The Plant Journal (1993) 4(2), 403–410

TECHNICAL ADVANCE

A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers

Andrzej Konieczny[†] and Frederick M. Ausubel*
Department of Genetics, Harvard Medical School, 25 Shattuck St., Boston, MA 02115, USA, and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA

Summary

A set of mapping markers have been designed for *Arabidopsis thaliana* that correspond to DNA fragments amplifed by the polymerase chain reaction (PCR). The ecotype of origin of these amplified fragments can be determined by cleavage with a restriction endonuclease. Specifically, 18 sets of PCR primers were synthesized, each of which amplifies a single mapped DNA sequence from the Columbia and Landsberg *erecta* ecotypes. Also identifed was at least one restriction endonuclease for each of these PCR products that generates ecotype-specific digestion patterns. Using these co-dominant cleaved amplified polymorphic sequences (CAPS), an *Arabidopsis* gene can be unambiguously mapped to one of the 10 *Arabidopsis* chromosome arms in a single cross using a limited number of F_2 progeny.

Introduction

Genetic maps consisting primarily of restriction fragment length polymorphic (RFLP) markers are being constructed for a variety of plants including maize (Burr and Burr, 1991; Helentjaris, 1987), barley (Tragoonrung *et al.*, 1992), wheat (Kam-Morgan and Gill, 1989), rice (McCouch *et al.*, 1988), tomato (Bernatsky and Tanksley, 1986), lettuce (Landry *et al.*, 1987), *Brassica oleraceae* (Slocum *et al.*, 1990) and *Arabidopsis thaliana* (Chang *et al.*, 1988; Nam *et al.*, 1989). RFLP maps are well-suited to mapping newly cloned DNA sequences. However, most plant genes are first identified by mutation. Mapping such a mutation on to a pre-existing RFLP map is a lengthy procedure requiring the isolation of DNA from individual F_2 plants or F_3 families, and performing DNA blot analysis using each of the RFLP markers as a hybridization probe.

Received 18 December 1992; revised 4 March 1993; accepted 10 March 1993.
*For correspondence at the Massachusetts General Hospital (fax +1 617 726 5949).
[†]Present address: ImmuLogic Pharmaceutical Corporation, 610 Lincoln Street, Waltham, MA 02154, USA.

Recently, techniques based on the polymerase chain reaction (PCR) (Mullis and Faloona, 1987) have been used in addition to or in place of traditional RFLP markers in genetic analysis (Cox and Lehrach, 1991). In contrast to traditional RFLP markers, PCR-generated markers can be scored using a small sample of DNA without the use of radioactivity and without the time-consuming DNA blotting procedure. One widely used PCR-based approach involves the use of single short PCR primers of arbitrary sequence (called RAPD primers for random amplified polymorphic DNA; Reiter *et al.*, 1992; Williams *et al.*, 1990). . A major advantage of RAPDs is that they provide large numbers of markers. On the other hand, because the amplification of a specific sequence or sequences using a RAPD primer is frequently sensitive to PCR conditions, including template concentration, it can be difficult to correlate results obtained by different research groups (Devos and Gale, 1992). A second limitation of the RAPD method is that it usually cannot distinguish heterozygotes from one of the two homozygous genotypes (Williams *et al.*, 1990). Finally, RAPD primers frequently amplify more than one sequence, which can complicate the analysis (Riedy *et al.*, 1992). In this report we describe a simple PCR-based strategy to map a gene to one of the 10 *Arabidopsis* chromosome arms which is not subject to the limitations of the RAPD method.

Results

Polymorphisms in amplified sequences revealed by restriction endonuclease digestion

As illustrated in Figure 1(a), the PCR-based strategy that we used to map *Arabidopsis* genes is based on a method used previously to detect RFLPs in cereals (Tragoonrung *et al.*, 1992; Weining and Langridge, 1991; Williams *et al.*, 1991). The method utilizes amplified DNA fragments that are digested with a restriction endonuclease to display an RFLP. To facilitate the development of a set of PCR markers that could be used to map a mutation to one of the 10 *Arabidopsis* chromosome arms, we first identified a set of *Arabidopsis* genes that had already been mapped and sequenced and then used these mapped sequences to design PCR primers.

As of 4 March, 1992, 26 of the 254 *Arabidopsis* genomic or cDNA sequences in GenBank had assigned map positions in AAtDB (an *Arabidopsis thaliana* database;

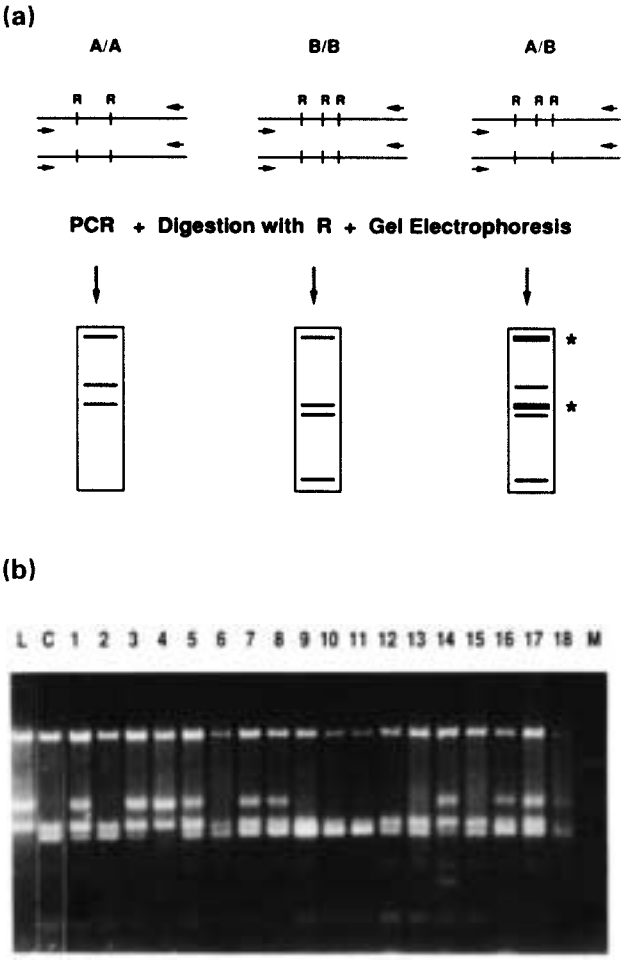

Figure 1. Generation and visualization of *Arabidopsis* ecotype-specific CAPS markers.

(a) Unique sequence primers are used to amplify a mapped DNA sequence from two different inbred ecotypes, A/A and B/B and from the heterozygote A/B. The amplified fragments from ecotypes A/A and B/B contain two and three recognition sites, respectively, for restriction endonuclease R. In the case of the heterozygote A/B, two different PCR products will be obtained, one which is cleaved three times by R and one which is cleaved twice. When fractionated by agarose or acrylamide gel electrophoresis, the PCR products digested by R from individual plants will give readily distinguishable patterns as seen in the lower portion of part (a). The asterisks indicate bands that will appear as doublets.

(b) Segregation of the *PG11* locus in F_3 progeny from a Col *rps2-101C* × Landsberg *erecta* cross. DNA from pooled F_3 plants (lanes 1–18) was amplified using *PG11* primers. PCR products were digested with *Bfa*I, electrophoresed on a 2.5% agarose gel, and stained with ethidium bromide. The two left-most lanes correspond to Ler (L) and Col (C) parents. DNA size markers (510, 390, 344, and 298 bp) are shown in lane M. In general, the doublet bands are readily apparent in the heterozygous plants (lanes 1, 3, 5, 7, 8, 14, 16–18) serving to verify the heterozygous restriction pattern. On the other hand, the observed stiochiometry of the two intermediate sized bands for some of the presumptive Col/Col homozygous plants (lanes 9,10,11) was unexpected. Although we have no simple explanation for this result, it is likely due to an artifact of *Bfa*I digestion since unexpected stiochiometries were not observed with other restriction endonuclease digests.

Cherry *et al.*, 1992). Among these 26, the following were chosen for further analysis because they were widely dispersed in the *Arabidopsis* genome (Table 1): *GAPB* (Shih *et al.*, 1991), *ADH* (Chang and Meyerowitz, 1986), *GPA1* (Ma *et al.*, 1990), *GAPC* (Shih *et al.*, 1991), *GAPA* (Shih *et al.*, 1991), *GL1* (Oppenheimer *et al.*, 1991), *BGL1* (Dong *et al.*, 1991), *GA1* (Sun *et al.*, 1992), *AG* (Yanofsky *et al.*, 1990), *DHS1* (Keith *et al.*, 1991), *ASA1* (Niyogi and Fink, 1992), *DFR* (Shirley *et al.*, 1992), and *LFY3* (Weigel *et al.*, 1992). Four additional sequences were obtained from other laboratories: *PVV4* (Van der Straeten *et al.*, 1992), *NCC1* (Crawford, personal communication), *m246* (Ma, personal communication), and *PG11* (Galant and Goodman, personal communication). A final DNA sequence, corresponding to RFLP marker *m429* (Chang *et al.*, 1988) was obtained by sequencing 316 bp of an end probe from a 110 kb YAC on chromosome II identified during a walk to the *DET2* gene (Nagpal and Chory, unpublished data; see Experimental procedures for details). These 18 mapped DNA sequences are distributed along all five *Arabidopsis* chromosomes: four on chromosome I, three on chromosome II, four on chromosome III, four on chromosome IV, and three on chromosome V (Table 1).

A pair of oligonucleotide primers (shown in Table 1) was synthesized for each of the DNA sequences listed in Table 1. When possible, primers were chosen such that the PCR products would include introns, to maximize the possibility of finding polymorphisms. When these primers were used to amplify DNA fragments from Columbia (Col) and Landsberg *erecta* (Ler) DNA, each pair of primers amplified a single major PCR product and the size of each PCR product was as predicted from the nucleotide sequence (Table 2), with the single exception of *GAPA*, in which case the Ler product was approximately 10 bp larger than the Col product.

To identify RFLPs between the amplified Ler and Col sequences, the primary PCR products were subjected to digestion with a panel of restriction endonucleases. Enzymes were chosen based on the DNA sequence of the amplified product which, in most cases, was known for at least one of the two ecotypes. Typically, each PCR product was digested with 25 restriction enzymes, which cut at least twice within the sequence. If none of the enzymes detected a polymorphism, a second series of digestions was performed with 25 different enzymes until a polymorphism was found or no more enzymes (out of a panel of 83) were available. The results are shown in Table 2. We named these DNA markers 'CAPS' for cleaved amplified polymorphic sequences.

Figure 1(b) shows the segregation of the CAPS marker *PG11* in the progeny from a Col X Ler cross. In the experiment shown in Figure 1(b), the template DNA for PCR analysis was isolated from pooled F_3 plants derived by selfing individual F_2 plants. However, we also found

Table 1. Map positions and sequences of PCR primers for *Arabidopsis* CAPS markers

Chr.	Gene, marker	Map position[a]	Forward primer[b]	Reverse primer[b]	Enzyme(s)
I	*PVV4*	25.2	GTT TGA AAG TGT AGA TGT AAC GAC	GGT TGT GTT TTG CTA GCA TC	*Bsa*AI
	NCC1	47.9	GTC CTA TCT CTA CGA TGT GGA TG	AAG TTA TAA GGC ATT AGA ATC ATA ATC	*Rsa*I
	GAPB	92.6	TCT GAT CAG TTG CAG CTA TG	GGC ACT ATG TTC AGT GCT G	*Bfa*I, *Dde*I
	ADH	13.7	GCG TGA CCA TCA AGA CTA AT	AAA AAT GGC AAC ACT TTG AC	*Xba*I, *Bfa*I
II	*m246*	14.1	TGA AGA GCT ATC CGA GAT GG	GCT TGA ACT CCT CCT CCT TC	*Mae*III
	GPA1	56.2	GGG ATT TGA TGA AGG AGA AC	ATT CCT TGG TCT CCA TCA TC	*Afl*III
	m429	69.6	TGG TAA CAT GTT GGC TCT ATA ATT G	GGC AGT TAT TAT GAA TGT CTG CAT G	*Scr*FI
III	*GAPC*	1.0	CTG TTA TCG TTA GGA TTC GG	ACG GAA AGA CAT TCC AGT C	*Eco*RV
	GAPA	30.9	CAC CGT GAT CTA AGG AGA GCA AG	TGT GCT CAA CCA AAC TTA GCC	*Dde*I
	GL1	44.5	ATA TTG AGT ACT GCC TTT AG	CCA TGA TCC GAA GAG ACT AT	*Taq*I
	BGL1	87.1	TCT TCT CGG TCT ATT CTT CG	TTA TCA CCA TAA CGT CTC CC	*Rsa*I, *Sau*3A, *Afl*III
IV	*GA1*	10.5	AAG CTT CGA ACT CAA GGT TC	CCG GAG AAT CGT ACG GTA C	*Bsa*BI
	AG	47.5	CAA CAG GTT TCT TCT TCT TCT C	CAA ACA CCA TTT AAT CTT GAC A	*Xba*I
	PG11	67.6	CGC AAC TAA CCA CAC ATT AC	AGT GAA ATT CAC CAG CAT G	*Bfa*I
	DHS1	88.7	CAA GTG ACC TGA AGA GTA TCG	AGA GAG AAT GAG AAA TGG AGG	*Dde*I, *Bsa*AI, *Mbo*III
V	*ASA1*	17.2	CTT ACT CCT GTT CTT GCT TAC	CCT CTA GCC TGA ATA ACA GAA C	*Bcl*I
	DFR	63.4	AGA TCC TGA GGT GAG TTT TTC	TGT TAC ATG GCT TCA TAC CA	*Bsa*AI
	LFY3	90.9	TAA CTT ATC GGG CTT CTG C	GAC GGC GTC TAG AAG ATT C	*Rsa*I

[a]The map positions given are those listed in version 1.3 of AAtDB, an *Arabidopsis* data base described in Cherry *et al.* (1992).
[b]Primers are shown 5′ to 3′.

Table 2. Restriction enzymes that detect polymorphisms in Landsberg *erecta* (Ler) and Columbia (Col) CAPS amplification products

Chr.	Gene, marker	Size of PCR product (bp)	Enzyme	Restriction fragments in base pairs[a]
I	*PVV4*	1064	*Bsa*AI	Ler: 706; 311; 47 Col: 753; 311
	NCC1	970[b]	*Rsa*I	Ler: 920[b]; 50[b] Col: 870[b]; 50[b]
	GAPB	1481	*Bfa*I	Ler: 850[b]; 360[b]; 212; 58 Col: 1211; 212; 58
			*Dde*I	Ler: 350[b]; 284; 255[b]; 225; 174 Col: 605; 284; 225; 174
	ADH	1291	*Xba*I	Ler: 1097; 262 Col: 1291
			*Bfa*I	Ler: 849[b]; 250[b]; 127; 65 Col: 1099; 127; 65
II	*m246*	1354	*Mae*III	Ler: 1122; 232 Col: 1354
	GPA1	1594	*Afl*III	Ler: 1385; 209 Col: 705[b]; 680; 209
	m429	316	*Scr*FI	Ler: 216[b]; 100[b] Col: 316
III	*GapC*	1148	*Eco*RV	Ler: 713; 390[b]; 340[b] Col: 735; 713
	GAPA	761/771[c]	*Dde*I	Ler: 240[b]; 190[b]; 178; 100; 33; 19; 10 Col: 420; 178; 100; 33; 19; 10
	GL1	519	*Taq*I	Ler: 372; 100; 47 Col: 298[b]; 100; 74[b]; 47
	BGL1	1269	*Rsa*I	Ler: 785; 485 Col: 785; 340[b]; 105[b]
			*Sau*3A	Ler: 1269 Col: 875; 395
			*Afl*III	Ler: 494; 434; 258; 84 Col: 494; 344; 258; 150; 84
IV	*GA1*	1196	*Bsa*BI	Ler: 1196 Col: 707; 527
	AG	1366	*Xba*I	Ler: 1073; 293 Col: 1366
	PG11	1293	*Bfa*I	Ler: 644; 353; 296 Col: 644; 296; 263[b]; 90[b]
	DHS1	1668	*Dde*I	Ler: 1620; 48 Col: 1491; 129; 48
			*Bsa*AI	Ler: 1668 Col: 1188; 480
			*Mbo*II	Ler: 520; 480; 173; 141; 84[b]; 81 Col: 1084; 173; 141; 81
V	*ASA1*	1728	*Bcl*I	Ler: 686; 553[b]; 489[b] Col: 1042; 686
	DFR	1143	*Bsa*AI	Ler: 609; 318; 216 Col: 609; 534
	LFY3	1330	*Rsa*I	Ler: 855; 236; 126; 78; 35 Col: 708; 236; 147; 126; 78; 35

[a]Only those fragments that are large enough to be readily visible are listed.
[b]Denotes approximate size of the fragment.
[c]The Ler product is approximately 10 bp larger than the Col product.

that the DNA isolation procedure described by Dellaporta *et al.* (1983) could be used to isolate DNA from a single *Arabidopsis* leaf that was of sufficient quality to perform the CAPS mapping procedure. We generally obtained 1–5 μg of DNA from a single leaf which is enough DNA for at least 50 PCR reactions. We have successfully amplified the desired PCR products using as little as 5 ng of DNA obtained from this mini-preparation procedure (Glazebrook and Ausubel, unpublished data).

We used the experimental protocol described in detail in the Experimental procedures in developing the CAPS mapping procedure. However, a variety of modifications can be made in the protocol. For example, we have successfully used 40 ng of each PCR primer instead of 200 ng (Glazebrook and Ausubel, unpublished data) and it is possible that even less primer could be used. In addition, instead of using 15 units of each of the restriction endonucleases to digest the PCR products, some of the endonucleases have been used successfully at lower concentrations (10 units of *Eco*RV, *Xba*I, *Rsa*I; 8 units of *Sau*3A; 4 units of *Dde*I; 3 units of *Taq*I; Glazebrook and Ausubel, unpublished data). On the other hand, we found that at least 15 units of *Bfa*I were required to digest the *PG11* PCR product to completion.

Frequency of CAPS polymorphisms

We detected 20 polymorphic changes in approximately 5227 nucleotides that comprised the recognition sites surveyed for polymorphisms in this study. This suggests that on average, Ler and Col genes differ once in every 261 (5227/20) bp. This is an oversimplification, however, because polymorphisms are not randomly distributed in the *Arabidopsis* genome. For example, the *ADH* locus contains 10 bp changes between Ler and Col concentrated in a 220 bp region, but only 13 total changes within a total of 2900 bp that have been sequenced in both ecotypes (Hanfstingl and Ausubel, unpublished data). Most of the polymorphisms (19 out of 20) that we detected were probably due to single base pair changes. One insertion/deletion of approximately 10 nucleotides was found in the *GAPA* gene.

Large insertion/deletion markers and microsatellite-based PCR markers

In addition to CAPS, we also identified PCR-based mapping markers that correspond to relatively large deletions and insertions such as the 5 kb internal region of the retrotransposon *Ta1-1* which is present in single copy on chromosome IV of Landsberg *erecta*, but which is absent from Columbia (Voytas *et al.*, 1990). PCR primers corresponding to *Ta1-1* (ACAGCATCCTGAAGACCTCG and TCGTTGATCGACTTAGTATC) amplified a 435 bp sequence from Ler DNA but failed to amplify any sequence from Col DNA (data not shown). We decided not to pursue this type of PCR marker further because it is not possible to distinguish plants that are homozygous from those that are heterozygous for the amplified sequence.

A third class of PCR markers that we identified was based on PCR amplification across tandem repeats of dinucleotide sequences called microsatellites (Hearne *et al.*, 1992). Microsatellites occur frequently and randomly in most eukaryotic DNAs and display polymorphisms due to variations in the number of repeat units. Screening the *Arabidopsis* sequences deposited in GenBank for different combinations of dinucleotide repeats revealed the presence of at least 10 repeats with a total length of 20 bp or more. We synthesized PCR primers that amplified 150–200 bp fragments that contained the following four microsatellite sequences: a (GA)10 repeat in *GBF3* (Schindler *et al.*, 1992); a $(AC)_8AG(AC)_4$ repeat in the non-coding region of a *myb*-homologous gene (Oppenheimer *et al.*, 1991); a $(AT)_{14}$ repeat in the intron of the gene encoding a basic chitinase (Samac *et al.*, 1990); and a (AT)3AA(AT)10 repeat in the *ATS1A* gene (Krebbers *et al.*, 1988). PCR products were resolved on 8% acrylamide gel. Among these four, only the *ATS1A* repeat was found to be polymorphic between Col and Ler (data not shown). Although we did not pursue the use of microsatellites further, we think that they may serve as highly informative markers in the future once more of them have been identified and mapped (Ecker, personal communication).

Discussion

The 18 sets of PCR-based CAPS primers described in this paper can be used to rapidly assign map positions to newly identified *Arabidopsis* genes. Given the limited number of CAPS markers currently available, however, subsequent analysis using traditional RFLP markers is needed to determine a map position accurate enough to initiate a chromosomal walk. There are several potential sources of new CAPS markers, which would allow the use of CAPS for the generation of high resolution maps. Sources of new CAPS markers include newly identified (and mapped) genes, DNA sequences around mapped RFLPs (where a polymorphic site is already known to exist), or DNA sequences at the ends of YAC clones (e.g. m429 marker on chromosome II). We anticipate that the results presented in this paper will lead others to generate and use additional CAPS. We also anticipate that 'second generation' CAPS markers may eventually supplant some of the original ones. For examples, the *PG11* CAPS marker requires the use of *Bfa*I, an expensive and not particularly robust restriction endonuclease. In the future, *PG11* could be replaced by another CAPS marker that maps nearby.

AAtDB (An *Arabidopsis thaliana* Database, Cherry, *et al.*, 1992) at the Department of Molecular Biology, Massachusetts General Hospital and the *Arabidopsis* Stock Centers at Ohio State University and at The University of Nottingham could serve as clearing houses for the distribution of CAPS-related information.

The maximum distance between any two of the 18 CAPS markers described in this paper is 46.2 cM (*DFR* and *ASA1*). Although it is not possible to determine with certainty the distance between a particular marker and the end of the chromosome on which it resides, the relatively high density of the *Arabidopsis* RFLP map makes it unlikely that any of the flanking CAPS markers on any of the five chromosomes are significantly more than 25.2 cM (in the case of *PVV4*) from an end. Therefore, the farthest away a gene of interest could be from one of the 18 CAPS markers is 25.2 cM.

To carry out the CAPS mapping procedure in the most efficient manner, it is helpful to calculate the minimum number of F_2 progeny from a Col × Ler cross that must be screened for segregation of the CAPS markers to establish linkage to one of them. Based on the binomial distribution, we calculate that a minimum of only 28 F_2 plants are required to establish linkage to one of the 18 CAPS markers at the 95% confidence limit. Because the CAPS markers are co-dominant, analysis of 28 F_2 plants is equivalent to analyzing 56 chromosomes.

To develop a CAPS marker from a known sequence, we routinely designed primers that amplified 1.5–2.5 kb sequences. In some cases, once a polymorphism was identified by testing a panel of restriction enzymes, the primers were redesigned (based on the sequence) to better visualize the polymorphism. For example, the original set of *GAPA* primers preferentially amplified DNA sequences from the Ler ecotype, making it difficult to unambiguously identify heterozygous plants. An additional problem with the original *GAPA* primers was that the polymorphic bands were too close in size to distinguish Ler and Col amplified products with a high degree of confidence. The *GAPA* primers listed in Table 1 eliminated both of these problems.

A major advantage of CAPS is that they are co-dominant genetic markers; that is, different digestion patterns are obtained for plants that are homozygous or heterozygous for the parental alleles. However, as illustrated with the original *GAPA* primers that we designed (see above), a particular set of primers may preferentially amplify only one of the two parental sequences when both are present in a heterozygous plant. Therefore, it is critical, to use an equal mixture of the two homozygous parental DNAs or DNA isolated from a known heterozygous plant as a control.

The CAPS primers described in this paper were specifically designed for the *Arabidopsis* Columbia and Landsberg *erecta* ecotypes. However, six of the PCR markers reported here were tested using the Landsberg *erecta* and Niederzenz ecotypes and three of them detected polymorphisms (Chory, personal communication).

To date, the CAPS method has been used to map several genes including *RPS2* (resistance to *Pseudomonas syringae*; Yu *et al.* 1993), *pad*1, *pad*2 and *pad*3 (phytoalexin deficient mutants; Glazebrook and Ausubel, unpublished data) and *trp4* (a tryptophan biosynthetic pathway mutant; Keith, personal communication).

Experimental procedures

Plant material

Arabidopsis plants used in this study were the Columbia and Landsberg *erecta* ecotypes. For the cross described in Figure 1(b), Ler was used as the female parent and the Columbia *rps2-101C* mutant (Yu *et al.*, 1993) was used as the pollen donor.

DNA isolation

For large scale PCR reactions, DNA was isolated from whole plants using the procedure described in Ausubel *et al.* (1993). A slightly modified version of the method described by Dellaporta *et al.* (1983) was used to isolate DNA from one to three *Arabidopsis* leaves for CAPS mapping. Briefly, leaves were frozen in liquid nitrogen, ground to a powder in a mortar and pestle, and transferred to 1.5 ml microcentrifuge tubes. Extraction buffer (500 μl) containing 100 mM Tris pH 8, 50 mM EDTA pH 8, 500 mM NaCl and 10 mM mercaptoethanol was added followed by 35 μl of 20% SDS. The samples were incubated at 65°C for 10 min and 130 μl of 5 M potassium acetate was added. After 5 min incubation at 0°C, the precipitate was pelleted for 10 min at 15 000 ***g***, the supernatant was transferred to a 2 ml tube containing 640 μl isopropyl alcohol, and 60 μl 3 M sodium acetate, mixed, and incubated at −20°C for 10 min. The precipitated DNA was centrifuged at 15 000 ***g*** for 15 min and redissolved in 200 μl of 50 mM Tris pH 8.0, 10 mM EDTA. This solution was centrifuged at 15 000 ***g*** or 5 min to remove insoluble material, and the supernatant was transferred to a tube containing 20 μl 3 M sodium acetate and 440 μl ethanol. After incubation at −20°C for 10 min, DNA was pelleted by centrifugation at 15 000 ***g*** for 5 min, and washed with 70% ethanol. Pellets were dried and dissolved in 50 μl water.

PCR amplification of Arabidopsis *genomic DNA*

PCR primers were designed using the computer program PRIMER, (Version 0.5, May 1991, by Stephen E. Lincoln, Mark J. Daly and Eric S. Lander, Whitehead Institute for Biomedical Research). Large scale PCR reactions were performed using reagents from Boehringer (10× buffer), Pharmacia (deoxynucleotides), and Promega (Taq polymerase). The reactions were carried out in 100 μl that contained 0.125 mM of each deoxynucleotide, 0.5 μg of each primer, 2.5 U of Taq polymerase in 2.5 μl, and 50–100 ng of *Arabidopsis* genomic DNA. Conditions for the amplification were as follows: 30 sec at 95°C; annealing for 30 sec at 56°C; polymerization for 3 min at 72°C. The cycle was repeated 50 times. For mapping purposes, the reaction was scaled down to 10 μl final volume, which contained 0.125 mM each of four deoxynucleotides, 0.2 μg of both primers, 0.25 U of Taq polymerase and approximately 50 ng of *Arabidopsis* DNA (1 μl of the 50 μl obtained from a miniprep). Usually a premix containing deoxy-

nucleotides, Taq polymerase buffer, primers and Taq polymerase was prepared for 30 reactions and transferred to 0.5 ml microcentrifuge tubes to which various DNAs were added. Other PCR conditions were the same as for the large scale reaction.

Restriction enzyme digestions and analysis of PCR products

To identify restriction endonucleases that would generate a polymorphism, reactions were carried out in 10 μl final volume containing 3.5 μl of PCR product, 1 μl of 10× concentrated restriction enzyme buffer, and 1–3 units of the appropriate restriction endonuclease. The digestion products were analyzed on 1.5 % agarose gels for all markers except *m429*, *PG11* and *LFY3*, which were run on 2.5% gels. Ler and Col products were run in adjacent lanes to visualize polymorphisms. For mapping, restriction endonuclease digestion was carried out in the same tube as the PCR reaction. Restriction enzyme mix (10 μl) containing 2 μl of restriction enzyme buffer and 15 U of enzyme was added to each tube and incubated for 2 h at the temperature optimal for the particular enzyme activity. Four microliters of 6× loading dye (Ausubel *et al.*, 1993) was added to each tube prior to electrophoresis.

DNA sequencing

Partial DNA sequence of an end probe from a 110 kb YAC clone (mapped to chromosome II) was obtained by sequencing of an inverse PCR product provided by J. Chory and P. Nagpal (Salk Institute, La Jolla) using the *fmol* DNA Sequencing system (Promega) according to the manufacturer's instructions.

Acknowledgments

We thank Guo-Liang Yu for DNA samples, Richard Gardner for help in developing the microsatellite markers, Fumiaki Katagiri for helpful discussions, and Hong Ma, Nigel Crawford, Paul Gallant and Howard Goodman for providing DNA sequences prior to publication. We also thank Jane Glazebrook for help in developing the *GL1* marker, for help in refining the mapping procedure, and for help in preparation of the manuscript. This work was funded by USDA-CRGO grant No. 91-37300-6384 and by a grant from Hoechst AG to Massachusetts General Hospital.

References

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1993) *Current Protocols in Molecular Biology*. New York: Greene Publishing Associates/ Wiley Interscience.

Bernatzky, R. and Tanksley, S.D. (1986) Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. *Genetics*, **112**, 887–898.

Burr, B. and Burr, F.A. (1991) Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. *Trends Genet.* **7**, 55–60.

Chang, C. and Meyerowitz, E.M. (1986) Molecular cloning and DNA sequence of *Arabidopsis thaliana* alcohol dehydrogenase gene. *Proc. Natl Acad. Sci. USA*, **83**, 1408–1412.

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

Cherry, J.M., Cartinhour, S.W. and Goodman, H.M. (1992) AAtDB, An *Arabidopsis thaliana* Database. *Plant Mol. Biol. Rep.* **10**, 308.

Cox, R.L. and Lehrach, H. (1991) Genome mapping: PCR based meiotic and somatic cell hybrid analysis. *BioEssays*, **13**, 193–198.

Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.* **1**, 19–21.

Devos, K.M. and Gale, M.D. (1992) The use of random amplified polymorphic DNA markers in wheat. *Theor. Appl. Genet.* **84**, 567-572.

Dong, X., Mindrinos, M., Davis, K.R. and Ausubel, F.M. (1991) Induction of *Arabidopsis* defense genes by virulent and avirulent *Pseudomonas syringae* strains and by a cloned avirulence gene. *Plant Cell*, **3**, 61–72.

Hearne, C.M. Ghosh, S. and Todd, J.A. (1992) Microsatellites for linkage analysis of genetic traits. *Trends Genet.* **8**, 288–294.

Helentjaris, H. (1987) A genetic linkage map for maize based on RFLPs. *Trends Genet.* **3**, 217–221.

Kam-Morgan, L.M.W. and Gill, B.S. (1989) DNA restriction fragment length polymorphism: a strategy for genetic mapping of D genome of wheat. *Genome*, **32**, 724–732.

Keith, B., Dong, X., Ausubel, F.M. and Fink, G.R. (1991) Differential induction of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase genes in *Arabidopsis thaliana* by wounding and pathogenic attack. *Proc. Natl Acad. Sci. USA*, **88**, 8821–8825.

Krebbers, E., Seurinck, J., Herdies, L., Cashmore, A.R. and Timko, M.P. (1988) Four genes in two diverged subfamilies encode the ribulose-1,5-biphosphate carboxylase small subunit polypeptides of *Arabidopsis thaliana*. *Plant Mol. Biol.* **11**, 745–759.

Landry, B.S., Kessel, R., Farrara, B. and Michelmore, R.W. (1987) A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. *Genetics*, **116**, 331–337.

Ma, H., Yanofsky, M.F. and Meyerowitz, E.M. (1990) Molecular cloning and characterization of *GPA1*, a G protein a subunit gene from *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA*, **87**, 3821–3825.

McCouch, S.R., Krochert, G., Yu, Z.H., Wang, Z.Y., Kush, G.S., Coffman, W.R. and Tanksley, S.D. (1988) Molecular mapping of rice chromosomes. *Theor. Appl. Genet.* **76**, 815–829.

Mullis, K.B. and Faloona, F. (1987) Specific synthesis of DNA in vitro via a polymerase chain reaction. *Methods Enzymol.* **155**, 355–350.

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

Niyogi, K.K. and Fink, G.R. (1992) Two anthranilate synthase genes in *Arabidopsis*: defense related regulation of the tryptophan pathway. *Plant Cell*, **4**, 721–733.

Oppenheimer, D. G., Herman, P.C., Sivakumaran, S., Esch, J. and Marks, M.D. (1991) A *myb* gene required for leaf trichome differentiation in *Arabidopsis* is expressed in stipules. *Cell*, **67**, 483–493.

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

Riedy, M.F., Hamilton III, W.J. and Aquadro, C.F. (1992) Excess of non-parental bands in offspring from known primate pedigrees assayed using RAPD PCR. *Nucl. Acids Res.* **20**, 918.

Samac, D.A., Hironaka, C.M., Yallaly, P.E. and Shah, D.M. (1990) Isolation and characterization of the genes encoding basic and acidic chitinase in *Arabidopsis thaliana*. *Plant Physiol.* **93**, 907–914.

Schindler, U.M., Menkens, A.E., Beckman, H., Ecker, J.R. and Cashmore, A.R. (1992) Heterodimerization between light regulated and ubiquitously expressed *Arabidopsis* GBF bZIP protein. *EMBO J.* **11**, 1261–1273.

Shih, M.-C., Heinrich, P.C. and Goodman, H.M. (1991) Cloning and chromosomal mapping of nuclear genes encoding chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenase. *Gene*, **104**, 133–138.

Shirley, B.W., Hanley, S. and Goodman, H.M. (1992) Effects of ionizing radiation on a plant genome: analysis of two *Arabidopsis transparent testa* mutations. *Plant Cell*, **4**, 333–347.

Slocum, M.K., Figdore, S.S., Kennard, W.C., Suzuki, Y.K. and Osborn, T.C. (1990) Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleraceae*. *Theor. Appl. Genet.* **80**, 57–64.

Sun, T.-P., Goodman, H.M. and Ausubel, F.M. (1992) Cloning the *Arabidopsis GA1* locus by genomic subtraction. *Plant Cell*, **4**, 119–128.

Tragoonrung, S., Kanizin, V., Hayes, P.M. and Blake, T.K. (1992) Sequence-tagged-site-facilitated PCR for barley genome mapping. *Theor. Appl. Genet.* **84**, 1002–1008.

Van Der Straeten, D., Rodrigues-Pousada, R.A., Villarroel, R., Hanley, S., Goodman, H.M. and Van Montagu, M. (1992) Cloning, genetic mapping, and expression analysis of an *Arabidopsis thaliana* gene that encodes 1-aminocyclopropane-1-carboxylate synthase. *Proc. Natl Acad. Sci. USA.* **89**, 9969–9973.

Voytas, D.F., Konieczny, A., Cummings, M.P. and Ausubel, F.M. (1990) The structure, distribution and evolution of the Ta*1* retrotransposable element family of *Arabidopsis thaliana*. *Genetics*, **126**, 713–721

Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F. and Meyerowitz, E.M. (1992) *LEAFY* controls floral meristem identity in *Arabidopsis*. *Cell*, **69**, 843–859.

Weining, S. and Langridge, P. (1991) Identification and mapping of polymorphisms in cereal based on the polymerase chain reaction. *Theor. Appl. Genet.* **82**, 209–216.

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

Williams, M.N.V., Pande, N., Nair, M., Mohan, M and Bennet, J. (1991) Restriction fragment length polymorphism analysis of polymerase chain reaction products amplified from mapped loci of rice (*Oryza sativa* L.) genomic DNA. *Theoret. Appl. Genet.* **82**, 489–498.

Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldman, K.A. and Meyerowitz, E.M. (1990) The protein encoded by the *Arabidopsis* homeotic gene *AGAMOUS* resembles transcriptional factors. Nature, **346**, 35–39.

Yu, G.L., Katagiri, F. and Ausubel, F.M. (1993) *Arabidopsis* mutations at the *RPS2* locus result in loss of resistance to *Pseudomonas syringae* strains expressing the avirulence gene *avrRpt2*. *Mol. Plant–Microbe Interact.*, in press.