RESEARCH ARTICLE

Cloning the Arabidopsis GA1 Locus by Genomic Subtraction

Tai-ping Sun,1 Howard M. Goodman, and Frederick M. Ausubel2

Department of Genetics, Harvard Medical School, and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114

Arabidopsis thaliana ga1 mutants are gibberellin-responsive dwarfs. We used the genomic subtraction technique to clone DNA sequences that are present in wild-type Arabidopsis (ecotype Landsberg erecta, Ler) but are missing in a presumptive ga1 deletion mutant, ga1-3. The cloned sequences correspond to a 5.0-kb deletion in the ga1-3 genome. Three lines of evidence indicated that the 5.0-kb deletion in the ga1-3 mutant is located at the GA1 locus. First, restriction fragment length polymorphism mapping showed that DNA sequences within the 5.0-kb deletion map to the GA1 locus. Second, cosmid clones that contain wild-type DNA inserts spanning the deletion in ga1-3 complemented the dwarf phenotype when integrated into the ga1-3 genome by Agrobacterium tumefaciens-mediated transformation. Third, we identified molecular lesions in four additional ga1 alleles within the 5.0-kb region deleted in mutant ga1-3. One of these lesions is a large insertion or inversion located within the most distal intron encoded by the GA1 locus. The three other lesions are all single base changes located within the two most distal exons. RNA gel blot analysis indicated that the GA1 locus encodes a 2.8-kb mRNA. We calculated a recombination rate of $10^{-5}$ cM per nucleotide for the GA1 region of the Arabidopsis genome.

INTRODUCTION

Cloning a plant gene that is defined solely by a mutant phenotype can be an arduous task. Chromosome walking, which is used to isolate a gene that is linked to a restriction fragment length polymorphism (RFLP) marker, is currently laborious and can be impeded by unclonable sequences or stretches of repetitive DNA. Gene tagging systems that use engineered heterologous transposable elements to generate random mutations over an entire genome are being developed but are not yet available for most plant species (Jones et al., 1989; Martin et al., 1989; Masson and Fedoroff, 1989). In the case of Arabidopsis, Agrobacterium tumefaciens T-DNA has been used successfully to tag a variety of genes (Feldman et al., 1989; Feldman, 1991). However, individual transformed plants have to be generated for each T-DNA insertion. In the case of Arabidopsis with a genome size of $10^8$ bp (B. Hauge and H.M. Goodman, unpublished data), over 100,000 independent transformants (Clarke and Carbon, 1979) are required to achieve 95% probability of having an insert every 2 kb in the genome, assuming an average of 1.4 random inserts per plant (Feldman, 1991).

Recently, our laboratory developed a technique called "genomic subtraction" that provides a useful addition to the chromosome walking and gene tagging approaches for cloning genes corresponding to a mutant phenotype (Straus and Ausubel, 1990). The procedure identifies DNA sequences that are present in a wild-type organism (target sequences) but are missing in a homozygous deletion mutant. Genomic subtraction involves multiple cycles of hybridization between wild-type DNA and an excess of biotinylated deletion mutant DNA. In each cycle, avidin-coated beads are added to remove common sequences present in both wild-type and mutant DNA samples. The remaining DNA after several cycles of subtraction should be enriched for sequences that are absent in the deletion mutant. Following several cycles of subtraction, adapters are ligated to the ends of the remaining DNA fragments that are not bound to avidin beads and the DNA fragments are amplified by the polymerase chain reaction (PCR) for further analyses.

Genomic subtraction was developed using a defined 5.0-kb deletion of Saccharomyces cerevisiae, which has a genome size of $\sim 2 \times 10^7$ bp. When contemplating the extension of genomic subtraction to clone plant genes, it was necessary to take into account the fact that genomic subtraction requires self-annealing of the target DNA sequences. Because the concentration of target DNA sequences decreases as the genome size increases, we initially sought to determine whether the technique could be successfully used with Arabidopsis, the plant with the smallest known genome (Meyerowitz and Pruitt, 1985). When we initiated this project, there were no known deletion mutants of Arabidopsis. However, one presumptive deletion had been described at the Arabidopsis GA1 locus (Koornneef et al., 1983a).
Figure 1. Genetic Map of the Arabidopsis GA1 locus in cM × 10⁻² (Koornneef et al., 1983a). The alleles have been renamed as described in Methods. The extent of the presumptive deletion in gai-3 is indicated by the heavy horizontal line.

Arabidopsis gai mutants are nongerminating, gibberellin (GA)-responsive, male-sterile dwarfs whose phenotype can be converted to wild type by repeated application of GA (Koornneef and van der Veen, 1980). GAs, especially GA1 and GA4, are diterpenoid plant growth hormones required for seed germination, leaf expansion, stem elongation, flowering, and fruit set in Arabidopsis (Graebe, 1987). As diagrammed in Figure 1, Koornneef et al. (1983a) constructed a fine-structure genetic map of the Arabidopsis GA1 locus by using nine independently isolated gai alleles. Three of these gai alleles (ga1-2, ga1-3, ga1-4) were generated by fast neutron bombardment and six (ga1-1, ga1-6, ga1-7, ga1-8, ga1-9, ga1-10) were generated by ethyl methane sulfonate mutagenesis. One of the fast-neutron-generated mutants, ga1-3, failed to recombine with the six alleles indicated in Figure 1 and was classified as an intragenic deletion (Koornneef et al., 1983a).

Based on the genetic analysis at the GA1 locus, we used genomic subtraction to clone the Arabidopsis GA1 locus using sheared biotinylated DNA from mutant gai-3 and Sau3A-digested nonbiotinylated DNA from the wild-type parent, ecotype Ler. To determine the efficiency of the genomic subtraction technique for genomes the size of Arabidopsis, we first carried out a reconstruction experiment in which biotinylated Arabidopsis DNA was used to subtract the same DNA preparation from which an approximately single-copy level of Ad-2 DNA had been added. After four cycles of subtraction, the starting materials and the end products of each cycle were ligated with Sau3A adaptors, amplified by PCR, and analyzed on a 1.5% agarose gel. Figure 2 shows that Ad-2-specific bands could be detected in PCR-amplified DNA starting at the end of the third cycle of subtraction.

The amplified unbound DNA from each cycle of subtraction was ligated into the Smal site of pUC13. Figure 3 shows the degree of enrichment of Ad-2 DNA after each of three cycles of subtraction. Approximately 4000 colonies containing cloned Sau3A fragments from each cycle were plated onto nitrocellulose filters and hybridized with 32P-labeled Ad-2 DNA. In a separate experiment, 200 to 400 colonies containing cloned Sau3A fragments from each cycle were plated per filter to facilitate counting the number of clones that contained Ad-2 sequences (data not shown). Before subtraction, 0.1% of the cloned Sau3A fragments hybridized to the Ad-2 DNA probe. On average, there was a 10-fold enrichment of the Ad-2 DNA per cycle in each of the first three cycles of subtraction, and, after four cycles of subtraction, more than 95% of the colonies contained Ad-2 sequences.

Figure 2. PCR-Amplified Unbound DNA from the Genomic Subtraction Reconstruction Experiment. PCR-amplified products were size fractionated on a 1.5% agarose gel containing 0.5 μg/mL ethidium bromide. Lane 1, starting material before subtraction; lanes 2 to 5, unbound DNA after first to fourth cycles of subtraction; lane 6, Sau3A-digested Ad-2 DNA.
Cloning the Arabidopsis \textit{GA1} Locus

**BEFORE SUBTRACTION**

**FIRST CYCLE**

**SECOND CYCLE**

**THIRD CYCLE**

**Figure 3.** Enrichment of Ad-2 DNA in the Reconstruction Experiment.

Autoradiogram of colony hybridization experiment is shown for filters containing \(~\text{4000}}\) recombinant colonies after each cycle of subtraction probed with \(^{32}\text{P}\)-labeled Ad-2 DNA.

To determine the extent of the deletion in \textit{ga1-3} DNA identified by pGA1-1, pGA1-1 DNA was used as a hybridization probe to identify larger genomic fragments of wild-type Arabidopsis DNA. Two such clones, \( \lambda \text{GA1-3} \), identified in a \( \lambda \text{FIX} \) library of Ler genomic DNA, and pGA1-4, identified in a \( p\text{OCA18} \) cosm id library of Columbia genomic DNA, are shown in Figure 5. A subclone of \( \lambda \text{GA1-3} \), pGA1-2 (Figure 5), was used to probe a DNA gel blot containing HindIII-digested DNA from wild-type Ler and from several \textit{ga7} mutants. As shown in Figure 4B and diagrammed in Figure 5, pGA1-2 hybridized to four HindIII fragments (1.0 kb, 1.2 kb, 1.4 kb, and 5.6 kb) in wild-type DNA that were absent in \textit{ga7} DNA. The deletion mutation results in an extra HindIII fragment (4.2 kb) in \textit{ga7} DNA. These results and additional DNA gel blot analyses (data not shown) showed that the deletion in \textit{ga7} is 5 kb, corresponding to 0.005\% of the Arabidopsis genome.

**RFLP Mapping and Complementation Analysis**

Several lines of evidence indicated that the 5.0-kb deletion in mutant \textit{ga7} identified by genomic subtraction corresponds to the \textit{GA1} locus. First, RFLP mapping analysis carried out by the procedure detailed in Nam et al. (1989) using \( \lambda \text{GA1-3} \) (Figure 5) as a hybridization probe showed that \( \lambda \text{GA1-3} \) maps

![Figure 4](image_url)

**Figure 4.** Detection of Deletions and Insertions (or Inversion) in \textit{ga7} and \textit{ga1} DNA, Respectively.

(A) Autoradiogram of DNA gel blot probed with the 250-bp Sau3A fragment from pGA1-1. The blot contains 1 \( \mu \text{g} \) HindIII-digested DNA isolated from Landsberg erecta, lane 1, and from three \textit{ga7} mutants, \textit{ga7-3} (1.3 kb), and \textit{ga7-2}, lane 4.

(B) The same blot shown in (A) probed with the 6-kb fragment from pGA1-2 that covers the entire deleted region in \textit{ga7-3} (Figure 5). The arrows indicate altered HindIII fragments in \textit{ga7-3} (4.2 kb) and \textit{ga1} (1.3 and 3.3 kb).

**Cloning the Arabidopsis \textit{GA1} Locus by Genomic Subtraction**

Following the successful completion of the Ad-2 reconstruction experiment, we used genomic subtraction, as described in Methods, to isolate the Arabidopsis \textit{GA1} locus using biotinylated \textit{ga7} DNA and Sau3A-digested DNA from the wild-type parent, ecotype Ler. A control experiment was carried out in parallel in which biotinylated \textit{ga1} DNA was used to subtract Sau3A-digested Ler DNA that contained an approximately single-copy level (0.5 \( \mu \text{g} \)) of Ad-2 DNA. In accordance with the previous reconstruction experiment, Ad-2–specific bands could be detected in the PCR-amplified subtraction products after the third cycle of subtraction (data not shown). In the actual subtraction experiment, DNA fragments remaining after five cycles of subtractive hybridization were amplified by PCR and cloned in \( p\text{UC13} \). Six individual clones were radiolabeled and used as hybridization probes for DNA gel blot analyses. As shown in Figure 4A, one of these six clones, pGA1-1, contained a 250-bp Sau3A fragment that hybridized to a 1.4-kb HindIII fragment in DNA isolated from wild-type Ler DNA and from the \textit{ga1} and \textit{ga1} mutants but did not hybridize to \textit{ga7} DNA.
Figure 5. Physical Map of the Arabidopsis GA1 Locus.

The heavy horizontal line is a HindIII restriction map. The location of a 5-kb deletion in ga1-3 is indicated by the hatched box. The horizontal lines above the restriction map indicate the extent of the sequences contained in the λ clone λGA1-3, the plasmid pGA1-2, and the cosmid clone pGA1-4. The diagram below the horizontal lines depicts the location of introns (lines) and exons (open boxes) of the GA1 gene and the locations of the insertion or inversion mutation in ga1-2 and the point mutations in ga1-6, ga1-7, and ga1-8 alleles. The ga1 locus of ga1-2 was interrupted by unknown DNA between nucleotides 159 and 161 (with a nucleotide C missing) in the intron. The single-nucleotide changes in ga1-6, ga1-8, and ga1-7 are at nucleotides 200 and 205 of the penultimate exon and at nucleotide 72 of the last exon, respectively.

to the telomere proximal region at the top of chromosome 4 (data not shown), consistent with the location to which the GA1 locus had been mapped previously by Koornneef et al. (1983b).

Second, the cosmid clone pGA1-4 (Figure 5), which contains a 20-kb insert of wild-type (ecotype Columbia) DNA spanning the deletion in ga1-3, complemented the dwarf phenotype of the ga1-3 mutant. pOCA18, the parental vector of pGA1-4, is a cosmid cloning vector that contains the right and left borders and the "overdrive" sequences of Agrobacterium T-DNA required for efficient transfer of cloned DNA into plant genomes. It also carries a chimeric neomycin phosphotransferase gene that confers kanamycin resistance in transgenic plants (Olszewski et al., 1988). Agrobacterium strain LBA4404 containing pGA1-4 was used to infect root explants of mutant ga1-3, and kanamycin-resistant (Km<sup>r</sup>) transgenic plants were selected as described (Valvekens et al., 1988). One hundred and thirty Km<sup>r</sup> plants (T1 generation) were regenerated that set seeds in the absence of exogenous GA. Fifty to 300 seeds from each of four different T1 plants showed 100% linkage of the GA1<sup>+</sup> and Km<sup>r</sup> phenotypes, which segregated approximately 3:1 in relation to the GA1<sup>−</sup> kanamycin-sensitive phenotype (T2 generation). As shown in Figure 6, all T2 generation GA1<sup>+</sup>/Km<sup>r</sup> transgenic plants tested did not require exogenous GA for normal growth. Germination, stem elongation, and seed set were the same in the transgenic plant as in the wild-type plant without exogenous GA treatment.

As shown in lanes 4 and 5 of Figure 7A, DNA gel blot analysis, using the 6-kb fragment from pGA1-2 as a probe, indicated that both the endogenous mutant ga1-3 locus (4.2-kb HindIII fragment) and the transgene wild-type (Columbia) GA1 locus (5.0-, 1.4-, 1.2-, and 1.0-kb HindIII fragments) are present in two independent T3 transgenic lines. Further DNA gel blot analysis, shown in lanes 3 and 4 of Figure 7B, using pOCA18, which contains the T-DNA border sequences as a probe, showed that only two border fragments were present in the genomes of both transgenic plants. These results indicated that a copy of the wild-type GA1 transgene was integrated at a single locus in the genomes of both transgenic plants.

Identification of Four Other ga1 Alleles within the Region Deleted in ga1-3

To obtain unequivocal evidence that the 5.0-kb deletion in ga1-3 corresponds to the GA1 locus, we showed that four additional

Figure 6. Phenotypes of 6-Week-Old Arabidopsis Landsberg erecta Plants.

From left to right, a ga1-3 mutant, a transgenic ga1-3 plant containing the 20-kb insert from pGA1-4 (Figure 5), and a wild-type Landsberg erecta plant. Seeds of transgenic ga1-3 and wild-type plants were germinated on agarose plates containing 1 x Murashige and Skoog (1962) salts and 2% sucrose (MS plates) with or without kanamycin. Seeds of the mutant ga1-3 were soaked in 100 μM GA<sub>3</sub> for 4 days before being germinated on MS plates. Seven-day-old seedlings were transferred to soil. To show the dwarf phenotype, no additional GA<sub>3</sub> was given to the mutant ga1-3 after germination.
of PCR products from ga1-2 DNA (data not shown) indicated that the ga1 gene of ga1-2 mutant is interrupted by at least 3.4-kb DNA of unidentified origin within the last intron defined by the partial cDNA clone (Figure 5). The mutation could be due to either an insertion or an inversion event. RFLP mapping of the unknown “insertion” DNA would be needed to distinguish between these two possibilities. DNA gel blot analyses, using pGA1-2 (Figure 4B) and pGA1-4 (data not shown) as probes, showed that there are no detectable deletions or insertions in ga1-4 DNA.

The ethyl methane sulfonate-induced ga1-6, ga1-7, and ga1-8 alleles are located at or near the ga1-2 allele on the genetic map (Figure 1). Direct sequencing of PCR products amplified from ga1-6, ga1-7, and ga1-8 mutant DNA templates revealed single nucleotide changes in the 1.2-kb HindIII fragment in each of the three mutants (Figure 5). Mutant ga1-7, which defines one side of the genetic map, contained a single nucleotide change in the most distal exon of the GA1 gene. Mutants ga1-6 and ga1-8 contain single base changes in the penultimate exon. The single nucleotide changes in mutants ga1-6, ga1-7, and ga1-8 result in missense mutations, which is consistent with their leaky phenotypes (Koornneef et al., 1983a). It is unlikely that the base changes observed in mutants ga1-6, ga1-7, and ga1-8 are PCR artifacts or are due to the highly polymorphic nature of the GA1 locus because the 1.2-kb HindIII fragments amplified and sequenced from ga1-1 and ga1-9 both had the wild-type sequence. Moreover, the PCR products were sequenced directly and the sequence analysis was carried out twice, using the products of two independent amplifications for each allele examined.

The GA1 Locus Encodes a 2.8-kb mRNA

Figure 8 shows that the GA1 locus encodes a 2.8-kb mRNA detected by RNA gel blot analysis using the 0.9-kb partial GA1 cDNA as a hybridization probe. As expected, this RNA could not be detected in the deletion mutant ga1-3. The amount of RNA in each lane is the same as shown in lanes 1, 2, and 4 using the CAB gene as a probe. The decreased hybridization of the CAB probe in lane 3 is due to low expression of the CAB gene in the siliques. The GA1 locus was expressed at an extremely low level as reflected by the length of the exposure time needed for the blots: more than 3 weeks for GA1 mRNA versus 2 min for the CAB mRNA. The low level of expression of the GA1 locus was also reflected in the fact that the frequency of GA1 clones in the siliqua cDNA library was .5 x 10\(^{-5}\). The level of expression of the GA1 gene in immature siliques (lane 3) seemed to be slightly higher than that in whole plants (lanes 1 and 2). This is consistent with the observation that the maximum ent-kaurene synthesizing activity was found in immature siliques (Zeevaart, 1986). However, a marker gene with constitutive expression in different tissues needs to be used for quantitative analysis in future studies.

\[ \text{ga1 alleles contain alterations from the wild-type sequence within the region deleted in ga1-3 in the order predicted by the genetic map. To aid in this analysis, a partial GA1 cDNA clone (0.9 kb), containing poly(A) and corresponding to the 1.2-kb HindIII fragment (Figure 5), was isolated from a cDNA library constructed from RNA isolated from siliques (seed pods) of Arabidopsis ecotype Columbia (J. Giraudat and H.M. Goodman, unpublished data). Four exons and three introns in the 1.2-kb HindIII fragment in the GA1 locus (Figure 5) were deduced by comparison of the cDNA and genomic DNA sequences (data not shown). The identification of this cDNA clone showed that the 1.2-kb HindIII fragment is located at the 3' end of the GA1 gene and suggested that the mutations in the ga1 alleles 2, 3, 6, 7, and 8 should also be located at the 3' end of the GA1 gene (Figures 1 and 5).}

In addition to ga1-3, two other mutants, ga1-2 and ga1-4, were generated by fast neutron mutagenesis (Koornneef et al., 1983a). As shown in Figure 4B, the 1.2-kb HindIII fragment in ga1-2 DNA was replaced by two new fragments of 1.3 and 3.3 kb without alteration of the adjacent 1.4- and 5.6-kb fragments. Further DNA gel blot analysis and DNA sequencing
One of the seven clones contained a highly repetitive sequence, three appeared to be single-copy Arabidopsis sequences that hybridized to both Ler and gat-3 DNA, and three contained inserts that only hybridized to the preparation of Ler DNA used in the subtraction experiment but did not hybridize to other preparations of Ler DNA.

DISCUSSION

Genomic Subtraction for Isolating Genes from Arabidopsis

In cloning the GA1 locus by genomic subtraction, only one in 20 cloned Sau3A fragments following five cycles of subtraction corresponded to the 5-kb deletion in the gat-3 mutant. This level of enrichment was insufficient to use the amplified product directly as a hybridization probe for screening an Arabidopsis genomic library. Instead, individual clones were analyzed by DNA gel blot hybridization. Because no further enrichment was achieved after the fifth cycle, it is possible that the contaminating Arabidopsis fragments had sequence-specific features that prevented them from being subtracted, despite the fact that they hybridized to Arabidopsis DNA on a DNA gel blot at normal stringency. The contaminating sequences that only hybridized to the particular preparation of Ler DNA used in the subtraction experiment most likely originated from contaminating microorganisms because the Arabidopsis plants used in these experiments were not grown under sterile conditions. To eliminate foreign DNA, it is clearly desirable to isolate wild-type DNA from plants grown under sterile conditions and to use disposable tubes and pipettes for DNA purification.

In carrying out genomic subtraction, we found that it was convenient to carry out a parallel reconstruction experiment using biotinylated deletion mutant DNA to subtract Sau3A-digested wild-type DNA containing a small amount of Sau3A-digested adenovirus DNA (as described in Results). Specific adenovirus DNA fragments can be directly visualized on an ethidium bromide-stained gel in the amplified DNA products after the third cycle of subtraction. This offers a quick indication of the efficiency of the subtraction.

Genomic subtraction is not labor intensive, and the results reported here indicate that genomic subtraction could be routinely used to clone other nonessential Arabidopsis genes if a method were available for generating deletions at reasonable frequency. In addition to the fast-neutron-induced deletion in mutant gat-3 (Dellaert, 1980; Koornneef et al., 1982, 1983a), x-ray and γ-ray irradiation have also been shown to cause short viable deletions in Arabidopsis at the chl3 (Wilkinson and Crawford, 1991), tt3 (B. Shirley and H.M. Goodman, unpublished data), and glt (Oppenheimer et al., 1991) loci. Experiments are in progress to systematically test...
several mutagens to determine the optimal mutagenesis conditions for generating deletions in the Arabidopsis genome at high frequency.

Interestingly, the ga7-3 mutant in which we identified a 5.0-kb deletion at the GA1 locus was only backcrossed once to wild type before being used for genetic analysis (M. Koornneef, personal communication). Because genomic subtraction only identified DNA sequences that map to the GA1 locus, it seems likely that the fast neutron mutagenic treatment used to generate ga7-3 did not result in the generation of relatively large deletions at many unlinked loci.

The GA1 Locus of Arabidopsis

The linkage map of Arabidopsis is ~600 cM (S. Hanley and H.M. Goodman, unpublished data), and the genome size is ~10⁸ bp. Therefore, on average, the recombination frequency per nucleotide is ~6 × 10⁻⁶ cM. We used the recombination frequency between different ga7 alleles reported by Koornneef et al. (1983a) (assuming only reciprocal crossovers, i.e., no gene conversions) to calculate the recombination frequency within the ga7 locus. The reported recombination frequency between ga7-6 and ga7-8 alleles and ga7-1 is 0.5 × 10⁻² cM (Figure 1). Our DNA sequence analysis showed that the mutations in ga7-6 and ga7-1 and in ga7-8 and ga7-1 are separated by 432 and 427 bp, respectively (data not shown). This calculation suggests that the recombination frequency in the GA1 locus is ~10⁻⁵ cM per nucleotide, in good agreement with the average recombination frequency of the Arabidopsis genome. The calculation also suggests that the extent of the entire GA1 locus defined by mutants ga7-4 and ga7-7 is ~7 kb, more than sufficient to accommodate the detected 2.8-kb GA1 mRNA.

The biosynthesis of GAs in the Arabidopsis ga7 mutants is blocked at the conversion of geranylgeranyl pyrophosphate (GGPP) to ent-kaurene (Barendse et al., 1986; Zevenaar, 1986). GGPP is a branch point metabolite that also serves as a precursor of other diterpenes and tetraterpenes, such as the phytol chain of chlorophyll and carotenoids. Because ent-kaurene is the first committed intermediate in GA biosynthesis, it is likely that the GA1 gene, required for the formation of ent-kaurene, is a point of regulation for GA biosynthesis (Graebe, 1987). Indeed, the biosynthesis of ent-kaurene has been shown to occur preferentially in rapidly developing tissues, such as immature seeds, shoot tips, petals, and stipules near the young elongating internodes in some species (Chung and Coolbaugh, 1986; Moore, 1989; Moore and Coolbaugh, 1991).

Isolation of the Arabidopsis GA1 gene will allow us to investigate the regulation of GA biosynthesis in response to environmental and developmental signals. Future studies on the localization and pattern of expression of the GA1 gene should provide information on the site and the regulation of GA biosynthesis at an early stage in the pathway.

Extension of Genomic Subtraction to Plants with Complex Genomes

To extend genomic subtraction to organisms with complex genomes such as maize, it is desirable to improve the efficiency of the hybridization steps by increasing the rate of reassociation. It would also be desirable to eliminate the need to reassociate small amounts of the final product. Attempts are currently in progress to modify the protocol by carrying out the hybridization reactions in the presence of dextran sulfate to increase the reassociation kinetics (D. Straus, personal communication) and by ligating adapters onto the Sau3A-digested wild-type DNA before subtraction rather than afterward to permit the amplification of single-stranded sequences (D. Straus, personal communication; M. Mindrinos, B. Shirley, and F.M. Ausubel, unpublished data). The genomic subtraction protocol used successfully to clone the Arabidopsis GA1 locus does not contain these modifications and is not likely to be very efficient for isolating DNA sequences corresponding to small deletions from organisms with complex genomes. Other groups have also published subtractive hybridization procedures that employ multiple rounds of subtraction (Welcher et al., 1986; Bjourson and Cooper, 1988; Wieland et al., 1990). However, the reported levels of enrichment obtained with these procedures were less than what we obtained with genomic subtraction. It is likely that these procedures will also require modification before they can be used to isolate DNA sequences corresponding to small deletions from organisms with complex genomes.

METHODS

Plant Materials

Arabidopsis ga7 mutants were obtained from M. Koornneef. In agreement with Dr. Koornneef (Agricultural University, Wageningen, The Netherlands), we have renamed ga7 alleles based on recommendations promulgated at the Third International Arabidopsis Meeting (East Lansing, MI, 1987). ga7-7 to ga7-10 correspond to the alleles NG5, 659, 31.89, 29.9, 29.423, d352, 6027, A428, NG4, and d69, respectively, described previously in Koornneef et al. (1983a). Growth conditions of ga7 mutants for DNA isolation and seed harvesting were the same as previously described (Koornneef et al., 1983a).

Preparation of Plant DNA for Genomic Subtraction Experiments

Ler DNA and ga7-3 mutant DNA were isolated from 4- to 6-week-old plants and purified by CsCl gradient centrifugation as described elsewhere (Auszubel et al., 1990). The concentration of the purified DNA was determined by measuring the UV absorbance at 260 nm. To confirm the DNA concentration, we compared both Ler and ga1-3 DNAs with bacteriophage λ DNA of known concentration on an agarose gel.
Sonication and Biotinylation of Presumptive Deletion Mutant DNA

Procedures for sonication and photobiotinylation were as described elsewhere (Straus and Ausubel, 1990) with slight modifications. Two hundred micrograms of DNA isolated from the presumptive deletion mutant was resuspended in 2 mL of 10 mM N-(2-hydroxyethyl)-piperazine-N'-3-propanesulfonic acid (EPAPS), pH 8.0, 1 mM EDTA, pH 8.0 (1 x EE buffer) and sonicated in a 15-mL conical plastic tube until the average size of the DNA was ~3 kb as determined by agarose gel electrophoresis. We used a sonicator (model No. W-375, Ultrasonics, Farmsdale, NY) set at continuous output mode setting No. 5 for 5 to 10 sec. We do not recommend using DNA fragments of average size less than 1 kb. It is important that the presumptive deletion DNA be uniformly biotinylated. The procedure described below only results in the addition of approximately one biotin moiety per 100 to 400 bp (Forster et al., 1985). The sonicated DNA was aliquoted to four 2-mL microcentrifuge tubes, boiled for 2 min, and precipitated with ethanol. It is important to aliquot the DNA solution in several flat-bottomed 2-mL microcentrifuge tubes to ensure even and maximum illumination during the next step. Each DNA sample was resuspended in 50 μL dH2O, and 50 μL of 2 μg/μL photoactivatable biotin (model No. K1030; Clontech, Palo Alto, CA) was added to each tube. The tubes with tops open were placed in a styrofoam float in an ice-water bath and placed 10 cm from a sunlamp for 15 min. The solution was dark brown following this treatment. The aliquots of biotinylated DNA were pooled into one tube, and 20 μL 1 M Tris-HCl (pH 9.0) was added. DNA was extracted several times with equal volumes of water-saturated 1-butanol until the butanol phase was colorless. Biotinylated DNA was precipitated with ethanol and resuspended in 72 μL 1 x EE (2.5 μg/μL, assuming 90 % recovery). The DNA pellet should be dark brown.

Sau3A Restriction Digestion of Wild-Type DNA for Genomic Subtraction

Five micrograms of Ler DNA and Ad-2 DNA (Bethesda Research Laboratories) were digested with 20 units of Sau3A (New England Biolabs, Beverly, MA), respectively, for 3 to 6 hr at 37°C to ensure complete restriction digestion.

Genomic Subtraction

Genomic subtraction experiments were performed similarly to those described previously (Straus and Ausubel, 1990). In the first cycle of subtraction, 0.5 μg of Sau3A-digested wild-type Arabidopsis DNA was mixed with 12.5 μg of biotinylated deletion mutant DNA in a 0.5-mL microcentrifuge tube. In the case of the reconstruction experiment, 0.5 μg of Sau3A-digested Ler DNA plus 0.5 ng of Ad-2 DNA (Bethesda Research Laboratories) were hybridized with 12.5 μg of the same preparation of Ler DNA that had been sheared and photobiotinylated. An end-labeled synthetic oligonucleotide (84-mer) with 50,000 cpm (Cerenkov counts) was added as a tracer to ensure complete recovery after each cycle of subtraction. The DNA mixture was boiled in 10 to 15 μL of 2 x EE for 1 min, lyophilized, resuspended in 3 μL dH2O, and mixed with 1 μL 4 M NaCl, 0.16 M (Na)EPPS (pH 8.0), 20 mM EDTA (pH 8.0). Hybridization was carried out for at least 20 hr in a 65°C oven. Two microliters of yeast tRNA (20 μg/μL) and 100 μL EEN buffer (0.5 M NaCl, 1 x EE) was added by pipetting up and down to thoroughly mix. One milliliter of Fluoricon avidin polystyrene assay particles (IDEXX Corp., Portland, ME) was washed twice in 1 x EE to remove azide and free avidin, resuspended in 0.9 mL EEN, and stored at 4°C. One hundred microliters of washed avidin-coated particles was added to the DNA solution, and the sample was incubated at room temperature for 30 min. The mixture was then transferred to an Ultrafree-MC filter unit (model No. UFC3 OGV 00; Millipore Corp., Bedford, MA) and spun 10 sec in a microcentrifuge. The supernatant was transferred to a fresh 1.5-mL microcentrifuge tube. The particles on the filter were washed with 0.2 mL EEN, and the supernatants were pooled, precipitated, and counted using Cerenkov counts.

Recovery should be greater than 70% except for this first cycle. In the first cycle, recovery may be lower due to the presence of unincorporated label used to phosphorylate the tracer DNA. The DNA pellet from the first cycle was resuspended in 6 μL 1 x EE and transferred to a 0.5-mL microcentrifuge tube. Five percent (0.3 μL) of this DNA was saved and diluted into 5 μL 1 x EE for further analysis. For the second and subsequent cycles of subtraction, 10 μg of biotinylated deletion mutant DNA was added to the remaining DNA sample from the previous cycle. Subtractive hybridization was carried out as described for the first cycle except that no additional tRNA was added. In the case of the reconstruction experiment, 5% of the starting materials was saved and diluted in 5 μL 1 x EE. In the actual subtraction experiment for the GAI locus, 10% of the unbound DNA was reserved and diluted into 5 μL 1 x EE at the end of the fifth and sixth cycles.

Addition of Sau3A Adaptors

Before analyzing unbound DNA fragments further, a fraction of this DNA was treated at 80°C, ethanol precipitated, and ligated to Sau3A adaptors, as described previously (Straus and Ausubel, 1990), with indicated modifications. After several cycles of subtraction, the unbound DNA from the last cycle was resuspended in 5 μL 1 x EE. In the case of the reconstruction experiment, 10 ng Sau3A-digested Ad-2 DNA (as control), 2 μL of the reserved samples from the starting material and the first three cycles of subtraction, and 0.5 μL (10%) of the unbound DNA from the fourth cycle were added to 50 μL of 1 M NaCl, 1 x EE, 400 μg/mL yeast tRNA in 1.5-mL microcentrifuge tubes, respectively, and incubated at 80°C for 30 min. In the actual subtraction experiment for the GAI locus, 2 μL of the reserved samples from the first four cycles, 5 μL of the reserved samples from the fifth and sixth cycles, and 0.5 μL of the DNA from the seventh cycles were treated at 80°C.

Polymerase Chain Reaction Amplification of Unbound DNA after Subtractive Hybridization

DNA with Sau3A adaptors was amplified as described (Straus and Ausubel, 1990) except that 0.2 μg of phosphorylated primer was used for each 50 μL polymerase chain reaction (PCR). Five microliters of each PCR was analyzed on 1.5% agarose gels. For cloning individual unbound fragments, PCR-amplified DNA fragments were incubated with Klenow to create blunt ends and ligated into the Smal site of pUC13.

Isolation of Corresponding Genomic and GAI cDNA Clones

λGAI-3 was isolated from a Landsberg erecta genomic library constructed in λFIX (Voytas et al., 1990) using 32P-labeled pGAI1 as probe. pGAI-2 was obtained by ligating a 6-kb Sall-EcoRI fragment from λGAI-3 into Sall plus EcoRI-cleaved pBluescript II SK (Stratagene).
pGA1-4 was isolated from a genomic library of Arabidopsis ecotype Columbia DNA constructed in the binary vector pOCA18 (Olszewski et al., 1988). The 0.9-kb GA1 cDNA clone was isolated by screening a cDNA library, constructed from RNA isolated from siliques (seed pods) of Arabidopsis ecotype Columbia (J. Giraudat and H.M. Goodman, unpublished data), using 32P-labeled 6-kb SalI-EcoRI DNA from pGA1-2 as a hybridization probe.

DNA Gel Blot Analyses

One microgram of HindIII-digested DNA from Arabidopsis was fractionated on 1% agarose gels, transferred to GeneScreen membrane (Du Pont/New England Nuclear), and hybridized with gel-purified, 32P-labeled DNA fragments, as described (Church and Gilbert, 1984).

Agrobacterium tumefaciens-Mediated Transformation of Arabidopsis Root Explants

The transformation procedure was as described previously (Valvekens et al., 1988) with slight modifications as indicated. To prepare roots of gal-3 for transformation, gal-3 seeds were germinated in liquid media containing 1 × Murashige and Skoog (1962) salts and 2% sucrose (MS media) with 100 μM GA3. Media were changed weekly. Two weeks after germination, 1 × MS without gibberellin (GA) was used. pGA1-4 (Figure 5) and three other clones (data not shown) isolated from the pOCA18 library were introduced into Agrobacterium LBA4404 by electroporation (Ausubel et al., 1990). The stability of the inserts in these cosmid clones in LBA4404 was tested by hybridizing DNA gel blots containing HindIII-digested DNA from LBA4404 carrying different clones with 32P-labeled plasmid DNA. Among four clones isolated from the pOCA18 library, two clones were very unstable in LBA4404, in these cosmid clones in LBA4404 was tested by hybridizing DNA gel blots containing HindIII-digested DNA from LBA4404 carrying different clones with 32P-labeled plasmid DNA. Among four clones isolated from the pOCA18 library, two clones were very unstable in LBA4404, indicating that it is important to monitor the stability of plasmids used in Agrobacterium before transformation.

A fresh overnight culture of LBA4404 carrying pGA1-4 was used to infect root explants of 5-week-old gal-3 plants. Kanamycin-resistant (Km') transgenic plants were regenerated as described (Valvekens et al., 1988). To score for GA'/Km' and GA- kanamycin-sensitive segregation, seeds of transgenic plants were placed on MS agar plates containing kanamycin (50 μg/mL). GA'/Km' seeds germinated within 3 days and developed into green seedlings. Nongerminating seeds after 8 days were transferred onto MS plates containing 100 μM GA3 and 50 μg/mL kanamycin. GA- kanamycin-sensitive seeds germinated in the presence of GA3 but became white seedlings.

RNA Isolation and RNA Gel Blot Analysis

Poly(A)+ RNA of 4-week-old and 5-week-old plants was prepared from the entire plant except the roots, and silique RNA was prepared from immature siliques plus some flower buds and stems as previously described (Ausubel et al., 1990). Approximately 2 μg RNA from each sample was treated with glyoxal, size-fractionated on a 1% agarose gel (Maniatis et al., 1982), transferred to GeneScreen membranes, and hybridized with a 32P-labeled 0.9-kb EcoRI DNA fragment from the GA1 cDNA and a 32P-labeled 1.65-kb EcoRI fragment containing the Arabidopsis CAB gene (AB165) (Leutwiler et al., 1986).

DNA Sequence Analyses

DNA sequences of GA1 genomic DNA and cDNA were obtained using the dideoxy method (Ausubel et al., 1990) with Sequenase (U.S. Biochemical Corp.) and both single- and double-stranded DNA templates. The gal-2 DNA was restricted with HindIII and was ligated under conditions that favor the intramolecular ligation (Innis et al., 1990). The inverse PCR (Ausubel et al., 1990; Innis et al., 1990) was carried out to amplify the junctions of the insertion or inversion in gal-2 DNA using 20 to 30 ng of HindIII-digested and religated gal-2 DNA. The inverse PCR amplified fragments were cloned into the pSK vector, and DNA sequences were obtained by using double-stranded DNA templates. The 1.2-kb HindIII fragments were amplified from 10 ng of genomic DNA isolated from Ler and gal-1, gal-6, gal-7, gal-8, and gal-9 by PCR (Ausubel et al., 1990). For DNA sequence analyses, single-stranded DNA templates corresponding to each of the 1.2-kb HindIII fragments were obtained by single-primer reamplification (Ausubel et al., 1990) and purified by selective ethanol precipitation with 2 M ammonium acetate (Innis et al., 1990).

ACKNOWLEDGMENTS

We thank Maarten Koornneef for Arabidopsis gal mutants, Donald Straus for help with the genomic subtraction technique, Susan Hanley for RFLP analysis, Gabor Lazar and Ursula Hanfstdingl for help with Agrobacterium-mediated transformation, and Brian Seed for oligonucleotide synthesis. We also thank Dan Voytas, Neil Olszewski, Jerome Giraudat, and Elliot Meyerowitz for providing Arabidopsis genomic and cDNA libraries. We are grateful to Jean Greenberg for helpful comments on the manuscript and to Rachel Hyde for help in preparing the manuscript. This work was supported by USDA-CRGO Grant No. 91-37300-6384 and by a grant from Hoechst AG to Massachusetts General Hospital.

Received November 26, 1991; accepted December 23, 1991.

REFERENCES


