Downloaded from rstb.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Advantages of Arabidopsis for Cloning Plant Genes

Caroline Dean

Phil. Trans. R. Soc. Lond. B 1993 **342**, 189-195
doi: 10.1098/rstb.1993.0146

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Phil. Trans. R. Soc. Lond. B* go to: **http://rstb.royalsocietypublishing.org/subscriptions**

This journal is © 1993 The Royal Society

Downloaded from rstb.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Advantages of Arabidopsis for cloning plant genes

CAROLINE DEAN

Cambridge Laboratory, AFRC IPSR, John Innes Centre, Colney, Norwich NR4 7UJ, U.K.

SUMMARY

Arabidopsis thaliana (Arabidopsis) has been adopted by a large number of plant biologists as the model organism in which to use a molecular genetic approach to dissect many plant processes. The small genome size of Arabidopsis and the concentration of effort on this plant species has led to an international collaboration to construct a physical map of the genome. The physical map will greatly facilitate gene isolation without the need for each laboratory to initiate chromosome walks. This, in combination with the development of efficient insertional mutagenesis systems, will mean that in the next few years the number of genes cloned from Arabidopsis will increase enormously.

1. INTRODUCTION

Many processes fundamental to the development and functioning of flowering plants remain very poorly understood. Of special interest are processes such as the transition from vegetative development to flowering, the perception of and reaction to environmental stimuli, and the molecular recognition involved in plant–pathogen interactions. Until recently, these processes have not been analysed using molecular genetic techniques. This situation, however, is changing rapidly. There is now considerable emphasis on the isolation of genes using molecular genetic approaches in species such as maize, snapdragon, tomato, wheat, rice and most notably, Arabidopsis. Arabidopsis has many advantages for such an analysis:

1. It is a very small plant with a rapid generation time (6 weeks from seed to seed). This makes large mutagenesis screens and genetic studies quite feasible. Each plant also produces large numbers (up to 10 000) of seed.
2. The haploid genome size is estimated to be approximately 100 Mb (Hwang *et al.* 1991), similar to that of the nematode *Caenorhabditis elegans*. This is considerably smaller than most other well studied plant systems. It also has a very low level of repeated DNA sequences (Leutwiler *et al.* 1984; Pruitt & Meyerowitz 1986).
3. There are over 360 mapped restriction fragment length polymorphism (RFLP) markers (Chang *et al.* 1988; Nam *et al.* 1989; R. Whittier, personal communication) This translates to an average distance from a known marker to any point in the genome of around 70 kb (Meyerowitz *et al.* 1991). Randomly amplified polymorphic DNA (RAPD) markers (Williams *et al.* 1990) are also being employed and currently 252 of these markers are mapped onto the Arabidopsis genetic map (Reiter *et al.* 1992).
4. The visible marker map currently has 117 markers distributed over the five chromosomes. The number of newly identified mutations, induced with a variety of mutagens, is increasing at a considerable rate. Newly identified loci are being mapped, either relative to other visible markers, using the 19 mapping strains with multiple markers on each chromosome developed by M. Koornneef (Koornneef *et al.* 1987; Meyerowitz *et al.* 1991), or relative to RFLP, RAPD or cleared amplified polymorphic sequence (CAPS) markers (Konieczny & Ausubel 1993).
5. A transformation system using the soil bacterium *Agrobacterium tumefaciens* is available. Many *Agrobacterium* vector systems using a number of different antibiotic or herbicide resistance markers have been developed to select for transformants carrying the integrated T-DNA (the part of the *Agrobacterium* Ti plasmid that is integrated into the plant chromosome). A transformation procedure where a culture of *Agrobacterium* is mixed with root explants (Valvekens *et al.* 1988), followed by regeneration of intact plants from the transformed root material, is being used by many laboratories. A non-tissue-culture procedure based on imbibing seed in a solution of *Agrobacterium* has also been successful for Feldmann & Marks (1987).

2. CHROMOSOME WALKING AND PHYSICAL MAPPING

(a) Chromosome walking experiments to clone genes based on phenotype and map position

The increasing density of RFLP/RAPD markers and the small genome size with relatively few repeated DNA sequences mean that gene isolation using chromosome walking is considerably easier in Arabidopsis than in most other plant species. During the past couple of years many chromosome walking experiments have been initiated, with the aim of cloning various loci. These walks are generating local physical maps at

Phil. Trans. R. Soc. Lond. B (1993) **342**, 189–195
Printed in Great Britain

© 1993 The Royal Society

The Royal Society is collaborating with JSTOR to digitize, preserve, and extend access to *Philosophical Transactions: Biological Sciences.* STOR® www.jstor.org

Downloaded from rstb.royalsocietypublishing.org

Table 1. *Examples of specific loci being cloned by chromosome walking*

locus	chr	locus	chr
ABI1 (abscisic acid insensitivity)	4	*FCA* (late flowering)	4
ABI2 (abscisic acid insensitivity)	5	*FVE* (late flowering)	2
ABI3 (abscisic acid insensitivity)	3	*FWA* (late flowering)	4
AXR1 (auxin resistance)	1	*GI2* (late flowering)	1
AXR2 (auxin resistance)	3	*FAE1* (deficient in fatty acid elongation)	4
CO (late flowering)	5	*TTG* (glabrous, lack of anthocyanins)	5
EIN2 (ethylene insensitive)	5	*ARA1* (arabinose sensitivity)	4
ETR (ethylene resistance)	1	*MS1* (male sterility)	5
GA2 (gibberellic acid deficiency)	1	*RPM1* (resistance to a *Pseudomonas* pathovar)	3
GAI (gibberellic acid insensitive)	1	*RPT1* (resistance to a *Pseudomonas* pathovar)	3
DET1 (de-etiolated)	4	*RPT2* (resistance to a *Pseudomonas* pathovar)	4
DET2 (de-etiolated)	2	*RPP5* (resistance to a *Peronospora* pathovar)	4
FAD3 (fatty acid desaturation)	2	*TFL1* (terminal flower)	5

various regions around the genome. The walks initially used cosmid libraries (Hauge *et al.* 1991; Bleecker 1991) constructed in the cosmid vectors, pCIT30 (Yanofsky *et al.* 1990) or Lorist (Cross & Little 1986). More recently, the availability of several yeast artificial chromosome (YAC) libraries has meant that most chromosome walking experiments employ YAC clones since the inserts are on average three times larger than those in cosmid clones. More than 30 chromosome walking projects in over 20 laboratories are currently known to be in progress (Meyerowitz *et al.* 1991). Some examples are summarized in table 1.

The general strategy that has been adopted for chromosome walking experiments in Arabidopsis is to generate a YAC contig covering the region containing the locus of interest, in combination with fine mapping of the locus within that region (outlined in figure 1). The locus to be cloned (b) is initially mapped relative to other visible markers (e.g. A and C). These flanking markers are then used to select recombination events close to b in F2 progeny from a cross between two parents (e.g. Landsberg *erecta* and Columbia) which are polymorphic at the DNA level. DNA is isolated from pooled progeny plants carrying recombination events between b and A or b and C. RFLP markers mapping to the same region are hybridized to Southern blots of DNA from these recombinants. Recombination points are mapped relative to these DNA markers by monitoring the switch of the fragment pattern on the Southern blots from a homozygous pattern to a heterozygous pattern (i.e. both Landsberg *erecta* and Columbia alleles present). The analysis of the recombination break points determines the map position of b with respect to the RFLP markers.

YAC clones corresponding to the RFLP markers flanking b are then identified by using the RFLP markers as probes on YAC libraries. End-probes from these YAC clones are generated and hybridized back to the YAC libraries to find overlapping YAC clones. The direction of the walk towards b is determined by utilizing the recombination events between A and b or C and b and end-probes from the YAC clones which reveal polymorphisms between Landsberg *erecta* and Columbia. Recombination points again are mapped relative to these DNA markers by monitoring the switch from the homozygous to the heterozygous pattern. Once the walks are oriented, they are extended towards b in order to build one contiguous region covering b.

As soon as the locus of interest has been localized to one or part of one YAC clone, DNA from this YAC clone is subcloned into vectors, which can be used in plant transformation experiments. This can be achieved by subcloning the YAC directly or using it as a probe to a library of Arabidopsis DNA cloned in a plant transformation vector. The subclones are then used to transform the mutant Arabidopsis line and identify clones which complement the mutant phenotype. Final identification of the open reading frame is through further rounds of complementation with smaller subclones, in conjunction with sequence and transcript analysis. The first two successful chromosome walks have now been described (Arondel *et al.* 1993; Giraudat *et al.* 1992). Many more have localized the locus to one cosmid clone and are in the process of final rounds of complementation with transcript and sequence analysis.

3. AN OVERLAPPING LIBRARY OF THE ARABIDOPSIS GENOME

(*a*) *Generation of an overlapping cosmid library*

To facilitate future cloning and avoid the need for each laboratory to do chromosome walks to loci of interest, Hauge *et al.* (1991) initiated a project several years ago, to generate an overlapping cosmid library of the Arabidopsis genome. They adopted the same non-directed strategy used for the *C. elegans* genome project by Coulson, Sulston and co-workers (Coulson *et al.* 1986). Random cosmid clones (average size 40 kb) were 'fingerprinted' by digestion with HindIII (which cut each cosmid an average of 15 times), labelled with ^{32}P dATP and subsequently cleaved with Sau3A. The labelled fragments were then separated by polyacrylamide gel electrophoresis and the banding patterns entered into the computer using a scanning densitometer and an image-processing package (Sulston *et al.* 1989). The regions of overlap were determined and the clones were assembled into contigs (Coulson *et al.* 1986; Sulston *et al.* 1988). Seven

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Downloaded from rstb.royalsocietypublishing.org

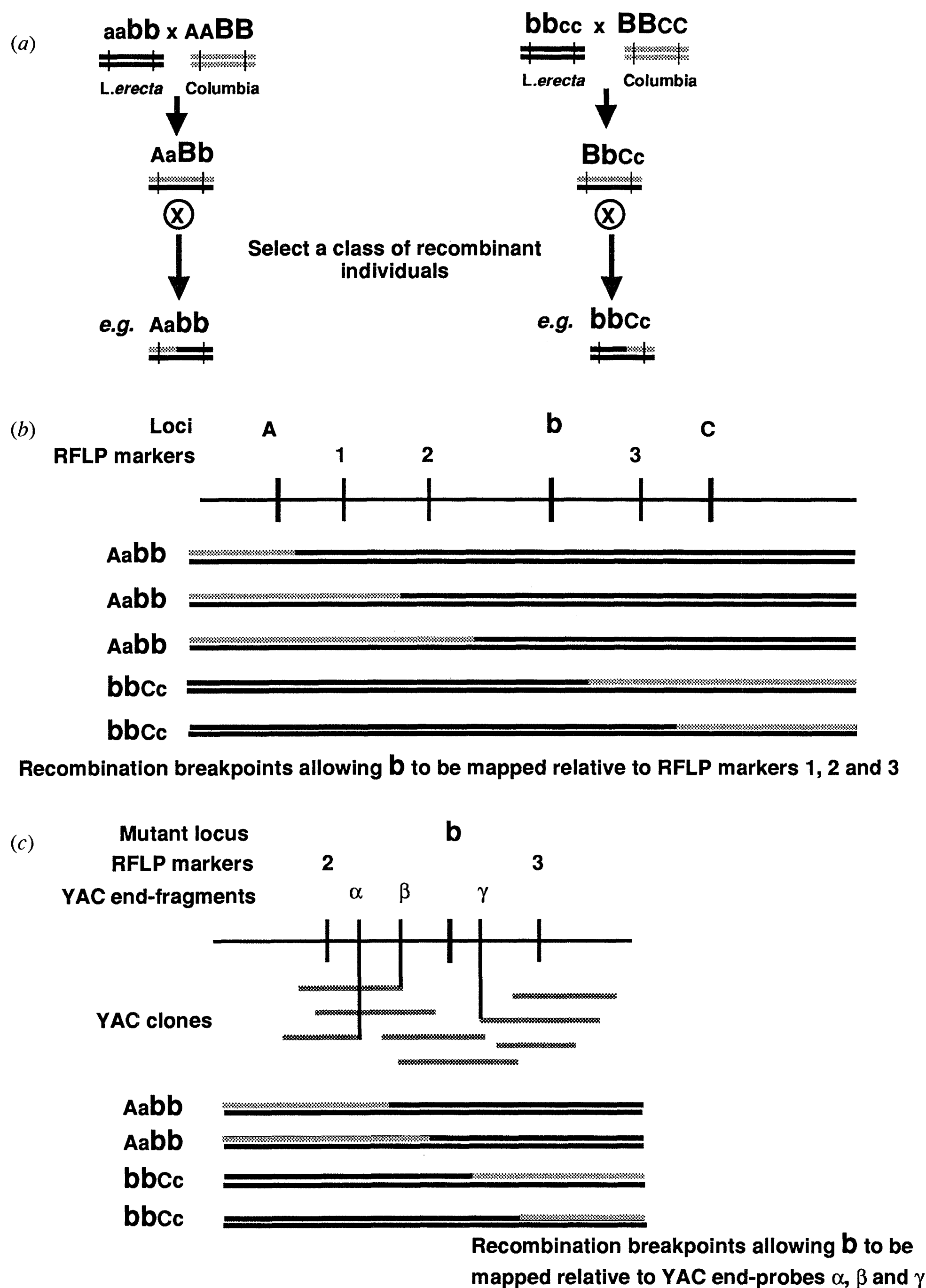

Figure 1. Chromosome walking to a locus. (*a*) Crosses are made to isolate recombinant individuals between the locus of interest **b** and flanking loci A and C. These crosses can be set up with either the flanking markers in coupling (as shown in the figure) or in repulsion (aa**BB** × AA**bb**). The mutant, for example in the Landsberg *erecta* background (black lines), is crossed to a wild-type line of a different ecotype, e.g. Columbia (stippled line), which is polymorphic at the DNA level. Resulting F1 plants are allowed to self pollinate and recombinant individuals are selected for in the F2 generation. In the figure only one class of possible recombinants is indicated. (*b*) The recombinant individuals are subjected to an RFLP analysis. Recombination break points for five different recombinants are depicted, although it should be noted that the break points can only be defined to an area between two markers in such an analysis. The distribution of the recombination events in the different individuals allows the RFLP markers 1, 2 and 3 used in this analysis to be mapped with respect to the loci A, **b** and C. (*c*) YAC clones corresponding to the RFLP markers 2 and 3 are isolated. End-fragments of some of these YAC clones are generated to serve as probes for the isolation of overlapping YAC clones to form a contiguous region covering the locus. Furthermore, end-probes which reveal an RFLP between the two Arabidopsis ecotypes (α,β,γ) are used to obtain a higher resolution map position of **b** relative to the YAC end fragments using recombinant individuals.

Downloaded from rstb.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

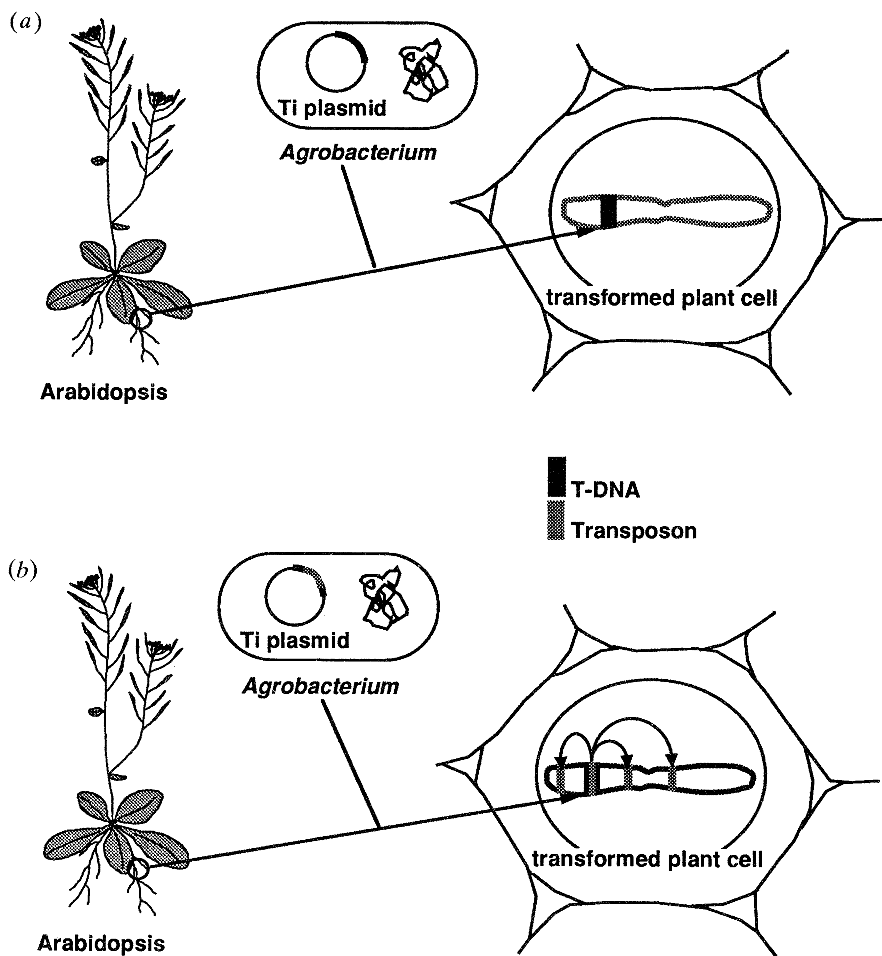

Figure 2. Insertions of T-DNA and heterologous transposable elements into the genome of Arabidopsis. (*a*) Random integration of the T-DNA into the Arabidopsis genome after infection of plant cells with *Agrobacterium tumefaciens*. (*b*) Random integration of a T-DNA containing a transposable element into the Arabidopsis genome. Upon excision and reinsertion of the transposon, new random integration sites are obtained. Most re-insertion sites are linked to the T-DNA.

hundred and fifty contigs have been assembled after analysis of 17000 clones. Cosmid contigs cover 91–95% of the genome. Based on this work, Hauge and Goodman estimate the genome size to be approximately 100 Mb (Hwang *et al.* 1991). Taking into account a genome size of 100 Mb and a 91–95% coverage of the genome in cosmid contigs, the average size of each of the identified cosmid contigs is approximately 120–130 kb. The number of cosmids fingerprinted represent 8–10 genome equivalents, so a continuation of this approach to reduce the number of contigs is no longer fruitful. The difficulties of obtaining map closure using cosmid clones are thought to be caused by regions of eukaryotic genomes being unclonable in *E. coli* (Coulson *et al.* 1991). It may also be due to the relatively large overlap (35%) required before a cosmid is included in a contig or the presence of clones which contain few or no HindIII sites, since a minimum number of bands is required in the statistical analysis to detect overlapping clones (Hauge *et al.* 1991).

(*b*) *Generation of an overlapping* YAC *library of the Arabidopsis genome*

In addition to the effort described above, there is a large international collaboration to generate an overlapping yeast artificial chromosome (YAC) library of the Arabidopsis genome. Using YAC clones provides two advantages over cosmid clones. First, YAC clones carry much larger inserts and second, it is thought likely that there will be fewer unclonable sequences in yeast than in *E. coli* (Coulson *et al.* 1991). As well as providing cloned regions of all the genome, a representative set of YAC clones, arrayed in order, could provide a minimal set of clones covering the Arabidopsis genome with which accurately to map new probes.

This project has taken advantage of the relatively extensive framework of mapped markers and characterized repeated DNA sequences to position YAC clones on the genetic map (Meyerowitz *et al.* 1991). The three YAC libraries used in the initial stages of this project were constructed from DNA isolated from the Arabidopsis ecotype Columbia or the Landsberg *erecta* mutant *abi-1* and fractionated using either BamHI partial digestion (EG and *abi* libraries; Grill & Somerville 1991*a*) or random shear of high-molecular-mass Arabidopsis DNA (EW library; Ward & Jen 1990). A derivative of pYAC4 (Burke *et al.* 1987), pYAC 41, carrying T3 bacteriophage promoters flanking a BamHI cloning site was the vector used for

Downloaded from rstb.royalsocietypublishing.org

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

two of the libraries (Grill & Somerville 1991*a*). Each of these YAC libraries contains between 2100 and 2300 YAC clones. The average insert size has been established to be approximately 150 kb for the two Columbia libraries. Therefore, these libraries should contain approximately three genome equivalents. A new library of 2300 clones, constructed using EcoRI partial digests cloned into pYAC4 (Burke *et al.* 1987), is now available and this has an average insert size of 250 kb (Ecker 1990).

The strategy adopted by the different research groups involved in the linking of the YAC clones was to hybridize all the available RFLP markers (Chang *et al.* 1988; Nam *et al.* 1989), in the region being focused on by each group, to at least one of the available YAC libraries. After YAC clones corresponding to all mapped markers were identified, walks would then be initiated to link up the YAC clones hybridizing to adjacent RFLP markers. Extensive chromosome walking on the top half of chromosome 4 (R. Schmidt, J. West & C. Dean, unpublished data) has shown that this would be a labour intensive way to link up all the YAC contigs covering a whole chromosome. Hybridization of new probes to the YAC libraries appears to be the most efficient method to link existing contigs. For this reason, attention has been focused on mapping new markers onto the RFLP map. This objective is made easier by the availability of recombinant inbred lines (Reiter *et al.* 1992; Lister & Dean 1993). Many hundreds of new markers are currently being mapped by a number of groups internationally. In addition, hybridization of representative cosmids from the cosmid contigs and sequenced cDNA clones to the YAC libraries should provide clones to identify YAC clones extending and therefore joining the existing YAC contigs.

In April 1991, the first results from the laboratories participating in the international collaboration to generate an overlapping YAC library were compiled and presented in Hwang *et al.* (1991). This initial effort resulted in approximately 30% of the Arabidopsis genome being represented in mapped YAC clones. Furthermore, a number of RFLP markers (25% of the markers tested) were shown to be physically linked by one or more YAC clones. Since then work has progressed and for example, there is currently 80% coverage of the top halves of chromosomes 4 and 5 (R. Schmidt, J. West & C. Dean, unpublished data).

(*c*) *Sequencing of the Arabidopsis genome*

The final aim of the Arabidopsis physical mapping project will be the complete sequence analysis of the nuclear and organellar genomes. The sequencing of the chloroplast and mitochondrial genomes have been initiated (Meyerowitz *et al.* 1991) and 160 kb of unique sequence in numerous small linkage groups has already been established for the mitochondrial genome (Gröncer *et al.* 1991). Large scale sequencing of the nuclear genome has already been started by H. Goodman and two 40 kb cosmid clones have been sequenced (Hauge *et al.* 1991). There are also international programmes to fund large scale sequencing of genomic regions (within the EC ESSA programme coordinated by M. Bevan, Norwich, U.K.) and random cDNA clones (C. Somerville and colleagues, Michigan State University, U.S.A.; M. Caboche and colleagues, GDP, France). Databases to store all the information produced from the mapping and the sequence analysis have been established (Cherry *et al.* 1992; Meyerowitz *et al.* 1991).

4. DEVELOPMENT OF INSERTIONAL MUTAGENESIS SYSTEMS IN ARABIDOPSIS

Insertional mutagenesis is a powerful strategy for the isolation of genes where little or nothing is known about the biochemical function of the gene product. There is a very large effort to develop an efficient insertional mutagenesis system in Arabidopsis. Two types of mutagenic element have been used, the T-DNA from the Ti plasmid of *Agrobacterium tumefaciens* and transposons. Over 10 000 lines carrying an introduced T-DNA insertion have been generated using the non-tissue culture transformation procedure (Feldmann & Marks 1987). Many of these result in a mutation and these lines have provided a vital resource with which to clone a number of important genes (Feldmann 1991).

In addition to T-DNA mutagenesis, transposon systems have also been developed. Transposons have the advantage over T-DNA tagging in that they can continue to excise and re-integrate after the transformation step and thus each line has the potential to generate many more mutations. Also any somaclonal mutations caused by the transformation steps can be discarded before transposition events are selected. The endogenous transposon systems in Arabidopsis have only recently been discovered and are relatively poorly characterized (Konieczny *et al.* 1991; Peleman *et al.* 1991; Tsay *et al.* 1993). For this reason, a number of heterologous transposons from either maize or snapdragon (reviewed in Bhatt & Dean 1992) have been introduced into Arabidopsis using *Agrobacterium* transformation. Most effort has focused upon the maize transposons *Ac (Activator)* and *Ds (Dissociation)*. The behaviour and activity of these elements has been extensively characterized (Schmidt & Willmitzer 1988; Dean *et al.* 1992; Bancroft *et al.* 1992; Swinburne *et al.* 1992; Grevelding *et al.* 1992; Keller *et al.* 1992; Bancroft & Dean 1993*a*). These elements preferentially transpose to linked sites in the genome (Bancroft & Dean 1993*b*). In a tagging experiment targetting a particular locus, it is therefore beneficial to start with a line carrying a transposon in a nearby position in the genome. For this reason a large number of T-DNA insertions carrying various modifications of the *Ac* and *Ds* elements are being mapped onto the Arabidopsis RFLP map. This is achieved by using inverse PCR (Ochman *et al.* 1989) to amplify the plant DNA flanking the T-DNA insertion carrying the transposon. The IPCR fragment is then cloned, used to detect polymorphisms between the parents used to generate the recombinant inbred lines, for example, Landsberg *erecta* and Columbia (Lister & Dean 1993) and then mapped using the recombinant inbred lines. At least

Downloaded from rstb.royalsocietypublishing.org

50 T-DNA integrations carrying transposons will soon have been mapped. This will enable tagging experiments to be initiated from linked sites anywhere in the genome.

Transposon-induced mutations have recently been isolated in Arabidopsis (Bancroft *et al.* 1993*c*). The first, *drl-1*, severely alters the development of all major organs of the plant. Another, *wlc-1*, has wavy leaves and cotyledons and alters the flowering time of the plant. Normally, in short photoperiods, Arabidopsis flowering is delayed. However, the flowering of *wlc-1* is not delayed (I. Bancroft & C. Dean, unpublished data). These mutants and others, continuously being isolated, are currently being analysed further.

This concentration of effort and easy access to the ever increasing amount of information should quickly ensure that gene cloning in Arabidopsis is no longer a limiting factor in the dissection of complex plant processes.

I thank all members of my laboratory for providing unpublished results and Dr David Flanders for help in preparation of the figures and critical reading of the manuscript.

REFERENCES

Arondel, V., Lemieux, B., Hwang, I., Gibson, S., Goodman, H. & Somerville, C. 1992 Map based cloning of a gene controlling omega-3 fatty acid desaturation in Arabidopsis. *Science, Wash.* **258**, 1353–1356.

Bancroft, I., Bhatt, A.M., Sjodin, C., Scofield, S., Jones, J.D.G. & Dean, C. 1992 Development of an efficient two-element transposon tagging system in *Arabidopsis thaliana*. *Molec. gen. Genet.* **233**, 449–461.

Bancroft, I. & Dean, C. 1993*a* Factors affecting the excision frequency of the maize transposable element *Ds* in *Arabidopsis thaliana*. *Molec. gen. Genet.* **240**, 65–72.

Bancroft, I. & Dean, C. 1993*b* Transposition pattern of the maize element *Ds* in *Arabidopsis thaliana*. *Genetics* **134**, 1221–1229.

Bancroft, I., Jones, J. & Dean, C. 1993 Heterologous transposon tagging of the *DRL1* locus in Arabidopsis. *Pl. Cell* **5**, 631–638.

Bhatt, A.M. & Dean, C. 1992 Development of tagging systems in plants using heterologous transposons. *Curr. Opin. Biotechnol.* **3**, 152–158.

Bleecker, A. 1991 Genetic analysis of ethylene responses in *Arabidopsis thaliana*. In *Molecular biology of plant development* (ed. G. I. Jenkins & W. Schuch), pp. 149–158. Cambridge: The Company of Biologists.

Burke, D.T., Carle, G.F. & Olson, M.V. 1987 Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science, Wash.* **236**, 806–812.

Chang, C., Bowman, J.L., DeJohn, A.W., Lander, E.S. & Meyerowitz, E.M. 1988 Restriction fragment length polymorphism linkage map for *Arabidopsis thaliana*. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6856–6860.

Cherry, J.M., Cartinhour, S.W. & Goodman, H.M. 1992 AAtDB, An *Arabidopsis thaliana* database. *Pl. molec. Biol. Rep.* **10**, 308.

Coulson, A., Sulston, J., Brenner, S. & Karn, J. 1986 Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7821–7825.

Coulson, A., Waterston, R., Kiff, J., Sulston, J. & Kohara, Y. 1988 Genome linking with yeast artificial chromosomes. *Nature, Lond.* **335**, 184–186.

Coulson, A., Kozono, Y., Lutterbach, B., Shownkeen, R., Sulston, J. & Waterston, R. 1991 YACs and the *C. elegans* genome. *BioEssays* **13**, 413–417.

Cross, S.H. & Little, P.F.R. 1986 A cosmid vector for systematic chromosome walking. *Gene* **4**, 9–13.

Dean, C., Sjodin, C., Page, T., Jones, J. & Lister, C. 1992 Behaviour of the maize transposable element *Ac* in *Arabidopsis thaliana*. *Plant J.* **2**, 69–81.

Ecker, J.R. 1990 PFGE and YAC analysis of the Arabidopsis genome. *Methods* **1**, 186–194.

Feldmann, K.A. & Marks, M.D. 1987 *Agrobacterium*-mediated transformation of germinating seeds of *Arabidopsis thaliana*: a non-tissue culture approach. *Molec. gen. Genet.* **208**, 1–9.

Feldmann, K.A. 1991 T-DNA insertion mutagenesis in Arabidopsis: mutational spectrum. *Plant J.* **1**, 71–82.

Giraudat, J., Hauge, B.M., Valon, C., Smalle, J., Percy, F. & Goodman, H.M. 1992 Isolation of the Arabidopsis *ABI3* gene by positional cloning. *Pl. Cell* **4**, 1251–1261.

Grevelding, C., Becker, D., Kunze, R., von Menges, A., Fantes, V., Schell, J. & Masterson, R. 1992 High rates of *Ac/Ds* germinal transposition in Arabidopsis suitable for gene isolation by insertional mutagenesis. *Proc. natn. Acad. Sci. U.S.A.* **89**, 6085–6089.

Grill, E. & Somerville, C. 1991*a* Construction and characterization of a yeast artificial chromosome library of Arabidopsis which is suitable for chromosome walking. *Molec. gen. Genet.* **226**, 484–490.

Grill, E. & Somerville, C. 1991*b* Development of a system for efficient chromosome walking in Arabidopsis. In *Molecular biology of plant development* (ed. G. I. Jenkins & W. Schuch), pp. 5762. Cambridge: The Company of Biologists.

Grönger, P., Unseld, M., Eckert, U. & Brennicke, A. 1991 Sequence analysis of a mitochondrial genome of *Arabidopsis thaliana* A progress report. In *Third international congress of plant molecular biology, molecular biology of plant growth and development*, Tucson, Arizona. Abstract no. 20–50.

Hauge, B.M., Hanley, S., Giraudat, J. & Goodman, H.M. 1991 Mapping the Arabidopsis genome. In *Molecular biology of plant development* (ed. G. I. Jenkins & W. Schuch), pp. 4556. Cambridge: The Company of Biologists.

Hwang, I., Kohchi, T., Hauge, B.M. *et al.* 1991 Identification and map position of YAC clones comprising one-third of the Arabidopsis genome. *Plant J.* **1**, 367–374.

Keller, J., Lim, E., James, D.W. Jr & Dooner, H. 1992 Germinal and somatic activity of the maize element *Activator* (*Ac*) in Arabidopsis. *Genetics* **131**, 449–459.

Koornneef, M., van Eden, J., Hanhart, C.J., Stam, P., Braaksma, F.J. & Feenstra, F.J. 1983 Linkage map of *Arabidopsis thaliana*. *J. Hered.* **74**, 265–272.

Koornneef, M., Hanhardt, C.J., van Loenen-Martinet, E.P. & van der Veen, J.H. 1987 A marker line, that allows the detection of linkage on all Arabidopsis chromosomes. *Arabidopsis Inf. Serv.* **23**, 46–50.

Koornneef, M. 1990 Linkage map of *Arabidopsis thaliana* In *Genetic maps, locus maps of complex genomes*, Book 6 (*Plants*) (ed. S. J. O'Brien), pp. 6.94–6.96. New York: Cold Spring Harbor Laboratory Press.

Konieczny A., Voytas, D.F., Cummings, M.P. & Ausubel, F.M. 1991 A superfamily of *Arabidopsis thaliana* retrotransposons. *Genetics* **127**, 801–809.

Konieczny, A. & Ausubel, F.M. 1993 A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* **4**, 403–410.

Leutwiler, L.S., Hough-Evans, B.R. & Meyerowitz, E.M.

Downloaded from rstb.royalsocietypublishing.org

1984 The DNA of *Arabidopsis thaliana*. *Molec. gen. Genet.* **194**, 15–23.

Lister, C. & Dean, C. 1993 Recombinant inbred lines of Arabidopsis for mapping phenotypic and RFLP markers. *Plant J.* (In the press.)

Meyerowitz, E., Dean, C., Flavell, R. *et al.* 1991 *The multinational coordinated Arabidopsis thaliana genome research project. Progress report: Year one.* Publ. 9160. Washington, D.C.: National Science Foundation.

Nam, H.-G., Giraudat, J., den Boer, B., Moonan, F., Loos, W.D.B., Hauge, B.M. & Goodman, H. 1989 Restriction fragment length polymorphism linkage map of *Arabidopsis thaliana*. *Pl. Cell* **1**, 699–705.

Ochman, H., Gerber, A.S. & Hartl, D.L. 1988 Genetic application of an inverse polymerase chain reaction. *Genetics* **120**, 621–623.

Peleman, J., Cottyn, B., Van Camp, W., Van Montagu, M. & Inze, D. 1991 Transient occurrence of extrachromosomal DNA of an *Arabidopsis thaliana* transposon-like element, *Tat1*. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3618–3622.

Pruitt, R.E. & Meyerowitz, E.M. 1986 Characterization of the genome of *Arabidopsis thaliana*. *J. molec. Biol.* **187**, 169–184.

Reiter, R.S., Williams, J.G.K., Feldmann, K.A., Rafalski, J.A., Tingey, S.V. & Scolnik, P.A. 1992 Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1477–1481.

Schmidt, R. & Willmitzer, L. 1989 The maize autonomous element *Activator* (*Ac*) shows a minimal germinal excision frequency of 0.2–0.5% in transgenic *Arabidopsis* plants. *Molec. gen. Genet.* **220**, 11–24.

Sulston, J., Mallett, F., Staden, R., Durbin, R., Horsnell, T. & Coulson, A. 1988 Software for genome mapping by fingerprinting techniques. *Comput. applic. Biosci.* **4**, 125–129.

Sulston, J., Mallett, F., Durbin, R. & Horsnell, T. 1989 Image analysis of restriction enzyme fingerprint autoradiograms. *Cabios* **5**, 101–105.

Swinburne, J., Balcells, L., Scofield, S.R., Jones, J.D.G. & Coupland, G. 1992 Elevated levels of *Ac* transposase mRNA are associated with high frequencies of *Ds* excision in Arabidopsis. *Pl. Cell* **4**, 583–595.

Tsay, Y.-F., Frank, M.J., Page, T., Dean, C. & Crawford, N.M. 1993 Identification of a mobile endogenous transposon in *Arabidopsis thaliana*. *Science, Wash.* **260**, 342–344.

Valvekens, D., van Montagu, M. & van Lijsebettens, M. 1988 *Agrobacterium tumefaciens*-mediated transformation of Arabidopsis root explants using kanamycin selection. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5536–5540.

Ward, E.R. & Jen, G.C. 1990 Isolation of single-copy-sequence clones from a yeast artificial chromosome library of randomly-sheared *Arabidopsis thaliana* DNA. *Pl. molec. Biol.* **14**, 561–568.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. & Tingey, S.V. 1990 DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucl. Acids Res.* **18**, 6531–6535.

Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.D., Feldmann, K.A. & Meyerowitz, E.M. 1990 The protein encoded by the Arabidopsis homeotic gene *agamous* resembles transcription factors. *Nature, Lond.* **346**, 35–39.