amounts of plant material can be obtained from an F_2 plant, F_3 populations derived from single F_2 lines are more frequently used for further analysis. These can be grown in soil or *in vitro*. The F_3 population must be completely collected to ensure that the true genotype of the F_2 is reflected.

An alternative to selfing the F_1 generation is to conduct a back-cross, which can then be used to establish a recombinant line as with the 'selfing strategy'. In all cases the selection strategy must be clearly defined prior to crossing and the markers chosen such that identification can be easily carried out regardless of the overall ecotypic background of the plant. This may be difficult to ensure for some mutations, and may need further investigation during recombinant generation. The direction of the initial cross also will affect recombination frequency (Vizir and Korol, 1990), as will whether the mutations are in the coupling or repulsion phase.

The lines chosen for RFLP mapping will depend on the plant species. They should be sufficiently diverse to enable the identification of RFLPs but sufficiently close to facilitate crossing. In the case of *Arabidopsis*, Landsberg *erecta* has been commonly used for mutagenesis and the ecotypes Wassilewskija, Columbia, Estland and Niederzenz have been used extensively for constructing RFLP and RAPD maps. Other ecotypes from distant geographical origins may also be useful since they may be more divergent and therefore easier for polymorphism identification. When selecting lines for RFLP analysis it can also be beneficial to select lines that have already been used for mapping purposes since the degree of polymorphism has been established and the combining of mapping data is easier.

An alternative approach to mapping is to use recombinant inbred (RI) populations. These have been successfully used to map markers to the mouse, maize (Burr and Burr, 1991) and *Arabidopsis* genomes (Reiter *et al.*, 1992; Lister and Dean, 1993). In RI populations, plants are crossed to produce an F_1, then individual F_1 plants are taken through seven generations of selfing. Each generation is taken from a single seed descent and repeated until the level of heterozygosity is minimal (F_8 or greater). These lines therefore have stably inherited multiple breakpoints between the original parents, and can be used to accurately map multiple molecular markers.

B. DNA Isolation for RFLP Analyses

1. Growth of plants/DNA preparation

Once selection of recombinants has been conducted, bulking of plants for DNA isolation (usually an F_3 population from an F_2 individual) can be conducted in soil or *in vitro*. The following protocol has been successfully used to grow large numbers of plants for RFLP analysis in *Arabidopsis*.

2. Arabidopsis in vitro *plant growth*

(1) Seeds are sterilised in 70% (v/v) ethanol (2 min), followed by 5% (v/v) sodium hypochlorite, 0.5% (w/v) – (SDS) (15 min), and then washed in at least three changes of sterile distilled H_2O.
(2) The seeds are then plated on to medium – as used for the initial stages of root culture (Czako and Marton, 1992) – which contains 1 × MS salts, 3% (w/v) sucrose, 0.018% (w/v) KH_2PO_4, vitamin B_1 0.1 mg ml^{-1}, vitamin B_6 0.01 mg ml^{-1}, glycine 0.02 mg ml^{-1}, nicotinic acid 0.01 mg ml^{-1}, folic acid 0.005 mg ml^{-1}, biotin 0.01 mg ml^{-1}, no growth regulators and 0.8% (w/v) agar, and kept at 25°C 2 000 lux, continuous illumination.
(3) After 1 week the seedlings should have germinated and started development. They are then transferred to 250 ml flasks containing 25 ml of the same culture medium without agar. Any contaminated seedlings are discarded.
(4) The flasks are then maintained at 2000 lux, 70–80 rpm, for 1–2 weeks, or until sufficient growth has occurred for harvesting. Different recombinants/genetic backgrounds tend to exhibit differential rates of development in liquid culture.
(5) Plant material is harvested and dried on 3MM (Whatman) paper, prior to freezing in liquid N_2 and storing at –80°C. It is necessary to ensure that plant material is quite dry before freezing, otherwise large tissue–ice clumps form, which makes grinding for extraction very difficult.

3. DNA isolation

The method described here is based on that of Dellaporta *et al.* (1983). The reader is also referred to Chapter 1 in this volume.

(1) Using liquid N_2, freeze and grind 1–3 g of plant material.

(2) Transfer the ground material to a 50 ml tube containing 15 ml of extraction buffer (100 mM Tris–HCl, pH 8, 50 mM EDTA, 500 mM NaCl, 10 mM β-mercaptoethanol) at 65°C. Add 1 ml of 20% (w/v) SDS. Incubate at 65°C for 10 min.
(3) Add 5 ml of 5 M potassium acetate, and gently mix the tube by inversion. Incubate at 0°C for at least 20 min.
(4) Centrifuge the tubes (4°C, 4500 rpm) for 10 min, then pour the supernatant through muslin into a fresh tube containing 10 ml of isopropanol. Incubate at –20°C for at least 20 min.
(5) Centrifuge (4500 rpm) for 10 min. Carefully pour off the supernatant and drain the pellet. Redissolve the pellet in 3 ml of TE (10 mM Tris–HCl, pH 8, 1 mM EDTA).
(6) Extract using 1 ml of equilibrated phenol (10 mM Tris–HCl, pH 8): chloroform:isoamyl alcohol (25:24:1).
(7) Precipitate the DNA by adding sodium acetate to 0.3 M and 2.5 volumes of ethanol. Centrifuge (13 000 rpm) for 10 min.
(8) Redissolve the DNA pellet in 1 ml of TE. Add 10 μl of RNAase (10 mg ml^{-1}) and incubate at 37°C for 15 min.
(9) Add an equal volume of cetyltrimethylammonium bromide (CTAB) extraction buffer (0.2 M Tris–HCl, pH 8.0, 0.05 M Na_2EDTA, 2.0 M NaCl, 2.0% (w/v) CTAB) and incubate at 65°C for 15 min.
(10) Cool to room temperature and extract with 1 ml of chloroform:isoamyl alcohol (24:1).
(11) Precipitate the DNA by adding sodium acetate to 0.3 M and 2.5 volumes of ethanol. Centrifuge (13 000 rpm) for 10 min. Resuspend the pellet in 400 μl of TE.
(12) Transfer the suspended DNA to a microcentrifuge tube and precipitate it with an equal volume of isopropanol. Wash the pellet with 70% (v/v) ethanol. Resuspend the DNA in 300 μl of TE. The DNA can be accurately measured by fluorometry using Heochst 33258 (Hoeffer TK100 dedicated mini fluorometer).

C. Detection of RFLPs

RFLPs are identified by restriction digestion of genomic DNA, size separation by gel electrophoresis, probing using a labelled marker and identified by autoradiography or non-isotopic detection. Initially when working with any mapping population or new markers it is essential to identify which restriction enzymes will give polymorphisms in the parental populations. The number of restriction enzymes needed to identify polymorphisms will depend on the extent of genetic diversity between the two mapping populations. The amount of prescreening required will also depend on whether an RFLP map is to be constructed for the whole genome or whether polymorphisms are required in a particular region using specific probes.

1. RFLP Presceening in Arabidopsis

(1) Cut 1–2 μg of genomic DNA from both parental RFLP mapping lines using a range of six base pair restriction enzymes (e.g. *Sty*I, *Eco*RI, *Eco*RV, *Bam*HI, *Hind*III, *Bsc*I; ~ 10 U μg^{-1}).
(2) Separate by gel electrophoresis (0.5 × TBE; 0.7–0.8% (w/v) agarose).

(3) Pretreat the gel with 0.2 M HCl (15 min), followed by 1.5 M NaCl, 0.5 M NaOH (2 × 15 min) and 1.5 M NaCl and 0.5 M Tris–HCl, pH 7.5 (30 min). Transfer to Hybond-N$^+$ nylon membrane (Amersham) using 10 × SSC (20–24 h). The DNA should be fixed by ultraviolet irradiation (150 mJ-wet membrane; BioRad GeneLinker) and baked for 2 h at 80°C.
(4) Prehybridise the filters for 0.5–1 h at 65°C in 1 M NaCl, 10% (w/v) dextran sulfate and 1% (w/v) SDS.
(5) Screen with RFLP markers in the region of interest. Probes can be constructed by random priming, nick translation or RNA transcription (Sambrook *et al.*, 1989). Hybridise (overnight (o/n) 65°C) in the same solution with the addition of 50–100 μl of 10 mg ml^{-1} of herring sperm (Sigma) per millilitre of hybridisation solution.
(6) For homologous probes, wash the filters three times (15 min, 65°C) with 2 × SSC buffer; and 1% (w/v) SDS, and twice (20–30 min, 65°C) with 0.1 × SSC buffer and 0.1% (w/v) SDS. Rinse with 2 × SSC buffer and expose to X-ray film (Kodak).
(7) Restriction enzymes should be selected for RFLP mapping that give clearly discernible polymorphisms in the two mapping populations and can be identified in a heterozygous state.

D. Mapping

Once the restriction enzymes have been selected for RFLP mapping, DNA from the recombinant populations can be analysed. On each gel, parental controls should be run to aid in polymorphism identification. The genotypes of the recombinants can then be established based on both parental RFLP patterns. These data can then be used in computer-aided mapping analysis (MAPMAKER; Lander *et al.*, 1987), in small-scale analyses to identify markers close to genes, or to establish the directions of chromosome walks. The accuracy of mapping experiments depends on the sizes of populations that are used (Koornneef and Stam, 1992). In the case of the *Arabidopsis* genome, if a region of 20 cm is to be mapped, 100 recombinants in this region should allow a recombination event to occur approximately every 30 kb. This assumes that recombination is random between the two markers; however, this is known to not be the case and will depend greatly on gene location and the physical parameters of the chromosome around the region in question.

Once markers have been identified that lie close to the gene of interest, chromosome walks can be commenced to clone the gene based on position.

E. Alternatives to RFLPs

An alternative type of molecular marker that has been used to identify closely linked probes for positional cloning is amplified fragment length polymorphisms (AFLPs) (Keygene, The Netherlands; Perkin Elmer). This approach involves selectively amplifying different loci across the genome and looking for polymorphisms between pools of related individuals, which vary in the presence of a specific locus. Both DNA pools are digested with a rare and a frequent cutting enzyme, followed by the ligation of adaptors to both ends. PCR primers which are based on the adaptors, the linking 'restriction site' sequence and a random 3 bp sequence (at the 3′ end) are then used. This random

sequence alters the specificity of the PCR and limits amplification. By the use of different primers with various 'random' sequences, different genomic fragments will be amplified. The AFLPs can then be analysed using sequencing gels. Approximately 100–150 fragments representing different loci can be analysed on one gel, thus permitting the simultaneous study of a large number of loci (Thomas *et al.*, 1995). Bands that appear polymorphic between the two pools can be isolated and tested for linkage analysis to the gene being cloned. The use of such large numbers of markers achievable using AFLPs means that positional cloning can be carried out by targeted 'chromosome landing' (Tanksley *et al.*, 1995) rather than chromosome walking.

IV. CHROMOSOME WALKING

A. Walking Libraries

Libraries that are to be used for walking must contain in excess of three overlapping haploid genomes. Most walking strategies use libraries containing large inserts, so that each step of a walk can be maximised. The favoured libraries have tended to be YACs, since they can contain very large inserts (up to 1 Mb), although the P1 bacteriophage cloning system (Sternberg, 1990; Ioannou *et al.*, 1994) has been used to clone fragments of the order of 100–300 kb. Another type of library that has recently been used, often as an additional tool to YAC libraries, is that of bacterial artificial chromosomes (BACs). The BAC system uses vectors based on the *Escherichia coli* fertility (or F-factor) plasmid to construct artificial chromosomes that are propagated in *E. coli*. These are able to maintain up to 1 Mb of DNA (Shizuya *et al.*, 1992), and since they are maintained at low copy number they appear to be relatively stable; they also have the advantage that, unlike YACs, they can be easily isolated away from the host chromosomal DNA (Kim *et al.*, 1992).

The main characteristics of YAC vectors (Burke *et al.*, 1987) are that they can be maintained in *E. coli* as plasmids and in *Saccharomyces cerevisiae* as linear chromosomes. As chromosomes they are inherently more stable than plasmids, and can maintain a significantly larger DNA insert. A series of pYAC vectors have been constructed based on Burke *et al.* (1987), with various cloning sites and some with RNA polymerase-binding sites which flank them (Schlessinger, 1990; Charlieu *et al.*, 1991) (Table 3.1). In most cases the cloning site falls within the ochre suppressor *SUP4*, which serves as an interruptable marker; clones maintained in an *ade2–ochre* host which are expressing the suppressor form white colonies, whilst those with the marker interrupted by an insert produce red colonies. The *SUP4* marker does appear to be leaky; therefore, some YAC clones can appear white despite the presence of stable inserts.

The pYAC vectors are highly chimeric in origin, containing: a bacterial selectable marker (ampicillin resistance) and origin of replication; yeast autonomous replicating sequence (ARS), selectable markers and centromere; and telomeric sequences from *Tetrahymena* macronuclear rDNA. Prior to cloning, the pYAC vectors require restriction digestion (the specific enzyme depends on the vector used) to generate three fragments, a left and right arm and a 'linking fragment', which is later lost (Fig. 3.2). Each YAC arm carries a yeast-selectable marker (*TRP1*, left/centromere arm; *URA3*, right arm) to ensure that YAC clones contain both arms and are stably maintained. Therefore, after ligation and transformation into yeast, clones are selected based on their red coloration and ability to grow on selection medium (Kiwi medium) (Ausubel *et al.*, 1988; Putterill *et al.*, 1993).

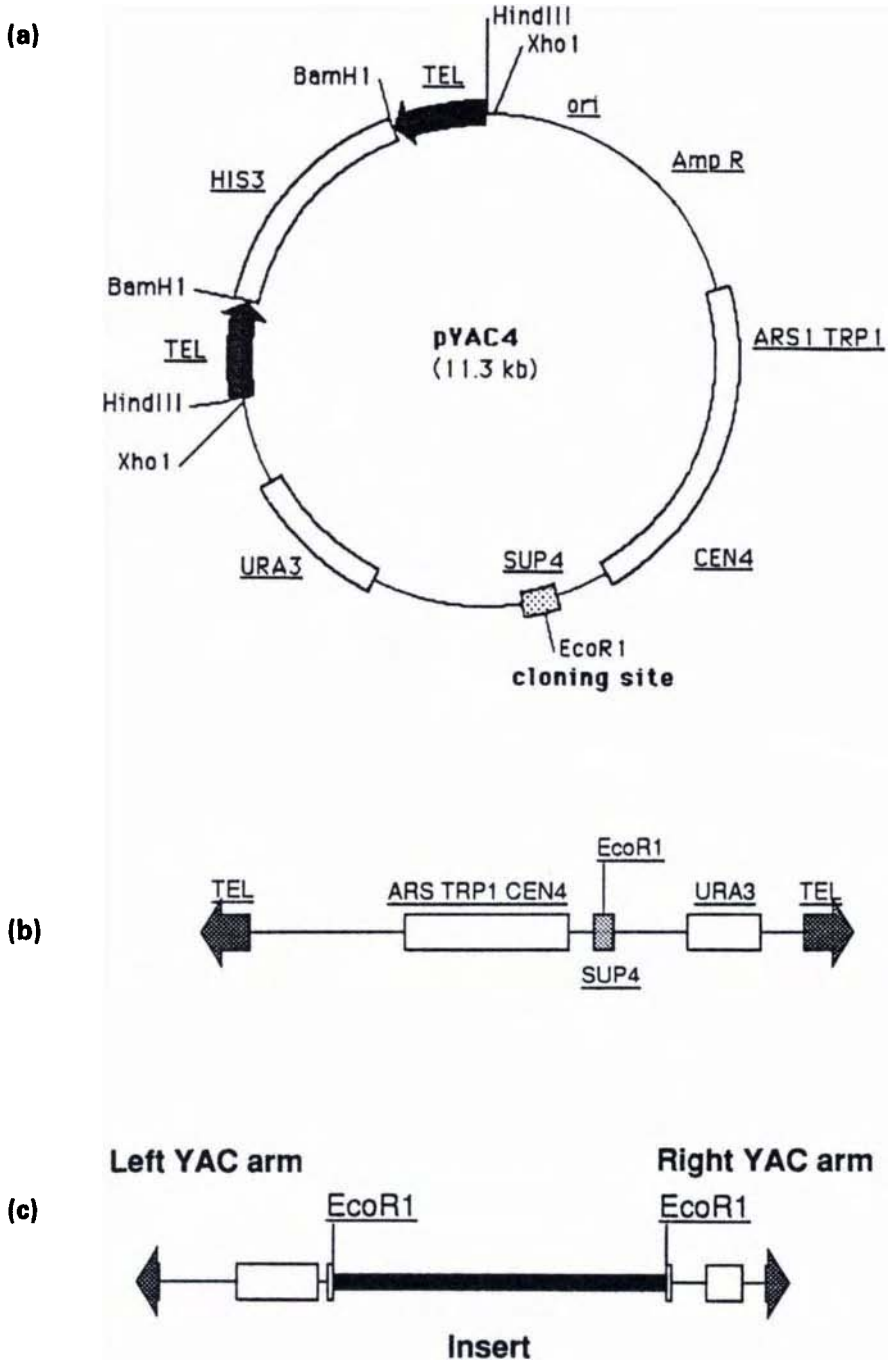


FIG. 3.2. Structure of the pYAC vector and the cloning strategy to form a linear artificial chromosome. (a) pYAC4 vector. (b) pYAC4 cut with *Bam*HI to remove *HIS3* linker and generate a linear artificial chromosome. (c) Linear artificial chromosome with an insert incorporated into the *SUP4* gene after digestion with *Eco*RI.

TABLE 3.1. pYAC vectors.

Vector	Cloning site	Characteristics	Reference for vector (e.g. YAC libraries)
pYAC2	*Sma*I	Blunt ends	Burke *et al.* (1987)
pYAC3	*Sna*BI	Cohesive ends	Ward and Jen (1990)
pYAC4	*Eco*RI	–	Ecker (1990)
pYAC41	*Bam*HI	T3-R T3-L	Grill and Sommerville (1991)
pYAC45	*Bam*HI	T7-R T3-L	Grill and Sommerville (1991)
pYAC5	*Not*I	–	Burke *et al.* (1987)
pJS9 and pJS98	Various	T7-L T7-R	Shero *et al.* (1991)
pYAC4 Neo(NRN)	*Eco*RI	*Not*1 excision site T3-L T7-R G418 resistance	Charlieu *et al.* (1991)

Although the techniques for YAC library construction and handling have been pioneered in animal molecular genetics, namely with the mouse and human genome projects, YAC libraries have now been constructed for a variety of plant species, for example *Arabidopsis* (Ecker, 1990; Ward and Jen, 1990; Grill and Sommerville, 1991), carrot (Guzman and Ecker 1988), maize (Gupta and Hoo, 1991; Edwards *et al.*, 1992), rice (Umehara *et al.*, 1995), sugar beet (Delfavero *et al.*, 1994) and tomato (Martin *et al.*, 1992). The upper cloning range of YAC vectors has not been fully determined, but is of the order of 1–1.5 Mb (Guzman and Ecker, 1988; Schlessinger, 1990). However, to date, the plant YAC libraries constructed have tended to have much smaller inserts than this, and average insert sizes are typically about 150 kb (Ward and Jen, 1990; Grill and Sommerville, 1991) and 250 kb (Ecker, 1990). This differs significantly in scale as compared to human and mouse YAC libraries (Connelly *et al.*, 1991; Larin *et al.*, 1991), and may reflect the problems associated with DNA isolation from higher plants, although a more recent *Arabidopsis* library has been reported with an average insert size of 450 kb (Creusot *et al.*, 1994). In most cases DNA for library preparation is isolated in an agarose matrix to protect it from shearing damage, is partially, or fully digested using rare restriction cutters (see also Chapter 1), and size fractioned by pulsed-field gel electrophoresis (Anand *et al.*, 1989; McCormick *et al.*, 1989), then cloned into an appropriate pYAC vector. The size separation may be repeated before the ligation products are transformed into competent yeast cells (Burgers and Percival, 1987; Rech *et al.*, 1990); AB1380 (*MATα*, *Ψ⁺*, *ura3*, *trpl*, *his5*, *ade2–1*, *can1–100*, *lys2–1*) (Burke *et al.*, 1987) has been commonly used as the maintenance yeast strain.

B. Screening YAC Libraries

Chromosome walking relies upon high-density screening strategies; often several libraries will be screened to identify a number of positive clones, which helps to confirm the integrity of a walk. Such screening approaches can include hybridisations to high-density colony blots or PCR screens.

1. Colony blots

(a) Preparing filters. Colony blots are constructed by plating out colonies at high density on Kiwi medium (Table 3.2), which is a selective medium that lacks uracil and tryptophane. This can be conducted by hand or by robots. Several different approaches to constructing high-density plates are used depending on the initial array of the library. For libraries which have been ordered into 96-well formats, these are best plated out in multiples of 96. The densities used depend on the size of colony and the accuracy of transfer, but can range from 768 to 2304 on a 9 x 11 cm filter. Colonies are initially transferred on to Kiwi medium, and allowed to grow (30°C, o/n; duration of incubation depends on the number of colonies used, each must grow evenly but remain distinct). The colonies are then transferred on to a master filter (Hybond-N; Amersham) by placing a piece of premoistened nylon membrane on to the colonies. This is then used to make further copies of the library. The master is first incubated, colony side uppermost, on a solidified Kiwi medium plate (2 h, 30°C), then placed on to 3MM paper (colonies uppermost) on a glass plate. Copy filters are then made by placing prewetted nylon filters on to the master filter and pressing down on to the filter using further sheets of 3MM paper and a glass plate to help make an even 'sandwich'. After making six copy filters it is necessary to place the original back on to the Kiwi medium plate (5–10 min, 30°C) to allow it to rehydrate; 50–100 copies can be prepared from each master filter. Each filter is then placed on to solidified Kiwi medium plate (colonies uppermost, 30°C, 1–2 days) to permit colony growth. Filters are then prepared for hybridisation based on Brownstein *et al.* (1989) and Coulson *et al.* (1988).

TABLE 3.2. *Kiwi medium* (1 litre) (Ausubel *et al.*, 1988; Putterill *et al.*, 1993).

Component	mg l^{-1}
Adenine	45.0
L-Arginine	22.0
Aspartic acid	108.0
L-Glutamic acid	108.0
L-Histidine	22.0
Leucine	65.0
L-Lysine	32.0
L-Methionine	22.0
L-Phenylalanine	54.0
L-Serine	405.0
Threonine	216.0
L-Tyrosine	32.0
L-Valine	16.0
Component	**g l^{-1}**
Yeast nitrogen base without amino acids	6.7
Glucose	20.0
Casamino acids	11.0

pH to 6.8 using approximately one pellet of NaOH per litre.
Agar medium is solidified with 2.2% (w/v) Bacto-agar.

(b) Filter processing. Bulk processing of filters is recommended due to the extent of screening required for colony linking and the relative ease of large-scale processing. Filters once processed can be stored for in excess of 6 months (dry, room temperature).

(1) Yeast cells are spheroplasted by a 10 min pretreatment on filter paper moistened with 1 M sorbitol, 20 mM EDTA, 10 mM Tris–HCl, pH 8.0, and 0.8% (w/v) dithiothreitol (DTT), followed by incubation on filter paper sealed in Petri dishes that have been impregnated with 1 M sorbitol, 20 mM EDTA, 10 mM Tris–HCl, pH 8.0, 1 mg ml^{-1} of Novozyme 234 (Lysing enzyme; Sigma) and 1% (v/v) β-mercaptoethanol (o/n, 37°C).

(2) Filters are then lysed and processed by placing them, in succession, on 3MM paper moistened with the following solutions:

10% (w/v) SDS	10 min, room temperature
0.5 M NaOH, 1.5 M NaCl	10 min, room temperature
0.5 M Tris–HCl, pH 7.4, 1.5 M NaCl	10 min, room temperature
0.5 M Tris–HCl, pH 7.4, 1.5 M NaCl	10 min, room temperature
2 × SSC buffer	10 min, room temperature

(3) Filters are then dried (room temperature) and baked at 80°C for 2 h.

(c) Library screening.

(1) Immediately prior to hybridisation the filters should be lightly washed using distilled water to remove any yeast debris remaining on the filters after processing. Prehybridise the filters (50% (v/v) formamide, 5× Denhardt's solution (0.5% (w/v) Ficoll 400, 0.5% (w/v) bovine serum albumin, 0.5% (w/v) polyvinylpyrrolidone), 3× SSPE buffer, 0.5% (w/v) SDS; 42°C) for 1 h.

(2) Probes, labelled by random priming, nick translation or riboprobes (Sambrook *et al.*, 1989), can be hybridised overnight using the same hybridisation solution with the addition of 10 μg ml^{-1} of herring sperm. Most walks to date have utilised ^{32}P labelling, although non-radioactive screening protocols are now available (Amersham).

(3) Care should be taken when washing filters after hybridisation since some background signal is required for locating the colony position. Wash for 20 min at 55°C once in 3× SSC buffer and 0.1% (w/v) SDS and once in 0.1× SSC buffer and 0.1% (w/v) SDS. If low levels of background are present, colony positioning can be assisted by the use of ^{35}S-labelled YAC probes, which hybridise to all colonies.

Colony identification can be a problem with high-density filters; therefore, it is advisable for each hybridisation to use at least two copies of the library which have been plated in opposite orientations. Positives can then be compared and determined by matches across the filters. Putative positive colonies then need to be confirmed by further 'low-density' colony blots, if specific identification is not possible; or by pulsed-field gel electrophoresis (PFGE) of the individual YAC clones, Southern blotting and hybridisation with the corresponding clone. PFGE also serves to size the insert whilst confirming its nature.

2. PCR screening

This type of screening requires sequence knowledge and corresponding oligonucleotide primers to the target site, but has the advantage that large numbers of clones can be analysed very easily, which is of particular importance with large genomes (Edwards *et al.*, 1992). Most procedures for PCR screening utilise pooled lysates of YAC clones and some form of pyramidal matricing, which allows for colony identification after a number of rounds of PCR (Heard *et al.*, 1989; Kwiatkowski *et al.*, 1990; Rosenthal, 1992). The identification of flanking clones or 'continuing steps' relies upon the sequencing of YAC fragments using the primers and designing new primers accordingly; such approaches can be automated to a high degree due to developments in direct PCR sequencing.

C. Sizing of YACs

1. Yeast chromosome preparation

(1) YAC clones should be streaked out on to fresh Kiwi medium, and a single colony used for liquid culture (25 ml).
(2) Cultures should be grown to late exponential phase, which takes approximately 24 h (30°C, 200 rpm). Harvest by centrifugation (300 rpm, 3 min).
(3) Wash the pellet in 1 ml of 1 M sorbitol, recentrifuge in a microcentrifuge tube (13 000 rpm, 1 min) and resuspend cells in 300 µl of SCEM (1 M sorbitol, 0.1 M sodium citrate 0.05 M EDTA, pH 7, 2.4 µl ml^{-1} of β-mercaptoethanol, 10 mg ml^{-1} of Novozyme 234 (Lysing enzyme; Sigma) (1 h, room temperature). Centrifuge at 6000 rpm for 5 s (microcentrifuge tube) to loosely pellet the cells.
(4) Resuspend the pellet in 1 M sorbitol to 100 µl total volume. Add 100 µl of 1% (w/v) SeaPlaque agarose (Flowgen) made in 1 M sorbitol and kept at 50°C. Mix gently but thoroughly. Dispense 100 µl aliquots into a PFGE plug former or the base of a Petri dish (if no commercial former is available). Allow to set on ice (10–20 min).
(5) Incubate the blocks in 1 ml of 0.5 mg ml^{-1} of proteinase K in 100 mM EDTA, pH 8, and 1% (w/v) sarkosyl at (50°C). Replace the buffer (1 ml) after 4 h and incubate o/n at 50°C.
(6) Wash plugs with 50 mM EDTA and store in this state (50 mM EDTA, 4°C). Before use, plugs should be washed as follows: three washes in TE at 50°C for 30 min, three times in TE at room temperature for 30 min and once in 0.5 × TBE (or an appropriate buffer for the electrophoresis) at room temperature for 30 min.

An alternative to this method that has been successfully used to make high-quality YAC chromosomal preparations is the yeast genomic DNA plug kit (Bio Rad).

2. Restriction digest of YACs

If the DNA samples are to be used for restriction digest they should be treated with two 20 min incubations (50°C) of 1 mM phenylmethylsulfonyl fluoride (PMSF) in TE prior to washing (see Section C.1, step (6)), to ensure inactivation of any residual proteinase K. After washing (step (6)) they should be given an additional incubation in 1 × restriction buffer, approximately 100 µl/100 µl plug (0°C, 30 min), then incubated in 1 × restriction

buffer containing approximately 50 units of enzyme for 30 min prior to incubation for digestion. This serves to ensure that the salts and enzyme penetrate completely into the block prior to digestion.

3. PFGE of YACs

PFGE should be conducted to separate a size range of 50–500 kb (see Chapter 4). Lambda concatamers and yeast chromosomal plugs can be used as markers. Figure 3.3 shows separations achieved using the transverse alternating-field electrophoresis (TAFE) GeneLine 2 electrophoresis system (Beckman) (370 mA, 23 s pulse, 12 h; 390 mA, 30 s pulse, 6 h; 390 mA, 22 s, 18 h).

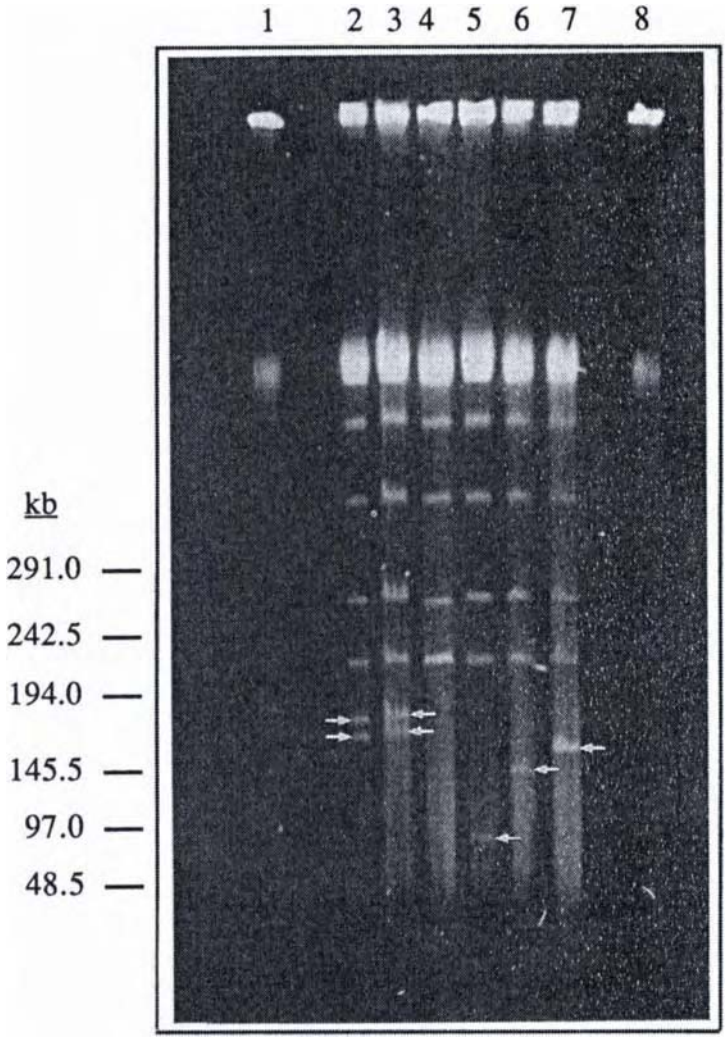


FIG. 3.3. TAFE gel electrophoresis of YAC clones. YACs are marked with arrows. In lanes 2 and 3 two YAC inserts can be seen: these are thought to have originated due to a double cloning event. Run conditions were 370 mA for 12 h using a pulse of 23 s, 390 mA for 6 h with a pulse of 30 s and, finally, 390 mA for 18 h using a pulse of 22 s (1.0% LE agarose, 15°C, 0.25 × TBE). Lanes 1 and 8: lambda concatamer standards. Lanes 2–7: YAC clones from the Grill and Sommerville (1991) library.

4. Southern hybridisation of YACs

To ensure complete transfer of DNA to nylon, YAC DNA must be nicked prior to blotting. This can be achieved by ultraviolet treatment or by depurination (0.25 N HCl, 20 min). If ultraviolet nicking, the DNA should be saturated with ethidium bromide stain (0.5 $\mu g\ \mu^{-1}$, 30 min), and nicked (60 mJ; BioRad Crosslinker) prior to visual inspection and photography.

The DNA can then be denatured (1.5 M NaCl, 0.5 M NaOH; 30 min, room temperature) and transferred to nylon (1.5 M NaCl, 0.25 M NaOH; >24 h, room temperature). The capillary blot should be held together with minimal weight, since excessive weights tend to compress the gel, making DNA transfer difficult. Hybridisations can then be carried out as for normal Southern blots (Sambrook *et al.*, 1989).

D. Linking YAC Clones

To make a step in a walk and link clones together, probes need to be prepared from the YACs and used for library screening. This is most efficiently conducted by preparing probes from the ends of the YACs. However, since in most cases the integrity, i.e. whether a YAC may be chimeric, and its orientation will be unknown, it will usually be necessary to prepare end-probes from both ends and from many YACs to produce an accurate contig map of a region.

1. End-probe production

Several different protocols have been developed to clone or amplify the ends of YACs for probe preparation, although the need for such approaches may be diminished with time as new YAC cloning vectors with RNA polymerase-binding sites flanking the cloning site are used more extensively. However, a number of laboratories have found riboprobe generation from YACs to be problematical, possibly because of the difficulties in obtaining large amounts of high-quality DNA for RNA transcription. The approaches used to generate YAC end-probes include PCR-based approaches (inverse PCR (IPCR) (Ochman *et al.*, 1988), vectorette PCR (Riley *et al.*, 1990; Copley *et al.*, 1991) and thermal asymmetric interlaced (TAIL) PCR (Liu and Whittier, 1995) and the use of degenerate sequences (Lu *et al.*, 1995)) and plasmid rescue. The DNA that can be used for such protocols can be of intact chromosomal size (see Section C.1) or can be prepared by minipreparation methods in which the maximum size will be approximately 25–50kb (commercial minipreparation versions are also available; Puragene, FlowGen).

2. Total yeast minipreparation DNA isolation

(1) Spheroplast yeast cells as described in the yeast chromosome preparation (see Section C.1). Resuspend the pellet in 400 µl of 0.4% (w/v) SDS, 75 mM Tris–HCl, pH 7.5, and 40 mM EDTA, pH 8.0. Incubate at 65°C for 30 min.
(2) Cool on ice, add 100 µl of 4 M potassium acetate and incubate on ice for 30 min.
(3) Centrifuge (13 000 rpm, 10 min), collect the supernatant and precipitate using 0.3 M sodium acetate and 2.5 volumes of ethanol (ice, 10 min). Centrifuge (13 000 rpm, 10 min), and wash the pellet with 70% (v/v) ethanol.

(4) Resuspend the pellet in 200 μl of TE. Once dissolved, centrifuge (13 000 rpm, 5 min) to remove any residual SDS.
(5) Transfer supernatant to a fresh tube and extract once with equilibrated phenol (10 mM Tris–HCl, pH 8): chloroform: isoamyl alcohol (25:24:1) and once with chloroform:isoamyl alcohol (24:1).
(6) Reprecipitate the DNA using 0.3 M sodium acetate and 2.5 volumes of ethanol (ice, 10 min). Wash in 70% (v/v) ethanol, and redissolve in 50 μl of TE.
(7) Minipreparation YAC DNA can be used for restriction digestion, Southern blotting or end-probe preparation.

3. Protocols for end-probe preparation

(a) ***IPCR.*** Modified from Ochman *et al.* (1988), this involves the restriction digestion of the YAC clones, self-ligation to form circles, followed by restriction digestion in the YAC arms, to create a linear fragment of insert which is flanked by the YAC sequence (Fig. 3.4). The PCR is then conducted using primers based on the flanking YAC sequences so that the insert is amplified.

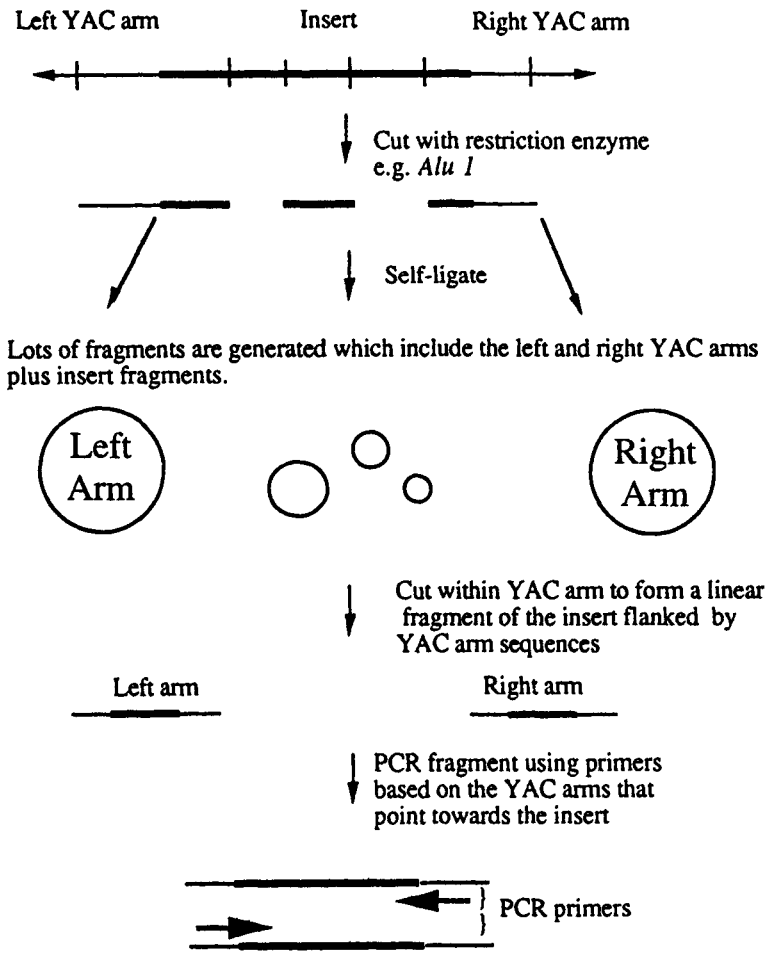


FIG. 3.4. Schematic diagram of IPCR to generate specific YAC end-probes.

(1) Restrict 0.5 μg total yeast DNA with 10 units of *AluI** (2–4 h, 37°C) in 200 μl.
(2) Extract the restricted DNA once with equilibrated phenol (10 mM Tris–HCl, pH 8), once with equilibrated phenol (10 mM Tris–HCl, pH 8):chloroform:isoamyl alcohol (25:24:1) and once with chloroform:isoamyl alcohol (24:1).
(3) Precipitate using 0.3 M sodium acetate and 2.5 volumes of ethanol (–20°C, 30 min). Wash pellet in 70% (v/v) ethanol and resuspend in 50μl of TE.
(4) Add 80 μl of 5× ligation buffer (0.25 M Tris, pH 7.6, 50 mM $MgCl_2$, 5 mM DTT, 5mM ATP, 25% (w/v) polyethylene glycol-8000), 269 μl of water and 1 μl of T4 DNA ligase. Incubate at 15°C (o/n).
(5) Heat the inactivate ligase (65°C, 15 min); repeat the ethanol precipitation (step (3)). Resuspend in 50 μl of TE.
(6) Purify the DNA through a Sepharose CL 6B spin column (equilibrated with TE). After passing through the column adjust the sample volume to 60 μl.
(7) Digest 30 μl aliquots with 5 units of *NheI* for the left arm, and 5 units of *SspI* for the right arm (30°C, 2 h; 37°C).
(8) Use 10 μl of the DNA in a 100 μl PCR mix containing, 1 × buffer (10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% (w/v) gelatine), 0.2 mM dNTPs (dATP, dCTP, dGTP, dTTP), 0.4 μM of each primer, 2.5 units of Taq DNA polymerase and 30 μl of mineral oil. Amplify using 30 cycles of 94°C for 1 min; 57°C for 1 min; 72°C for 2 min; and a final cycle of 72°C for 10 min.
(9) Analyse 10 μl of the reaction by gel electrophoresis (1.5% (w/v) agarose; 0.5 × TBE).

In some peoples' hands, greater reproducibility and specificity have been achieved by 'hot-starting' the PCR. For hot-start PCR the reactions are heated to 94°C for 5–10 min, to ensure total template denaturation, prior to the addition of Taq DNA polymerase (Ruano *et al.*, 1991). In some cases the incorporation of 1.0 μl of 'Perfect Match' (Stratagene) into the PCR can also help to avoid non-specific amplification products.

Primers. These are designed based on the sequence of the YAC vector arms (left or right according to the end required). Figure 3.5 shows the sequence surrounding the cloning site for pYAC41. Examples of suitable primers (Grill and Sommerville, 1991) are shown in Table 3.3. Nested primers can be used to generate increased levels of product and to ensure specificity of the amplification (Putterill *et al.*, 1993). However, gel purification (1.5% (w/v) agarose) of the fragment is still advisable to prevent non-specific amplifications.

(b) Vectorette PCR. An alternative to IPCR is vectorette or bubble PCR (Copley *et al.*, 1991; Riley *et al.*, 1990; Wilson *et al.*, 1992). This technique is similar to IPCR except that a mismatched 'priming linker' is added to the insert after initial digestion (Fig. 3.6). This linker, in combination with the YAC arm sequence, is then used to construct primers to amplify the insert. Vectorette PCR has advantages over IPCR since fewer manipulations are required, and modifications can be made when the linker is added, e.g. an RNA polymerase site can be incorporated in the linker so that RNA probes can be made for subsequent screening.

* Alternative restriction enzymes can also be used depending on the pYAC vector (e.g. left end — *AluI*, *DraI*, *EcoRV*, *HincII*, *PvuII* and *RsaI*; right end — *AluI*, *FspI*, *HincII*, *NruII* and *RsaI*). To determine which enzymes are suitable for specific clones, the YAC clones can be digested with a range of enzymes and analysed by Southern blotting with the corresponding arm (Grill and Sommerville, 1991). This will determine the expected size of PCR products. To permit subsequent use as probes and efficient PCR amplification, the products should be a minimum of 400 bp and a maximum of 2 kb, including pYAC vector sequences.

1 40
AAGTT TGATA TTTGA TTGTA AAATA CCGTA ATATA TTTGC
41 80
ATGAT CAAAA GGCTC AATGT TGACT AGCCA GCATG TCAAC
81 120
CACTA TATTG ATCAC CGATA TATGG ACTTC CACAC CAACT
121 160
AGTAA TATGA CAATA AATTC AAGAT ATTCT TCATG AGAAT
161 200
GGCCC AGCTG TTCCT TCAAC ATTTG AGAGA TCTGA AATGA
201 240
CCTCG AGTCT TTGAT TTAAA AGTGA AACGC GCTGA TTGAT
241 280
CTCTA AGTAA CCTCT TGTTG CTTGG TATAT GGGTT CTAAC
281 320
TGCGG TTCTG ACCAC ATAAT TTCTG GAGAG TCTAA AACTG
321 360
ATCCA TGTAG ATTGA TATTT ATCCT CAAAA TGAAT AGCTC
361 400
CCCTA TACTT TTCAT TATAT CTTCT TTGCT CATAG AAACT
401 440
TTACC GCTGT ATGCA ATTTC TTGTG GTATA TCCTG AGTAT
441 480
CTTCA ATAGT ATTAT CCACA AGCTC TTCGA AAAGT GAAAT
481 520
TTTTA CGCTT TGAGC AATTG CATGA GCTCA TGTTT GACAG
521 560
CTTAT CATCG ATAAG CTTTA ATGCG GTAGT TTATC ACAGT
561 600
TAAAT TGCTA ACGCA GTCAG GCACC GTGTA TGAAA TCTAA
601 640
CAATG GCTCA TCGTC ATCCT CGGCA CCGTC ACCCT GGATG
641 680
CTGTA GGCAT AGGCT TGGTT ATGCC GGTAC TGCCG GGCCT
681 720
CTTGC GGGAT ATCGT CCATT CCGAC AGCAT CGCCA GTCAC
721 760
TATGG CGTGC TGCTA GCGCT ATATG CGTTG ATGCA ATTTC
761 800
TATGC GCACC CGTTC TCGGA GCACT GTCCG ACCGC TTTGG
801 840
CCGCC GCCCA GTCCT GCTCG CTTCG CTACT TGGAG CCACT
841 880
ATCGA CTACG CGATC ATGGC GACCA CACCC GTCCT GTGGA
881 920
TCAAT TCTTT AGTAT AAATT TCACT CTGAA CCATC TTGGA
921 960
AGGAC CGGAT AATTA TTTGA AATCT CTTTT TCAAT TGTAT
961 1000
ATGTG TTATG TAGTA TACTC TTTCT TCAAC AATTA AATAC
1001 1040
TCTCG GTAGC CAAGT TGGTT TAAGG CGCAA GACTT TAATT
1041 1080
TATCA CTACG GAATT CGCGG CCGCA ATTAA CCCTC ACTAT
1081 1120
AGGGA TCCCT TTAGT GAGGG TTAAT TGCGG CCGCG AATTC

1121 1160
CGTAA TCTTG AGATC GGGCG TTCGA CTCGC CCCCG GGAGA
1161 1200
TTTTT TTGTT TTTTT ATGTC TCCAT TCACT TCCCA GACTT
1201 1240
GCAAG TTGAA ATATT TCTTT CAAGC TCTAC GCCGC ACGCA
1241 1280
TCGTG GCCGG CATCA CCGGC GCCAC AGGTG CGGTT GCTGG
1281 1320
CGCCT ATATC GCCGA CATCA CCGAT GGGGA AGATC GGGCT
1321 1360
CGCCA CTTCG GGCTC ATGAG CGCTT GTTTC GGCGT GGGTA
1361 1400
TGGTG GCAGG CCCCG TGGCC GGGGG ACTGT TGGGC GCCAT
1401 1440

CTCCT TGCAT GCACC ATTCC TTGCG GCGGC GGTGC TCAAC
1441 1480
GGCCT CAACC TACTA CTGGG CTGCT TCCTA ATGCA GGAGT
1481 1520
CGCAT AAGGG AGAGC GTCGA CCGAT GCCCT TGAGA GCCTT
1521 1560
CAACC CAGTC AGCTC CTTCC GGTGG GCGCG GGGCA TGACT
1561 1600
ATCGT CGCCG CACTT ATGAC TGTCT TCTTT ATCAT GCAAC
1601 1640
TCGTA GGACA GGTGC CGGCA GCGCT CTGGG TCATT TTCGG
1641 1680
CGAGG ACCGC TTTCG CTGGA GCGCG ACGAT GATCG GCCTG
1681 1720
TCGCT TGCGG TATTC GGAAT CTTGC ACGCC CTCGC TCAAG
1721 1760
CCTTC GTCAC TGGTC CCGCC ACCAA ACGTT TCGGC GAGAA
1761 1800
GCAGG CCATT ATCGC CGGCA TGGCG GCCGA CGCGC TGGGC
1801 1840
TACGT CTTGC TGGCG TTCGC GACGC GAGGC TGGAT GGCCT
1841 1880
TCCCC ATTAT GATTC TTCTC GCTTC CGGCG GCATC GGGAT
1881 1920
GCCCG CGTTG CAGGC CATGC TGTCC AGGCA GGTAG ATGAC
1921 1960
GACCA TCAGG GACAG CTTCA AGGAT CGCTC GCGGC TCTTA
1961 2000
CCAGC CTAAC TTCGA TCACT GGACC GCTGA TCGTC ACGGC
2001 2040
GATTT ATGCC GCCTC GGCGA GCACA TGGAA CGGGT TGGCA
2041 2080
TGGAT TGTAG GCGCC GCCCT ATACC TTGTC TGCCT CCCCG
2081 2120
CGTTG CGTCG CGGTG CATGG AGCCG GGCCA CCTCG ACCTG
2121 2160
AATGG AAGCC GGCGG CACCT CGCTA ACGGA TTCAC CACTC
2161 2200
CAAGA ATTGG AGCCA ATCAA TTCTT GCGGA GAACT GTGAA
2201 2240
TGCGC AAACC AACCC TTGGC AGAAC ATATC CATCG CGTCC
2241 2280
GCCAT CTCCA GCAGC CGCAC GCGGC GCATC CCCCC CCCCC
2281 2320
TTTCA ATTCA ATTCA TCATT TTTTT TTTAT TCTTT TTTTT
2321 2360
GATTT CGGTT TCTTT GAAAT TTTTT TGATT CGGTA ATCTC
2361 2400
CGAAC AGAAG GAAGA ACGAA GGAA GGAGC ACAGA CTTAG
2401 2440
ATTGG TATAT ATACG CATAT GTAGT GTTGA AGAAA CATGA
2441 2480
AATTG CCCAG TATTC TTAAC CCAAC TGCAC AGAAC AAAAC
2481 2520
CTGCA GGAAA CGAAG ATAAA TCATG TCGAA AGCTA CATAT
2521 2560
AAGGA ACGTG CTGCT ACTCA TCCTA GTCCT GTTGC TGCCA
2561 2600
AGCTA TTTAA TATCA TGCAC GAAAA GCAAA CAAAC TTGTG
2601 2640
TGCTT CATTG GATGT TCGTA CCACC AAGGA ATTAC TGGAG
2681 2680
TTA

FIG. 3.5. The DNA sequence of the region of pYAC41 flanking the cloning site (compiled by modifying Buddy Brownstein's version of pYAC4 by Gibson; Kuhn and Ludwig, 1994). This sequence has been compiled as a 'best guess' of this region. The pYAC41 insert sequence is marked in bold; this is inserted into the original *Eco*R1 site. The 'new' *Bam*H1 cloning site is underlined.

Commercial kits for vectorette PCR are available from Cambridge Research Biochemicals, Cheshire, UK. These provide linkers to suit a range of restriction sites. This can be very useful in providing flexibility in the range of restriction enzymes that can be used for the initial YAC/insert digests to obtain an insert of the appropriate size for PCR and library screening. An alternative 'home-made' cassette is described below. This incorporates a T7 RNA polymerase site in it to facilitate riboprobe preparation, but is suitable only with *Hind*III – digested DNA (although the *Hind*III site on the linker can be filled and used for blunt end cloning).

Cassette preparation:

(1) The cassette linker can be made using the following oligonucleotides:

Top strand (5′ ... 3′) (57 bp: 4 bp overhang, 12 bp matched, 29 bp mismatched, 12 bp matched):

AGC TTG CAT AAG GAT GCG TGA CTG ATC AAG ATT GTA CGA AGA CGA GAC ATG CGT CTC

Bottom strand (5′ ... 3′) (53 bp: 12 bp matched, 29 bp mismatched, 12 bp matched):

GAG ACG CAT GTC ATC ATA ATA CGA CTC ACT ATA GGG AGA TCC ATC CTT ATG CA

(2) The oligonucleotides are pooled in equal concentrations, heated to 94°C and left at room temperature to cool to allow the matched sequences to anneal. Annealing the 'cassette' results in a linker with a *Hind*III site at one end and a blunt end at the other (Fig. 3.6).

Cassette primer. A 29 bp oligonucleotide is needed to act as the 'cassette' primer. This is complementary to the 3′ ... 5′ sequence which is generated by first-strand synthesis from the 'known' primer, in this case the YAC arm sequence. This ensures that amplification must utilise the YAC sequence primer.

5′ TCA TCA TAA TAC GAC TCA CTA TAG GGA GA 3′

TABLE 3.3. Primers for the amplification of YAC arms.

Code	End amplified	Sequence (5′... 3′)
EG1	Left	GGC GAT GCT GTC GGA ATG GAC GAT A
EG2	Left	CTT GGA GCC ACT ATC GAC TAC GCG ATC
D71	Left	TCC TGC TCG CTT CGC TAC TT
C78	Left	GCG ATG CTG TCG GAA TGG AC
D72	Left	CAC TAT CGA CTA CGC GAT CA
C77	Left	GTG ATA AAC TAC CGC GAT CA
EG3	Right	CCG ATC TCA AGA TTA CGG AAT
EG4	Right	TTC CTA ATG CAG GAG TCG CAT AAG
C69	Right	CTG GGA AGT GAA TGG AGA CAT A
C70	Right	AGG AGT CGC ATA AGG GAG AG
C72	Right	CGA GTC GAA CGC CCG ATC TC
C71	Right	AGA GCC TTC AAC CCA GTC AG

DNA preparation. Total genomic yeast DNA can be used, either prepared as chromosomal sized DNA (see Section C.1) or as minipreparations (Section D.2). The following details the use of chromosomal preparations:

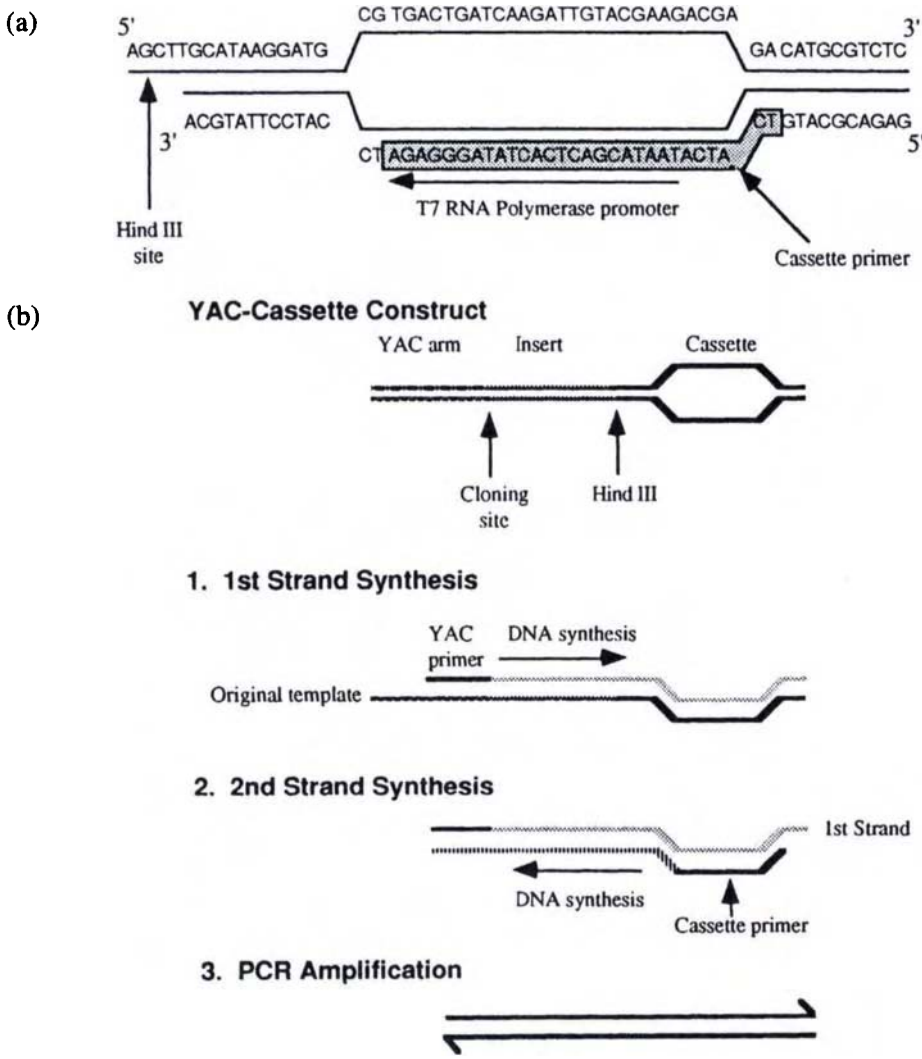


FIG. 3.6. Schematic diagram of vectorette/bubble PCR. (a) Cassette linker. (b) The cassette is ligated on to the restriction digested YAC; the insert is then amplified using primers designed from the YAC arm and cassette.

(1) After preparing and washing yeast DNA plugs (see Section C.1), digest one plug using 60 units of *Hind*III (o/n, 37°C).
(2) Remove *Hind*III buffer and wash the plugs with 100 μl ligation buffer.
(3) Replace the buffer with 100 μl of fresh ligation buffer. Heat to 65°C (5–15 min) to melt the plug. Cool to 37°C and add rATP to 10 mM, 20 pmol of annealed cassette and 2 units of T4 DNA ligase (o/n, 16°C).
(4) Dilute ligated sample with 400 μl of distilled H_2O.
(5) Use 5 μl of the diluted DNA sample in a 100 μl PCR mix containing the following: 10 mM Tris–HCl, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1 mg/ml^{-1} of gelatine, pH 8.3; 0.2 mM dNTPs; 0.4 mM of each primer; and 2.5 units of Taq DNA polymerase. Amplify by 38 cycles of 94°C for 1 min; 55°C for 2 min; 72°C for 3 min; and a final cycle of 72°C for 10 min.

(6) Analyse or purify, for further 'nested PCRs', the fragments by gel electrophoresis (1.4% w/v agarose, 0.5 × TBE).
(7) Inserts can be purified from the YAC arms and linker, by digestion at (or close to) the cloning site and the *Hind*III 'cassette' site.
(8) End-probes for library screening can be produced after removing the YAC arm either by gel purification and random primer labelling, based on Feinberg and Vogelstein (1983), or by riboprobe preparation from the T7 RNA polymerase site.

*(c) **Plasmid end-rescue**.* Both of the PCR approaches described earlier are very effective; however, they are governed by the occurrence of the appropriate restriction sites in the 'insert' DNA, and in most cases tend to give end-probes of 500–1000 bp. An alternative approach which generally gives larger end-probes is plasmid rescue. For plasmid rescue, the YAC clones are digested with *Xho*I, self-ligated and then transformed into *E. coli.* The YAC arm and insert fragment are maintained due to the bacterial origin of replication and the selectable marker (ampicillin resistance) present in the pYAC vector. The cloning efficiency of plasmid rescue tends to be low; therefore, high transformation efficiencies, as obtained using electroporation, are vital for this procedure to be successful.

(1) Digest 2 μg of total YAC DNA with *Xho*I (alternative enzymes are *Nde*I, *Nde*I or *Xho*I + *Sal*I*) in 25 μl (4 h, 37°C).
(2) Extract once with equilibrated phenol (10 mM Tris–HCl, pH 8), equilibrated phenol (10 mM Tris–HCl, pH 8):chloroform:isoamyl alcohol (25:24:1) and with chloroform:isoamyl alcohol (24:1).
(3) Precipitate using 0.3 M sodium acetate and 2.5 volumes of ethanol (–20°C, 30 min). Wash pellet in 70% (v/v) ethanol.
(4) Resuspend in 100 μl of ligation buffer (containing 1 mM ATP). Add 1 unit of T4 DNA ligase. Incubate at 15°C (o/n).
(5) Repeat phenol extractions (step (2)) and ethanol precipitation (step (3)). Resuspend pellet in 5 μl of distilled H_2O.
(6) Transform 3 μl aliquots of the DNA sample into electrocompetent *E. coli* cells, e.g. DH5α. Select on LB medium containing 50 μg ml^{-1} of ampicillin.

4. Verification of end-probes

Once isolated from the vector sequences, end-probes can be used to rescreen the libraries to identify flanking clones. However, prior to screening it is necessary to confirm that the probes correspond to the starting YAC clone and that they map to the appropriate region of the genome.

(1) The initial confirmation can be carried out by hybridisation to colony blots or Southern blots of YAC clones from the region. The use of Southern blots has the advantage that the RFLP patterns generated by the YAC and genomic DNA can be compared to ensure that rearrangement has not occurred in the YAC clone and to

*Plasmid rescue was initially only possible for the left YAC arm since this arm carries the bacterial origin of replication and a bacterial selectable marker (AmpR). The application of this approach has now been extended by the construction of two rescue vectors (Hermanson *et al.*, 1991) which can be inserted into the YAC vector by homologous recombination (Cellini *et al.*, 1991). Both carry a bacterial origin of replication and selectable markers, thus both ends can be 'retrofitted' and then used for plasmid rescue. The use of these vectors also extends the application of the end-rescue process for the left arm, since additional restriction enzymes can be used, thus removing the necessity for an *Xho*I site near the cloning site in the insert DNA. Alternative YAC vectors (Shero *et al.*, 1991) also allow recovery of both YAC insert arms.

establish whether the end-probe is single copy. The greater the number of YACs that are isolated in a given region the greater the certainty of the walk, since they can all be used to confirm hybridisations in a given region. This is very important since in a number of cases, end-probes have been found to contain repetitive sequences or represent different regions of the chromosome due to the chimeric nature of the YAC (Putterill *et al.*, 1993).

(2) Hybridisation to recombinants in the region also provides additional evidence of the integrity of the clone and the location of the clone. This is advisable since a high proportion (possibly 40%) of YAC clones have been identified as chimeric in most of the libraries constructed to date; these may constitute both double insertion or recombination events (Green *et al.*, 1991).

E. Positioning of the Walk

At numerous stages in a walk it is necessary to identify 'where the walk is and what direction it is moving in'. This can be achieved in part by using a number of YAC libraries and relying on the verification of clones by their hybridisation to other clones in the region, but mainly requires the hybridisation of clones to recombinants flanking the gene. Recombinants are a vital prerequisite to determining map positioning and delimiting the region containing the gene. The final stages of many chromosome walks rely upon the complementation of the mutant phenotype with the putative wild-type gene sequence or the identification of the corresponding cDNA by Northern blot analysis, etc. The effective use of recombinants can accurately define the gene position and thus greatly reduce the numbers of transformations required for complementation analysis or screenings for Northern blotting. This may be of particular importance if the mutation in question cannot be complemented.

The locality of the gene of interest is usually established once a walk has been completed between two marker points which are known to flank the gene. However, this region may be quite large, and constitute several YACs. End-probes from both sides of specific YACs can be used to identify RFLPs in the recombinant populations. These then serve to position the YACs on the physical/RFLP maps. Once RFLPs have been identified using the ends of YACs, internal probes can be used to further define the region. The limitations of this approach are the numbers of recombinants available for mapping and the ease of identifying RFLPs in the mapping populations.

1. Complementation analysis

Verification of the gene cloning has tended to be carried out by complementation analyses (Herman and Marks, 1989; Giraudat *et al.*, 1992; Bent *et al.*, 1994; Mindinos *et al.*, 1994). The use of YAC insert-derived probes for hybridisation to recombinants around the gene of interest can significantly limit the scale of complementation experiments necessary. Clones for use in complementation analysis can be isolated by two main approaches:

(1) Subcloning the YACs in smaller pieces into a vector suitable for transformation.
(2) By screening libraries with the YACs to identify clones for transformation.

2. Subcloning YACs

Problems can be encountered when subcloning YACs due to the low amounts of DNA present compared to the total yeast DNA, although this may have been removed to some

extent by new YAC vectors for which the copy number can be altered (Smith *et al.*, 1990) and by improved YAC isolation procedures (Maule *et al.*, 1994). To be of use a YAC 'mini-library' must be representative of the whole of the YAC. This results in the dilemma of which vector to use for subcloning. Cosmid vectors have the advantage that they can be used directly for transformation; however, they appear to be non-randomly propagated in *E. coli*, leading to high levels of cloning bias (Hauge and Goodman, 1992). Lower levels of cloning bias have been seen in lambda libraries. It may therefore be necessary to prepare libraries initially using lambda to ensure that all sequences are fully represented.

One procedure for YAC subcloning to make a 'mini-library' is detailed below, although alternative versions of this have been successfully described (Giraudat *et al.*, 1992; Hauge and Goodman, 1992):

(1) Size separate YACs by PFGE using 1.0% (w/v) low gelling temperature (LGT) agarose gel (SeaChem; FlowGen) in 0.5 × TBE. Use approximately (100 μl) DNA blocks for each YAC.
(2) Stain the gel using ethidium bromide (0.5 μg ml^{-1}), and destain as necessary. Quickly (in order to minimise ultraviolet damage) photograph the gel, and use this to identify the migration range of the YACs.
(3) Excise the YAC bands, removing any excess agarose.
(4) Equilibrate the agarose/YAC plug in 1× *Sau3A* buffer (100 μl × 2, 30 min, 4°C).
(5) Replace with fresh buffer containing 50 units of *Sau3A*. Incubate for 4 h (37°C).
(6) Rinse the plug in 1 ml of TE (3 × 20 min, 4°C), and place it in 50 μl of TE.
(7) Melt the plug at 65°C (20 min) to inactivate *Sau3A*. Transfer to 37°C.
(8) Set up ligation using *Bam*HI-digested, phosphatased vector, e.g. pATX (20 ng of vector, 2 mM ATP, 1 × ligase buffer) and varying aliquots of the YAC digest (1, 3 and 10 μl). Incubate at 15°C (o/n).
(9) Melt digests and introduce into *E. coli* by transformation of competent cells or electroporation.
(10) Recombinant colonies can then be analysed by hybridisation to YAC colonies and digests.

3. Screening existing libraries

An alternative to YAC subcloning is to identify clones corresponding to the YAC inserts from existing libraries. For *Arabidopsis*, two libraries have been used extensively. These are a cosmid library (Olszewski *et al.*, 1988) and an *Agrobacterium* library (Lazo *et al.*, 1991). The *Agrobacterium* clones have the advantage of being able to be directly used for transformation, whilst the cosmids require prior mobilisation into *Agrobacterium* by triparental mating using a helper strain.

V. CONFIRMING GENE CLONING

A. Transformation

Two approaches can be used to conduct complementation analyses. These are (1) direct transformation into the homozygous mutant line and screening for complementation of the mutant phenotype and (2) transformation into a wild-type line (often of an alternative

ecotype), followed by crossing to the homozygous mutant. These plants are then allowed to self, and the F_2 populations screened for mutant versus wild-type appearance and for the presence, and complementation, of the cloned gene/T DNA (Giraudat *et al.*, 1992). The former approach is most direct; however, its application will depend on the type of mutation and the ease of transforming the mutant plant.

The transformation protocol will depend on the individual plant species. In the case of *Arabidopsis*, several protocols have been published (Lloyd *et al.*, 1986; Sheikholeslam and Weeks, 1987; Schmidt and Wilmizer, 1988; Valvekens *et al.*, 1988; Chaudhury and Signer, 1989; Damm *et al.*, 1989; Bechtold *et al.*, 1993) with varying transformation frequencies observed between different ecotypes. High transformation frequencies have been observed with most ecotypes using a variation to the Valvekens *et al.* (1988) protocol, which incorporates 5 mg l^{-1} of silver thiosulfate into the media for seed germination and callus generation (Clarke *et al.*, 1992), and by vacuum in filtration (Bechtold *et al.*, 1993).

B. Alternative Strategies for Characterising the Location of the Gene

1. cDNA screening

The final stages of positional cloning may be hastened by a number of approaches, including the analysis of cDNAs. The region for complementation analyses, and possibly the gene itself, may be identified by mapping cDNAs from tissues in which the gene may be expressed on to specific YACs which map to the region. This type of analysis can be conducted in several ways. One approach is by screening Northern blots of the mutant and wild-type plants with specific YAC or cosmid probes (Giraudat *et al.*, 1992). Another method is by differential screening of cDNA libraries to look for the absence or shift of a specific message; or alternatively by the construction of sets of cDNA clones which correspond to specific YACs (Elvin *et al.*, 1990; Lovett *et al.*, 1991; Parimoo *et al.*, 1991; Geraghty *et al.*, 1993).

2. Sequencing

As part of chromosome walks, some researchers have commenced sequencing of all end-probes that have been generated from YACs. This serves to ensure that the end-probes correspond to the expected sequence, allows the preparation of PCR primers and may assist in gene identification. This is particularly true for *Arabidopsis*, since large numbers of *Arabidopsis* cDNA sequences are now being deposited in databases.

VI. CONCLUSIONS FROM CHROMOSOME WALKING

Cloning based on position, by the techniques associated with chromosome walking, tends to be lengthy and labour intensive. The ease of such approaches will be increased as physical and genetic maps are improved and there is a greater linking of libraries. Other factors such as new series of YAC vectors which allow increased copy number (Smith *et al.*, 1990), 'retrofitting' to allow easier end-probe cloning (Hermanson *et al.*, 1991; Shero *et al.*, 1991), the development of new procedures for library screening (Heard *et al.*, 1989; Kwiatkowski *et al.*, 1990; Rosenthal, 1992), the use of YACs as hybridisation

probes (Bancroft *et al.*, 1992) and the development of other cloning vectors e.g. BACs will all make the task of chromosome walking simpler in future.

An alternative approach to clone genes based on the mutant phenotype which is proving considerably quicker is gene tagging. This is described briefly in the next section.

VII. GENE TAGGING

A. Introduction

Gene tagging is a means to clone genes based upon their mutant phenotype, and has been successfully used to clone a number of *Arabidopsis* genes (Marks and Feldmann, 1989; Koncz *et al.*, 1990; Van Lijsebettens *et al.*, 1991a; Ahmad and Cashmore, 1993). Tagging can be achieved in two ways, one by using the transfer of DNA (T-DNA) from the Ti plasmid of *Agrobacterium* as a tag and the other by the use of transposons as mobile molecular tags.

T-DNA tagging utilises the ability of the bacterium *Agrobacterium tumefaciens* to stably transfer sections of the Ti plasmid (Zambryski, 1988), which it carries, into a plant genome by illegitimate recombination (Mayerhofer *et al.*, 1991). The regions of the Ti plasmid, the T-DNA, that are transferred during transformation are flanked by 25 bp inverted repeats; the sequence between the repeats can, however, be varied. Transfer is achieved due to virulence genes present on the Ti (or Ri) plasmid, although it is not necessary for them to be *trans* to the T-DNA to permit transfer. Once transformed, the T-DNA remains stably incorporated at the insertion site. Studies in a number of transgenic plants have suggested that T-DNA is preferentially inserted into regions of active transcription (Koncz *et al.*, 1989), thus enhancing the likelihood of tagging genes rather than non-coding DNA sequences. If the tag is inserted into a transcribed region, gene function may be disrupted and a mutation may be observed, thus serving as a means for studying plant developmental biology.

T-DNA tagging has also been used as an approach for defining the promoter and enhancer sequences of genes. T-DNA cassettes have been constructed with promotorless reporter genes, or with only weak promoters, and used for tagging. In such cases a defined mutant phenotype may not be exhibited; however, evidence that a transcription unit has been tagged can be seen by the expression of the reporter genes. For example, promoter tagged lines have been identified by the expression of APH (3′) II enzyme activity using the aminoglycoside phosphotransferase II (*aph* (3′)II) T-DNA construct (Koncz *et al.*, 1989); tissue-specific expression and enhancer sequences have been identified using β-glucuronidase GUS fusions (Topping *et al.*, 1991; Topping and Lindsey, 1995).

The applications of tagging as a means of identifying mutations and gene cloning depend on the efficiency of mutation; this in turn depends on the frequency of non-coding DNA and highly repetitive DNA. The reported frequency for T-DNA mutagenesis in *Arabidopsis* is extremely variable, ranging between 0.5 and 20% (Feldmann and Marks, 1987; Feldmann, 1991; Van Lijsebettens *et al.*, 1991b; Koncz and Schell, 1992), although in most cases insufficient numbers of plants have been analysed in detail to determine the true mutagenic frequency of the T-DNA. Promoter fusions have been reported at a rate of approximately 30% in tobacco and *Arabidopsis* (Koncz *et al.*, 1989). For plants such as

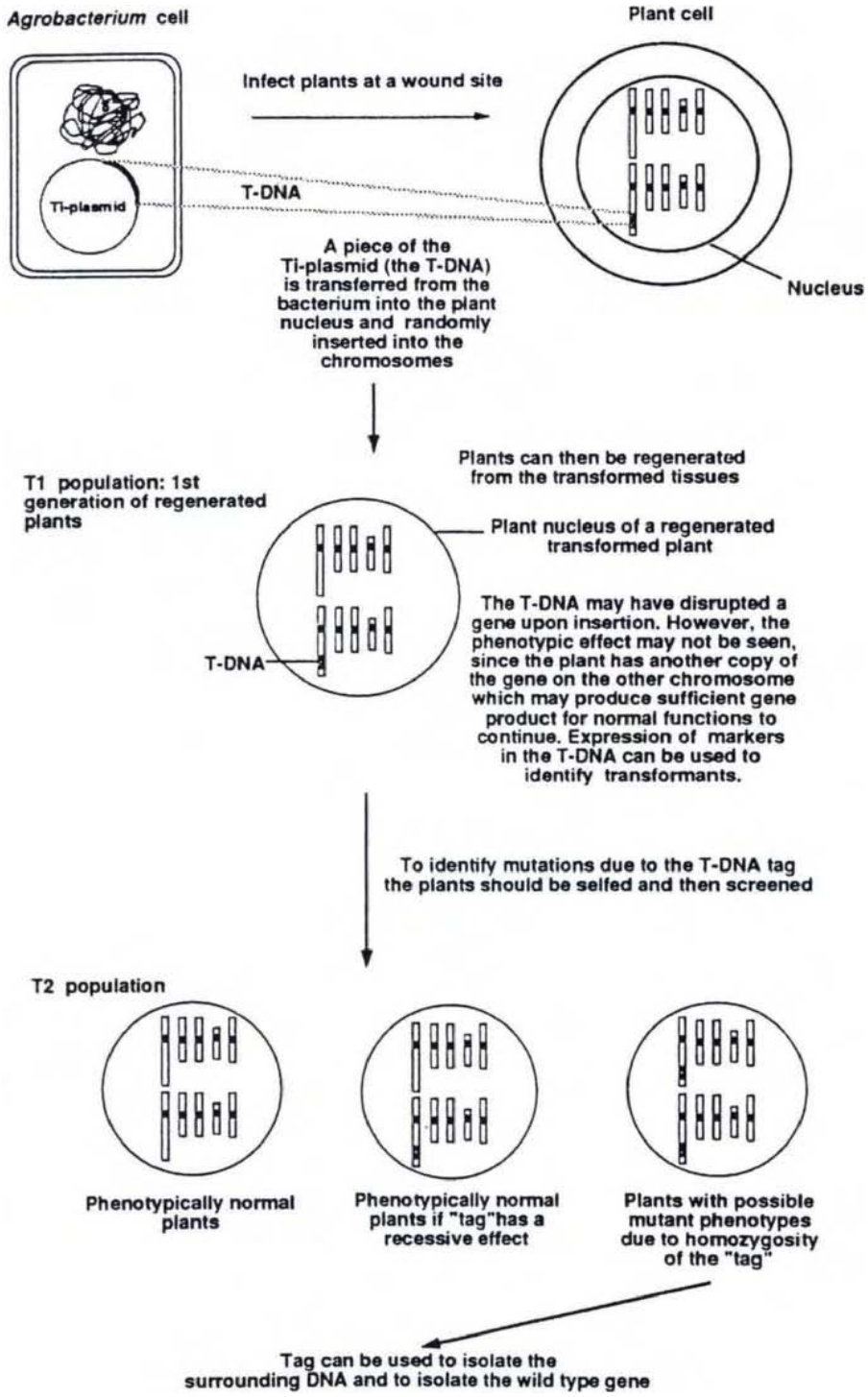


FIG. 3.7. Schematic representation of T-DNA tagging.

Arabidopsis with low (10–14%) levels of repetitive DNA, tagging provides a very powerful means to study gene function and to clone genes that have low levels of expression which may be involved in complex developmental pathways.

B. Approaches for Transformation

Various vectors and transformation approaches have been used for T-DNA tagging in *Arabidopsis*. The frequency of tagging mutagenesis and the reproducibility of these techniques has been shown to vary depending on the ecotype used and the procedure adopted. Three main protocols have been described: one utilises a tissue culture-based system of transformation (Koncz *et al.*, 1989), one involves direct seed transformation (Feldmann and Marks, 1987) and the last uses vacuum infiltration (Bechtold *et al.*, 1993).

1. Tissue culture transformation

This approach is probably the most commonly used method for plant transformation in many species. It involves the generation of callus from tissue segments, which is then inoculated using a specific *Agrobacterium* strain and maintained on antibiotic-containing medium to ensure bacterial removal. Transformed calluses are then used to regenerate shoots or embryos. Seed is then collected from these regenerants. If the T-DNA insert has been engineered to contain a selectable marker, e.g. the hygromycin-resistance gene (Koncz *et al.*, 1989), seedlings can be easily screened using this marker for the presence of the T-DNA. Plants derived from the T_1 seed population will be heterozygous for the insert; therefore, to identify recessive mutations based upon the T-DNA tags it is necessary to allow them to self, and homozygous plants must be screened in the following generations (Fig. 3.7). This has been successfully used to carry out high-frequency (2.2%) tagging in *Arabidopsis* (Koncz and Schell, 1992).

Tissue culture T-DNA tagging protocols are effective ways of generating mutations resulting from molecular tags. This has been demonstrated by the cloning of a nuclear gene encoding a chloroplast protein (Koncz *et al.*, 1990) that results in *pale* mutants (*chlorata*, *ch-2*). However, such approaches are highly labour-intensive, and the transformation and regeneration process alone may result in a high frequency of somaclonal variation which may obscure the true T-DNA mutation frequency. Such procedures are therefore difficult for large-scale mutagenesis.

Tissue culture tagging strategies have also been described which use leaf discs as an alternative starting material for transformations. This approach resulted in a T-DNA mutation frequency of 0.5–1%; tagging was verified by the cloning of a gene which alters leaf morphology (Van Lijsebettens *et al.*, 1991b). This form of transformation approach may minimise somaclonal variation due to the lack of a callus stage and mild hormone treatment during shoot regeneration. Using this approach, 2–5% of recessive mutations were found to be due to the tissue culture procedure (Van Lijsebettens *et al.*, 1991a,b).

2. In planta *transformation*

Seed transformation, as established by Feldmann and Marks, 1987, is an alternative approach to transformation in which the tissue culture stages are totally avoided. The exact process and timing of transformation is uncertain, but it is thought to involve the passage of T-DNA into the seed through the micropyle and into the shoot meristem dur-

ing imbibition of the seed and integument. The *Agrobacterium* material thought to be kept in the interstitial spaces and maintained as the plant develops. Transformation may occur at some point during sporogenesis, gametogenesis or fertilisation (Feldmann, 1991). Seed transformation avoids the problem of somaclonal variation, although the frequency of transformation is low and the experimental conditions appear particularly sensitive to variation, making reproducibility difficult. Since this approach seems to be highly susceptible to experimental variables, Feldmann (1991) recommends the use of several treatments and an initial pre-screen to identify the pool with the greatest transformation efficiency. It also appears to be highly ecotype-dependent and vector specific. Successful transformations have to date only been obtained using the ecotype Wassilewskija and the C58C1Rif strain containing the co-integrate 3850:1003 Ti plasmid; strain LBA4404 containing six different binary vectors failed to give any transformants (Feldmann, 1991).

In a study to assess the feasibility of seed transformation, a very high level of mutagenesis was described (approximately 20%) (Feldmann, 1991). This study generated more than 8000 transformants and in excess of 1000 putative mutants. These exhibited a broad spectrum of phenotypes ranging from seedling lethals, reduced fertility, size and pigment variation and form mutants. By molecular analysis of 50 of these mutants, Feldmann (1991) suggested that the tagging frequency was the order of 80%. In many cases the T-DNA has been shown to be inserted as concatamers of direct or inverted repeats (Feldmann and Marks, 1987; Feldmann *et al.*, 1989; Errampalli *et al.*, 1991). Verification of this approach for tagging has been obtained by the cloning of a number of genes, including trichome development (Marks and Feldmann, 1989) and agamous (Yanofsky *et al.*, 1990). However, from analysis of embryo-lethal frequencies in these populations (Errampalli *et al.*, 1991) and from other investigators using these mutants (Koncz *et al.*, 1992), the frequency of mutations linked to T-DNA tags seems to be significantly lower than previously estimated.

This approach therefore has considerable application for a large-scale mutagenesis scheme; however, the frequency of transformation is relatively low, making the process time-consuming and labour-intensive. An *in planta* approach which utilises transformation of the reproductive meristem has been described (Chang *et al.*, 1990).

An alternative non-tissue culture approach has also been described for *Arabidopsis* (Bechtold *et al.*, 1993). This procedure utilises vacuum infiltration of *Arabidopsis* plants with a suspension culture of *Agrobacterium* cells. Transformation is thought to occur late during floral development since the transformants generated were hemizygous and contained different T-DNA inserts. A large-scale mutagenesis programme is currently in progress using this approach (Bechtold *et al.*; 1993), which will yield data determining the overall efficiency of the process for T-DNA tagging mutagenesis.

VIII. ESTABLISHING LINKAGE BETWEEN THE MUTATION AND THE T-DNA TAG

Once a mutation has been identified in a transformed line it is necessary to show Mendelian linkage of the mutant phenotype with the T-DNA tag. This can be done by selfing the mutant and analysing the offspring, but is more robust if the cross is conducted to a wild-type stock, as continual selfing can result in the co-fixing of the T-DNA insert

(Walbot, 1992). For a recessive mutation the cross can be conducted using a homozygote mutant (mm) and a wild-type (MM) plant. The F_1 progeny are then selfed for a test of 3:1 segregation of the normal:mutant phenotype. The F_1 progeny can also be crossed to a mutant type to test for a 1:1 mutant:wild-type ratio. Following the segregation pattern of the T-DNA insert is best conducted by looking for expression of a selectable marker in the tag. If this cannot be carried out due to the nature of the tag, screening can be conducted by hybridising Southern or dot blots of genomic DNA from the progeny, with the T-DNA. Close linkage of 1–2 cM must be obtained as supporting evidence of tagging.

IX. CLONING THE REGIONS FLANKING THE TAGGED GENE

Once close linkage has been established between the tag and the mutation, the regions flanking the T-DNA tag can be isolated and used to obtain the full clone. Several different approaches have been used to clone the flanking DNA, and are as follows.

A. Plasmid Rescue

A procedure for rescuing the regions that flank the T-DNA tag is plasmid rescue (Koncz *et al.*, 1989; Koncz *et al.*, 1990; Behringer and Medford, 1992). This approach has been outlined previously for cloning YAC end-probes in Section D.3c. It involves restriction digestion of genomic DNA from the tagged line using an enzyme which does not cleave within the tag, self-ligation and then transformation of competent *E. coli* cells. Behringer and Medford (1992) describe the use of *Sal*I digests to clone flanking regions of the Ti plasmid 3850:1003, as used by Feldmann and Marks (1987). The nature of these procedures means that the transformation efficiency is low; therefore, the transformation is best conducted by electroporation using highly competent cells ($> 5 \times 10^8$).

B. Other Approaches

PCR approaches (IPCR, vectorette PCR) as outlined in Section D.3a,b can be used to clone the regions flanking the T-DNA tag. An alternative procedure which was used in the cloning of the GLABROUS1 gene (*GL1*) (Marks and Feldmann, 1989) is to prepare a genomic library from the T-DNA tagged line. This library can then be screened using the T-DNA. The clones that are identified can then be used to screen a genomic library from a non-transformed plant to clone the intact gene. This approach is very effective although it tends to be quite time-consuming.

Problems have been frequently encountered when trying to rescue the T-DNA tag together with the flanking plant genomic DNA. One possible explanation for this has been cytosine methylation of the DNA in the transgenic plants, which causes DNA rearrangements during plasmid rescue (Mandal *et al.*, 1994). Mandal *et al.* found that these effects could be avoided if the methylation was inhibited in the transgenic plants by azacytidine treatment or by cloning into a modified cytosine restriction deficient strain of *E. coli*.

C. Verification of Cloning

Once the putative complete gene has been cloned, confirmation can be established by complementation analysis (see Section E.1) (Herman and Marks, 1989; Giraudat *et al.*,

1992). Alternative evidence can be obtained by looking for differential expression patterns of the clone, in specific tissues of the mutant and wild-type plants.

X. CONCLUSIONS

The relative ease of T-DNA tagging mutagenesis techniques permits the large-scale generation of tagged populations for mutant identification and gene tagging (Koncz *et al.*, 1989; Feldmann, 1991). Such techniques will permit the rapid analysis of many genes whose function can be determined on the basis of mutant phenotype. This will make the molecular analysis of many areas of developmental biology feasible for detailed study. However, it is generally considered that the mutations identified by T-DNA tagging can only be 'loss of function' mutants. This means that a large number of developmental genes and embryolethal genes may not be accessible by this route, and alternative approaches such as positional cloning may be necessary.

ACKNOWLEDGEMENTS

I would like to thank Dr B. J. Mulligan and Dr P. J. Tighe for their assistance in reading this manuscript.

REFERENCES

Ahmad. M. and Cashmore, A. R. (1993). *Nature* **366**, 162–166.
Anand, R., Villasante, A. and Tyler-Smith, C. (1989). *Nucl. Acids Res.* **17**, 3425–3433.
Arondel, V., Lemieux, B., Hwang, I. A. S., Goodman, H. M. and Somerville, C. R. (1992). *Science.* **258**, 1353–1354.
Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1988). *In 'Current Protocols in Molecular Biology'.* **2**, 13.0.1–13.13.9. John Wiley, Chichester.
Ballabio, A. (1993). *Natural (Genetics)* **3**, 277–279.
Bancroft, I., Westphal, L., Schmidt, R. and Dean, C. (1992) *Nucl. Acids Res.* **20**, 6201–6207.
Bechtold, N., Ellis, J. and Pelletier, G. (1993) *C. R. Acad. Sci.* **316**, 1194–1199.
Behringer, F. J. and Medford, J. I. (1992). *Plant Mol. Biol. Rep.* **10**, 190–198.
Bender, W., Spierer, P. and Hogness, D. S. (1983). *J. Mol. Biol.* **168**, 17–33.
Bent, A. F., Kunkel, B. N., Dahlbeck, D., Brown, K. L., Schmidt, R., Giraudat, J., Leung, J. and Staskawicz, B. J. (1994). *Science* **265**, 1856–1860.
Brownstein, B. H., Silvermann, G. A., Little, R. D., Burke, D. T., Korsmeyer, S. J., Schlessinger, D. and Olson, M. V. (1989). *Science* **244**, 1348–1351.
Burgers, P. M. J. and Percival, K. (1987). *Anal. Biochem.* **163**, 391–397.
Burke, D. T., Carle, G. F. and Olson, M. V. (1987). *Science* **236**, 806–812.
Burr, B. and Burr, F. A. (1991). *Trends Gen.* **7**, 55–60.
Butler, R., Ogilvie, D. J., Elvin, P., Riley, J. H., Finniear, R. S., Slynn, G., Morten, J. E. N., Markham, A. F. and Anand, R. (1992). *Genomics* **12**, 42–51.
Cellini, A., Lacatena, R. M. and Tocchini-Valentini, G. P. (1991). *Nucl. Acids Res.* **19**, 997–1000.
Chang, C. and Meyerowitz, E. M. (1991). *Curr. Opin. Biotechnol* **1**, 178–183.
Chang, C., Bowman, J. L., De John, A. W., Lander, E. S. and Meyerowitz, E. M. (1988). *Proc. Natl Acad. Sci. USA* **85**, 6856–6860.
Chang, S.-S., Park, S.-K. and Nam, H.-G. (1990). Fourth International Conference on *Arabidopsis* Research, Vienna, 2–5 June 1993, p. 28.

Charlieu, J. P., Roizes, G., Marcais, B. and Bellis, M. (1991). *Bio Techniques* **10**, 317–318.
Chaudhury, A. M. and Signer, E. R. (1989). *Plant Mol. Biol. Rep.* **7**, 258–265.
Clarke, M. C., Wei, W. and Lindsey, K. (1992). *Plant Mol. Biol. Rep.* **10**, 178–189.
Connelly, C., McCormick, M. K., Shero, J. and Hieter, P. (1991). *Genomics* **10**, 10–16.
Copley, C. G., Boot, C., Bundell, K. and McPheat, W. L. (1991). *Bio Technology* **9**, 74–79.
Coulson, A., Waterston, R., Kiff, J., Sulston, J., Kohara, Y. (1988). *Nature* **335**, 184–186.
Creusot, F., Fouilloux, E., Dron, M., Lafleuriel, J., Picard, G., Billault, A., Le Paslier, Cohen, D., Chabouté, M. E., Durr, A, Fleck, J., Gigot, C., Camilleri, C., Bellini, C., Caboche, M. and Bouchez, D. (1994). *Plant J.* **8**, 763–770.
Czako, M. and Marton, L. (1992). *In. 'Arabidopsis: The Compleat Guide'* (D. Flanders, and C. Dean eds). AFRC PMB Programme. http://genome-www.stanford.edu/Arabidopsis/
Damm, B., Schmidt, R. and Willmitzer, L. (1989). *Mol. Gen. Genet.* **217**, 6–12.
Delfavero, J., Weyens, G., Jacobs, M., Edwards, K. E. and Vauterin, M. (1994). *Theor. Appl. Genet.* **88**, 449–453.
Dellaporta, S. L., Wood, J. and Hicks, J. B. (1983). *Plant Mol. Biol. Rep.* **4**, 19–21.
Devos, K. M. and Gale, M. D. (1992). *Theor. Appl. Gen.* **84**, 567–572.
Ecker, J. R. (1990). *Methods: A Companion Methods Enzymol.* **1**, 186–194.
Edwards, K., Thompson, H., Edwards, D., de Saizieu, A. and Sparks, C. (1992). *Plant Mol. Biol.* **19**, 299–308.
Elvin, P., Slynn, G., Black, D., Graham, A., Butler, R., Riley, J., Anand, R. and Markham, A. F. (1990). *Nucl. Acids Res.* **18**, 3913–3917.
Errampalli, D., Patton, D., Castle, L., Mickelson, L., Hansen, K., Schnall, J., Feldmann, K. and Meinke, D. (1991). *Plant Cell.* **3**, 149–157.
Fabri, C. O. and Schäffner, A. R. (1994). *Plant J.* **5**, 149–156.
Feinberg, A. P. and Vogelstein, B. (1983). *Anal. Biochem.* **132**, 6–13.
Feldmann, K. A. (1991). *Plant J.* **1**, 71–82.
Feldmann, K. A. and Marks, M. D. (1987). *Mol. Gen. Genet.* **208**, 1–9.
Feldmann, K. A., Marks, M. D., Christianson, M. L. and Quatrano, R. S. (1989). *Science.* **243**, 1351–1354.
Ganal, M. W., Young, N. D. and Tanksley, S. D. (1989). *Mol. Gen. Genet.* **215**, 395–400.
Geraghty, M. T., Brody, L. C., Martin, L. S., Marble, M., Kearns, W., Pearson, P., Monaco, A. P., Lehrach, H. and Valle, D. (1993). *Genomics* **16**, 440–446.
Giovannoni, J. J., Wing, R. A., Ganal, M. W. and Tanksley, S. D. (1991). *Nucl. Acids Res.* **19**, 6553–6558.
Giraudat, J., Hauge, B., Valon, C., Smalle, J. and Parcy, F. (1992). *Plant Cell.* **4**, 1251–1261.
Green, E. D. and Olson, M. V. (1990). *Science* **250**, 94–98.
Green, E. D., Riethman, H. C., Dutchik, J. E. and Olson, M. V. (1991). *Genomics* **11**, 658–669.
Grill, E. and Sommerville, C. (1991). *Mol. Gen. Genet.* **226**, 484–490.
Gupta, M. and Hoo, B. (1991). *Plant Mol. Biol.* **17**, 361–369.
Guzman, P. and Ecker, J. R. (1988). *Nucl. Acid Res.* **16**, 11091–11105.
Hauge, B. M. and Goodman, H. (1992). *In* 'Methods in *Arabidopsis* Research'. (C. Koncz, N.-H. Chua, and J. Schell, eds). 191–223. World Scientific, Singapore.
Hauge, B. M., Hanley, S. M., Cartinhour, S., Cherry, J. M. and Goodman, H. M. (1993). *Plant J.* **3**, 745–754.
Heard, E., Davies, B., Feo, S. and Fried, M. (1989). *Nucl. Acids Res.* **17**, 5861.
Helentjaris, T., Slocum, M., Wright, S., Schaefer, A. and Nienhuis, J. (1986). *Theor. Appl. Gen.* **72**, 761–9.
Herman, P. L. and Marks, M. D. (1989). *Plant Cell* **1**, 1051–1055.
Hermanson, G. G., Hoekstra, M. F., McElligott, D. L. and Evans, G. A. (1991). *Nucl. Acids Res.* **19**, 4943–4948.
Hwang, I., Kohchi, T., Hauge, B. M., Goodman, H. M., Schmidt, R., Cnops, G., Dean, C., Gibson, S., Iba, K., Lemieux, B., Arondel, V., Danhoff, L. and Somerville, C. (1991). *Plant J.* **1**, 367–374.
Ioannou, P. A., Amemiya, C. T., Garnes, J., Kroisel, P. M., Shizuya, H., Chen, C., Batzer, M. A. and Dejong, P. J. (1994). *Nature Gen.* **6**, 84–89.

Kim, U.-J., Shizuya, H., de Jong, P. J., Birren, B. and Simon, M. I. (1992). *Nucl. Acids Res.* **20**, 1083–1085.
Kleine, M., Michalek, W., Graner, A., Herrmann, R. G. and Jung, C. (1993). *Mol. Gen. Genet.* **240**, 265–272.
Klein-Lankhorst, R. M., Vermunt, A., Weide, R., Liharska, T. and Zabel, P. (1991). *Theor. Appl. Gen.* **83**, 108–114.
Koncz, C. and Schell, J. (1992). *In* 'Methods in *Arabidopsis* Research' (C. Koncz, N.-H. Chua and J. Schell, eds). 224–273. World Scientific, Singapore.
Koncz, C., Martini, N., Mayerhoffer, R., Koncz-Kalman, Z., Körber, H., Rédei, G. P. and Schell, J. (1989). *Proc. Natl Acad. Sci. USA* **86**, 8467–8471.
Koncz, C., Mayerhofer, R., Konczkalman, Z., Nawrath, C., Reiss, B., Redei, G. P. and Schell, J. (1990). *EMBO J.* **9**, 1337–1346.
Koncz, C., Nemeth, K., Redei, G. P. and Schell, J. (1992). *Plant Mol. Biol.* **20**, 963–976.
Konieczy, A. and Ausubel, F. M. (1993). *Plant J.* **4**, 403–410.
Koornneef, M. (1986). 'Genetic Maps 1987: A Compilation of Linkage and Restriction Maps of Genetically Studied Organisms'. Cold Spring Harbour.
Koornneef, M. and Stam, P. (1987). *Arabidopsis Inf. Serv.* **25**, 35–40.http://genome-www.stanford.edu/Arabidopsis/ais/1987/koorn-1987-aabuy.html
Koornneef, M. and Stam, P. (1992). *In* 'Methods in *Arabidopsis* Research (C. Koncz, N. -H. Chua and J. Schell eds). 83–99. World Scientific, Singapore.
Kuhn, R. M and Ludwig, R. A. (1994) *Gene* **141**, 125–127.
Kwiatkowski, T. J., Zoghbi, H. Y., Ledbetter, S. A., Ellison, K. A. and Chinault, A. C. (1990). *Nucl. Acid Res.* **18**, 7191–7192.
Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E. and Newberg, L. (1987). *Genomics* **1**, 174–181.
Larin, Z., Monaco, A. P. and Lehrah, H. (1991). *Proc. Natl Acad. Sci. USA* **88**, 4123–4127.
Lazo, G. R., Stein, P. A. and Ludwig, R. A. (1991). *Bio/Technology* **9**, 963–967.
Lister, C. and Dean, C. (1993). *Plant J.* **4**, 745–750.
Liu, Y. G. and Whittier, R. F. (1995). *Genomics* **25**, 674–681.
Lloyd, A. M., Barnason, A. R., Rogers, S. G., Byrne, M. C., Fraley, R. T. and Horsch, R. B. (1986). *Science* **234**, 464–467.
Lovett, M., Kere, J. and Hinton, L. M. (1991). *Proc. Natl Acad. Sci. USA* **88**, 9628–9632.
Lu, X. H., Ruezinsky, D. M. and Giovannoni, J. J. (1995). *Plant Mol. Biol. Rep.* **13**, 369–376.
McCormick, M. K., Shero, J. H., Cheung, M. C., Kan, Y. W., Hieter, P. A. and Antoarakis, S. E. (1989). *Proc. Natl Acad. Sci. USA* **86**, 9991–5.
Mandal, A., Sandgren, M. and Palva, E. T. (1994). *In Vitro Cell. Dev. Biol.-Plant* **30P**, 204–209.
Marks, M. D. and Feldmann, K. A. (1989). *Plant Cell* **1**, 1043–1050.
Martin, G. B., Ganal, M. W. and Tanksley, S. D. (1992). *Mol. Gen. Genet.* **233**, 25–32.
Martin, G. B., Brommonschenkel, S. H., Chunwongse, J., Frary, A., Ganal, M. W., Spivey, R., Wu, T., Earle, E. D. and Tanksley, S. D. (1993). *Science* **262**, 1432–1436.
Maule, J. C., Porteous, D. J. and Brookes, A. J. (1994). *Nucl. Acids Res.* **22**, 3245–3246.
Mayerhofer, R., Koncz, K. R., Nawrath, C., Bakkeren, G., Crameri, A., Angelis, K., Redei, G. P., Schell, J., Hohn, B. and Koncz, C. (1991). *EMBO J.* **10**, 697–704.
Mindinos, M., Katagiri, F., Yu, G-L. and Ausubel, F. M. (1994). *Cell* **78**, 1089–1099.
Nam, H.-G., Giraudet, J., den Boer, B., Moonan, F., Loos, W. D. B., Hauge, B. M. and Goodman, H. M. (1989). *Plant Cell* **1**, 699–705.
Ochman, H., Gerber, A. S. and Hartl, D. L. (1988). *Genetics* **120**, 621–623.
Olszewski, N. E., Martin, F. B. and Ausubel, F. M. (1988). *Nucl. Acids Res.* **16**, 10765–10787.
Papp, A., Rougvie, A. E. and Ambros, V. (1991). *Nucl. Acids Res.* **19**, 623–630.
Parimoo, S., Patanijali, S. R., Shukali, H., Chaplin, D. D. and Weissman, S. M. (1991). *Proc. Natl Acad. Sci. USA* **88**, 9623–9627.
Pepper, A., Delaney, T., Washburn, T., Poole, D. and Chory, J. (1994). *Cell* **78**, 109–116.
Pruitt, R. E. and Meyerowitz, E. M. (1986). *J. Mol. Biol.* **187**, 169–183.
Putterill, J., Robson, F., Lee, K. and Coupland, G. (1993). *Mol. Gen. Genet.* **239**, 145–157.
Rech, E. L., Dobson, M. J., Davey, M. R. and Mulligan, B. J. (1990). *Nucl. Acids Res.* **18**, 1313.

Reiter, R. S., Williams, J. G. K., Feldmann, K. A., Rafalski, J. A., Tingley, S. V. and Scholnik, P. A. (1992). *Proc. Natl Acad. Sci. USA* **89**, 1477–1481.
Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. C. and Markham, A. F. (1990). *Nucl. Acids Res.* **18**, 2887–2890.
Rosenthal, A. (1992). *Trends Biotechnol.* **10**, 44–48.
Ruano, G., Brash, D. E. and Kidd, K. K. (1991). *Amplifications* **Dec.**, 1–6.
Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). 'Molecular Cloning: A Lab Manual'. Cold Spring Harbour.
Schlessinger, D. (1990). *Trends Gen.* **6**, 248–258.
Schmidt, R. and Wilmizer, L. (1988). *Plant Cell Rep.* **7**, 583–586.
Schmidt, R., Putterill, J., West, J., Cnops, G., Robson, F., Coupland, G. and Dean, C. (1994). *Plant J.* **5**, 735–744.
Schumacher, K., Ganal, M. and Theres, K. (1995). *Mol. Gen. Genet.* **246**, 761–766.
Segal, G., Sarfatti, M., Schaffer, M. A., Ori, N., Zamir, D. and Fluhr, R. (1992). *Mol. Gen. Genet.* **231**, 176–185.
Sheikholeslam, S. N. and Weeks, D. P. (1987). *Plant Mol. Biol.* **8**, 291–298.
Shero, J. H., McCormick, M. K., Antonarakis, S. E. and Hieter, P. (1991). *Genomics* **10**, 505–508.
Shizuya, H., Birrien, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y. and Simon, M. (1992). *Proc. Natl Acad. Sci. USA* **89**, 8794–8797.
Smith, D. R., Smyth, A. P. and Moir, D. T. (1990). *Proc. Natl Acad. Sci. USA* **87**, 8242–8246.
Sternberg, N. (1990). *Proc. Natl Acad. Sci. USA* **87**, 103–107.
Tanksley, S. D., Young, N. D., Paterson, A. H. and Bonierbale, M. W. (1989). *Bio/Technology* **7**, 257–264.
Tanksley, S. D., Ganal, M. W. and Martin, G. B. (1995). *Trends Genet.* **11**, 63–68.
Thomas, C. M., Vos, P., Zabeau, M., Jones, D. A., Norcott, K. A., Chadwick, B. P. and Jones, J. D. G. (1995). *Plant J.* **8**, 785–794.
Topping, J. F. and Lindsey, K. (1995). *Transgenic Res.* **4**, 291–305.
Topping, J. F., Wei, W. and Lindsey, K. (1991). *Development* **112**, 1009–1019.
Umehara, Y., Inagaki, A., Tanoue, H., Yasukochi, Y., Nagamura, Y., Saji, S., Otsuki, Y., Fujimura, T., Kurata, N. and Minobe, Y. (1995). *Mol. Breeding* **1**, 79–89.
Valvekens, D., Van Montagu, M. and Van Lijsebettens, M. (1988). *Proc. Natl Acad. Sci. USA* **85**, 5536–5540.
Van Lijsebettens, M., den Boer, B., Hernalsteens, J.-P. and Van Montague, M. (1991a). *Plant Sci.* **80**, 27–37.
Van Lijsebettens, M., Van der Haeghen, R. and Van Montagu, M. (1991b). *Theor. Appl. Gen.* **81**, 277–284.
Vijayraghavan, U., Siddiqi, I. and Meyerowitz, E. (1995). *Genome* **38**, 817–823.
Vizir, I. Y. and Korol, A. B. (1990). *Heredity.* **65**, 379–383.
Walbot, V. (1992). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43**, 49–82.
Ward, E. R. and Jen, G. C. (1990). *Plant Mol. Biol.* **14**, 561–568.
Weining, S. and Langridge, P. (1991). *Theor. Appl. Gen.* **82**, 209–216.
Wilson, Z. A., Tighe, P. J., Dawson, J. and Mulligan, B. J. (1992). *In Arabidopsis: The Compleat Guide* (D. Flanders and C. Dean, eds). AFRC PMB Programme. http://genome-www.stanford.edu/Arabidopsis/
Woo, S. S., Rastogi, V. K., Zhang, H. B., Paterson, A. H., Schertz, K. F. and Wing, R. A. (1995). *Plant Mol. Biol. Rep.* **13**, 82–94.
Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A. and Meyerowitz, E. M. (1990). *Nature* **346**, 35–39.
Zambryski, P. (1988). *Annu. Rev. Genet.* **22**, 1–30.

4 *Agrobacterium*-Mediated Transformation

RICHARD WALDEN

Max-Planck-Institut für Züchtungforschung, Carl-von-Linné-Weg 10, D-50829 Köln 30, Germany

METHODS IN PLANT BIOCHEMISTRY Vol. 10b
ISBN 0-12-461021-8

I. INTRODUCTION

Tumour formation on wounded dicotyledonous plants by *Agrobacterium* has been one of the most studied examples of interspecies gene transfer. The investigation of the molecular basis of tumour formation has been significant for at least two reasons. First, the mechanisms underlying tumour establishment and sustained neoplasmic growth have provided valuable insight into how the plant cell responds to plant growth substances. Second, the process by which a specific piece of plasmid DNA, the transferred or T-DNA, is transferred to the plant genome, while itself an intriguing phenomenon, has led to the development of strategies for the production of transgenic plants.

Plant transformation mediated by *Agrobacterium* results in the generation of plants which can contain, within the nuclear genome, single unrearranged inserts of defined sequence. This characteristic has been considered as advantageous in comparison with transgenic plants prepared using naked DNA uptake where multimeric and rearranged DNA inserts may be obtained. *Agrobacterium*-mediated gene transfer has become a routine method with many dicotyledonous species, and it is now clear that *Agrobacterium* can also transfer DNA to the cells of monocotyledonous plants, raising the possibility that it may also prove useful in transforming cereals, which make up the world's most important crop plants. Hence, *Agrobacterium*-mediated transformation might be considered as a general means of inserting foreign DNA into the cells of all plant species. The difficulty in such a widespread application of *Agrobacterium*-mediated transformation, however, remains in regenerating whole plants from the initial transformed cell.

In this chapter I will attempt, in broad terms, to describe the background to *Agrobacterium*-mediated transformation, concentrating on those aspects relevant to the use of *Agrobacterium* to create transgenic plants. My aim is not to discuss the process of transformation in detail – this has been ably done by others recently (Zambryski, 1992; Gelvin, 1993; Zupan and Zambryski, 1995) – but rather to describe the implications and the practicalities of generating transgenic plants.

II. *AGROBACTERIUM* AND THE TI PLASMID

The establishment of a plant cell line containing the T-DNA is the culmination of a complex and still little understood interaction between bacteria normally present in the rhizosphere and a wounded plant cell. *Agrobacterium tumefaciens* incites the growth of tumours, whereas infection by the closely related *Agrobacterium rhizogenes* results in a disease called hairy root. The ability of *Agrobacterium* to form tumours, or hairy roots, is dependent on the presence, within the bacteria, of an approximately 200 kb plasmid, the tumour-inducing Ti plasmid, or root-inducing Ri plasmid. The processes of transformation carried out by *A. tumefaciens* and *A. rhizogenes* are similar. For simplicity, Ti and Ri plasmids can be divided into three functional domains relevant in plant transformation.

A. The Origin of Replication and Plasmid Maintenance Functions

The Ti or Ri plasmid is maintained in *Agrobacterium* as a low, or single, copy plasmid. The plasmid is lost from populations of bacteria growing at temperatures above 30°C. This has the practical consequence that the plasmid itself is difficult to work with when compared with high-copy *Escherichia coli* plasmids, and the bacteria need to be cultured at 28°C in order for the Ti (or Ri) plasmid to be maintained.

B. The Virulence or *vir* Region

The *vir* region encodes proteins responsible not only for sensing wounded plant cells but also for mediating in the transformation process itself, and mutations in this region generally reduce the transforming ability of the host bacterium. The *vir* region spans approximately 30 kb, and is subdivided into six or seven complementation groups depending on the strain of the bacteria: *vir*A, B, C, D, E, F and G.

C. Transferred or T-DNA

This is the segment of DNA which is transferred from the bacterium to the genome of the plant cell and is responsible for tumour or hairy root formation. Depending on the strain of bacteria, the T-DNA comprises one (nopaline strains of *A. tumefaciens*), two (*A. rhizogenes*) or three (octopine strains of *A. tumefaciens*) T-DNAs. The T-DNA encodes enzymes which short-circuit the normal auxin and cytokinin biosynthetic pathways of the plant cell, resulting in uncontrolled cell proliferation. T-DNA is physically delimited by 25 bp almost perfect repeats, the so-called border repeats. Transfer of T-DNA is polar, being initiated at the right border repeat. The right border sequence in turn is flanked by 'overdrive', or an equivalent, which is a sequence acting to optimise the production of T-DNA transfer intermediates. The genes normally encoded by the T-DNA are not necessary for the transfer process itself and can be replaced by foreign DNA without apparently interfering with the process of transformation. However, the presence of the border sequences is required to both define and delimit the plasmid sequences which are transferred to the plant genome.

III. CULTURE OF *AGROBACTERIUM*

A. tumefaciens can be cultured and handled in essentially the same manner as *E. coli* except that all culturing is carried out at 28°C. The culture medium we use is YEB (Table 4.1) solidified with 1.5% Bacto-agar. The bacteria generally are 'stickier' than *E. coli*, so streaking for single colonies may require some care. The antibiotics used for selection depend on the bacterial strain and plasmids they contain (see Chapter 5), but the routine concentrations used are shown in Table 4.2.

IV. *AGROBACTERIUM*-MEDIATED TRANSFORMATION

The transformation process itself, which is mediated by genes encoded by the bacterial chromosome as well as those encoded by the Ti plasmid, can be divided into several stages (Fig. 4.1).

A. Bacterial Recognition of the Wounded Plant Cell

Tumour formation is initiated by the *Agrobacterium* recognising and responding to factors such as phenolics, sugars, amino acid compounds and secondary plant metabolites released by wounded plant cells. These stimulate chemotaxis of the bacteria to the wound site and activate the molecular events underlying T-DNA transfer. Conversely, the

Agrobacterium produces plant hormones which, if the wounded tissue itself has not already done so, trigger cell division, thus making the recipient cell 'competent' for transformation. Direct contact between *Agrobacterium* and the wounded plant cell is required for transformation to proceed. The production of sugar-binding proteins and surface exopolysaccharides by the bacteria and the presence of cellulose microfibrils on the plant cell surface serve to optimise contact between the bacterium and the plant cell.

TABLE 4.1. Bacterial media.

YEB		LB	
Bactoyeast extract	1 g l^{-1}	Bactoyeast extract	5 g l^{-1}
Bactobeef extract	5 g l^{-1}	Bactotryptone	10 g l^{-1}
Peptone	5 g l^{-1}	NaCl	10 g l^{-1}
Sucrose	5 g l^{-1}	pH 7.0	
$MgSO_4$	2 × 10^{-3} mol l^{-1}		
pH 7.2			

TABLE 4.2. Selective concentrations of antibiotics used in bacterial culture.

Antibiotic	Selective concentrations		Stock concentration (mg ml^{-1})
	E. coli	*A. tumefaciens*	
	(μg ml^{-1})	(μg ml^{-1})	
Ampicillin	50	–	25
Carbenicillin	100	100	50
Kanamycin	100	100	100
Tetracycline	15	2.5	12.5[a]
Rifampicin	100	100	50[b]
Streptomycin	12.5	300	900
Spectinomycin	50	100	50

[a]Stock in methanol.
[b]Stock in ethanol.

B. Activation of the *vir* Genes

The products of the *vir* region mediate the process of T-DNA transfer to the plant cell. The process is activated by the *vir*A and *vir*G products. Both *vir*A and *vir*G are constitutively expressed; *vir*A encodes a transmembrane protein kinase which is not only autophosphorylated but also upon recognising signals released by the wounded plant cell phosphorylates the *vir*G product. The phosphorylated *vir*G product in turn acts as a transcriptional activator of itself and the remaining *vir* genes.

C. Appearance of T-DNA Transfer Intermediates

Transfer of the T-DNA from the Ti plasmid to the plant cell starts by the appearance in the bacteria of single-stranded nicks in the left and right border sequences of the T-DNA, followed by the appearance of a single-strand copy of the T-DNA – the T strand – which

is thought to be the T-DNA transfer intermediate. The *vir*D1 and *vir*D2 products are essential for the appearance of the T strand. *vir*D2 encodes an endonuclease activity, and is required for the generation of nicks within the border sequences and single-strand DNA synthesis; *vir*D1 encodes a helicase activity which is also required for T-strand production. The product of *vir*C1 appears to interact with overdrive to enhance production of the T strand, but whether or not the *vir*C1 product interacts directly with *vir*D2 is not known. The T strand is coated within the bacterium by the *vir*E product, which is a single-stranded DNA-binding protein. In addition, the *vir*D2 product is covalently bound to the 3′ end of the T strand. These bound proteins most probably act to protect the T strand from degradation within the bacteria.

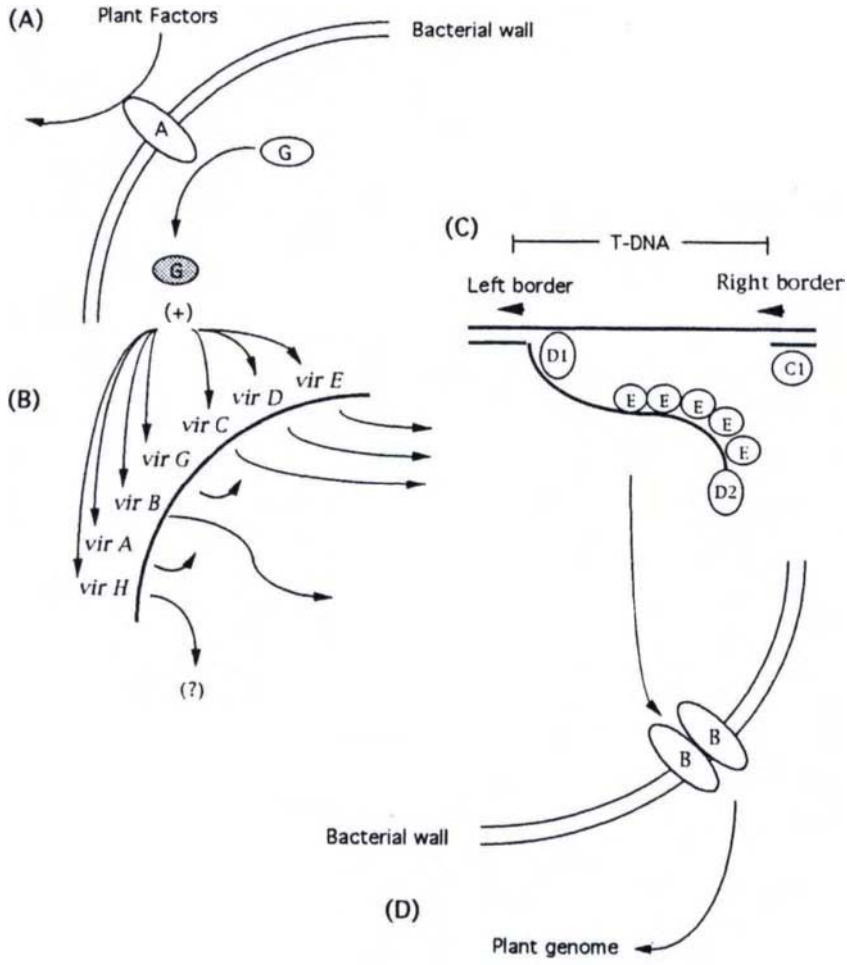


FIG. 4.1. Schematic diagram of the T-DNA transfer process. (A) Plant factors released by wounded cells are sensed by *vir*A, which phosphorylates *vir*G. (B) Phosphorylated *vir*G activates expression of the *vir* region. (C) The *vir* products mediate T-strand formation and (D) the T strand is transferred to the plant cell via the *vir*B products.

D. Transfer of the T Strand to the Plant Cell

The process by which the T strand complexed with *vir*E and *vir*D2 is transferred to the plant cell is not understood, although there are compelling reasons to suspect that the process is similar to bacterial conjugation. The production of nicks is the initiation step of both T-strand transfer and DNA transfer during conjugation; sequences involved with both processes are similar and, indeed, a conjugational transfer origin has been found to serve as an origin for plant cell transformation. The process of transfer itself is mediated by proteins encoded by the *vir*B operon, many of which are membrane associated and bear similarities to proteins involved in the process of conjugation.

E. Integration of T-DNA into the Plant Genome

Our understanding of how the T strand passages through the plant cell and integrates within the genome remains fragmentary. The *vir*D2 and *vir*E products which are bound to the T strand both contain signals which can target the protein to the plant cell nucleus, and it has been suggested that these act also to target the T-strand–protein complex. Once within the nucleus the *vir*D2 endonuclease activity could act to nick the plant genome, thus allowing invasion of the T strand. By comparing plant genomic DNA sequences before and after integration of the T-DNA, it has been proposed that the right T-DNA border 'scans' regions of the genome for a sequence of limited homology and subsequently uses base pairing to anchor itself to the genome, while resident plant DNA is displaced by the remaining T strand. DNA synthesis, strand repair and ligation of the T-DNA into the genome may be carried out by host enzymes. During the whole process flanking host DNA may be deleted or duplicated. Significant, however, is the finding, using T-DNAs containing a marker gene lacking a promoter sequence, that the T-DNA preferentially inserts into regions of the genome which may be transcribed or replicated. This may be explained by the genomic DNA in such cases being more amenable to the invasion of the T-DNA than when quiescent.

Once integrated into the genome, the T-DNA is stable and can segregate as a single Mendelian marker. Unrearranged inserts can be obtained, although tandems and partial inserts of T-DNA are often observed. It appears that the pattern of integration varies according to the combination of host bacteria and vector, as well as the transformation protocol used.

Following the natural transformation process, the genes carried by the T-DNA become expressed, and their products deregulate the internal cellular levels of auxin and cytokinin as well as the response of the cell to these growth substances. It is intriguing that the promoters responsible for this are bacterial, yet they have been found to be finely tuned to direct expression in plant cells.

From our current knowledge of the natural process of transformation, several points are directly relevant to optimising the process for the production of transgenic plants:

(1) *Induction of the transformation process.* The process of T-strand generation may be primed in the bacteria prior to application to the plant tissue by addition to freshly growing cultures of any of the factors (phenolics, sugars, etc.) known to be released from the wounded plant cell.

(2) *Optimisation of virA/virG activity.* As the *vir*A and *vir*G products control expression of the *vir* region, these play a prime role in the efficiency of the transformation process. For example, *vir*A has a precise pH optimum (5.2–5.8) which may have an effect on host range. In addition, overexpression of *vir*A and *vir*G prod-

ucts achieved either by employing overexpression mutants or high gene copy number serves to increase both host range and virulence.

(3) *T-strand formation.* The presence (or absence) of an 'overdrive' sequence in different bacterial strains may affect the efficiency of T-strand formation. This becomes of particular importance in vector construction (see Chapter 5) where elements from Ti plasmids from differing bacterial strains may be exchanged.

(4) *Target cell competence.* The ability to transform cells efficiently is apparently linked to the target cell undergoing cell division. In using tissue explant inoculation this process may be initiated prior to exposure to the bacteria by incubation in the presence of the appropriate growth hormones to initiate callus formation from the wound sites.

V. THE GENERATION OF TRANSGENIC PLANTS BY *AGROBACTERIUM*-MEDIATED TRANSFORMATION

Central to the production of transgenic plants is the initial requirement of *Agrobacterium* to be able to transfer its DNA to a wounded plant cell, followed by the ability of that cell to regenerate into a plant. Currently three general methods of *Agrobacterium*-mediated transformation are available: protoplast co-cultivation, explant inoculation and vacuum infiltration. The first two are described in detail below; the third is described in Section VII.

A. Protoplast Co-cultivation

This involves the isolation of mesophyll protoplasts followed by their cultivation in liquid culture with a diluted, freshly grown population of bacteria. Co-cultivation is allowed to proceed for approximately 2 days, after which the majority of bacteria are separated from the protoplasts by differential centrifugation, and washed protoplasts are cultured further in the presence of an antibiotic, which acts to stop further multiplication of the bacteria. Callus growth is induced by culture of the protoplasts in appropriate levels of auxin and cytokinin. Manipulation of the ratios of these growth substances is subsequently used to induce shoot formation from callus, and shoots are later removed and cultured further to induce root growth. The key to this approach is the growth of protoplasts for several days following isolation so that cell wall formation is initiated. This is important to allow attachment of *Agrobacterium* to the protoplasts. While an involved technique, with a transformation frequency of 10–20%, co-cultivation allows for the generation of extremely large numbers of individual transgenic cells. This has proven useful in studying the fate of the T-DNA strand within the plant cell during the transformation process, the mechanism of T-DNA insertion as well as generating populations of transgenic cells for screening in T-DNA insertional mutagenesis. However, a drawback with this technique is that it requires a protoplast system which can form callus, and ultimately plants, with high efficiency, and this currently limits the technique to plants of the solanaceous species.

1. Protocol: protoplast– Agrobacterium *co-cultivation* (modified from Depicker *et al.*, 1987)

- *Day 1.* Protoplast preparation (Negrutiu *et al.*, 1987). Leaf tissue from 6- to 8-week-old sterile SR1 tobacco plants is cut into approximately 1 cm^3 pieces, and 5 g

of this material is added to 20 ml of protoplasting solution comprising 1.5% cellulase R10 and 0.5% macerozyme in 0.4 M K3 medium (Table 4.3) (Nagy and Maliga, 1976). The mixture is incubated overnight in the dark at 26°C.

- *Day 2.* Digested leaf material is shaken briefly to release protoplasts and sieved through a 100 μm mesh and divided amongst 12 ml plastic centrifuge tubes. The tubes are spun at room temperature in a swing-out bucket centrifuge at about 3000 rpm for 6 min. The intact protoplasts band at the top of the tube, and the underlying solution and cell debris are carefully removed. The protoplasts are resuspended in 10 ml of 0.4 M K3 medium, mixed gently and the procedure repeated. Protoplasts are collected by centrifugation and the yield estimated microscopically with a cell counter. Routinely, this method yields 10^6 protoplasts per 10 ml of protoplasting solution. Using a pipette, 10^6 protoplasts are transferred into 94 mm Petri dishes, to give a final concentration of 10^6 protoplasts per 10 ml, and cultured in a final solution of 5 ml of 0.4 M K3 medium with auxin (1.0 mg^{-1}) and kinetin (0.2 mg^{-1}).
- *Day 4.* A fresh colony of *Agrobacterium* is used to inoculate 3 ml of YEB medium containing a single antibiotic, which ensures maintenance of the transformation vector within the bacterial population. The culture is incubated with shaking for 20 h at 28°C.
- *Day 5.* The 3 ml culture of *Agrobacterium* is used to inoculate 50 ml of YEB medium containing the selective antibiotic, and this is cultured further with shaking at 28°C for 20 h.
- *Day 6.* The density of the bacteria is estimated by measuring the optical density at OD_{600}. The amount of bacteria in microlitres to be added to 10^6 protoplasts is calculated as $100/1.8 \times OD_{600}$. This results in approximately 100 bacteria being inoculated for every protoplast. Co-cultivation is carried out at 26°C in the dark for 48 h.
- *Day 8.* Protoplasts are washed in W5 medium (see Table 4.3c) (or seawater) and then centrifuged at 700 rpm for 3 min in a swing-out bucket rotor centrifuge. The protoplasts are carefully resuspended in W5 medium, and 10^6 protoplasts are transferred to 94 mm Petri dishes and cultured in 10 ml of 0.4 M K3 medium containing 500 mgl^{-1} of claforan with kinetin (0.2 mgl^{-1}) and auxin (1 mgl^{-1}). At this point, selection for the growth of transgenic cells can be applied. Culture is carried out in the light for 2 days at 26°C, and then the protoplasts are embedded in 'low-melting' agarose. For each 10^6 protoplasts, embedding is carried out in a 14 mm Petri dish by adding, dropwise, 10 ml of hand-warm 1.6% agarose in 0.2 M K3 medium. The resulting solidified agar is cut into quarters and transferred to a 145 mm Petri dish to which 20 ml of W5 medium containing selective antibiotics and hormones is added. Culture is carried out at 26°C in the light, with the medium being changed every 2 weeks.

B. Explant Inoculation

This technique has proven to be the simplest and most versatile method in generating transgenic tissue. Here, tissue explants (leaf, stem, root, hypocotyl or tuber) are incubated with a fresh culture of *Agrobacterium* containing the transformation vector. Following the incubation period the majority of the bacteria are removed by washing, and the explants placed on agarose-solidified media containing an antibiotic to curtail further growth of *Agrobacterium*, a selective pressure to select growth of transgenic cells as well as growth substances at concentrations which induce callus formation. Callus formation is induced in the cells at the periphery of the explant, i.e. those that were initially in contact with *Agrobacterium*. Callus thus formed is removed, cultured further and shoot formation induced. Shoots are then removed and placed on root-induction medium. The

method of explant inoculation, in one form or another, has proven to be widely applicable to almost all dicotyledonous species, and is certainly the technique of choice when the aim of the experiment is routine transformation.

TABLE 4.3. Plant culture media.

(a) 0.4 M K3 medium.

NaH_2PO_4	150	mg l^{-1}	Microelements[a]	1	ml l^{-1}
$CaHPO_4$	50	mg l^{-1}	Vitamins[b]	1	ml l^{-1}
$CaCl_2$	900	mg l^{-1}	Inositol	100	mg l^{-1}
KNO_3	2500	mg l^{-1}	Xylose	250	mg l^{-1}
NH_4NO_3	250	mg l^{-1}	Sucrose[c]	136.9	g l^{-1}
$(NH_4)_2SO_4$	134	mg l^{-1}	Fe/EDTA[d]	5	ml l^{-1}
$MgSO_4$	250	mg l^{-1}			

Adjust to pH 5.6 with KOH

[a] Microelement stock solution:

KI	0.83 g l^{-1}
H_3BO_3	6.2 g l^{-1}
$ZnSO_4 \cdot 7H_2O$	10.6 g l^{-1}
$CaSO_4 \cdot 5H_2O$	25 mg l^{-1}
$Na_2MoO_4 \cdot 2H_2O$	0.25 g l^{-1}
$CoCl_2 \cdot 6H_2O$	25 mg l^{-1}

[b] Vitamin stock solution:

Glycine	400 mg/200 ml
Thiamine–HCl	20 mg/200 ml
Nicotinic acid	400 mg/200 ml
Pyridoxine HCl	400 mg/200 ml

[c] Sucrose: 136.9 g l^{-1} of sucrose = 0.4 M

[d] Fe in EDTA solution:
5.57 g l^{-1} of $FeSO_4 \cdot 7H_2O$
7.45 g l^{-1} of Na_2EDTA

(b) MS medium.

Macroelements	NH_4NO_3	1650 mg l^{-1}
	KNO_3	1900 mg l^{-1}
	$CaCl_2 \cdot 2H_2O$	440 mg l^{-1}
	$MgSO_4 \cdot 2H_2O$	370 mg l^{-1}
	KH_2PO_4	170 mg l^{-1}
Microelements	Na_2EDTA	37.3 mg l^{-1}
	$FeSO_4 \cdot 7H_2O$	27.8 mg l^{-1}
	H_3BO_3	6.2 mg l^{-1}
	$MnSO_4 \cdot H_2O$	16.9 mg l^{-1}
	$ZnSO_4 \cdot 7H_2O$	8.6 mg l^{-1}
	KI	0.83 mg l^{-1}
	$Na_2MoO_4 \cdot 2H_2O$	0.25 mg l^{-1}
	$CaSO_4 \cdot 5H_2O$	25 μg l^{-1}
	$CoCl_2 \cdot 6H_2O$	25 μg l^{-1}
Vitamins	*myo*-Inositol	100 mg l^{-1}
	Thiamine hydrochloride	0.1 mg l^{-1}
	Glycine	2.0 mg l^{-1}
	Nicotinic acid	0.5 mg l^{-1}
	Pyridoxine hydrochloride	0.5 mg l^{-1}

Sucrose 10 g l^{-1} (pH 5.8)

Solidified with 10 g l^{-1} of agar

Continued

(c) W5 medium.

154 mM NaCl
25 mM $CaCl_2$
5 mM KCl
5 mM glucose

pH 5.6–6.0

(d) *Arabidopsis* culture medium.

		ARA	ARA I	ARA II	ARA III
MS salts	(g l^{-1})	4.6	4.6	4.6	4.6
Sucrose	(g l^{-1})	30	30	30	30
Agar	(g l^{-1})	9	9	9	9
myo-Inositol	(mg l^{-1})	100	100	100	100
MES[a]	(mg l^{-1})	500	500	500	500
Thiamine	(mg l^{-1})	20	20	20	20
Nicotinic acid	(mg l^{-1})	1	1	1	1
Pyridoxine	(mg l^{-1})	1	1	1	1
Biotin	(mg l^{-1})	1	1	1	1
BAP[b]	(mg l^{-1})	–	1	1	–
NAA[c]	(mg l^{-1})	–	0.1	0.1	–
2,4-D[d]	(mg l^{-1})	–	–	0.5	–
2iP[e]	(mg l^{-1})	–	–	–	0.5
IAA[f]	(mg l^{-1})	–	–	–	0.15

Media are adjusted to pH 5.7 with KOH prior to autoclaving

[a] 4 morpho-lino-ethanol-sulfonic acid
[b] 6 benzylaminopurine
[c] naphthalene acetic acid
[d] 2,4-dichlorophenoxyacetic acid
[e] 2-isopentenyadenine
[f] indole-3-acetic acid

1. Protocol: explant inoculation I – tobacco leaf discs

- *Day 1.* Inoculate 5 ml of YEB containing the selective antibiotic with a single colony of *Agrobacterium* containing the transformation vector. Incubate at 28°C with shaking for 48 h.
- *Day 3.* Taking sterile plant material (SRI tobacco) grown for 6–8 weeks on MS medium, cut leaf discs of approximately 1 cm^3, place in 20 ml of liquid MS medium and add 200 µl of the fresh bacterial culture. Incubate for 10 min at room temperature. Carefully remove leaf discs from solution, shaking excess solution free, and lay topside down on solidified MS medium and incubate for 2 days at 26°C.
- *Day 5.* Transfer leaf discs to fresh MS medium containing claforan (500 mgl^{-1}), kinetin (0.2 mg l^{-1}), auxin (1.0 mg l^{-1}) and antibiotic to select for growth of transgenic cells. If at this point there has been a large growth of bacteria, the leaf discs can be washed in liquid MS medium containing 500 mg l^{-1} of claforan to remove excess bacteria. Leaf material is then incubated at 26°C, 16 h day length, for 1 week

and then transferred to a new plate of MS medium containing claforan, kinetin, auxin and selective antibiotic.

- *Day 28.* Between 3 and 4 weeks following incubation, callus should have formed at the periphery of the leaf discs, and shoots will begin to appear. Once these are 0.5–1.0 cm in size they can be removed and placed on MS medium with claforan but lacking auxin and kinetin; roots will begin to form.
- *Days 35–40.* Once sufficient root formation has occurred, the plantlets can be removed to pots and transferred to a greenhouse for further growth and genetic analysis.

2. *Protocol: explant inoculation II* – Arabidopsis *root inoculation* (Kemper *et al.*, 1992)

- Surface-sterilised *Arabidopsis* (ecotypes Columbia and Nossen) are germinated on ARA medium (Table 4.3) and either cultured on solid ARA medium for 4–5 weeks, or two to three plants are cultured in 50 ml of liquid ARA medium with shaking for 2–3 weeks under a 16 h light/8 h dark regimen at 20–22°C.
- Root tissue is removed to Whatman 3MM paper, cut into 1–2 cm pieces and gently abraded with a scalpel. These are then placed on ARA II medium and incubated in the light at 20–22°C for 4 days.
- A single colony of *Agrobacterium* containing the transformation vector is used to inoculate 5 ml of LB medium containing an antibiotic selective for the presence of the transformation vector and incubated for 20 h at 28°C with shaking.
- Bacteria are collected by centrifugation at 6000 rpm for 10 min and washed twice in ARA medium. The final bacterial pellet is then resuspended in 10 ml of ARA medium.
- The *Arabidopsis* root segments are placed in the solution of *Agrobacterium* for 20 min, removed, blotted dry on 3MM paper and transferred to ARA I medium; they are then cultured at 22°C until a light halo of bacteria is seen surrounding the root segments (3–4 days).
- The root segments are collected and washed twice with ARA medium containing either claforan (500 $\mu g\ ml^{-1}$) or vancomycin (500 $\mu g\ ml^{-1}$).
- Roots are then transferred to ARA I medium containing the appropriate antibiotic to select for growth of transgenic tissue and cultured in a 16 h light/8 h dark regimen at 22°C. Root segments are placed on fresh medium at 10 day intervals.
- Green callus should form at the periphery of the root segments, and when they are approximately 2 mm in size they are placed on ARA I medium containing either claforan or vancomycin (300 $\mu g\ ml^{-1}$) but lacking selective antibiotic for further callus growth and shoot induction.
- Shoots are transferred to ARA III medium containing selective antibiotics (at this point claforan and vancomycin should no longer be necessary to suppress growth of remaining *Agrobacterium*).
- When the shoots are approximately 1–2 cm in size, their stems are cut directly below the leaf nodes and are placed in a solution of indole-3-butyric acid (IBA, 1 $mg\ ml^{-1}$) for a few minutes. Shoots are then transferred to solid ARA medium, and roots should appear within 14 days. If roots do not appear, the IBA shock is repeated.

- Rooted plantlets are cultured further on ARA medium and transferred to 20 cm test tubes stoppered with cotton wool. Subsequently, the seed pods are removed and seeds collected.

VI. INDUCTION OF THE GROWTH OF HAIRY ROOTS

Regeneration of plants from roots produced by *A. rhizogenes* has proven to be a useful means of regenerating transgenic plants using species recalcitrant to plant regeneration using more routine tissue culture protocols. Root induction is dependent on the action of the root-inducing (*rol*) genes located within the T-DNA of the resident Ri plasmid. Here, bacteria containing an Ri plasmid are inoculated on to wounded hypocotyls or stems. Roots are formed at the wound site, and these can be cultured further. Hairy root-derived cultures in themselves can be used as a source of material useful in the production of secondary metabolites (Hamill *et al.*, 1987). Hairy roots under appropriate culture conditions can be induced to form secondary embryos from which transgenic plants can be formed. Plants recovered may contain the T-DNA of the Ri plasmid as well as the T-DNA of a transformation vector containing any foreign DNA. This can be determined by Southern blot analysis as well as observing the morphology of the leaves, as normally the action of the *rol* genes results in leaf wrinkling. This approach has been used particularly in the transformation of legumes, such as *Lotus corniculatus*, which is a model system used to study the molecular basis of symbiotic nitrogen fixation (Petit *et al.*, 1987).

VII. *IN PLANTA* TRANSFORMATION OF *ARABIDOPSIS*

Passage of plant tissue through culture carries the risk of inducing a variety of genetic and phenotypic changes which are grouped under the general name of somaclonal variation. In routine transformation where a large number of individuals bearing the same DNA construct may be generated, somaclonal variation is of little concern, with obvious variants being discarded. Nevertheless, somaclonal variants can be a potential problem if transgenics, generated as part of a mutagenesis programme, are to be screened for novel phenotypes. This requirement has been thrown into sharper focus with the emergence of *Arabidopsis* as a model plant for genetic analysis and the use of T-DNA as an insertional mutagen. To overcome this, *in planta* transformation has been adopted. Initially, *in planta* transformation of *Arabidopsis* involved imbibing seeds (T_1) in a fresh culture of *Agrobacterium*. Following imbibition, seedlings were grown to maturity, and seeds resulting from selfing the plants (T_2) were germinated and screened to select for transformants (by resistance to kanamycin), and finally these were selfed, with the resultant seedlings (T_3) being screened for mutant phenotypes. Seedling inoculation has been used with great effect to produce T-DNA tagged populations of *Arabidopsis* (Feldmann *et al.*, 1994); however, the precise mechanism of transformation is not understood, and the technique has proven difficult to reproduce effectively in different laboratories.

Vacuum infiltration is an extension of the *in planta Arabidopsis* transformation approach, and has proven to be a simple and reproducible method of creating transgenics (Bechtold *et al.*, 1993). It involves introducing *Agrobacterium* into the plant under vacuum, allowing the plants to recover, flower and self, and, finally, testing the progeny

for transformants. Transformation frequencies range from 1 to 10 transformants per treated plant, making this technique useful not only for generating transgenic plants but also for creating T-DNA tagged populations.

1. Protocol: in planta *transformation* – Arabidopsis *vacuum infiltration (*Bechtold *et al.,* 1993*)*

- Six-week-old *Arabidopsis* plants (grown under 16 h per day photoperiod, 15°C night/25°C day) that have had the first set of flowers removed so that the second flowers are forming are used in the experiment.
- A fresh culture of *Agrobacterium* in YEB media containing selective antibiotic is grown to an optical density (OD 600) of 0.8 and collected by centrifugation. The bacteria are resuspended in infiltration medium containing 10 $\mu g\ l^{-1}$ of 6-benzylaminopurine (BAP) to one-third the original culture volume.
- The pots containing *Arabidopsis* are lightly watered, to keep the soil in place (alternatively the plants can be grown through a fine plastic mesh), and inverted, with the plants being immersed in the bacterial solution.
- A vacuum is applied in a vacuum desiccator (approximately 10^4 Pa, though this is variable depending on the equipment used) for about 20 min.
- Treated plants are removed, gently rinsed with water and then removed to a greenhouse to allow further growth and seed set.
- Seeds are collected and sterilised by a 2 min incubation in 70% ethanol followed by an 8 min incubation with 10% NaOCl and 0.1% sodium dodecyl sulfate, followed by repeated washing with sterile water and drying in a sterile flow hood.
- Seeds are mixed, in a 10 ml plastic centrifuge tube, with 2 ml of room temperature top agar (water containing 0.2% low-melting agarose) and then plated on to Petri dishes of SM media containing 50 $\mu g\ ml^{-1}$ of kanamycin.
- Petri dishes are incubated at 4°C for 48 h and then placed in a growth chamber to allow the seeds to germinate and grow. Scoring for transformants can take place in 10–14 days.

VIII. ANALYSING TRANSGENIC PLANTS

Generally, the analysis of transgenic plant material is based on phenotypic and genetic changes in comparison with the untransformed wild type. Routinely this is initiated by growth of transgenic cells in the presence of a selection, the resistance of which is encoded by the T-DNA. Such dominant selection normally is applied at the stages of callus and root formation. While suggesting that the plant material is indeed transformed, growth under selective conditions is not without its pitfalls. Caution is required when the selective agent, normally an antibiotic or a herbicide, is unstable under conditions used in tissue culture. Moreover, after prolonged callus growth, cells may 'escape' selection; this is particularly a problem when shoot formation is induced from callus.

Confirmation that candidate transgenics indeed contain foreign DNA relies on Southern blot analysis. In addition to confirming that the tissue is transformed, this allows an assessment of the organization of the T-DNA itself. A detailed consideration of the possible results arising from a Southern blot analysis of T-DNA transformed plant

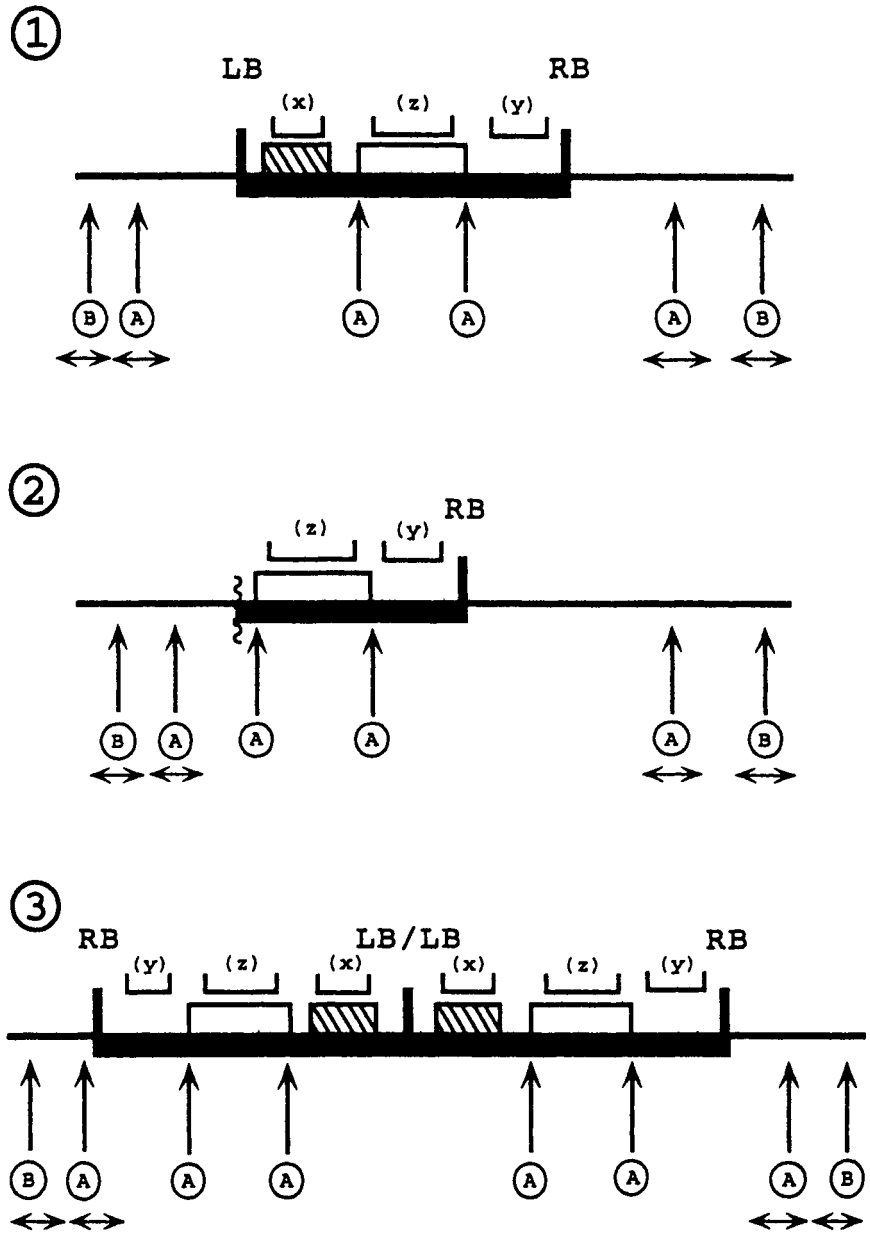


FIG. 4.2. Southern blot analysis of T-DNA inserts. The organisation of three different types of T-DNA routinely observed are shown: (1) a single unrearranged insert; (2) a partial insert; and (3) a tandem invert repeat. The T-DNA itself flanked by the left (LB) or right (RB) border sequences is represented by a thick line; flanking plant DNA by a thin line. Theoretical gene coding sequences are indicated by open and hatched boxes. Hybridisation probes representing differing regions of the T-DNA, (x), (y) and (z) are as indicated. Theoretical sites for restriction digestion (A) and (B) are shown; arrows beneath the sites indicate that the positions vary between different transformants.

tissue may be instructive. In Fig. 4.2, three possible scenarios are outlined: (1) a single, non-rearranged T-DNA insert; (2) a partial insert; and (3) an inverted tandem repeat. In each case, two hybridisation probes are used, (x) and (y) representing sequences located to the left and right portions, respectively, of the T-DNA. Digestion of the genomic DNA containing a single T-DNA insert with an enzyme that cuts at least once within the T-DNA will result in a single hybridising band in each case, regardless of whether probe (x) or (y) is used. Because the enzyme will also cut at an undetermined site within the genomic DNA, the size of the hybridising bands cannot be predicted and will vary between individual transgenics. Digestion of the genomic DNA with an enzyme cutting outside of the T-DNA (B) should produce a single band hybridising to both probes (x) and (y). The number of hybridising bands corresponds to the number of independent inserts. In the case of a partial insert, depending on the hybridisation probe and the restriction enzymes used, one border fragment will be missing. Considering that pseudo-border sequences may be present in the T-DNA, or that simple DNA strand breakage may truncate the T-DNA, this occurs with some regularity. It is to be noted that because of dominant selection normally applied during callus formation, the minimal amount of T-DNA present is likely to encode the dominant marker. Bearing in mind that T-DNA transfer and insertion into the plant genome proceeds from the right border in the direction of the left border it could be advantageous in vector design to locate the dominant marker at the left border to ensure recovery of the complete T-DNA. Hybridisation of DNA from a tandem inverted insert with probe (y) will produce two bands corresponding to the left and right border fragments, while probe (x) will produce a single hybridising band of predictable size. In this case, cutting the genomic DNA outside of the T-DNA will produce a single band greater than the sum of the size of a double T-DNA and which hybridises with both probes (x) and (y).

Often in Southern blot analysis, experimenters limit themselves to digesting genomic DNA with enzymes which cut within the T-DNA and hybridise with an internal hybridisation fragment (z). While indicating that a specific sequence is indeed present within the plant genomic DNA a brief look at Fig. 4.2 indicates that this reveals no information concerning the complexity of the insert. Moreover, such hybridisations carry the inherent danger of being unable to distinguish between *bona fide* plant DNA containing a T-DNA insert and contaminating DNA from inoculated *Agrobacterium*. This latter point is also the case where polymerase chain reaction (PCR) is used in an attempt to establish that plant material is indeed transformed.

Expression of a screenable marker, such as CAT, GUS or LUX (see Chapter 5), can also be used to indicate that plant tissue is transformed. At first, concern was raised that many plant-specific promoters are expressed, albeit at low levels, in *Agrobacterium*, suggesting that the enzyme activity seen might arise from contaminating bacteria. More recently, however, marker genes containing intron sequences unable to be processed in bacteria have been used to allay this concern.

Increasingly, plants are being transformed with genes whose products may disrupt the normal biochemistry of the cell, resulting in phenotypic changes when compared with untransformed controlled plants. Initially, publications described phenotypic changes scored from populations of primary transformants. However, it has become increasingly clear that primary transformants often display phenotypic changes arising from the effects of tissue culture. Such variation is generally not transmitted to progeny, hence it is good policy to use genetically segregating populations to demonstrate genetic linkage of the T-DNA to the changed phenotype. This is most easily carried out by crossing

primary transformants (T_1) with untransformed plants, selfing the progeny (T_2), determining which individuals of the subsequent generation (T_3) are homozygous, back-crossing these with untransformed plants and scoring the phenotype, segregating in all the resultant progeny.

IX. EXPRESSION OF GENE CONSTRUCTS IN TRANSGENIC PLANTS

Once integrated into the plant genome, foreign gene constructs containing a promoter functional in the plant cell become transcriptionally active. Routinely transferred gene constructs are expressed in a tissue-specific and developmentally regulated pattern associated with the promoter which has been used. However, once transformation strategies became established and were used to study expression directed by promoter derivatives, it became clear that differing independent transformants engineered to contain the same construct might express it at differing levels. This came to be known generally as the 'position effect', and, as a consequence, in experiments where transgenic plants are used to analyse the patterns of expression directed by differing promoter constructs it is generally necessary to measure gene expression in several independent transformants. Position effects are most likely the result of the transferred DNA coming under the influence of flanking plant DNA following integration. It could be envisaged that elements affecting gene expression, such as transcriptional enhancers or silencers, may play a role in this, or the higher-order structure of the chromatin might affect expression at different sites within the genome (Allen *et al.*, 1993).

Subsequently, most notably in analysing constructs which confer an easily screenable phenotype such as flower pigmentation, it has been found that both the expression of the introduced gene as well as the endogenous homologue may be silenced in transgenic tissue. This phenomenon has come to be known as 'co-suppression'. The phenomenon appears variable in that it apparently does not occur with all constructs. When it does occur it seems to be quite specific, only affecting homologous sequences. Three mechanisms have been proposed to explain co-suppression: (1) the unexpected production of antisense RNA; (2) methylation of the respective promoters, thus reducing transcription; and (3) pairing of repeated sequences so that transcription is prevented. However, the precise molecular basis of 'co-suppression' is currently not understood and remains a matter of some debate (Matzke and Matzke, 1995).

Finally, methylation may act to suppress expression of foreign genes in plant cells. Methylation has been correlated with silencing of gene expression during development, and apparently transgene expression may be inactivated by it as a result of either integration into a hypermethylated region of the genome or the susceptibility of the promoter in the construct to methylation. Interestingly, the suppression of expression by methylation can be developmentally reversible, and a hypermethylated allele appears to have a semi-dominant effect in inactivating a hypomethylated homologue (Meyer *et al.*, 1993).

X. *AGROBACTERIUM*-MEDIATED TRANSFORMATION – WHERE DO LIMITATIONS REMAIN?

A. Gene Targeting

The ability to direct a fragment of foreign DNA to a specific region of the plant genome could have two potential advantages over and above routine transformation. First, it could

allow transgenes to be placed in the same region of the genome as the resident homologous genes, and this may serve to confer additional controls on normal patterns of gene expression over and above those provided by the promoter linked to the transgene. Second, gene targeting could be used to 'knock out' or replace an endogenous gene with an engineered allele.

There have been two general types of experimental approach used to assess the potential of gene targeting using *Agrobacterium*-mediated transformation. One involves the repair of a mutant antibiotic resistance gene which has previously been stably inserted into a cell line by transformation. The other relies on the generation of a selectable phenotype should the incoming gene integrate into the endogenous homologue. In the former case, a tobacco line is engineered to contain a gene which normally produces growth in the presence of kanamycin but which is deleted at the 3′ end and hence is defective. Cells from this cell line are then transformed by protoplast co-cultivation with *Agrobacterium* containing a transformation vector encoding a kanamycin resistance gene deleted at the 5′ end. Calli were selected for growth in the presence of kanamycin. While the vast majority of the resultant kanamycin-resistant calli arose from the functional fusion of the 5′ deleted kanamycin gene to an active plant locus, 1.5% of the kanamycin-resistant calluses were found to be a result of fusion with the endogenous 3′ kanamycin resistance gene fragment (Offringa *et al.*, 1990). In the latter case, a non-engineered target sequence, acetolactate synthase (ALS), is used as a target for recombination. ALS catalyses the first step in valine, leucine and isoleucine biosynthesis and is the target of the herbicide chlorsulfuron. Selection for chlorsulfuron-resistant cells has led to the generation of cell lines which have mutations within the ALS gene, making them less sensitive to the herbicide. The resistant allele has a single amino acid change, and this confers resistance once transformed into protoplasts. Homologous recombination was investigated by transformation, again by protoplast co-cultivation, of a non-functional mutant allele into wild-type protoplasts and selecting for herbicide resistance. From seven independent chlorsulfuron-resistant clones, three had restriction patterns on Southern blot analysis, consistent with integration into the wild-type allele mediated by homologous recombination (Lee *et al.*, 1990). In both of the above-described cases, the frequency of recombination (approximately 10^{-3}–10^{-4}) is comparable with that seen by naked DNA uptake. The difficulty, however, with both of these approaches is that they are based on recovery, as a result of recombination, of a dominant selectable marker. Use of recombination in targeting plant genes is unlikely to allow such a selection. With this in mind, homologous recombination has recently been reported between a nuclear gene encoding a plant transcription factor and a T-DNA containing the gene, introduced *via* explant inoculation (Miao and Lam, 1995). This was the first report of a recombination event occurring at a non-selectable allele and it will be interesting to see whether this approach can be adapted for generalised gene targeting in plants.

B. *Agrobacterium*-Mediated Library Transfer

Although a variety of strategies have been developed to exploit the natural process of *Agrobacterium*-mediated gene transfer (see Chapter 5), the system lacks some of the flexibility of other eukaryotic transformation systems, for example of yeast. Currently, no means are available to use the system in a library transfer. This would be useful in the functional screening of cDNA libraries. While initial studies of the feasibility of transferring shotgun libraries to the plant genome have been described (reviewed in Walden,

1988), such approaches have not been widely adopted. Limitations in this appear to be the possible instability of the plant DNA within the bacteria as well as developing systems with a transformation frequency high enough to ensure complete library transfer to a population of plant cells followed by a feasible selection scheme, although rescue of genes by complementation in plant cells has been reported (reviewed in Gibson and Somerville, 1993).

C. Transfer of Large Segments of Genomic DNA

Using current *Agrobacterium*-mediated transformation strategies, only relatively small (< 40 kb) pieces of DNA can be transferred to the plant genome. This limitation results largely from difficulties in manipulating large DNA fragments into available transformation vectors. While for routine transformation this may not prove to be an obstacle, it is a problem in mutant complementation with candidate DNAs resulting from map-based cloning. Currently, each candidate DNA must be individually cloned into an *Agrobacterium*-based transformation vector for plant transformation, though a way around this might be provided by vectors containing the Cre–*lox* sequences (see Chapter 5). While there may be no theoretical limit to the size of DNA which may be transferred by the bacteria to the plant genome, the physical limitations of cloning, the random presence of pseudo-borders (i.e. sequences with similarity to the border repeats which may be nicked during T-strand formation) and the stability of an extremely large T-strand might limit the size of the DNA which can be transferred to the plant genome.

XI. CONCLUSIONS

In the last decade, plant transformation by *Agrobacterium*-mediated gene transfer has become routine for many plant species. While this chapter has described tobacco and *Arabidopsis* transformation in some detail, the concepts and ideas used have proven generally adaptable to other species. With this in mind, it is encouraging that a variation of *Agrobacterium* explant inoculation has proven effective in producing transgenic rice (Hiei *et al.*, 1994), raising the possibility that this approach may be useful in transforming other monocotyledonous plants.

Although involved, and generally requiring the manipulation of plant material in tissue culture, *Agrobacterium*-based transformation remains the only means of being able to transfer single non-rearranged sequences of DNA to the plant genome. Despite its wide-scale practical application as a means of generating transgenic plants, we still know surprisingly little concerning the transfer process itself. While we have begun to understand the molecular mechanisms of T-strand formation, the process by which the T-strand complex is transferred to the plant cell is little understood. In addition, our knowledge of how the T-strand complex ultimately passages through the plant cell to the nucleus and is integrated into the genome is fragmentary. Taking this into consideration, it may come as no surprise that *Agrobacterium*-mediated transformation, although generally applicable, may suffer from low efficiencies in many plant species. It may be expected that an increased understanding of the natural transformation process will lead to further improvements in optimising the process for practical purposes.

Taken together, however, the advances in *Agrobacterium*-mediated transformation in recent years have been such that the system might be considered applicable for most, if not all, purposes of plant transformation. Indeed, in some species, we may be reaching the stage that no further improvements in the system are necessary. This may allow us to concentrate less on improving transformation frequencies and more on understanding the process of transformation itself and transgene expression. With this in mind, possibly one of the most pressing issues concerning the practical application of transformation techniques is establishing the molecular basis of the variation in, and at times the inactivation of, transgene expression. This rather disconcerting phenomenon is likely to assume a profound importance now that transgenic plants, or their products, are entering the marketplace.

REFERENCES

Allen, G. C., Hall, G. E., Childs, L. C., Weissinger, A. K., Spiker, S. and Thompson, W. F. (1993). *Plant Cell* **5**, 603–613.

Bechtold, N., Ellis, J. and Palletier, G. (1993). *C. R. Acad. Sci. Paris, Life Sci.* **316**, 1194–1199.

Depicker, A. G., Herman, L., Jacobs, A., Schell, J. and Van Montagu, M. (1987). *Mol. Gen. Genet.* **210**, 477–484.

Feldmann, K., Malmberg, R. L. and Dean, C. (1994). *In 'Arabidopsis'* (E. M. Meyerowitz and C. R. Somerville, eds), pp. 137–172. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Gelvin, S. B. (1993). *In 'Transgenic Plants'*, Vol. I (S.-D. Kung and R. Wu, eds), pp. 49–88. Academic Press, New York.

Gibson, S. and Somerville, C. (1993). *TIBTECH* **11**, 306–313.

Hamill, J. D., Parr, A. J., Rhodes, M. J. L., Robins, R. J. and Walton, N. J. (1987). *Biotechnology* **5**, 800–804.

Hiei, Y., Ohte, S., Komari, T. and Kumashiro, T. (1994). *Plant J.* **3**, 271–282.

Kemper, E., Grevelding, G., Schell, J. and Masterson, R. (1992). *Plant Cell Rep.* **11**, 118–121.

Lee, K. Y., Lund, P., Lowe, K. and Dunsmuir, P. (1990). *Plant Cell* **2**, 415–425.

Matzke, M. A. and Matzke, A. J. M. (1995). *Plant Physiol.* **107**, 679–718.

Meyer, P., Heidmann, I. and Niedendorf, I. (1993). *Plant J.* **4**, 89–100.

Miao, Z.-H. and Lam, E. (1995). *Plant J.* **7**, 359–365.

Nagy, J. I. and Maliga, P. (1976). *Z. Pflanzenphysiol.* **78**, 453–455.

Negrutiu, I., Shillito, R., Potrykus, I., Biasini, G. and Sela, F. (1987). *Plant Mol. Biol.* **8**, 363–373.

Offringa, R., de Groot, M. J. A., Haagsman, H. J., Does, M. P., van der Elzen, P. and Hooykaas, P. J. J. (1990). *EMBO J.* **9**, 3077–3084.

Petit, A., Stougaard, J., Kühle, A., Marcker, K. A. and Tempé, J. (1987). *Mol. Gen. Genet.* **207**, 245–250.

Walden, R. (1988). *'Genetic Transformation in Plants'*. Open University Press, Milton Keynes.

Zambryski, P. (1992). *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **43**, 465–490.

Zupan, J. R. and Zambryski, P. (1995). *Plant Physiol.* **107**, 1041–1047.

This Page Intentionally Left Blank

5 Vector Systems for *Agrobacterium*-Mediated Transformation

RICHARD WALDEN

Max-Planck-Institut für Züchtungforschung, Carl-von-Linné-Weg 10, D-50829 Köln 30, Germany

METHODS IN PLANT BIOCHEMISTRY Vol. 10b
ISBN 0-12-461021-8

I. INTRODUCTION

Agrobacterium-mediated transformation is a convenient means of transferring a specific unrearranged segment of foreign DNA to the nucleus of the plant genome. In the decade in which vectors for use in *Agrobacterium*-mediated transformation have become available, a wide variety have been designed to exploit this. Initial vectors were relatively difficult to work with, being large and only partially characterised. More recently, improved vector design has generated defined series of vector plasmids containing differing components and cloning sites which can be manipulated with ease.

Vectors for use in *Agrobacterium*-mediated transformation have proven to be very versatile and have been used for a variety of purposes, for example:

(1) *Promoter analysis.* While transient expression assays using protoplasts can be used to dissect the functional domains of a promoter, analysis of the expression patterns of promoter derivatives linked to a marker gene in stable transgenics remains the only means of determining the regions of a promoter directing tissue-specific, or developmentally regulated, gene expression. In such cases, the T-DNA vector used contains a plant-specific selectable marker gene and an assayable marker gene either linked to a minimal promoter (i.e. a promoter fragment able to initiate transcription but lacking sequences conferring regulated gene expression) or lacking a promoter altogether. Promoter constructs, routinely deletion derivatives, are cloned upstream from the assayable marker gene, whose expression is then assayed in transgenic tissue.

(2) *The phenotypic effect of expression of a particular gene.* An ever-increasing number of genes are being transferred to plants, the expression of which may change the phenotype of the transformed host. Well-known examples of this include genes conferring herbicide and viral tolerance, although more recently genes determining a specific developmental, or biochemical, effect have been used. In such examples the T-DNA contains a plant-specific selectable marker and a promoter able to direct the expression of the test gene. Routinely, promoters considered as approaching constitutive in their pattern of expression are used to direct expression of the gene under investigation.

(3) *The modification of expression of endogenous genes.* In contrast to producing plants expressing a novel gene construct, transformation can be used to disrupt the expression of an endogenous gene as the result of the expression of antisense RNA, or by co-suppression. Here, constructs used in the transformation are the same as in the previous point, although often to ensure high levels of expression of the construct the promoter employed contains multiple transcriptional enhancers.

(4) *Delivery of semi-autonomous genetic elements.* In addition to the transfer of defined genes to the plant genome to assess the effect of their expression, plant transformation has been used extensively to dissect the structure and function of plant viruses – those with genomic DNA, or cDNA copies of RNA viruses, as well as transposable elements. In the first instance either complete copies of viral genomes and subgenomic derivatives have been transferred to the plant genome, or multimeric constructs have been used. Where the construct allows viral 'escape' following transfer to the plant cell by intramolecular recombination or transcription, infection of the plant host ensues. This phenomenon has come to be known as 'agroinfection' or 'agroinoculation', and is a convenient means of efficiently infecting large numbers of host plants. This has proven to be useful in testing for resistance or susceptibility to infection, initiating infection with viruses which can-

not be experimentally propagated by mechanical inoculation and demonstrating that indeed *Agrobacterium* can transfer T-DNA to monocotyledonous plants. The observation that transposable elements are mobile in a heterologous host following transfer by the T-DNA has served as the impulse in using transformation not only as a means of studying the mechanism of transposition itself but also to carry out transposon tagging in host plants where endogenous transposons are not well characterised.

(5) *Gene isolation using T-DNA as a gene tag* (see also Chapter 3). By virtue of its insertion into the plant genome the T-DNA can be considered as an insertional mutagen. The tendency of the T-DNA to insert preferentially into potentially transcribed regions of the genome will cause recessive gene-disruption-mutations at the site of insertion. This has been used extensively to create mutations in *Arabidopsis*. More recently, the concept has been further refined using a T-DNA engineered to contain multiple transcriptional enhancers so that following insertion into the plant genome, flanking genes are overexpressed, producing dominant mutations.

(6) *Complementation analysis*. In comparison with other eukaryotic systems, the use of complementation analysis in plant cells has been relatively limited. Nevertheless, increasingly, complementation of mutant lines as a final step in the process of map-based cloning is achieving practical importance. A variety of vectors can be used for this, although the current limit to this is that each candidate genomic DNA must be cloned into the vector individually prior to the process of transformation. Moreover, there are physical constraints on cloning large fragments of DNA into plasmid vectors.

Considering the wide variety of uses to which *Agrobacterium*-based transformation has been applied, it is clearly beyond the scope of a single chapter to describe in detail all the T-DNA vectors that have been designed and the markers developed to be used with them. This has been previously thoroughly reviewed (Walden, 1988; White, 1992). Rather, here I aim to describe, in broad terms, the individual classes of vectors and markers, and discuss the general considerations which need to be borne in mind when designing a transformation experiment and where the limitations in the experimentation may lie.

II. VECTORS FOR USE IN *AGROBACTERIUM*-MEDIATED TRANSFORMATION

The design of *Agrobacterium*-based transformation vectors capitalises on our understanding of the molecular basis of the natural transformation process carried out by *Agrobacterium*. While each vector system may vary in detail, they all share several characteristics. First, DNA to be transferred to the plant genome needs to be flanked by the T-DNA border sequences – or their equivalent. Second, the transfer process requires an *Agrobacterium* containing an active *vir* region. Third, it is helpful if the DNA to be transferred to the plant genome contains selectable markers, functional in plant cells, in order to select specifically for the growth of transgenic tissue. Broadly, *Agrobacterium*-based transformation vectors can be divided into two types: co-integrative and binary.

A. Co-integrative Vectors

Co-integrative vectors are based on wild-type Ti plasmids from which the T-DNA genes responsible for the tumour phenotype have been removed, or replaced by DNA from an

Escherichia coli plasmid. Foreign DNA to be transferred to the plant genome is cloned into an intermediate vector containing a region homologous to that contained between the borders of the Ti plasmid. Cloning foreign DNA into the intermediate vector is carried out using conventional techniques with *E. coli*. The intermediate vector, containing foreign DNA, is transferred to the *Agrobacterium* containing the co-integrative vector by conjugation. Routinely, the intermediate vector, based on an *E. coli* replicon, is not transferable by conjugation to *Agrobacterium*. Hence, conjugation proceeds via an intermediate bacterium containing helper plasmids encoding ColE1 transfer functions (Van Haute *et al.*, 1983). Once present within the *Agrobacterium* the intermediate vector and helper plasmids are lost from the bacterial population because they lack a functional origin of replication. However, before loss from the *Agrobacterium*, at a defined frequency, the intermediate vector has the opportunity to recombine with the co-integrative vector by virtue of the shared common region of homology. This results in the integration of the intermediate vector into the Ti plasmid, and this event can be selected for directly by use of an appropriate selectable marker carried by the intermediate vector (Figs 5.1 and 5.2).

1. Transfer of intermediate vectors to Agrobacterium *by triparental mating*

DNA to be transferred to the plant genome is first engineered, using conventional methods, into an intermediate vector in *E. coli*. The resultant plasmid is then transferred to *Agrobacterium* by triparental mating. This is most simply performed by mixing freshly growing bacteria containing the intermediate vector, the helper plasmids and the co-integrative vector on a non-selective bacterial plate, followed by selection for the growth of recombinants. During the selection process care needs to be taken to select against growth of parental bacteria. This is most easily carried out by multiple streaking of bacteria from single colonies on selective plates.

Procedure

- Inoculate 5 ml of LB medium containing selective antibiotics singly with a freshly streaked colony of the donor (i.e. the bacterium containing the intermediate vector) and helper *E. coli* strains and incubate overnight at 37°C with shaking.
- Inoculate 5 ml of LB medium containing selective antibiotic with *Agrobacterium* containing the co-integrative vector and incubate with shaking overnight at 28°C.
- *Either* pipette 100 μl of each fresh culture onto an LB plate and mix by spreading, *or* drop together 20 μl of each fresh culture on to an LB plate and incubate plates at 28°C overnight.
- Collect bacteria by sterile loop from the conjugation plate and streak on to LB plates containing antibiotics to select for growth of *Agrobacterium*-containing sequences introduced into the intermediate vector.

2. Isolation of total nucleic acids from Agrobacterium

Because of their size and low copy number, it is not feasible simply to isolate co-integrative vectors from *Agrobacterium* in an analogous manner to high copy number plasmids from *E. coli*. Hence, in order to confirm that transconjugants contain a co-integrative vector containing the construct of interest to be transferred to the plant cell it is necessary to carry out an isolation of total nucleic acid followed by Southern blot analysis.

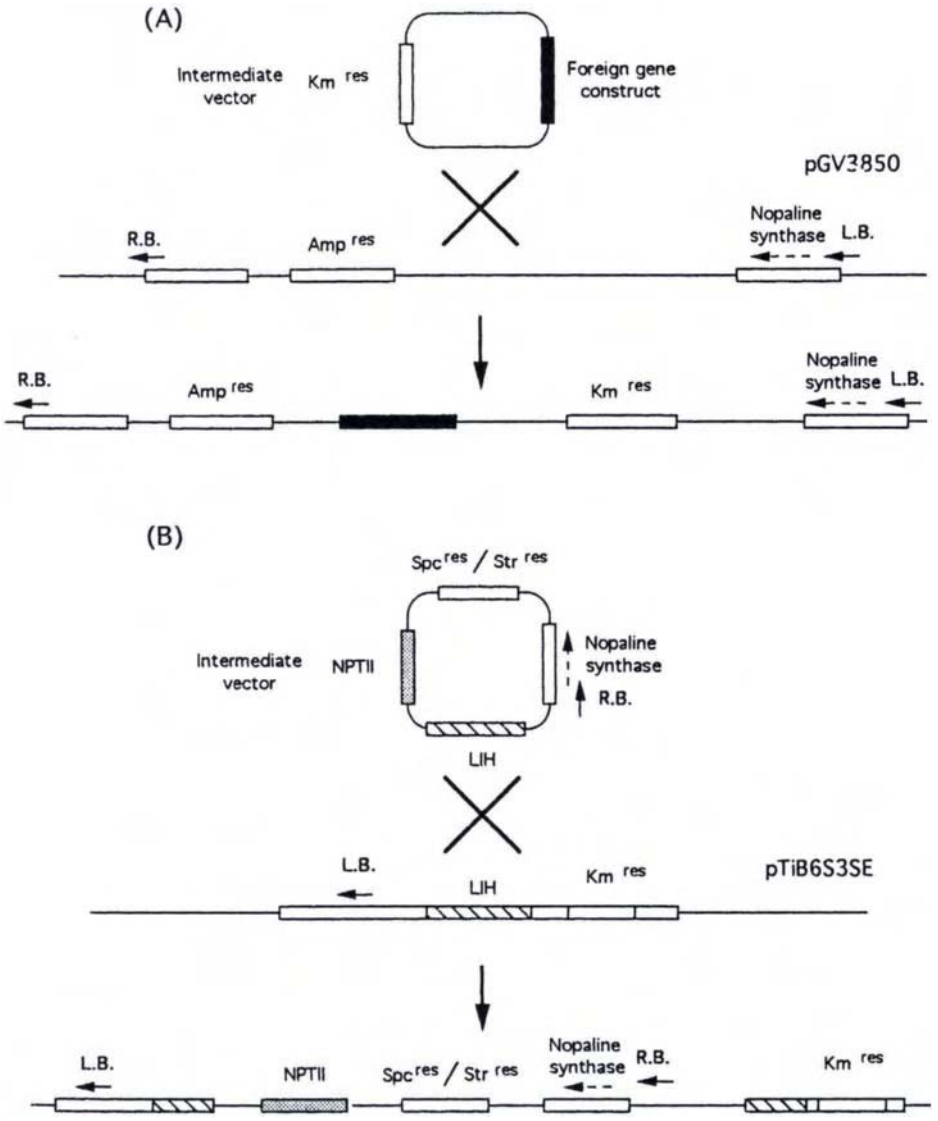


FIG. 5.1. Schematic representation of co-integrative vectors. (A) pGV3850 (Zambryski *et al.*, 1983). pGV3850 comprises a Ti plasmid in which the inner portion of the T-DNA is replaced by pBR322. At the left border (L.B.) is the nopaline synthase gene, whereas at the right border (R.B.) is a region encoding minor T-DNA transcripts. Intermediate vectors based on pBR322 and containing a kanamycin resistance gene are used in the system, with recombinants being selected for by growth on carbenicillin, encoded by the ampicillin resistance gene present on pGV3850, and kanamycin encoded by the intermediate vector. (B) The split end vector (SEV) system pTiB6S3SE (Fraley *et al.*, 1985). pTiB6S3 is an octopine Ti plasmid from which the right T-DNA has been removed and the majority of the left T-DNA has been replaced by a kanamycin resistance gene. The left border sequence remains as well as a region used for homologous recombination (LIH). The intermediate vector contains the same region (LIH) for homologous recombination, the right border sequence (R.B.), the nopaline synthase gene, a streptomycin/spectinomycin resistance gene allowing selection in bacteria and a kanamycin resistance gene functional in plant cells. Recombination between the region of homology carried by the intermediate and the co-integrative vectors results in the positioning of any foreign DNA to be transferred to the plant genome between the two border repeats and is selected for by growth of bacteria on streptomycin and kanamycin.

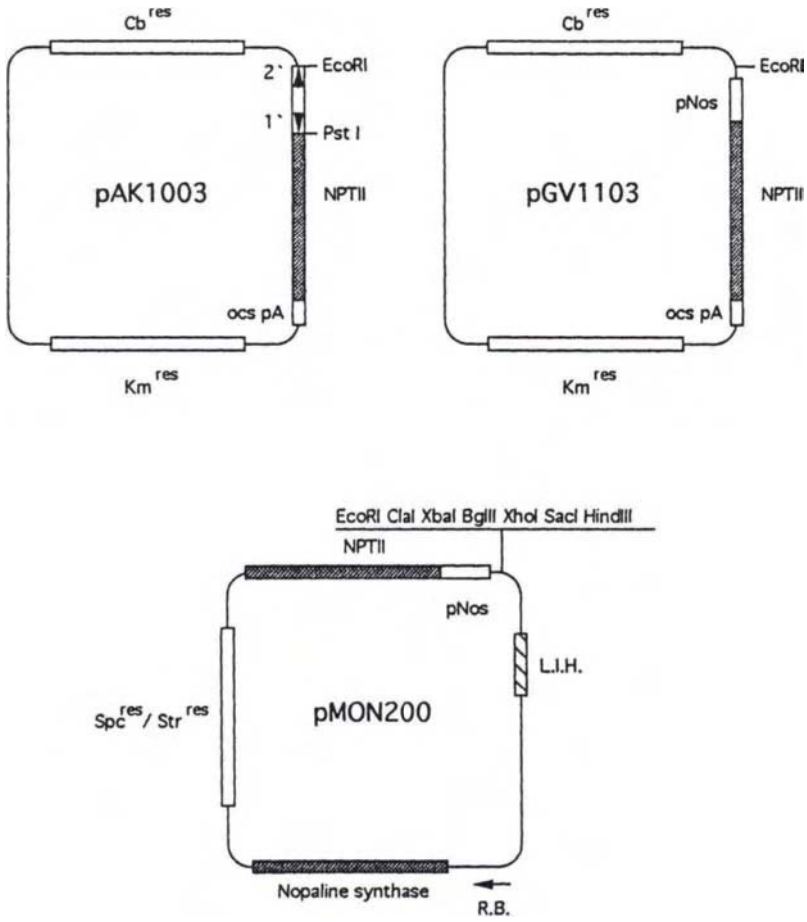


FIG. 5.2. Schematic representation of intermediate vectors. pAK1003 (Velten and Schell, 1985) is based on pBR322 and contains ampicillin and kanamycin resistance genes functional in bacteria and the 1′2′ dual promoter of pTiAch 5 arranged so that the 1′ promoter directs expression of the kanamycin resistance gene. An *Eco*RI site allows cloning of DNA fragments downstream from the 2′ promoter, and other vectors in this series are engineered to contain a variety of cloning sites. pGV1103 (Hain *et al.*, 1985) is related to pAK1003 but has the nopaline synthase promoter directing expression of the kanamycin resistance gene. pMON200 (Fraley *et al.*, 1985) is designed for use with the SEV system, and contains downstream from the right border sequence a nopaline synthase gene, a streptomycin resistance gene and a kanamycin resistance gene functional in plant cells. A multiple cloning site allows insertion of a variety of DNA fragments.

Procedure (modified from Dhaese *et al.*, 1979)

- Inoculate 2.5 ml of YEB medium with an appropriate selective antibiotic with a single bacterial colony and incubate with shaking at 28°C for at least 20 h.
- Transfer 1.5 ml of the culture to a plastic centrifuge tube and collect cells by centrifugation at 12 000 rpm for 5 min.
- Remove the supernatant and resuspend the cells well in 1 mM EDTA and 10 mM Tris–HCl (pH 8.0).
- Add 100 μl of 5% Sarkosyl in 1 mM EDTA and 10 mM Tris–HCl (pH 8.0), and mix.

- Add 150 μl of a 5 mg ml^{-1} pronase solution, mix and incubate at 37°C for 1 h.
- Add 500 μl of phenol: chloroform and, using a Gilson P1000 set at 1.0 ml, gently pass the mixture through a pipette tip to shear the viscous membrane mass present in the bacterial lysate.
- Centrifuge at 12 000 rpm for 5 min and remove the aqueous supernatant to a new 1.5 ml plastic centrifuge tube.
- Extract the aqueous phase twice by adding 500 μl of phenol: chloroform, mixing and centrifuging at 12 000 rpm for 5 min.
- Remove the aqueous phase from the phenol extraction to a new 1.5 ml plastic centrifuge tube and estimate the volume.
- Add 0.05 volume of 5 M NaCl and three new volumes of cold 95% ethanol and leave at –20°C for 1 h.
- Collect nucleic acid by centrifugation at 12 000 rpm for 10 min.
- Pour off the supernatant, wipe away the remaining ethanol with a tissue and resuspend the pellet in 50 μl water.
- Subsequent Southern blot analysis is carried out using standard protocols.

The process of engineering a co-integrative vector is somewhat involved because of both the size and the low copy number of the Ti plasmid. Despite this, co-integrative vectors have their advantages. Once the recombination event has occurred, the organisation of the foreign gene sequences within the T-DNA is stable (despite being flanked by regions of plasmid homology) and the Ti plasmid itself is stably maintained in the bacteria. This latter point is important when long periods of culture of the bacteria without direct selection for the presence of the plasmid are required, e.g. in explant or seedling inoculation.

B. Binary Vectors

Binary vectors are plasmids able to replicate in both *E. coli* and *Agrobacterium*. In addition to a selectable marker functional in both bacteria, they contain a plant-specific selectable marker gene and cloning sites located between T-DNA border sequences. Such vectors allow manipulation in *E. coli* followed by transfer into *Agrobacterium*. Transfer of the T-DNA carried by the binary vector to the plant genome is mediated by a Ti plasmid maintained in the host *Agrobacterium* from which the T-DNA and border sequences have been deleted. Routinely this helper Ti plasmid also encodes an antibiotic resistance gene, so its presence, through positive selection, is ensured. Binary vectors have proven to be both simple to use and versatile. Nevertheless, it needs to be borne in mind that some may not be readily maintained in the *Agrobacterium*, and constant selection may be required to ensure the continued presence of the binary vector in the bacterial population.

Routinely, binary vectors are based on the RK2 broad host range origin of replication and conjugation functions (Ditta *et al.*, 1980). A bacterial resistance gene allows for selection in both *E. coli* and *Agrobacterium*. Between the T-DNA border sequences are sequences allowing insertion of foreign DNA and selectable marker functional in plant cells (*Fig. 5.3*).

1. Transfer of binary vectors to Agrobacterium

Cloning foreign DNA in binary vectors is carried out in *E. coli* using standard recombinant DNA technology, and the resultant plasmid is transferred to *Agrobacterium* by con-

jugation (as described above), transformation or electroporation. We have found transformation to be the simplest method of transferring binary vectors to *Agrobacterium*, as it is rapid, overcomes the need for sequential selection following conjugation and does not require specialised equipment.

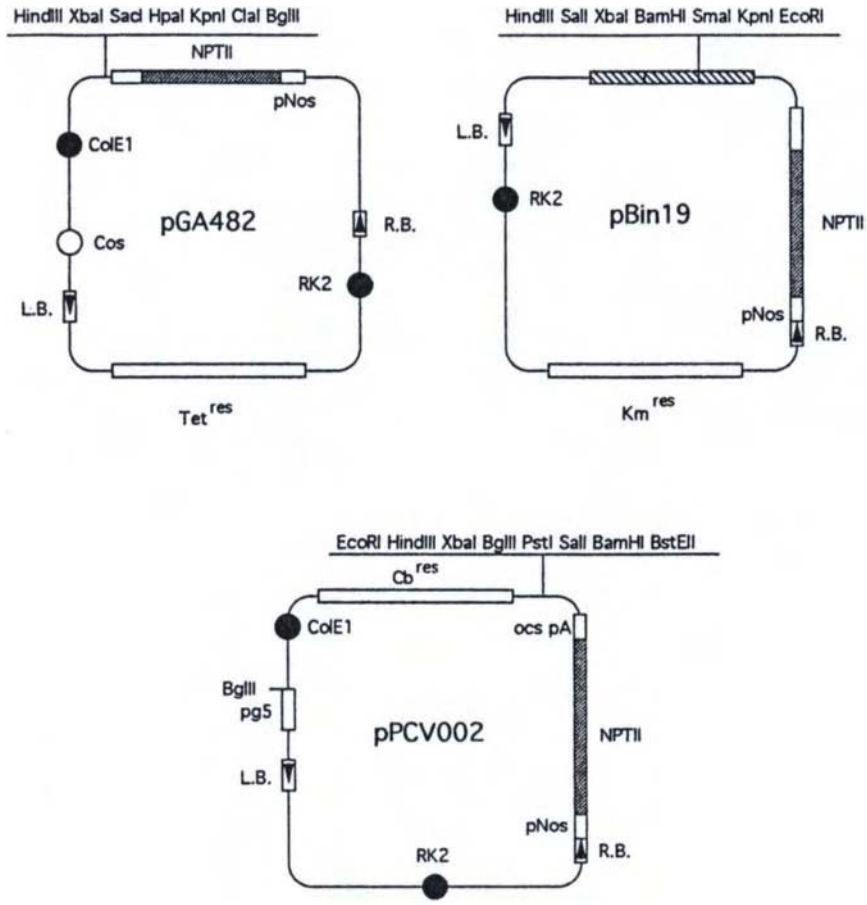


FIG. 5.3. Schematic representation of binary vectors. The pGA482 (An *et al.*, 1986) vector contains between the T-DNA border sequences a kanamycin resistance gene linked to the *nos* promoter upstream from a multiple cloning sequence. In addition to the ColEl origin of replication it contains a *cos* site to allow use of a λ packaging system. A tetracycline resistance gene allows selective maintenance in *Agrobacterium*. Outside of the T-DNA borders is a pRK2 origin of replication ori^{RK2} and origin of transfer ori^{T}. pBin19 (Bevan, 1984) is based on the broad host range replicon pRK252 and contains a kanamycin resistance gene functional in bacteria, T-DNA border sequences from pTiT37, a kanamycin resistance gene driven by the nopaline synthase promoter and multiple cloning sites located with the α complementary region of β-galactosidase. This allows selection of recombinant clones in appropriate strains of *E. coli* as being *lac*⁻ in the presence of 5-bromo-4-chloro-3 indoly β-D-galactoside (X-gal) and isopropyl-thiogalactoside (IPTG). The host *Agrobacterium* for pGA482 and pBin 19 is LBA4404, an octopine-type *Agrobacterium* containing a Ti plasmid from which the entire T-DNA has been deleted. pPCV002 (Koncz and Schell, 1986) contains between the border sequences a kanamycin resistance gene linked to the *nos* promoter, a multiple cloning sequence, an ampicillin resistance gene and an origin of replication functional in *E. coli* as well as a promoter derived from the *A. tumefaciens* T-DNA gene 5 promoter. It is maintained in *Agrobacterium* by origins of replication and mobilisation functions derived from RK2. The host bacteria for the PCV series of vectors is a C58 *Agrobacterium* derivative GV3101.

Protocol: transformation of Agrobacterium (modified from Ebert *et al.*, 1987)

- Inoculate 5 ml of YEP medium with a single colony of *Agrobacterium* containing the helper Ti plasmid and incubate with shaking overnight at 28°C.
- Add 2 ml of the overnight culture to 50 ml of YEP medium in a 250 ml flask which has been prewarmed to 28°C. Incubate with vigorous shaking (approximately 200 rpm) at 28°C until an optical density (OD_{600}) of 0.5–1.0 is obtained.
- Chill the culture on ice and pellet the bacteria by centrifugation at 6000 rpm for 5 min.
- Discard the supernatant and resuspend the cells in 1 ml of 20 mM $CaCl_2$ (ice cold) and dispense 0.1 ml aliquots to prechilled 1.5 ml plastic centrifuge tubes.
- The bacterial aliquots can either be used directly or frozen in liquid nitrogen and stored at –70°C.
- Add approximately 1 μg of plasmid DNA either to the freshly prepared bacteria and freeze directly in liquid nitrogen or to the frozen stored bacterial aliquot.
- Thaw the cells by incubation at 37°C for 5 min.
- Add 1 ml of YEP medium and incubate at 28°C for 2–4 h with gentle shaking.
- Collect bacteria by centrifugation for 30 s at 12 000 rpm and discard the supernatant. Resuspend the cells in 400 μl of YEP medium.
- Plate the cells in 200, 100 and 50 μl volumes on YEP agar plates containing antibiotic encoded by the binary vector (routinely carbenicillin (100 μg ml^{-1}) or kanamycin (100 μg ml^{-1}). Incubate at 28°C. Transformants appear 2–3 days following transformation.
- From a single colony, screen for the presence of the helper Ti plasmid and the binary vector by streaking on antibiotic plates containing appropriate antibiotics.

Protocol: introduction of DNA into Agrobacterium *by electroporation* (modified from Mattanovich *et al.*, 1989)

- A single colony is used to inoculate 3 ml of YEP medium and incubated for at least 16 h with shaking at 28°C.
- This culture is used to inoculate 300 ml of prewarmed YEP medium, and incubation is carried out at 28°C with shaking until an OD_{600} of 0.5 is reached.
- Chill the culture on ice and collect the bacteria by centrifugation at 6000 rpm for 10 min, and wash three times in 1 mM HEPES (pH 7.0) and once in 10% glycerol.
- Bacteria are resuspended in 3 ml of 10% glycerol, and 200 μl aliquots are frozen in a –70°C methanol bath and stored at –70°C.
- For electroporation, bacterial aliquots are thawed, and approximately 200 ng of plasmid DNA is added.
- From this suspension, 40 μl is placed in an electroporation cuvette with an electrode distance of 0.2 cm (Biorad).
- A single pulse of 2.5 kV using the 25 μF capacitor is used (Gene Pulser, Biorad).
- After the pulse, cells are transferred to 0.8 ml of YEP medium and incubated at 28°C for 1 h. Aliquots of this care spread on YEP medium containing selective antibiotic. Generally, the procedure produces 1×10^6–1.5×10^6 colonies per microgram of DNA.

2. Screening for the presence of binary vectors

Confirmation that the binary vector is indeed present in the bacteria can be achieved most easily using a modification of the alkaline lysis miniscreen procedure used with *E. coli* (Birnboim and Doly, 1979). It should be borne in mind that this procedure will also yield a variable amount of Ti plasmid, which can cause a background smear, and that the binary plasmid is present as a single or low-copy number in *Agrobacterium*. This latter point results in a low yield of DNA, when compared with miniscreens prepared from *E. coli*, and the enzymes used in the restriction analysis should be chosen so that they produce as few characteristic bands as possible to allow visualisation.

Protocol: Agrobacterium *miniscreen* (modified from Birnboim and Doly, 1979)

- Inoculate 3 ml of YEP medium containing the antibiotic whose resistance is encoded by the binary vector with a single bacterial colony and incubate with shaking for 20 h at 28°C.
- Transfer the culture to a 1.5 ml plastic centrifuge tube and pellet the bacteria by centrifugation at 12 000 rpm for 2 min.
- Remove the supernatant and resuspend the cells well in 0.1 ml of ice-cold 4 mg ml^{-1} lysozyme, 50 mM glucose, 10 mM EDTA and 25 mM Tris–HCl (pH 8.0), and incubate at room temperature for 10 min.
- Add 0.2 ml of 1% sodium dodecyl sulfate and 0.2 M NaOH, and mix well by inverting the tube vigorously. Incubate for a further 10 min at room temperature.
- Add 150 µl of 3M NaOH (pH 4.8). Mix well by inverting the tube.
- Centrifuge at 12 000 rpm for 3 min and pour the supernatant into a new 1.5 ml centrifuge tube.
- Fill the tube with cold 95% ethanol; mix by inverting the tube, and store at –20°C for 20 min.
- Collect the precipitate by centrifugation at 12 000 rpm for 10 min.
- Remove the supernatant and wash the pellet with 70% ethanol. Dry the pellet by vacuum desiccation and resuspend in 50 µl of 1 mM EDTA and 10 mM Tris–HCl (pH 8.0).

III. GENETIC MARKERS

Strictly speaking, a genetic marker is a gene construct with an ability, as a transgene, to confer a novel phenotype in a new host. By their very nature they are genetically dominant. Initially, the use of genetic markers in plant transformation was to fulfil two fundamental requirements. First, their expression set the criteria by which a plant cell was judged to be transformed by providing a novel phenotype, enzyme activity, or sequence that could be detected following DNA/RNA hybridisation. Second, they were used to assess expression levels directed by a particular promoter construct in transgenic tissue. While these points remain true today, both the range of markers available and the versatility of their use has expanded greatly.

The variety of plant genetic markers has been thoroughly reviewed recently (Bowen, 1993), and I do not intend to provide here an exhaustive list of markers. Rather, I aim to outline their general uses and discuss their application.

A. Selectable Markers

Selectable markers are used to select for the specific growth of transformed cells amongst a background of non-transformed individuals. This is important because transformation frequencies remain normally relatively low. Such markers allow growth, or at least viability, in the presence of the selective agent. Routinely, resistance to antibiotics or phytotoxins such as herbicides has been used. In the former case, antibiotic resistance genes derived from bacteria have been utilised (for example, kanamycin, hygromycin and streptomycin), whereas in the latter case genes encoding products which are more tolerant to herbicides, for example glyphosate, phosphinothricin and triazines, have been used. Generally, resistance or tolerance to a herbicide is provided by either the enzymatic modification of the selective agent or its sequestration, or by the overexpression of the target enzyme, or by reducing its affinity for the selective agent.

B. Negative Selection

In the previous example, selection is based on expression of a gene construct, and as such requires a promoter to be active at the time and location at which the selective pressure might be applied. Conversely, negative markers have been developed whereby promoters which are inactive at a specific time or location might be investigated. In this case, expression of the marker will lead to cell death. An example of this is provided by use of gene 2 of the T-DNA of *Agrobacterium tumefaciens* (Depicker *et al.*, 1989). The gene 2 product encodes an aminohydrolase, converting α-naphthalene acetamide to α-naphthalene acetic acid. α-Naphthalene acetic acid in high levels is toxic to plant cells, hence cells expressing the aminohydrolase should be unable to grow on media containing α-naphthalene acetamide. More recently, nucleases (Mariani *et al.*, 1990) and cell-specific toxins (Koltunow *et al.*, 1990) have been adopted for use. These latter markers have, when linked to tissue-specific promoters, been particularly useful ablating, or destroying, specific cell types, or tissues, due to their expression. This approach has been used most effectively to engineer male sterility (Mariani *et al.*, 1990).

C. Assayable or Screenable Markers

These types of marker genes encode enzyme activities which can be assayed with ease, either in plant extracts or *in situ*. While a simply assayed enzyme activity can be used to judge whether a plant cell or tissue is indeed transformed, the prime use of screenable markers is to judge differing levels of gene expression directed by specific promoter fragments. In the first instance, enzymes such as chloramphenicol transacetylase (CAT) and neomycin phosphotransferase (NPT II) were used to assess gene expression in either transfected protoplasts or intact tissue. Histochemical markers such as β-glucuronidase (GUS) or luciferase (LUX), or the green fluorescent protein (GFP) from jelly fish, are available and allow quite a precise measure of gene expression at a cellular level. The last

marker has the advantage that it can be visualised in the living cell in a non-destructive manner. Such markers have proven invaluable in the measurement of tissue-specific gene expression in transgenic plants.

Dominant selectable markers may also be used in screening for gene expression. For example, developing plant tissue grown in the presence of spectinomycin and not expressing a spectinomycin resistance gene will be bleached white due to the antibiotic disrupting chloroplast development. On the other hand, cells expressing the spectinomycin resistance gene will be green as a result of harbouring normal chloroplasts. Hence, tissue differentially expressing the spectinomycin resistance gene will be a mosaic of green and white tissue. An additional advantage of the use of this type of selection and screen is that both green and white tissue can be rescued and clonally propagated by transfer to media lacking spectinomycin.

D. Auxotrophic Markers

While auxotrophic mutants, i.e. those which have nutritional requirements, have been used extensively in bacterial and yeast genetics, in plants this is not the case, largely because of a lack of available mutants. Moreover, few can be propagated as homozygotes. In tobacco, a threonine dehydratase mutant requiring isoleucine and a nitrate reductase mutant have been complemented following transformation with a yeast gene or a tobacco gene, respectively (reviewed in Bowen, 1993). However, to date there have been no reports of gene rescue from a cloned library by complementation of a mutant phenotype. While advances in mutant generation and analysis in *Arabidopsis* will make this approach increasingly attractive, difficulties are likely to remain, not only in devising a means of efficiently transferring plant genomic libraries to plant cells but also in devising adequate selection schemes.

E. Morphological Markers

Initially, when plant transformation systems were being established, altered morphology proved to be a useful means of establishing that regenerated plant tissue was indeed transformed; for example, plants regenerated containing the *rol* genes of *Agrobacterium rhizogenes* display characteristic wrinkling. As more genes are isolated from the plant genome which modify normal patterns of plant development through hormone action, or through the activation/inactivation of differing developmental pathways via expression of homeotic genes, these could be used also as markers.

F. Pigmentation Markers

Anthocyanins producing pigmentation in both flowers, fruits and seeds are naturally powerful genetic markers. A detailed knowledge of the biochemistry of anthocyanin biosynthesis coupled with the cloning of the many genes involved has allowed the specific engineering of flower colour, particularly in the petunia. Antisense technology has been used extensively towards this end; however, it is particularly notable that simple visual screening of flower colour and patterning in transgenics has revealed curious effects of silencing of both transgenes and host genes. These effects have come to be known collectively as transinactivation, or co-suppression, and the molecular basis for them remains unclear. In

addition, the analysis of flower colour has revealed that transgene expression can be quite unstable. This was demonstrated by a mutant petunia line which normally produced white flowers but upon transformation with a maize dihydroflavonol reductase gene produced pale pink flowers (Meyer *et al.*, 1987). Detailed analysis of a large number of progeny, which normally would be considered to produce pink flowers, demonstrated that pigmentation failed to occur in a surprisingly large number of individuals, and this could be attributed to both methylation and, apparently, transactivation (Meyer *et al.*, 1993).

Not only the genes encoding the individual anthocyanin biosynthetic steps can be used to modulate pigmentation but also the transcriptional factors that control their expression. Two of these, encoded by the R and C gene families, have been cloned from maize. Cells expressing R and C alone do not synthesise anthocyanins; however, flavonoid pigments can be induced when both are active in the same cell and, interestingly, these factors appear functional in non-monocot cells, suggesting that they may be used widely in different cell types (Lloyd *et al.*, 1992).

IV. MARKERS FOR SPECIFIC PURPOSES

From the preceding discussion it should be clear that there is great flexibility in both the variety and uses of genetic markers which are currently available for use in plant cells. Indeed, it could be said that any gene whose product confers a novel phenotype on a transformed plant cell, be it biochemical, developmental or morphological, might be considered as being a genetic marker. It goes without saying that the choice and use of any marker depends largely on the aim of the experimenter, but several guidelines might be worth considering in adopting a marker for a specific purpose:

(1) *Availability of the resistance gene and selective agent.* It is obvious that without a required marker gene an experiment cannot proceed; however, it remains a sad fact that published marker genes, particularly from industrial sources, are often not freely distributed to the research community. Once obtained, the selective agent is also required, preferably in an affordable amount. This is particularly important in the use of herbicides as a selective agent. Spray application of herbicides may be of use in screening intact plants growing, for example, in a greenhouse; however, commercially available herbicide preparations may not be applicable with cells, tissue explants or whole plants grown in tissue culture because of the presence of toxic contaminants. In this case, access is required to the purified herbicide itself.

(2) *Stability of the selective agent.* To ensure adequate selection for the growth of transgenic tissue the selective chemical needs to be stable during the selection process. This is particularly important when callus or shoot formation in tissue culture takes time. Routinely under such circumstances, transfer of material to fresh selective conditions at 2- to 3-weekly intervals is recommended.

(3) *The marker used must allow selection, or screening, at the required stage of plant development*, i.e. at the single cell, callus or explant stage, as well as in intact plants. In this case the selective chemical needs to be active in cells at varying developmental stages. Similarly, should screening enzymatic activity be carried out in differing tissues, care needs to be exercised in ensuring access of the enzymatic substrate throughout the tissue.

(4) *Low background or inhibitory activity in the plant cell.* To be effective the enzymatic activity encoded by the marker gene needs to be unique in the plant cell and unaffected by factors present in the cell. Very few markers are free of this problem, although generally protocols have been developed to minimise these difficulties.

V. PROMOTERS USED TO EXPRESS PLANT-SPECIFIC MARKER GENES

To date, the promoters most often used to direct the expression of marker genes in plant cells have been derived from plant pathogens. This is largely because these promoters have been relatively well characterised and they are generally able to direct high levels of gene expression in the majority of the differing cell types in transgenic plants. While, as a general rule, this latter point holds true, it must be borne in mind that detailed studies have revealed that the activity of many promoters, including those considered as being constitutive, may be affected by changing the internal levels of growth substances as well as the circadian rhythm.

A. 35S RNA Promoter

The 35S RNA of cauliflower mosaic virus (CaMV) acts as both a replicative intermediate and as polycistronic RNA during CaMV infection, and accumulates to high levels in the normal host plants – the brassicas. The promoter, normally extending to approximately –400, has been found to direct high levels of expression in differing tissues in a wide variety of plants. Extensive dissection of the promoter has revealed that it comprises a collection of functional domains, each producing expression in differing tissues, which in the intact promoter combine to produce the constitutive mode of expression observed. For practical simplicity the promoter can be divided into three regions: 0 to –46, which contains the transcriptional start site and is often used as a minimal promoter (i.e. a promoter capable of only initiating transcription); –46 to –90, which contains the binding site of the general transcription factor TGAla; and –90 to –400, which acts as a transcriptional enhancer. The enhancer sequence has often been used to enhance levels of expression from heterologous promoters. Located outside of a promoter the enhancer will enhance transcription without changing the developmental expression of the promoter in question. The enhancer functions in an orientation- and position-independent manner. Multiplication of the enhancer results in concomitant enhancement of the expression obtained (for a review, see Sanfaçon, 1992).

B. Mannopine 1'2' Dual Promoter

Derived from the T-DNA of the *TiAch5* Ti plasmid, the divergent dual promoter directing the expression of *mas*1′ (mannopine reductase) *mas*2′ (mannopine conjugase) genes is located on a 479 bp intergenic region. The *mas*1′ and 2′ transcripts are the most abundant in transformed tissue, indicating high levels of transcription, and the close proximity of the two promoters initially suggested that both may be expressed at high levels and that position effects might not affect their relative levels of expression. The dual promoter system has proven useful in obtaining tissue where a specific construct is highly expressed (by linking the selective marker to the *mas*1′ promoter and the gene of interest to the *mas*2′

promoter (Vaeck *et al.* (1987)). Recent work, however, has shown that the relative levels of expression by the two promoters can vary (Peach and Veltin, 1991) and that expression from the promoters are tissue specific and wound and auxin responsive (Teeri *et al.*, 1989).

C. Nopaline Synthase Promoter

The product of the nopaline synthase gene of the T-DNA catalyses the condensation of α-ketoglutarate with arginine to form nopaline. The *nos* promoter has been often used to direct gene expression because it is active in most tissues. Nevertheless, detailed studies have shown that like the *mas*1′2′ promoter, it appears to be most active in the basal region of the plant, suggesting that it is auxin responsive. After flowering, the activity of the promoter in vegetative organs declines, whereas in flower tissue there is a general increase in activity (Mitra and An, 1989). More recently, it has been found also to be induced by wounding (An *et al.*, 1990).

VI. CONTROLLING TRANSGENE EXPRESSION

Routinely the expression of gene constructs used in plant transformation is controlled by constitutive (or, better, near-constitutive) promoters, as described above. However, there may be good reasons to want to limit the expression of a particular gene construct to a particular tissue at a specific developmental stage. This is of importance if the gene product has a detrimental effect on the host plant cell or changes its general growth characteristics, for example with genes encoding hormone biosynthetic genes. Many plant promoters bestowing tissue-specific gene expression have been isolated, but often they also direct low levels of expression in other tissues. In addition, it might be advantageous to induce directly the expression of transgenes by external treatment. Several means have been developed to do this, and each has involved the use of specific well-defined promoters (reviewed by Ward *et al.*, 1993):

(1) *Heat shock.* The expression of many plant genes can be modulated by a variety of environmental factors, and one of the most studied has been heat shock. Heat shock has been cited as a good means of specifically modulating gene expression as levels of induction can be high, induction can be obtained in a variety of plant systems and is tissue-independent, and induction can be reversible and does not greatly affect the growth or morphology of the treated plant (Ainley and Key, 1990). With this in mind, these workers developed an inducible expression cassette from a soya bean heat shock promoter. This construct allowed a 10-fold enhancement of gene expression in protoplast assays induced by culture at 40°C.

(2) *Chemical application.* An alternative means of inducing gene expression is by the application of chemicals, and an example of this is provided by the induction of the pathogenesis-related (PR) protein 1a promoter by salicylic acid (SA) or 2,6-dichloroisonicotinic acid (INA) (Williams *et al.*, 1992). This allows a specific gene to remain unexpressed in the absence of chemical induction and its induction by chemical application. An example of this is provided by a *Bacillus thuringiensis* endotoxin gene linked to the PR1a promoter and transferred to tobacco. Transgenic plants either treated or untreated with the chemical inducer were subjected to insect

predation. Plant lines treated with the chemical inducer sustained significantly less damage than those treated with water.

(3) *The Tn10 Tet repressor.* Both the previous examples involve treatment of the plant with an inducer which may have unexpected side-effects. Exogenous regulators which may affect processes which normally do not occur in the plant may overcome this by affecting only the transgene. With this in mind the Tet repressor of the bacterial transposable element Tn10 has been shown to function in plants (Gatz *et al.*, 1991). Here plants were transformed with the Tet repressor gene linked to the 35S RNA promoter. These plants, which express the repressor to a level of approximately 0.01% of total cell protein, were transformed once again with a GUS gene linked to a 35S RNA promoter into which had been engineered the Tet operator sequences. It was found that the presence of two tandem operator sequences within the promoter conferred a 50- to 80-fold repression in the presence of tetracycline.

A final means of engineering an 'off/on' switch controlling the expression of a gene utilises the Ac/Ds transposable element system. Here the transposable element Ds transposes only when an active Ac element is present within the genome. The Ds element can be cloned into a 5′ untranslated leader sequence of a marker gene between the promoter and the coding sequence. Once transferred into plants by transformation, under normal circumstances and in the absence of Ac, the presence of the Ds element disrupts the expression of the marker gene, and it is silenced. If these plants are crossed with a line containing Ac, in progeny containing Ac and Ds the latter has the potential of transposing from the leader sequence of the marker gene. Should transposition occur, expression of the marker gene can take place, and it is effectively switched on.

An example of the ability of this type of marker strategy is provided by a study of the effect of differing promoters linked to the Ac transposase gene (Swinbourne *et al.*, 1992). Here the Ds element was engineered between the 35S RNA promoter and the streptomycin resistance gene. Plants containing this construct are bleached when germinated and grown on media containing streptomycin. These plants were crossed with plants engineered to contain the Ac transposase gene linked to promoters of differing strengths. Progeny were sown on media containing streptomycin, and transposition was measured by the appearance of green patches, resulting from the functional streptomycin resistance gene on the cotyledons. In this manner it was found that the levels of transposase mRNA had a positive correlation with the frequency of excision of the Ds element.

VII. CONCLUSIONS

As we have seen, plant transformation vectors have been developed for a variety of purposes. Routinely, specific vectors have been designed for a particular purpose, and in designing a transformation experiment the individual needs to select the most appropriate vector. Generally, binary vectors have proven to be the most versatile and have been widely applied, although the stability in *Agrobacterium* of co-integrative vectors make them attractive for many purposes.

One of the greatest difficulties in designing a transformation experiment now is not in the transformation *per se*, but rather in determining possibly where, and to what level, a

specific gene may be expressed in the transformed tissue. As we have seen, detailed analysis of what have been considered to be constitutive promoters has indicated that they are prone to variation in expression level, depending on the developmental stage of the host plant cell. For many purposes this may not be an important consideration; however, it is increasingly becoming significant when trying to judge the effect of expression of a particular gene on plant growth and morphology. Certainly, ways and means have been developed in an attempt to control transgene expression, although these await widespread application.

The numbers of marker genes to be used in plant transformation continue to increase, and a number of dominant screenable or negative selection markers have passed into general use. One of the most important advances has been the development of histochemical markers allowing visualisation of cell-specific gene expression.

While vectors have been designed for most purposes, possibly the one area in which they have not been effective is in library transfer to plant cells. This is becoming increasingly important in gene rescue by mutant complementation. While pilot experiments have demonstrated the feasibility of library transfer with Ti plasmid based vectors (Prosen and Simpson, 1987; Lazo *et al.*, 1991), to my knowledge these have yet to be effectively exploited. Difficulties encountered in this goal are likely to include: (1) the relatively large size of Ti-based vector plasmids; (2) the stability of large regions of plant genomic DNA within *Agrobacterium*; and (3) the feasibility of transferring extremely large segments of DNA from *Agrobacterium* to the plant cell genome, notwithstanding the potential of the DNA itself having sequences able to function as border sequences. With this in mind, the potential of transferring large regions of DNA into a binary vector has been reported recently (Fuse *et al.*, 1995). In this case, binary vector sequences, flanked by *lox* sequences recognised by the bacteriophage Cre–*lox* recombination system are cloned between lamda phage arms. Chromosomal DNA can be cloned into the vector, and subsequently the vector can be transferred to a cre$^+$ bacteria, where the lambda arms are removed by recombination between the *lox* sites, converting the vector into a binary vector. This system has great potential, as it would allow the transfer of large sequences of DNA to binary vector systems without having to carry out laborious cloning steps.

Taken together, recent advances in plant transformation and vectors have been such that now virtually any plant can be transformed and foreign DNA expressed in any cell of the transformant. Possibly the most pressing questions now are what genes are worth transferring, and how do we isolate them.

REFERENCES

Ainley, W. M. and Key, J. L. (1990). *Plant Mol. Biol.* **14**, 949–967.
An, G., Watson, B. D. and Chiang, C. C. (1986). *Plant Physiol.* **81**, 301–305.
An, G., Costa, M. A. and Ha, S.-B. (1990). *Plant Cell* **2**, 225–233.
Bevan, M. W. (1984). *Nucl. Acids Res.* **12**, 8711–8721.
Birnboim, H. C. and Doly, J. (1979). *Nucl. Acids Res.* **7**, 1513–1523.
Bowen, B. A. (1993). *In* 'Markers for Plant Gene Transfer in Transgenic Plants' (S.-D. Kung and R. Wu, eds), Vol. 1, pp. 89–123. Academic Press, New York.
Depicker, A. G., Jacobs, A. M. and Van Montagu, M. (1989). *Plant Cell Rep.* **7**, 63–66.
Dhaese, P., De Greve, H., Decraemer, H., Schell, J. and Van Montagu, M. (1979). *Nucl. Acids Res.* **7**, 1837–1849.

Ditta, G., Stanfield, S., Corbin, D. and Helsinki, D. R. (1980). *Proc. Natl Acad. Sci. USA* **77**, 7347–7351.
Ebert, P. R., Ha, S. B. and An, G. (1987). *Proc. Natl Acad. Sci. USA* **84**, 5745–5749.
Fraley, R. T., Rogers, S. G., Horsch, R. B., Eicholtz, D. A., Flick, J. S., Fink, C. L., Hoffmann, N. L. and Sanders, P. R. (1985). *Biotechnology* **3**, 629–635.
Fuse, T., Kodama, H., Hayshida, N., Shinozaki, K., Nishimura, M. and Iba, K. (1995). *Plant J.* **7**, 849–856.
Gatz, C., Kaiser, A. and Wendenburg, R. (1991). *Mol. Gen. Genet.* **227**, 229–237.
Hain, R., Stabel, P., Czernilofsky, P., Steinbiss, H. H., Herrera-Estrella, L. and Schell, J. (1985). *Mol. Gen. Genet.* **199**, 160–168.
Koltunow, A. M., Truettner, J., Cox, K. H., Wallroth, M. and Goldberg, R. B. (1990). *Plant Cell* **2**, 1201–1224.
Koncz, C. and Schell, J. (1986). *Mol. Gen. Genet.* **204**, 383–396.
Lazo, G. R., Stein, P. A. and Ludwig, R. A. (1991). *Bio/Technology* **9**, 963–967.
Lloyd, A. M., Walbot, V. and Davis, R. W. (1992). *Science* **258**, 1773–1775.
Mariani, C., De Beuckeleer, M., Truettner, J., Leemans, J. and Goldberg, R. (1990). *Nature* **347**, 737–741.
Mattanovich, D., Rüker, F., da Camera Machado, A., Gaines, M., Regner, F., Steinkellner, H., Himmler, G. and Katinger, H. (1989). *Nucl. Acids Res.* **17**, 6747.
Meyer, P., Heidmann, I., Forkman, G. and Saedler, H. (1987). *Nature* **330**, 677–678.
Meyer, P., Heidmann, I. and Niedenhof, I. (1993). *Plant J.* **4**, 89–100.
Mitra, A. and An, G. (1989). *Mol Gen. Genet.* **215**, 294–299.
Peach, C. and Veltin, J. (1991). *Plant Mol. Biol.* **17**, 49–60.
Prosen, D. E. and Simpson, R. B. (1987). *Bio/Technology* **5**, 966–971.
Sanfaçon, H. (1992). *Can. J. Bot.* **70**, 885–899.
Swinbourne, J., Balcells, L., Scofield, S. R., Jones, J. D. G. and Coupland, G. (1992). *Plant Cell* **4**, 583–595.
Teeri, T., Levaslaiko, H., Franck, M., Votila, J., Heino, P., Palva, E., Van Montagu, M. and Herrera-Estrella, L. (1989). *EMBO J.* **8**, 343–350.
Vaeck, M., Reynaeuts, A., Hofte, H., Jansens, S., De Beuckeleer, M., Dean, C., Zabeau, M., Van Montagu, M. and Leemans, J. (1987). *Nature* **328**, 33–37.
Van Haute, E., Joos, H., Maes, S., Warren, G., Van Montagu, M. and Schell, J. (1983). *EMBO J.* **2**, 411–418.
Velten, J. and Schell, J. (1985). *Nucl. Acids Res.* **13**, 6981–6998.
Walden, R. (1988). 'Genetic Transformation in Plants'. Open University Press, Milton Keynes.
Ward, E. R., Ryals, J. A. and Miflin, B. J. (1993). *Plant Mol. Biol.* **22**, 361–366.
White, F. F. (1992). *In* 'Transgenic Plants' (S. D. Kung and R. Wu, eds). 15–48 Academic Press, New York.
Williams, S., Friedrich, L., Dincker, S., Carozzi, S., Kessmann, H., Ward, E. and Ryals, J. (1992). *Bio/Technology* **10**, 540–542.
Zambryski, P., Joos, H., Genetello, C., Leemans, J., Van Montagu, M. and Schell, J. (1983). *EMBO J.* **2**, 2143–2150.

6 Plant Viruses as Gene Vectors*

ROB GOLDBACH[1] and THOMAS HOHN[2]

[1]Department of Virology, Agricultural University Binnenhaven II, 6709 PD Wageningen, The Netherlands

[2]Friedrich Miescher Institut, PO BOX 2543, CH-4002, Basel, Switzerland

**Plant virus names and their abbreviations*:
ACMV, African cassava mosaic, *alias* CLV, cassava latent (bipartite Gemini); AlMV, alfalfa mosaic; BCTV, beet curly top (monopartite Gemini); BMV, brome mosaic (Bromo); BNYVV, beet necrotic yellow vein (Furo); BSMV, barley stripe mosaic (Hordei); BYDV, barley yellow dwarf (Luteo); BWYV, beet western yellows (Luteo) CaMV, cauliflower mosaic (Caulimo); CCMV, cowpea chlorotic mottle (Bromo); CMV, cucumber mosaic (Cucumo); CNV, cucumber necrosis (Tombus); CPMV, cowpea mosaic (Como); CyRSV, *Cymbidium* ringspot (Tombus); DSV, *Digitaria* streak (monopartite Gemini); GFLV, grapevine fanleaf (Nepo); MCMV, maize chlorotic mottle; MSV, maize streak (monopartite Gemini); ORSV, *Odontoglossum* ringspot (Tobamo); PEBV, pea early browning (Tobra); PPV, plum pox (Poty); PVX, potato X (Potex); RCNMV, red clover necrotic mosaic (Diantho); RTBV, rice tungro bacilliform (Badna); TBSV, tomato bushy stunt (Tombus); TCV, turnip crinkle (Carmo); TGMV, tomato golden mosaic (bipartite Gemini); TMV, tobacco mosaic (Tobamo); TRV, tobacco rattle (Tobra); TVMV, tobacco vein mottling (Poty); TYMV, turnip yellow mosaic (Tymo); WClMV, white clover mosaic (Potex); WDV, wheat dwarf (monopartite Gemini).

METHODS IN PLANT BIOCHEMISTRY Vol. 10b
ISBN 0-12-461021-8

I. INTRODUCTION

The establishment of homologous and heterologous gene libraries in *Escherichia coli* using derivatives of bacteriophage λ was the first successful example of the use of a virus vector for gene transfer. Double-stranded (ds) DNA papova viruses, e.g. bovine papilloma virus (BPV), which replicate as circular DNA, and retroviruses, which incorporate genes into host chromatin, later became important for introducing foreign genes into animal cells. The proof-reading mechanism (e.g. see Gluzman and Hughes, 1988) associated with the replication of these viruses as ds DNA leads to a relatively high stability of the transfected genes in target cells. In contrast, RNA viruses with the exception of the two α viruses Sinbis virus (Xiong *et al.*, 1989) and Semiliki forest virus (Liljeström and Garoff, 1991), have found only limited application.

Attempts to obtain plant virus vectors have been frustrated by the absence of true ds DNA viruses and viruses mediating DNA integration. Consequently, plant virus vectors are based on single-stranded (ss) DNA viruses, pararetroviruses and RNA viruses (B. Hohn *et al.*, 1987a; Gronenborn and Matzeit, 1989; Fütterer *et al.*, 1990; Ahlquist and Pacha, 1991; Joshi and Joshi 1991; Turpen and Dawson, 1992). However, the replication of such viruses can lead to template switching and hence to genome rearrangements. Transfected genes may be mutated or rapidly eliminated unless changes in virus replication, expression and other functions are avoided during vector construction. Elimination is in any case inevitable unless the inserted gene confers a selective advantage. On the other hand, since non-selective genes will not survive in hybrid plant virus derivatives released into the environment, this instability constitutes a biological containment. *Agrobacterium*-derived vectors mediating integration of foreign genes into plant chromatin or direct gene transfer (e.g. see Hohn and Schell, 1987) have been used to achieve stable transformation of plants.

A further distinction between plant and animal viruses of practical importance lies in their mechanisms of spreading. Whilst plant viruses make use of the plasmodesmata connections between cells (Hull, 1991; Lucas and Wolf, 1993), animal viruses progress by exo- and endocytosis. Because of the lack of plasmodesmata in plant cells in culture, infection with viruses is only accomplished by, for example, polyethylene glycol (PEG) treatment or electroporation; spread of virus within the host cell population does not occur.

II. PLANT INOCULATION WITH VIRUSES AND VECTORS

One widespread method for inoculating plants with both DNA and RNA viruses is leaf abrasion. This can also be used for naked nucleic acid. Alternatively, the DNA transfer mechanism of *Agrobacterium tumefaciens* (the agent of crown gall disease) can also be used, as shown initially with cauliflower mosaic virus (CaMV) (Grimsley *et al.*, 1986a; Hille *et al.*, 1986; B. Hohn *et al.*, 1987a) and potato spindle tuber viroid (Gardner *et al.*, 1986). At least 'one-and-a-bit-mers' of the virus genome cloned within the T-DNA borders of the Ti plasmid of *Agrobacterium* is introduced into the plant by bacterial infection. The viral genome escapes, probably in the nucleus, either by recombination or production of monomeric replicative intermediates, e.g. the 35S RNA from multimeric CaMV DNA (Grimsley *et al.*, 1986b).

Mechanical inoculation is not possible, for example, for viruses normally introduced directly into the phloem by insect vectors. Many of those viruses infect Gramineae, which are not considered to be hosts for the crown gall disease. The successful infection of maize plants with maize streak virus (MSV) via agroinfection (Grimsley *et al.*, 1987, 1988; B. Hohn *et al.*, 1987b; Boulton *et al.*, 1989) was therefore unexpected and shows that T-DNA transfer into Gramineae is in fact possible. This has overcome the barriers to the introduction of a number of additional viruses: *Digitaria* streak virus (DSV) into maize (Donson *et al.*, 1988), wheat dwarf virus (WDV) into wheat (Hayes *et al.*, 1988a; Woolston *et al.*, 1988) and rice tungro bacilliform virus (RTBV) into rice (Dasgupta *et al.*, 1991). Also for viruses that can be introduced by other means agroinfection was used, e.g. for tobacco mosaic virus (TMV) into tobacco (Turpen *et al.*, 1993). An alternative route for the introduction into plants of viral genomes that are not mechanically transmittable (e.g. luteoviral RNA; Young *et al.*, 1991) is protoplast transfection.

Agroinfection can also be used for the introduction of virus derivatives, e.g. mutants, vectors and hybrids. For example, *Nicotiana* leaf discs were transfected with a tomato golden mosaic virus (TGMV) vector (Hanley-Bowdoin *et al.*, 1988) and maize plants with an MSV vector (Shen and Hohn, 1991, 1992).

III. PARARETROVIRUS VECTORS

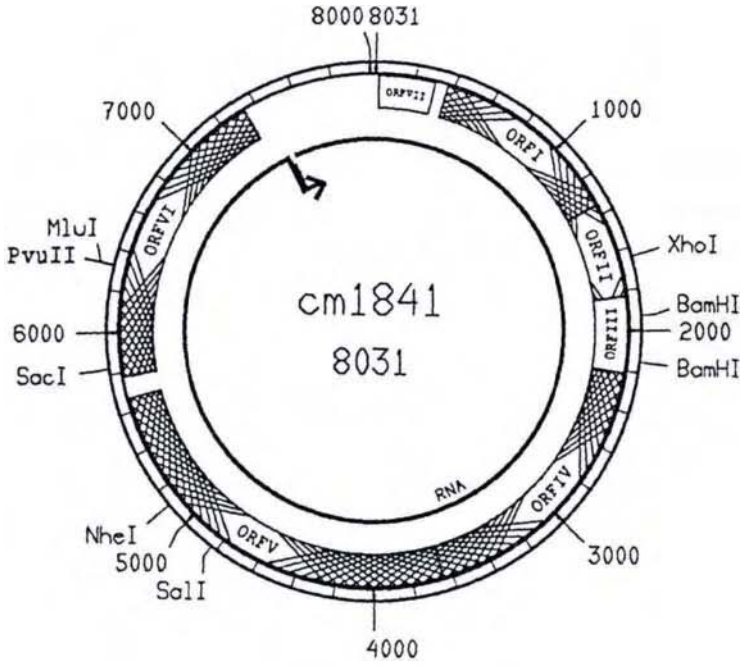
In the case of the pararetroviruses, as well as the retroviruses (Bonneville and Hohn, 1993), viral DNA is transcribed in the host nucleus into terminally redundant RNA; the DNA produced from the RNA in the cytoplasm by reverse transcription is transported into the nucleus either directly or via new infections. Whilst pararetrovirus DNA accumulates without replication as supercoiled circles, retrovirus DNA is inserted into host chromatin in a terminally redundant form. The two groups of plant pararetroviruses, icosahedral caulimoviruses and bacilliform badnaviruses, have genomes of about 8000 bp.

CaMV, a caulimovirus, is the only one that has so far been used as a vector. This virus has seven open reading frames (ORFs) densely packed into the pregenomic 35S RNA, all but one of which are probably translated from this polycistronic messenger (Hohn and Fütterer, 1991). Transmission of CaMV in nature is by aphids, and is mediated by the aphid transmission factor encoded by ORFII. This ORF is dispensable in the laboratory, where infection can be carried out by direct inoculation with virus particles or virus DNA (Dixon *et al.*, 1983), and the site can be used for foreign DNA.

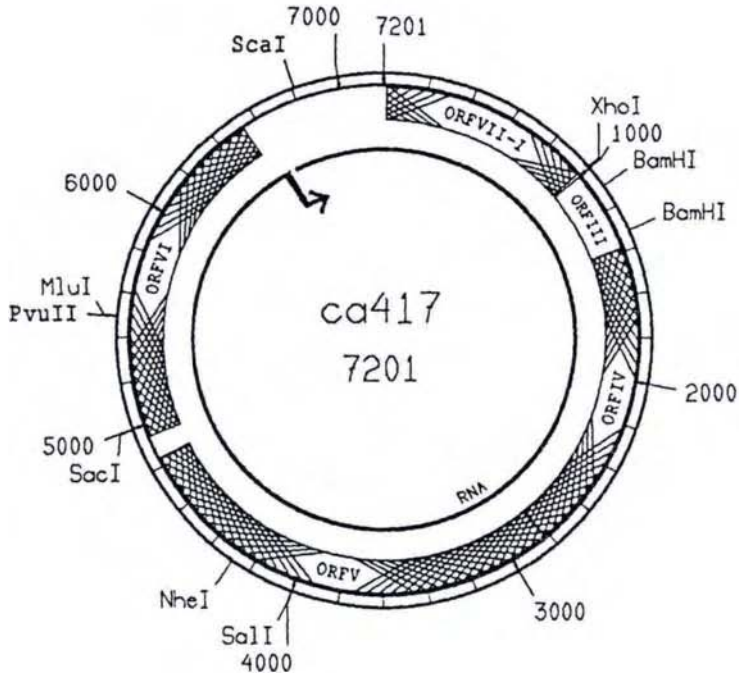
pCa-BB1, an ORFII replacement vector, contains an *Xho*I restriction site tightly placed between the stop codons of ORFI and ORFII; the ORFII is otherwise deleted. The coding ORF is adjusted to fit the *Xho*I site, with its start codon (preferably as part of the *Nco*I site C*CATG*G) directly adjacent to the 5′ insertion site and its last codon in phase with the ORFII stop codon. After amplification of the recombinant plasmid in *E. coli*, virus DNA (1–5 μg) excised from the vector is used to inoculate host plants. pCa417 (Fig. 6.1), derived from pCa-BB1 by fusing the start codon of ORFVII to ORFI, provides more space (800 bp) for foreign DNA. Further deletions, i.e. in the leader region, are possible, but lead to considerably retarded infections.

The bacterial dihydrofolate reductase (DHFR) coding region (240 bp) was cloned in pCa-BB1 and expressed systemically in turnips (Brisson *et al.*, 1984; Brisson and Hohn, 1986, 1988). This vector has also been used to clone and express Chinese hamster metallothionein (CHMTII; 200 bp) (Lefebvre *et al.*, 1987, 1990), human αD interferon (IFN;

(a)



(b)



XhoI
981 GAGAATAACC TCGAGGTGAA ATGGCTAATC 1010

500 bp) (deZoeten *et al.*, 1989) and the necrosis protein of beet necrotic yellow vein virus (BNYVV ORF N; 171 bp) (Jupin *et al.*, 1991). Infection with the latter elicited a necrotic response instead of the usual mosaic symptoms. The conclusion to be drawn so far from experiments with such vectors are:

(1) Illegitimate template switches of the nascent minus or plus DNA strands and negative selection frequently leads to loss of the DNA sequences inserted into CaMV. To minimise these recombination events, homologies at the insertion site should be avoided and conditions chosen that do not interfere with the virus life cycle. The genome of the vector should preferably not exceed the original virus genome length of 8000 bp. By deletion of ORFII and ORFVII, which are not needed for virus growth (Givord *et al.*, 1988), a maximum payload of 1000 bp can be achieved with about 300 bp overpackaging. Larger payloads might be possible with artificial bipartite CaMV genomes containing all essential functions and lacking recombination targets (Gronenborn, 1987; Hirochika and Hayashi, 1991).
(2) The unusual CaMV translation mechanism probably requires the tightly packed ORF organisation. Experience has shown that non-coding sequences between CaMV and passenger ORFs reduce the stability of the vector.
(3) The sequences inserted should not interfere with gene expression and replication of CaMV. For example, a new promoter might occlude the original CaMV promoters, an RNA polyadenylation signal would interfere with pregenomic RNA production, and an intron would be rapidly eliminated by the transcription/reverse transcription cycle (B. Hohn *et al.*, 1986).
(4) Gene products that inhibit virus replication directly or by host cell toxicity are undesirable.

The CaMV vector system has found application in whole plants rather than cell cultures, where it does not proliferate efficiently. Mature virus particles are seldom found in cell cultures (Paszkowski *et al.*, 1983; Rollo and Covey, 1985), although replicative intermediates have been found in infected protoplasts (Maule, 1985). It would appear that the choice between genome replication and translation is affected developmentally (Hohn and Fütterer, 1991). In protoplasts and derived cell cultures, a control mechanism appears to favour genome replication at the expense of particle formation; both processes depend on the same 35S RNA.

The host range of gene vectors may be increased by work on further plant pararetroviruses. For example, the badnavirus rice tungro bacilliform virus (Hay *et al.*, 1991; Qu *et al.*, 1991) could be of use with the Gramineae. The bacilliform capsid of this virus might reduce the problem of overpackaging.

FIG. 6.1. CaMV and a vector. (a) CaMV strain CM1841 is shown with its seven ORFs preceded by the intercistronic region, providing the leader of the CaMV large transcript. The large transcript (35S RNA) with its termonal redundancy is shown as the inner circle. ORFs are coding for movement factor (I), insect transmission (II), minor capsid protein (III), major capsid protein precursor (IV), pol-polyprotein, consisting of protease and polymerase/RNAaseH (V), and an unknown dispensible function (VII). (b) Ca417 is a derivative in which the AUG of the generally dispensible ORF VII is fused to ORFI and ORFII, which is dispensible for spreading in single plants, is deleted. The sequence around the *Xho*I cloning site is shown at the bottom with the ORFI stop and ORFIII start codons shown in bold letters.

IV. GEMINIVIRUS VECTORS

Geminiviruses, with ss circular DNA (2.5–3 kb), are characterised by twinned (geminate) icosahedral particle. Dicot plant geminiviruses are usually bipartite and transmitted by whiteflies, whereas those of monocots are monopartite and leafhopper transmitted. Beet curly top virus (BCTV) and tomato yellow leaf curl virus (TYLCV; Kheyr-Pour *et al.*, 1992) are exceptional in that they infect dicots but are monopartite and are transmitted by leafhoppers and whiteflies, respectively. Replication in single cells usually requires only component A of the bipartite geminiviruses; component B is required for systemic spreading and together with A for replication (e.g. see Davies and Stanley, 1989). The geminiviruses resemble animal polyomaviruses and SV40 viruses in their bidirectional expression strategy. The ds DNA forms in the nucleus contain two non-coding regions, one with the two polyadenylation signals and the other bearing the clockwise and counterclockwise promoters. RNAs coding for structural proteins are produced by transcription in one direction (consensually clockwise; Fig. 6.2) and RNAs required for DNA replication in the other direction.

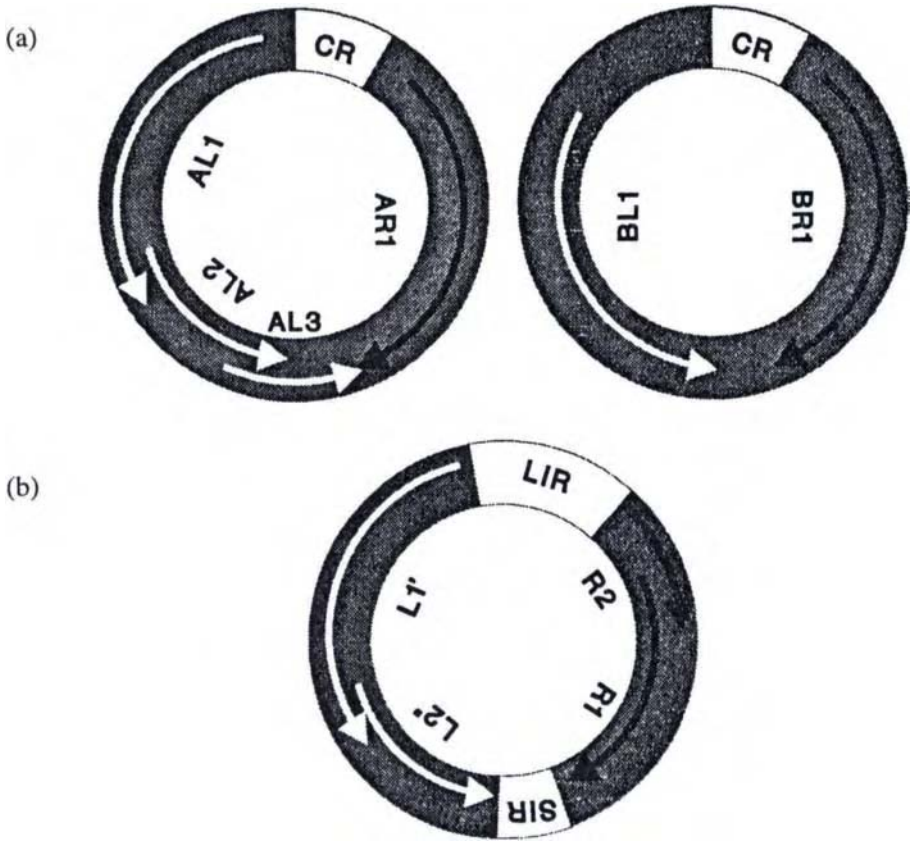


FIG. 6.2. Typical maps of geminiviruses. (a) Bipartite: AL1, replicase; AL2 transactivator; AL3, replication enhancement, AR1, coat protein; BR1, nuclear transport; BL1, movement protein; CR, common region. (b) Monopartite: L1′, L1″, replicase; R1, coat protein, R2 movement function; LIR, long intercistronic region, SIR, small intercistronic region. (Lazarowitz, 1992; Nouelry *et al.*, 1994.)

The coat proteins are not required for genome replication but are essential for insect transmission (Davies and Stanley, 1989) and for spreading within the plant of monopartite viruses (Lazarowitz *et al.*, 1989; Lazarowitz, 1992). Coat protein mutants thus can all replicate in single cells but only in the case of bipartite geminiviruses can they spread in the plant. Some coat protein replacement vectors for both dicots and monocots are given in Table 6.1.

TABLE 6.1. Geminivirus vectors.

Mono-/ bipartite	Virus	Target	Host	Payload	Reference
Bi	TGMV	Plants	*N. tabacum*	*CAT*, *GUS*, NPTII	Hayes *et al.* (1988b, 1989) Hanley-Bowdoin *et al.* (1988)
Bi	ACMV	Plants	*N. benthamiana*	*CAT*	Ward *et al.* (1988)
Mono	MSV	Protoplasts	*Z. mays*		Ugaki *et al.* (1991)
		Plants	*Z. mays*	*GUS*, DS1	Shen *et al.* (1994)
Mono	WDV	Protoplasts	*T. monococcum*	*CAT*, *GUS*, NPTII	Matzeit *et al.* (1991) Kammann *et al.* (1991)

A. Bipartite Geminivirus Vectors

An example of a systemically spreading bipartite geminivirus vector is the coat protein replacement vector of TGMV. Here, component A was agroinfected into plants already carrying component B (Hayes *et al.*, 1988b, 1989). With African cassava mosaic virus (ACMV), linearised coat protein replacement vectors of component A were inoculated jointly with component B (Ward *et al.*, 1988). Additional promoters could be inserted into these vectors to increase payload expression without affecting replication, e.g. the CaMV 35S promoter (Hayes *et al.*, 1989).

The need for systemic spreading is avoided with transient expression in leaf discs agroinfected with component A vectors (Hanley-Bowdoin *et al.*, 1988) or with plants transgenic for a component A vector. Here, component B is unnecessary as all cells contain and replicate the vector.

Systemically spreading vectors either shorter (Ward *et al.*, 1988; Etessami *et al.*, 1989) or longer (Hayes *et al.*, 1989) than the wild-type virus were subject to rearrangements that restored the original size. Vectors with a payload that restored the original size, e.g. the *CAT* ORF, were particularly stable. The basis of the size dependence in the absence of packaging is not known. In contrast, non-spreading vectors exceeding the normal genome size by as much as 4 kb (e.g. containing the long β-glucuronidase or neomycin phosphotransferase coding regions) were stable.

Paszkowski *et al.* (1993) found that ACMV DNA A can replicate in maize protoplasts (although like all bipartite geminiviruses it is not usually able to infect the Gramineae) with an efficiency of 10–20% of WDV, for which maize is a host. Although this shows some promise for heterologous virus vectors, a 1165 bp insertion of non-viral DNA into ACMV A DNA inhibited replication in maize but not in tobacco.

B. Monopartite Geminivirus Vectors

Coat replacement vectors of monopartite geminiviruses have been used in transient expression experiments with various plant tissues (Lazarowitz *et al.*, 1989; Töpfer *et al.*, 1989; Matzeit *et al.*, 1991). Payloads can also substitute for a small ORF of unknown

function upstream of the coat protein (Lazarowitz *et al.*, 1989; Ugaki *et al.*, 1991). Shuttle vectors replicating and expressing payload genes in both bacteria and Gramineae protoplasts have been created by incorporating an *E. coli* replicon (Kammann *et al.*, 1991; Ugaki *et al.*, 1991).

In MSV, the small intercistronic region (SIR) has been used as an insertion target by Shen and co-workers. This region of the MSV genome can be manipulated without loss of viral symptoms (Shen and Hohn, 1991). Viral symptoms of MSV with a 405 bp transposable element Ds1 which replicated in plant tissue only formed upon Ac element-mediated excision of the Ds1 (Shen and Hohn, 1992; Shen *et al.*, 1992). Agroinfection of MSV containing a 22 kb 35S *GUS* transcription unit within the SIR into the coleoptar node of 3-day-old seedlings resulted in *GUS* staining along the veins (Shen *et al.*, 1994). Less frequent transformation events and lighter staining occurred in the absence of the viral replicon. In this case, there was no systemic spread; instead the agroinfected cells gave rise to transformed lineages. Symptoms appeared only after reversions to wild type. Shen and Hohn (1994) finally showed that MSV vectors containing Ds1 replicated in the transfected plant cells but probably did not spread from cell to cell.

V. RNA VIRAL VECTORS

A. Infectivity of Cloned cDNA

For the development and exploitation of RNA viruses as gene vectors, it is crucial to obtain infectious clones (as cDNAs or as *in vitro* transcribed RNA copies) of their genome segments. This has been obtained for a growing number of plus-strand RNA viruses (but not minus-strand viruses; Table 6.2), following one of the following approaches:

TABLE 6.2. Infectious plant viral cDNAs.

Group	Virus	RNA No.	Reference
Alfamo	AIMV	3	Dore *et al.* (1990)
Bromo	BMV	1, 2, 3	Ahlquist and Janda (1984), Janda *et al.* (1987)
	CCMV	1, 2, 3	Allison *et al.* (1988)
Carmo	TCV		Heaton *et al.* (1989)
Como	CPMV	1, 2	Vos *et al.* (1988), Eggen *et al.* (1989)
Cucumo	CMV	1, 2, 3	Rizzo and Palukaitis (1990)
Diantho	RCNMV	1, 2	Xiong and Lommel (1991)
Furo	BNYVV	1, 2, 3, 4	Quillet *et al.* (1989), Ziegler-Graff *et al.* (1988)
Hordei	BSMV	α, β, γ	Petty *et al.* (1989)
Luteo	BYDV		Young *et al.* (1991)
	BWYV		Veidt *et al.* (1992)
Nepo	GFLV	1, 2	Viry *et al.* (1993)
Potex	PVX		Hemenway *et al.* (1990)
	WCIMV		Beck *et al.* (1990)
Poty	PPV		Riechmann *et al.* (1990)
	TVMV		Domier *et al.* (1989)
Sobemo	MCMV		Scheets *et al.* (1993)
Tobamo	TMV		Dawson *et al.* (1986), Meshi *et al.* (1986), Raffo and Dawson (1991)
Tobra	PEBV	1, 2	MacFarlane *et al.* (1991)
	TRV	1, 2	Hamilton and Baulcombe (1989), Angenent *et al.* (1989)
Tombus	CNV		Rochon and Johnston (1991)
	CRSV		Burgyan *et al.* (1990)
	TBSV		Hearne *et al.* (1990)
Tymo	TYMV		Weiland and Dreher (1989)

(1) Using a cDNA clone directly. This approach was used in the early beginning of trying to recover infectious virus from cloned copies and yielded only very limited results. Infectivity was reported for AIMV RNA 3 (Dore and Pinck, 1988; Dore *et al.*, 1990) and TMV RNA (Meshi *et al.*, 1986) and was proposed to be dependent on *in vivo* transcription from unidentified promoters.

(2) Using a cDNA clone fused to the CaMV 355 promoter. This approach results in a more efficient production of progeny virus, as reported first for BMV (Mori *et al.*, 1991) and beet necrotic yellow vein virus (BNYVV) (Commandeur *et al.*, 1991) and more recently for another seven plant viruses, belonging to different taxonomic groups (Boyer and Haenni, 1994).

(3) By agroinoculation, as achieved by beet western yellow virus (BWYV), a phloem-restricted, non-sap-transmittable luteovirus, using a cDNA clone provided with a ribozyme sequence to obtain *in vivo* run-off transcripts (Leiser *et al.*, 1992).

(4) As an *in vitro* transcript, using a linearised cDNA clone, provided with phage SP6, T7, T3 or *E. coli* RNA polymerase, in combination with their respective promoter. This approach is so far the most widely applied approach, and has been successfully used for 33 different viruses belonging to 17 different taxonomic groups (Table 6.2; Turpen and Dawson, 1992; Boyer and Haenni, 1994). For optimal infectivity it is essential that the *in vitro* transcripts are not only full-length, but also have (almost) authentic termini. In particular, extra, non-viral nucleotides at the 5′ terminus lead, in most cases, to a significant drop in infectivity whereas extra nucleotides at the 3′ terminus are better tolerated (Hohn and Goldbach, 1994; Boyer and Haenni, 1994). Notably, the T7 promoter and its corresponding polymerase have been used in *in vitro* transcription systems. This promoter can be trimmed to yield almost identical 5′ termini. It is observed that extra, non-viral sequences are always deleted in progeny RNAs. The genomic RNAs of some viruses, e.g. the como-, nepo-, poty- and sobemoviruses, possess a protein, denoted VPg, at their 5′ end. Nonetheless, *in vitro* transcripts devoid of this protein have been shown to be infectious, yielding progeny RNAs containing a VPg again.

TABLE 6.3. RNA virus vectors.

Group	Virus	Insertion Site	Target	Payload	Reference
Bromo	BMV	Coat protein gene	Protoplasts	*CAT*	French *et al.* (1986)
Como	CPMV	Small coat protein gene	Protoplasts	Epitope[b]	Usha *et al.* (1993)
Furo	BNYVV	25 kDa gene (RNA-3)	Inoculated leaf	*GUS*	Jupin *et al.* (1990)
Hordei	BSMV	b gene (RNAβ)	Protoplasts	*LUC*	Joshi *et al.* (1990)
Potex	PVX	Coat protein gene	Protoplasts	*GUS*	Chapman *et al.* (1992)
		Extra site (coat protein–promoter duplication)	Whole plant	*GUS*	Chapman *et al.* (1992)
Tobamo	TMV	Coat protein gene	Inoculated leaf	*CAT*	Takamatsu *et al.* (1987)
		Coat protein gene	Inoculated leaf	*CAT*	Dawson *et al.* (1989)
		Coat protein gene	Protoplasts	ENK	Takamatsu *et al.* (1990)
		Extra site[a]	Whole plant	NPTII	Donson *et al.* (1991)
		Extra site[a]	Whole plant	DHFR	Donson *et al.* (1991)
		Extra site[a]	Whole plant	ATS	Kumagai *et al.* (1993)
Tombus	TBSV	Coat protein gene	Inoculated leaf	*GUS/CAT*	Scholthof *et al.* (1993)

ENK, encephalin; ATS, α-trichosanthin.

[a] Vector TB2, see Fig. 6.3.

[b] Epitope from foot-and-mouth disease virus coat protein.

B. Constraints and Potentials of Plant RNA Viral Replacement Vectors

As for the caulimo- and geminiviruses, a gene replacement strategy has also been employed in the development of vectors from several plant RNA viruses (Table 6.3). In almost all cases, however, such gene replacement vectors provided of limited value, since they were not capable of systemic infection of host plants. This is due to the fact that most, if not all, genes in the genome of plant RNA viruses are indispensable. This not only holds for the viral replication genes or cell-to-cell movement gene, but even for the coat protein gene, which in many cases is indispensable for long-distance spread through the plant. Thus for a number of viruses, including brome mosaic virus (BMV) and TMV, the coat protein gene has been replaced by reporter genes (*CAT*, *GUS* or *LUC*) and shown to result in recombinant virus expressing successfully the incorporated foreign gene, but appeared limited to infection of single protoplasts or the inoculated leaf only (reviewed by Hohn and Goldbach, 1994). Though for tombusvirus tomato bushy stunt virus (TBSV) the coat protein gene is dispensable for systemic movement, exploitation of this site to express a *GUS* gene, led to greatly reduced expression levels in the upper leaves of inoculated plants (Scholthof *et al.*, 1993). This was explained by rapid deletion of the non-viral sequences, demonstrating the intrinsic genetic instability of viral RNA vectors.

C. Autonomous Gene Vectors Derived from TMV and PVX

So far, two RNA viruses, TMV and potato X virus (PVX), have been successfully developed as an autonomous replicating gene vector for plants, with enough genetic stability and general utility (Table 6.3). Both viruses have in common that they have a very broad host range, each spanning more than 500 plant species, lack a biological vector for spread, and are easily mechanically transmitted.

Both derived vectors (TMV, vector TB2; PVX, vector GC3; see Fig. 6.3) have in common that they are insertion vectors instead of replacement vectors, i.e. still possess all viral genes, and express payload genes from a duplicated coat protein gene promoter (Fig. 6.3). However, Raffo and Dawson (1991) designed mini-TMV replicons that might be used as vectors. Although the duplicated sequence led to some genome instability in the PVX construct, much of the viral RNA in systemically infected tissues has been shown to retain foreign gene insertions, especially in *Nicotania clevelandii* plants (Chapman *et al.*, 1992).

For TMV-derived vectors it was found that the duplicated promoter sequences should be divergent enough to make the construct 'recombination proof'. In vector TB2 this problem has been solved by substituting the TMV coat protein ORF by the foreign gene and adding the coat protein gene of another tobamovirus (*Odontoglossum* ringspot virus, ORSV), along with its (heterologous) subgenomic promoter (Donson *et al.*, 1991; Fig. 6.3). In addition to genetic instability caused by recombinational events, the accumulation of point mutations, due to the lack of any proof-reading mechanism in RNA replication, may be a second cause for inactivation of the (non-essential) payload genes. The experimental data obtained so far with both TMV and PVX vectors, however, indicate that high-level expression of foreign gene products from both TMV-based and PVX-based vectors is maintained for a long enough period of time to obtain significant accumulation of the desired product in the plant (Donson *et al.*, 1991; Chapman *et al.*, 1992; Grill, 1993; Kumagai *et al.*, 1993). The mutation rate in the (functionally not required) foreign gene sequences in TMV vectors has indeed been determined to be lower than predicted (Kearney *et al.*, 1993). It is not excluded that viral RNA vectors eventually lose their capacity to encode a functional, foreign protein upon serial passage, but this would

be rather an advantage in terms of biological containment than a disadvantage. A growing number of foreign genes have been expressed in a plant using either TMV vector TB2 or PVX vector GC3; not only reporter genes but also genes encoding, for example, α-trichosanthin, which was found to accumulate to approximately 2% of the total soluble plant protein (Kumagai *et al.*, 1993).

D. Helper-Dependent RNA Viral Vectors

A prerequisite for designing an autonomously replicating viral vector is that all *cis*- and *trans*-acting factors involved in genome replication, encapsidation and systematic spread are retained. Although this has been achieved for the TMV- and PVX-based vectors described above, it may be a major limitation in the development of further vectors from these or other viruses.

As an alternative, 'disarmed' vectors could be constructed for foreign gene expression in whole plants, which lack one or more essential functions and are only able to co-replicate with wild-type virus as a helper or in transgenic plants transformed with the required genes. In the latter case the properly transformed plant provides a bonus in the form of controlling field-released genetically modified viruses. The feasibility of this approach is demonstrated by the replication of engineered alfalfa mosaic virus (AlMV) RNA 3 molecules in tobacco plants transformed with the AlMV replication genes (Van der Kuyl *et al.*, 1991) and the complementation of spread-deficient TMV mutants in transgenic plants expressing the TMV movement protein (Deom *et al.*, 1987)

VI. THE STUDY OF DNA AND RNA REARRANGEMENTS USING PLANT VIRUS VECTORS

A. Introns

The precise excision in transfected plants of a leghaemoglobin intron presented in a CaMV vector not only verified the vector but was a proof for the reverse transcription of CaMV (B. Hohn *et al.*, 1986). McCullough *et al.* (1991) were able to measure the efficiency and accuracy of splicing and the effect of monocot and dicot introns on expression levels using *Agrobacterium*-mediated transfection of *Nicotiana benthamiana* leaf discs with hybrid TGMV geminivirus vector, which produced large amounts of pre-mRNA containing a variety of introns.

B. Recombination

Non-viable virus hybrids that become viable upon specific genome rearrangements can be used to study recombination. The precise construction of the one-and-a-bit-mers of caulimo- and geminiviruses mentioned above using redundant parts of different viruses allows the investigation of the recombination mechanism, recombination hot-spots, homology levels, etc. (Grimsley *et al.*, 1986b; Gal *et al.*, 1991). When the terminal redundancy was the promoter/polyadenylator region of CaMV, transcription led to the escape of pregenomic virus RNA (Grimsley *et al.*, 1986b). True recombination events could be detected by Gal *et al.* (1991) using transgenic *Brassica napus* containing CaMV one-and-a-bit-mers constructed so that CaMV could not escape by normal transcription. In this case, it was shown that mismatch repair was linked to the recombination process. Later, Gal *et al.* (1992) studied intermolecular recombination between transgenic CaMV

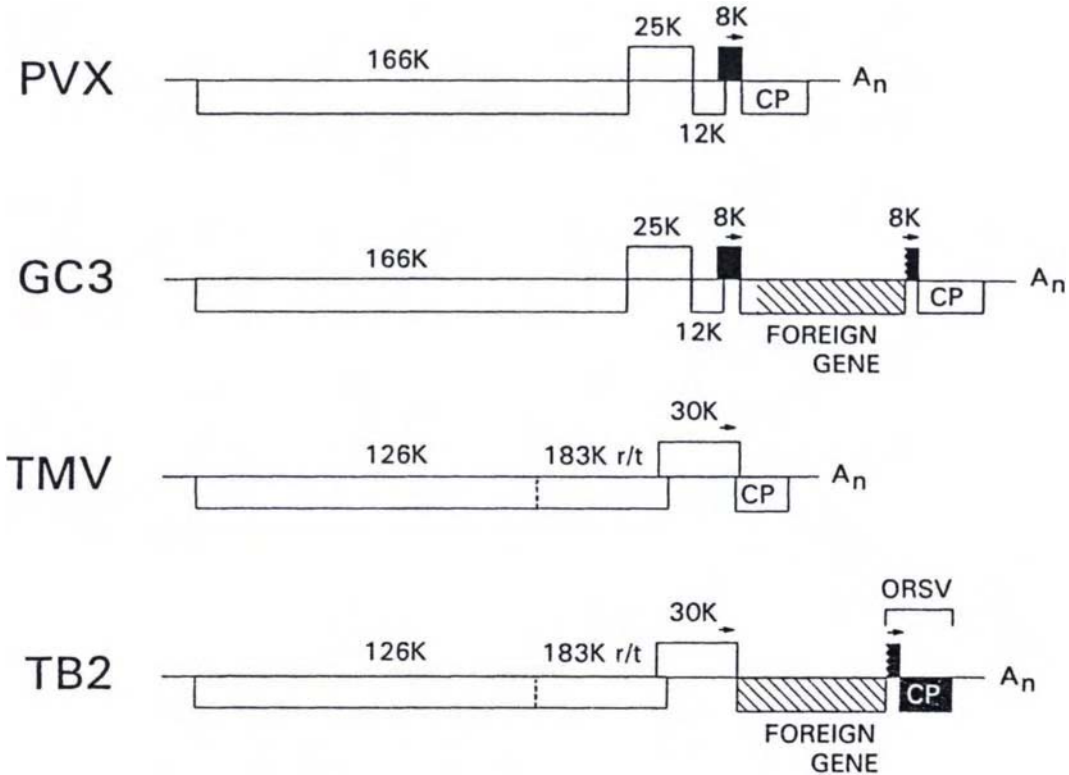


FIG. 6.3. Genetic map of TMV and PVX and their derived, autonomously replicating vectors (TMV, vector TB2; PVX, vector GC3). ORFs are indicated as open bars, and the sizes of their translation products in kilodaltons (K). The PVX 166K gene and 'triple gene block' (25K, 12K and 8K), as well as the TMV 126/183K and 30K genes are required for viral RNA replication and viral spread, and are therefore indispensible. Coat protein promoter sequences are indicated with arrows (located in the 8K ORF of PVX RNA, and in the 30K ORF of TMV). In vector GC3 the duplication of the coat protein gene promoter results in a duplicated region of 81 bases. In vector TB2, duplication of sequences is avoided by adding the coat protein gene of another tobamovirus, *Odontoglossum* ringspot virus (ORSV), along with its promoter (indicated in black).

ORFVI and supertransfected complementary virus sequences. RNA recombination, e.g. in bromoviruses (Bujarski and Kaesberg, 1986; Bujarski and Dzianott, 1991), and *cis*- and *trans*-acting factors in RNA replication (Ahlquist, 1992) have been studied using infectious transcripts from cloned RNA viruses.

C. T-DNA Transfer from *Agrobacterium*

A CaMV-based system to analyse individual *Agrobacterium* T-DNA transfer events, in which the complete T-DNA except the border sequences was replaced by the virus genome, was devised by Bakkeren *et al.* (1989). A viable replicon was produced by circularisation upon plant transfection. Replicon analysis showed that conservation of the remaining right-border sequences was higher than those of the left border. Small direct repeats found between some joined ends indicated the transport of linear DNA to the plant.

D. Transposition

Excision of transposable elements has been studied in both protoplasts (Laufs *et al.*, 1990) and whole plants (Shen *et al.*, 1992) using wheat dwarf virus (WDV) and MSV vectors, respectively, containing the maize transposable element Ac or its defective Ds derivatives. Ds excision was dependent upon the presence of Ac. Junction sequences remaining on the virus genomes after excision showed typical footprints.

E. Methylation

The effects of methylation on expression and replication were studied by Ermak *et al.* (1993) using ACMV (geminivirus) A DNA and derivatives in the presence of helper virus supplying an unmethylated replicase gene. Both transcription and replication of methylated DNA appeared to be inhibited; selection pressure for the maintenance of methylation-free replicons is thus supplied by two independent mechanisms.

VII. EPITOPE PRESENTATION

Surfaces of virus and retrotransposon capsids can be used to present peptides as epitopes for antibodies or receptors. Capsids of hepatitis B virus, polio virus, yeast Ty transposon and filamentous bacteriophages were used (for references, see Usha *et al.*, 1993). The particles so obtained can be used to produce efficient and low-cost vaccines. For this purpose, knowledge of the atomic resolution three-dimensional structure of the capsid is required to determine the insertion site to accept the epitope, and c-DNA clones must be available. The similarity in basic geometry of the plant comoviral and picornaviral capsid has been elegantly exploited by Lomonossoff and co-workers (Porta *et al.*, 1994) by demonstrating that cowpea mosaic virus (CPMV) particles can be used as an 'epitope-presenting system' for the production of vaccine against, for example, picornaviruses. Thus, an epitope derived from foot-and-mouth disease virus (FMDV) consisting of 25 amino acids (known as the 'FMDV loop' was inserted into the small coat protein (VP23) of CPMV at the appropriate site as selected by three-dimensional information on both

viruses (Usha *et al.*, 1993). This foreign sequence did not interfere with the carrier virus to assemble, nor with its ability to spread systemically through cowpea leaves. Likewise, a neutralising epitope of HIV-I, located in gp41, was incorporated into the CPMV small coat protein, yielding a CPMV recombinant which induced neutralising antibodies in mice (McLain *et al.*, 1994).

VIII. CONCLUSIONS

Plant viruses can potentially be engineered into gene vectors able to express desired genes in single plant cell (protoplast) systems or systemically through whole plants.

For some viruses, strategies have become available to design 'disarmed' vectors (only able to replicate in the presence of a helper virus or in transgenic plants that complement essential viral functions in *trans*) and vectors that are able to replicate and spread through plants independently.

Advantages of the virus vector transfection systems either for systemic spreading in whole plants, or for transient expression systems in plant protoplasts over stable plant transformation, are the ease of handling, the short time periods required and the replication of the vectors to high copy numbers that can result in expression levels much higher than in transgenic plants.

By the lack of any proof-reading mechanism and by the large number of RNA replication cycles (even in a single cell) one might suggest that payload genes in RNA virus and pararetrovirus vectors rapidly accumulate point mutations, leading to dysfunctioning of these non-essential inserts (Domingo and Holland, 1988). The (limited) experimental data obtained so far, however, do not support this view but indicate that, at least for the duration of a single protoplast batch or a single plant infection, functional proteins can be obtained in desirable amounts. It is not excluded that upon serial passage RNA vectors eventually lose their capacity to encode a functional protein, though this would merely provide an advantage in view of biological containment.

A second issue of concern is the risk of recombination involving viral vectors. Indeed, it has now been well documented that viral genomes including those of RNA viruses (King, 1988) are frequently the subject of recombinative events. One obvious mechanism for recombination in all classes of plant virus vectors relies on template switches of the nascent DNA and RNA strands. On one hand, non-viral inserts in viral vectors might become lost by recombination, due to the lack of selective pressure, whilst on the other hand, (new) completely functional (and pathogenic) viral genomes may arise from recombinational events involving (disarmed) RNA vectors, co-infecting (helper) viruses and possibly even viral transcripts from transgenes. In the former case (loss of inserts) the consequences seem to be only beneficial with respect to biological containment; in the latter case, however, original viruses or, in theory, even novel variant viruses may arise, which would lead to the undesired spread of pathogens (de Zoeten, 1991; Falk and Bruening, 1994). Model experiments showed that viable virus can be regenerated by recombination between a transfected and a transgenic mutant virus genome (Gal *et al.*, 1992; Greene and Allison, 1994). Although, so far no evidence for the generation of viruses with new properties has been obtained for such events it is clear that critical risk assessment analyses should be performed prior to possible release of viral vectors in agricultural practice.

ACKNOWLEDGEMENTS

We thank Pat King for considerable work in improving the English of this contribution, and the many authors who provided unpublished information, especially B. Hohn, J. Paszkowski, W. O. Dawson, K. Richards and G. Lomonossoff.

REFERENCES

Ahlquist, P. (1992). *Curr. Opin. Gen. and Dev.* **2**, 71–76.
Ahlquist, P. and Janda, M. (1984). *Mol. Cel. Biol.* **4**, 2876–2882.
Ahlquist, P. and Pacha, R. F. (1991). *Physiol. Plant.* **79**, 163–172.
Allison, R. F., Janda, M. and Ahlquist, P. (1988). *J. Virol.* **62**, 3581–3588.
Angenent, G. C., Posthumus, E. and Bol, J. F. (1989). *Virology* **173**, 68–76.
Bakkeren, G., Koukolikova-Nicola, Z., Grimsley, N. and Hohn, B. (1989). *Cell* **57**, 847–857.
Beck, D. L. O., Forster, R. L. S., Bevan, M. W., Boxen, K. A. and Lowe, S. C. (1990). *Virology* **177**, 152–158.
Bonneville, J. and Hohn, T. (1993). *In* 'Reverse Transcriptase' (N. Skalka and S. Goff, eds), pp. 357–390. Cold Spring Laboratory, Cold Spring Harbor.
Boulton, M. I., Buchholz, W. G., Marks, M. S., Markham, P. G. and Davies, J. W. (1989). *Plant Mol. Biol.* **12**, 31–40.
Boyer, J. and Haenni, A. L. (1994). *Virology* **198**, 415–426.
Brisson, N. and Hohn, T. (1986). *Methods Enzymol.* **118**, 659–668.
Brisson, N. and Hohn, T. (1988). *Methods Plant Mol. Biol.* 437–446.
Brisson, N., Paszkowski, J., Penswick, J. R., Gronenborn, B., Potrykus, I. and Hohn, T. (1984). *Nature* **310**, 511–514.
Bujarski, J. J. and Dzianott, A. M. (1991). *J. Virol.* **65**, 4153–4159.
Bujarski, J. J. and Kaesberg, P. (1986). *Nature* **531**, 528–531.
Burgyan, J., Nagy, P. D. and Russo, M. (1990). *J. Gen. Virol.* **71**, 1857–1860.
Chapman, S., Kavanagh, T. and Baulcombe, D. C. (1992). *Plant J.* **2**, 549–557.
Commandeur, U., Jarausch, W., Li, Y., Koenig, R. and Burgermeister, W. (1991). *Virology* **185**, 493–495.
Dasgupta, I., Hull, R., Eastop, S., Poggi-Pollini, C., Blackebrough, M. L., Boulton, M. I. and Davies, J. W. (1991). *J. Gen. Virol.* **72**, 1215–1221.
Davies, J. W. and Stanley, J. (1989). *TIG* **5**, 77–81.
Dawson, W. O., Beck, D. L., Knorr, D. A. and Grantham, G. L. (1986). *Proc. Natl Acad. Sci. USA* **83**, 1832–1836.
Dawson, W. O., Lewandowski, D. J., Hilf, M. E., Bubrick, P., Raffo, A. J., Shaw, J. J., Grantham, G. L. and Desjardins, P. R. (1989). *Virology* **172**, 285–292.
Deom, C. M., Oliver, M. J. and Beachy, R. N. (1987). *Science* **237**, 3284–3288.
de Zoeten, G. A., Penswick, J. R., Horisberger, M. A., Ahl, P., Schultze, M. and Hohn, T. (1989). *Virology* **172**, 213–222.
Dixon, L. K., Koenig, I. and Hohn, T. (1983). *Gene* **25**, 189–199.
Domier, L. L., Franklin, K. M., Hunt, A. G., Rhoads, R. E. and Shaw, J. G. (1989). *Proc. Natl Acad. Sci. USA* **86**, 3509–3513.
Domingo, E. and Holland, J. J. (1988). *In* 'RNA Genetics III' (E. Domingo, J. J. Holland, and P. Ahlquist, eds), pp. 3–36. CRC Press, Boca Raton.
Donson, J., Gunn, H. V., Woolston, C. J., Pinner, M. S., Boulton, M. I., Mullineaux, P. M. and Davies, J. W. (1988). *Virology* **162**, 248–250.
Donson, J., Kearney, C. M., Hilf, M. E. and Dawson, W. O. (1991). *Proc. Natl Acad. Sci. USA* **88**, 7204–7208.
Dore, J. and Pinck, L. (1988). *J. Gen. Virol.* **69**, 1331–1338.
Dore, J., Erny, C. and Pinck, L. (1990). *FEBS Lett.* **264**, 183–186.

Eggen, R., Verver, J., Wellink, J., De Jong, A., Goldbach, R. and Van Kammen, A. (1989). *Virology* **173**, 447–455.
Ermak, G., Paszkowski, U., Wohlmuth, M., Mittelsten-Scheid, O. and Paszkowski, J. (1993). *Nucl. Acids Res.* **21**, 3445–3450.
Etessami, P., Watts, J. W. and Stanley, J. (1989). *J. Gen. Virol.* **70**, 277–289.
Falk, B. W. and Bruening, G. (1994). *Science* **163**, 1395–1396.
French, R., Janda, M. and Ahlquist, P. (1986). *Science* **231**, 1294–1297.
Fütterer, J., Bonneville, J. and Hohn, T. (1990). *Physiol. Plant* **79**, 154–157.
Gal, S., Pisan, B., Hohn, T., Grimsley, N. and Hohn, B. (1991). *EMBO J.* **10**, 1571–1578.
Gal, S., Pisan, B., Hohn, T., Grimsley, N. and Hohn, B. (1992). *Virology* **187**, 525–533.
Gardner, R. C., Chanolens, K. R. and Owens, R. A. (1986). *Plant Mol. Biol.* **6**, 221–228.
Givord, L., Dixon, L. K., Rauseo-Koenig, I. and Hohn, T. (1988). *Ann. Inst. Pasteur/Virol.* **139**, 227–231.
Gluzman, Y. and Hughes, S. H. (1988). '*Viral Vectors*'. CSH 'Cold Spring Laboratory, Cold Spring Harbor'.
Greene, A. E. and Allison, F. (1994). *Science* **263**, 1423–1425.
Grill, L. K. (1993). *Agro-Food-Industry Hi-tech* **6**, 20–23.
Grimsley, N., Hohn, B., Hohn, T. and Walden, R. M. (1986a). *Proc. Natl Acad. Sci. USA* **83**, 3282–3286.
Grimsley, N., Hohn, T. and Hohn, B. (1986b). *EMBO J.* **5**, 641–646.
Grimsley, N., Hohn, T., Daubert, S. D. and Hohn, B. (1987). *Nature* **325**, 177–179.
Grimsley, N., Ramos, C., Hein, T. and Hohn, B. (1988). *Bio/Technology* **6**, 185–189.
Gronenborn, B. (1987). *Plant Gene Res.* **4**, 1–29.
Gronenborn, B. and Matzeit, V. (1989). *Cell Culture Somatic Cell Gen. Plants* **6**, 69–100.
Hamilton, W. D. O. and Baulcombe, D. C. (1989). *J. Gen. Virol.* **70**, 963–968.
Hanley-Bowdoin, L., Elmer, J. S. and Rogers, S. G. (1988). *Nucl. Acids Res.* **16**, 10511–10528.
Hay, J. M., Jones, M. C., Blackebrough, M. L., Dasgupta, I., Davies, J. W. and Hull, R. (1991). *Nucl. Acids Res.* **19**, 2615–2621.
Hayes, R. J., MacDonald, M. H., Coutts, R. H. A. and Buck, K. W. (1988a). *J. Gen. Virol.* **69**, 891–896.
Hayes, R. J., Petty, I. T. D., Coutts, R. H. A. and Buck, K. W. (1988b). *Nature* **334**, 179–182.
Hayes, R. J., Coutts, R. H. A. and Buck, K. W. (1989). *Nucl. Acids Res.* **17**, 2391–2403.
Hearne, P. Q., Knorr, D. A., Hillman, B. I. and Morris, T. J. (1990). *Virology* **177**, 141–151.
Heaton, L. A., Carrington, J. C. and Morris, T. J. (1989). *Virology* **170**, 214–218.
Hemenway, C., Weiss, J., O'Connell, K. and Tumer, N. E. (1990). *Virology* **175**, 365–371.
Hille, J., Dekker, M., Luttighuis, H., VanKammen, A. and Zabel, P. (1986). *Molec. Gen. Genet.* **205**, 411–416.
Hirochika, H. and Hayashi, K. (1991). *Gene* **105**, 239–241.
Hohn, B., Balázs, E., Rüegg, D. and Hohn, T. (1986). *EMBO J.* **5**, 2759–2762.
Hohn, B., Grimsley, N., Pisan, B. and Hohn, T. (1987a). *CIBA Found.* **133**, 185–195.
Hohn, B., Hohn, T., Grimsley, N., Boulton, M. I. and Davies, J. W. (1987b). *NATO Meeting Plant Mol. Biol. Ser.* **A140**, 459–468.
Hohn, T. and Fütterer, J. (1991). *Sem. Virol.* **2**, 55–70.
Hohn, T. and Goldbach, R. (1994). Vectors: Plant Viruses. *In* 'Encyclopedia of Virilogy'. 1536–1543. Academic Press Ltd, London.
Hohn, T. and Schell, J. (1987). *Plant Gene Res.* **4**, 1–348.
Hull, R. (1991). *Nucl. Acids Res.* **19**, 2615–2621.
Janda, M., French, R. and Ahlquist, P. (1987). *Virology* **158**, 259–262.
Joshi, R. L. and Joshi, V. (1991). *FEBS Lett.* **281**, 1–8.
Joshi, R. L., Joshi, V. and Ow, D. W. (1990). *EMBO J.* **9**, 2663–2669.
Jupin, I., Richards, K., Jonard, G., Guilley, H. and Pleij, C. W. A. (1990) *Virology* **178**, 273–280.
Jupin, I., Guilley, H., Richards, K. E. and Jonard, G. (1991). *EMBO J.* **11**, 479–488.
Kammann, M., Matzeit, V., Schmidt, B., Schell, J., Walden, R. M. and Gronenborn, B. (1991). *Gene* **104**, 247–252.
Kearney, C. M., Donson, J., Jones, G. E. and Dawson, W. O. (1993). *Virology* **192**, 11–17.

Kheyr-Pour, A., Bendahmane, M., Matzeit, V., Accotto, G. P., Crespi, S. and Gronenborn, B. (1992). *Nucl. Acids Res.* **19**, 6763–6769.
King, M. Q. (1988). *In* 'RNA Genetics II' (E. Domingo, J. J. Holland and P. Ahlquist, eds), pp. 149–166. (CRC Press Boca Raton.)
Kumagai, M. H., Turpen, T. H., Weinzettl, N., Della-Cioppa, G., Turpen, A. M., Donson, J., Hilf, M. E., Grantham, G. L., Dawson, W. O., Chow, T. P., Piatak, M. and Grill, L. K. (1993). *Proc. Natl. Acad. Sci. USA* **90**, 427–430.
Laufs, J., Wirtz, U., Kammann, M., Matzeit, V., Schaefer, S., Schell, J., Czernilovski, A. P., Baker, B. and Gronenborn, B. (1990). *Proc. Natl Acad. Sci. USA* **87**, 7752–7756.
Lazarowitz, S. G. (1992). *Crit. Rev. Plant Sci.* **11**, 327–349.
Lazarowitz, S. G., Pinder, A. J., Damsteegt, V. D. and Rogers, S. G. (1989). *EMBO J.* **8**, 1023–1032.
Lefebvre, D. D. (1990). *Plant Physiol.* **93**, 522–524.
Lefebvre, D. D., Miki, B. L. and Lalibert, J. F. (1987). *Bio/Technology* **5**, 1035–1056.
Leiser, R., Ziegler-Graff, V., Reutenauer, A., Herbach, E., Lemaire, O., Guilley, H., Richards, K. E. and Jonard, G. (1992). *Proc. Natl Acad. Sci. USA* **89**, 9136–9140.
Liljeström, P. and Garoff, H. (1991). *Bio/Technology* **9**, 1356–1361.
Lucas, W. J. and Wolf, S. (1993). *TICB* **3**, 309–315.
McCullough, A. J., Lou, H. and Schuler, M. A. (1991). *Nucl. Acids Res.* **19**, 3001–3009.
MacFarlane, S. A., Wallis, C. V., Taylor, S. C., Goulden, M. G., Wood, K. R., and Davies, J. W. (1991). *Virology* **182**, 124–129.
McLain, L. (1995). "Aids Research and Human Retroviruses" **11**, 327–334.
Matzeit, V., Schaefer, S., Kammann, M., Schalk, H. J., Schell, J. and Gronenborn, B. (1991). *Plant Cell* **3**, 247–258.
Maule, A. J. (1985). *Plant Mol. Biol.* **5**, 25–345.
Meshi, T., Ishikawa, M., Motoyoshi, F., Semba, K. and Okada, Y. (1986). *Proc. Natl Acad. Sci. USA* **81**, 1966–1970.
Mori, M., Mise, K., Kobayashi, K., Okuno, T. and Furusawa, I. (1991). *J. Gen. Virol.* **72**, 243–246.
Nouelry, A. O., Lucas, W. J. and Gilbertson, R. L. (1994). *Cell* **76**, 925–932.
Paszkowski, J., Shinshi, H., Koenig, I., Lazar, G., Hohn, T., Mandak, V. and Potrykus, I. (1983). 'Protoplasts', pp. 138–139. Birkhaeuser, Basel.
Paszkowski, U., Zhang, S., Potrykus, I. and Paszkowski, J. (1993). *J. Gen. Virol.* **74**, 2725–2729.
Petty, I. T. D., Hunter, B. G., Wei, N. and Jackson, A. O. (1989). *Virology* **171**, 342–349.
Porta, C., Spall, V. E., Woodland, J., Johnson, J. E., Barker, P. J. and Lomonossoff, G. (1994). *Virology*. Submitted. **202**, 949–955.
Qu, R., Bhattacharyya, M., Laco, G. S., de Kochko, A., Subba Rao, B. L., Kaniewska, M. B., Elmer, J. S., Rochester, D. E., Smith, C. E. and Beachy, R. N. (1991). *Virology* **185**, 354–364.
Quillet, L., Guilley, H., Jonard, G. and Richards, K. (1989). *Virology* **172**, 293–301.
Raffo, A. J. and Dawson, W. O. (1991). *Virology* **184**, 277–289.
Riechmann, J. L., Lain, S. and Garcia, J. A. (1990). *Virology* **177**, 710–716.
Rizzo, T. M. and Palukaitis, P. (1990). *Mol. Gen. Genet.* **222**, 249–256.
Rochon, D. M. and Johnston, J. C. (1991). *Virology* **181**, 656–665.
Rollo, F. and Covey, S. N. (1985). *J. Gen. Virol.* **66**, 603–608.
Scheets, K., Khosravi-Far, R. and Nutter, R. (1993). **193**, 1006–1009.
Scholthof, H. B., Morris T. J. and Jackson, A. O. (1993). *Mol. Plant-Microbe Interact.* **6**, 300–322.
Shen, W. and Hohn, B. (1991). *Virology* **183**, 721–730.
Shen, W. and Hohn, B. (1992). *Plant J.* **2**, 35–42.
Shen, W. and Hohn, B. (1994). *Plant J.* **5**, 227–236.
Shen, W., Das, S. and Hohn, B. (1992). *Molec. Gen. Genet.* **233**, 388–394.
Takamatsu, N., Ishikawa, M., Meshi, T. and Okada, Y. (1987). *EMBO J.* **6**, 307–311.
Takamatsu, N., Watanabe, Y., Yanagi, H., Meshi, T., Shiba, T. and Okada, Y. (1990). *FEBS Lett.* **269**, 73–77.
Töpfer, R., Gronenborn, B., Schell, J. and Steinbiss, H. H. (1989). *Plant Cell* **1**, 133–140.
Turpen, T. H. and Dawson, W. O. (1992). 'Transgenic plants: Fundamentals and Applications'. Marcel Dekker, New York.

Turpen, T. H., Turpen, A. M., Weinzettl, N., Kumagai, M. H. and Dawson, W. O. (1993). *J. Virol. Methods* **42**, 227–239.
Ugaki, M., Ueda, T., Timmermans, M. C. P., Vieira, J., Elliston, K. O. and Messing, J. (1991). *Res.* **19**, 371–378.
Usha, R., Rohll, J. B., Spall, V. E., Shanks, M., Maule, A. J., Johnson, J. E. and Lomonossoff, G. P. (1993). *Virology* **197**, 366–374.
Van der Kuyl, A. C., Neeleman, L. and Bol, J. F. (1991). *Virology* **183**, 731–738.
Veidt, I., Bouzoubaa, S. E., Ziegler-Graff, V., Guilley, H., Richards, K. and Jonard, G. (1992). *Virology* **186**, 192–200.
Viry, M., Serghini, M.A., Hans, F., Ritzenthaler, C., Pinck, M. and Pinck, L. (1993). *J. Gen. Virol.* **74**, 169–174.
Vos, P., Jaegle, M., Wellink, J., Verver, J., Eggen, R., Van Kammen, A. and Goldbach, R. (1988). *Virology* **165**, 33–41.
Ward, A., Etessami, P. and Stanley, J. (1988). *EMBO J.* **7**, 1583–1587.
Weiland, J. J. and Dreher, T. W. (1989). *Nucleic Acids Res.* **17**, 4675–4687.
Woolston, C. J., Barker, R., Gunn, H. V., Boulton, M. I. and Mullineaux, P. M. (1988). *Plant Mol. Biol.* **11**, 35–43.
Xiong, C., Lvis, R., Shen, P., Schlesinger, S., Rice, C. M. and Huang, H. V. (1989). *Science* **243**, 1188–1191.
Xiong, Z. and Lommel, S. A. (1991). *Virology* **182**, 388–392.
Young, M. J., Kelly, L., Larking, P. J., Waterhouse, P. M. and Gerlach, W. L. (1991). *Virology* **180**, 372–379.
Ziegler-Graff, V., Bouzoubaa, S., Jupin, I., Guilley, H., Jonard, G. and Richards, K. (1988). *Virol.* **69**, 2347–2357.

Index