Volume 16 Number 22 1988 Nucleic Acids Research

Specialized binary vector for plant transformation: expression of the *Arabidopsis thaliana* AHAS gene in *Nicotiana tabacum*

Neil E.Olszewski+, Frank B.Martin[1] and Frederick M.Ausubel

Department of Genetics, Harvard Medical School and Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114 and [1]Department of Applied Statistics, University of Minnesota, St Paul, MN 55108, USA

Received May 10, 1988; Accepted October 24, 1988

ABSTRACT

We constructed a cosmid vector, pOCA18, designed for transferring plant genomic libraries from Agrobacterium tumefaciens to plants. Clones from a genomic library of Arabidopsis thaliana DNA in pOCA18 were propagated stably in both Escherichia coli and A. tumefaciens. Clones from the pOCA18 A. thaliana library were used to construct transgenic Nicotiana tabacum plants; the DNA inserts were transferred intact in 10 out of 16 transgenic N. tabacum plants examined but were partially deleted in six others. Transgenic N. tabacum plants constructed with a mutant A. thaliana acetohydroxy acid synthase gene (from the pOCA18 library) that encodes an enzyme resistant to the herbicide chlorsulfuron were resistant to chlorsulfuron. A statistical analysis indicated that if the A. thaliana library contains 10^7 members and if 10^7 A. tumefaciens transconjugants containing the library were used to transform plant cells, then 2×10^4 transformed plant cells must be generated to have a 95% probability of constructing a transgenic plant carrying a specific DNA sequence from the A. thaliana library.

INTRODUCTION

Molecular cloning of genes from microorganisms is frequently facilitated by the ability to identify a specific gene from a genomic library by selecting for clones that phenotypically complement a mutant phenotype. This approach has not been used with multicellular eukaryotes because of the difficulty in obtaining sufficient numbers of transgenic organisms. However, recent progress in techniques designed to transfer DNA to plants suggests that cloning genes by phenotypic complementation of mutants should be feasible. This contention is supported by the recently reported success in rescuing a model gene encoding kanamycin resistance from the Arabidopsis thaliana genome by selecting for kanamycin resistant transgenic plants following "shotgun transfer" of a genomic library into Petunia cells (1).

A. thaliana has the smallest recorded genome of any plant, 7×10^7 bp (2), making A. thaliana an ideal source of genes for use in phenotypic complementation strategies. Moreover, it has recently been shown that A. thaliana can be used as a recipient host for Agrobacterium tumefaciens-mediated transformation (3).

Vectors designed to transfer DNA to plant cells have generally taken advantage of the relationship between the bacterial phytopathogen, A. tumefaciens, and its host plants. During pathogenesis, A. tumefaciens transfers a portion of its genome (T-DNA or transferred DNA) into the host plant's genome. In A. tumefaciens, T-DNA is located on the Ti or "tumor inducing"

plasmid; the Ti plasmid also contains a set of virulence (vir) genes required in trans for transfer of the T-DNA to plant cells (4,5,6). The T-DNA is flanked by 25 bp imperfect direct repeats called borders (for review see 7); the right border and an adjacent region located outside of the T-DNA, called "overdrive", are required for efficient transformation to occur (8,9,10,11,12). The left and right borders function in cis to delimit the DNA that is transferred.

Normally, expression of T-DNA genes in the plant host causes neoplastic tumor formation at the site of infection (for a recent review, see ref. 13). However, the intervening T-DNA sequences between the left and right borders are not required for DNA transfer to the plant cell and can be replaced by foreign DNA. Based on this observation, several plant transformation vectors have been constructed that lack intervening T-DNA genes but contain the T-DNA borders on a broad-host-range replicon (e.g., refs. 4,14,15). These vectors (called "binary" vectors) replicate in both E. coli and A. tumefaciens and can be mobilized between them. The DNA between the T-DNA borders of the binary vector is transferred at high frequency into a suitable plant host from an A. tumefaciens strain containing the vir genes on a mutagenized Ti plasmid in which the T-DNA has been deleted.

In this paper we describe the construction of a new binary vector, pOCA18, specifically designed for constructing genomic plant libraries. We showed that: 1) An A. thaliana genomic library constructed in pOCA18 and containing 20 kb inserts was stable in both E. coli and A. tumefaciens. 2) A. thaliana inserts in pOCA18 were transferred from A. tumefaciens into Nicotiana tabacum. 3) A pOCA18 A. thaliana genomic clone containing an herbicide-resistant acetohydroxy acid synthase gene conferred herbicide resistance to N. tabacum. In addition, we present a statistical analysis demonstrating that the problem of statistical dispersion occurring during transfer of libraries from E. coli into A. tumefaciens and during the subsequent transfer into plant cells (16) can be minimized by increasing the size of both the original genomic library and the number of A. tumefaciens exconjugates used to generate transgenic plants.

METHODS

Bacterial strains and media

All E. coli strains used, HB101 (17), SM10 (18), MC1061 (19) and MM294(pRK2013) (31) were grown on LB media (20). A. tumefaciens strain LBA4404(pAL4404) (21) was grown on Min A media (22) supplemented with 0.2% glucose and 100 μg/ml rifamycin. A. tumefaciens strain EHA101A is a chloramphenicol-resistant isolate of EHA101(pEHA101) (23). EHA101A was grown on modified LB media (0.5% NaCl) containing 50 μg/ml chloramphenicol. A. tumefaciens and E. coli strains containing pOCA18 or derivatives of pOCA18 were selected using 2.5 and 10 μg/ml of tetracycline, respectively.

Construction of pOCA18

pOCA18 was constructed using standard techniques (20). The region of pOCA18 enlarged in Fig. 1 was cloned into the unique EcoRI site of pRK290 (24); during this process the EcoRI site

was destroyed. The left border region is a synthetic oligonucleotide containing the octopine Ti-plasmid left border sequence (7). The oligonucleotide was synthesized with an EcoRI compatible end that does not regenerate an EcoRI site when cloned into an EcoRI site; the other end is a KpnI sticky end. The chimeric NOS:NPTI:OCS gene was constructed by ligating the NOS promoter from pLGV23neo contained on a 303 bp BclI fragment (25) to a 1.1 kb BclI to ApaI fragment from pLGVneo1103 (26). The ends of the NOS:NPTII:OCS gene were changed to ClaI and KpnI ends by the addition of linkers. The cos site from bacteriophage lambda was cloned between the SalI and PstI sites of πAN7 (27) as a 343 bp SfaNI to HincII fragment (λ positions 48359 to 200) following the addition of SalI linkers to the SfaNI end and PstI linkers to the HincII end. The right border fragment containing "overdrive" was purified from pLGV23neo as a 1.1 kb BclI to HindIII fragment and cloned between the BglII and HindIII sites of πAN7. πAN7 containing the cos site and right border was linearized by digestion with HindIII and cloned between the HindIII and ClaI sites to complete pOCA18.

Isolation of A. thaliana genomic DNA

Nuclei were purified from 4 week old A. thaliana plants, line GH50 (28) by a modification of the method of Olszewski et al. (29). One hundred grams of leaves were washed with cold distilled H_2O and chilled to 0°C. All of the following steps were carried out at 2°C to 4°C unless indicated. The tissue was chopped into small pieces, submerged in diethylether for 3 min, rinsed with cold distilled H_2O, homogenized in 3 volumes of nuclei isolation buffer (1 M sucrose, 10 mM Tris-HCl (pH 7.2), 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 400 μg/ml ethidium bromide) with a Brinkman Polytron PT20ST (Brinkman Instruments, Inc., Westbury, NY), and filtered through four layers of cheesecloth and one layer of Miracloth (Calbiochem-Behringer, La Jolla, CA). Inclusion of ethidium bromide in the isolation buffer reduced DNA degradation (30). The homogenate was subjected to centrifugation at 9,000 RPM for 15 min in a Beckman JA-10 rotor. The pellet was resuspended with a Dounce homogenizer in 50 ml of nuclei isolation buffer containing 0.5% Triton X-100, and subjected to centrifugation at 8,000 RPM for 10 min in a JS-13 rotor. This step was repeated except that the centrifugation step was at 6,000 RPM. The pellet was suspended in 10 ml nuclei isolation buffer containing 0.5% Triton X-100, layered over two discontinuous Percoll (Pharmacia Inc., Piscataway, NJ) gradients constructed in 30 ml Corex tubes containing 5 ml each of 35% and 60% (v/v) Percoll buffer (34.2 g sucrose, 1 ml 1M Tris-HCl (pH7.2), 0.5 ml 1M $MgCl_2$, 34 μl 2-mercaptoethanol, and Percoll to 100 ml) in nuclei isolation buffer, and subjected to centrifugation in a JS-13 rotor at 8,000 RPM for 20 min. Nuclei were harvested from between the 35% and 60% layers and diluted with 5 volumes of nuclei isolation buffer minus ethidium bromide. Nuclei were collected by centrifugation at 8,000 RPM for 10 min in a JS-13 rotor.

Genomic DNA was isolated from the purified nuclear pellet. The nuclear pellet was suspended in 10 ml of 250 mM sucrose, 10 mM Tris-HCl (pH 7.2), 5 mM $MgCl_2$. Following suspension, 11 ml of 10 mM Tris-HCl (pH 8.0), 20 mM EDTA, 2% N-laurosarkosine (w/v) and

200 μg/ml proteinase K was added, and the nuclei were incubated for 2 hr at 55°C. After the addition of 21 g of CsCl and 1 ml of 10 mg/ml ethidium bromide, DNA was banded by centrifugation at 45,000 RPM at 20°C for 36 hr in a Beckman 70.1 Ti rotor. Genomic DNA was collected, the ethidium bromide was partitioned into isopropanol saturated with 3 M CsCl, and the DNA was dialyzed against 10 mM Tris-HCl (pH 8.0), 1 mM EDTA. DNA isolated by this procedure was highly enriched for nuclear DNA but may still have contained some chloroplast DNA.

Construction of an A. thaliana genomic library

A cosmid library in pOCA18 was constructed by digesting A. thaliana genomic DNA with TaqI under conditions where partial digestion occurred and the mean size of the digestion products was 25 kb. Digestion products ranging in size from 15-30 kb were purified from low melting point agarose (International Biotechnologies, Inc., New Haven, CT) by phenol extracting melted gel slices. Size-selected fragments were ligated with a 4-fold molar excess of dephosphorylated ClaI digested pOCA18 at a final DNA concentration of 200 μg/ml. The ligated DNA was packaged into bacteriophage λ particles using Gigapack (Stratagene Cloning Systems, San Diego, CA) and transduced into E. coli HB101 or SM10 as recommended by the packaging extract manufacturer.

Identification of the A. thaliana AHAS gene

Plasmid pBL104 contains a 2.15 kb segment of the A. thaliana AHAS gene as an EcoRI fragment (B. Lalonde and G. Fink, personal communication). This fragment was gel purified, labeled by nick translation and used as a hybridization probe to identify A. thaliana genomic clones containing the AHAS gene (20). Four different AHAS genomic clones were identified when colony lifts containing 20,000 colonies from the pOCA18 A. thaliana genomic library were probed.

Mobilization of plasmids

Plasmids were mobilized into A. tumefaciens either in triparental matings using MM294(pRK2013) as the helper strain (31), or directly from the "mobilizing" strain SM10 (18). Mobilization frequency is expressed as the fraction of recipient cells which were transconjugates.

Plant transformation and tissue culture

Preconditioned N. tabacum cv Wisconsin 38 leaf discs were co-cultivated with A. tumefaciens according to Horsch et al. (32,33) except that the feeder cells were omitted. Transformed shoots were regenerated on solid MS media supplemented with 1.0 μg/ml benzyladenine, 0.1 μg/ml naphthalene acetic acid, 300 μg/ml kanamycin and 500 μg/ml Claforan (Hoechst-Roussel Pharmaceuticals, Inc. Somerville, NJ). Transformed shoots were rooted on solid media containing MS salts (Gibco Laboratories, Grand Island, NY) and 100 μg/ml kanamycin. In some experiments (indicated in the text), chlorsulfuron was substituted for kanamycin. Chlorsulfuron was generously provided by Dr. E. Beyer, Jr. of E. I. duPont de Nemours and Co., Wilmington, DE.

Chlorsulfuron and kanamycin sensitivity of callus growth from surface-sterilized leaf discs was determined on solid MS media supplemented with 2.0 μg/ml naphthalene acetic acid, 0.5 μg/ml benzyladenine and chlorsulfuron and kanamycin as indicated.

Isolation and analysis of DNA from transgenic plants

Total genomic DNA was prepared from transgenic plant leaves according to Dellaporta et al. (34). Genomic DNA (10 μg) was digested with EcoRI and subjected to electrophoresis in a 0.8% TBE agarose gel (20). Genomic DNA was transferred to Genescreen-plus (DuPont Co. NEN Products, Boston, MA) and hybridized according to the manufacturers instructions. Hybridization probes were synthesized using a nick translation kit (Boehringer Mannheim Biochemicals, Indianapolis, IN) according to the manufacturer. Autoradiography was performed at -70°C using an intensifying screen.

Plant genetics

Transformed plants were self pollinated and mature seeds were collected. Seeds were surface-sterilized by soaking in full strength Clorox containing 0.1% Tween 20 for 6 min and then rinsed 4 times with sterile H_2O. Sterile seeds were plated on solid media containing MS salts and the appropriate selective agent (kanamycin 200 μg/ml or chlorsulfuron 500 nM). Segregation of markers was scored when the first secondary leaf was just emerging on resistant seedlings.

RESULTS AND DISCUSSION

The construction of pOCA18 is described in Methods; a map of pOCA18 is shown in Fig. 1. pOCA18 contains the cis-signals necessary for efficient transfer of DNA into plants (i.e., the right border, "overdrive", and the left border sequences of the Ti plasmid). The region of pOCA18 that is transferred (T-DNA) into the plant genome contains a chimeric neomycin phosphotransferase (NPT-II) gene that confers kanamycin resistance to transgenic plants.

pOCA18 transforms tobacco

Tobacco (N. tabacum) leaf discs were co-cultivated with A. tumefaciens strain LBA4404 containing pOCA18 as described in Materials and Methods. In the presence of 300 μg/ml kanamycin, an average of 2.5 kanamycin resistant shoots were obtained per leaf disc. No kanamycin resistant shoots were obtained with leaf discs co-cultivated with LBA4404 not containing pOCA18. Kanamycin resistant shoots were rooted in the presence of 100 μg/ml of kanamycin, then transferred to soil (see Methods for details).

To determine whether pOCA18 T-DNA was present in the kanamycin resistant plants, a Southern blot containing EcoRI digested genomic DNA from three independent kanamycin resistant plants (regenerated from different leaf discs) and an untransformed control plant was probed with labeled pOCA18 (Fig. 2A). Hybridization to a 3.2 kb EcoRI fragment was observed with genomic DNA from the kanamycin resistant plants (Fig. 2A, lanes 1-3), whereas no detectable hybridization to control DNA was observed (Fig. 2A, lane C). Because pOCA18 contains only one EcoRI restriction site (see Fig. 1), EcoRI digestion of single pOCA18 inserts

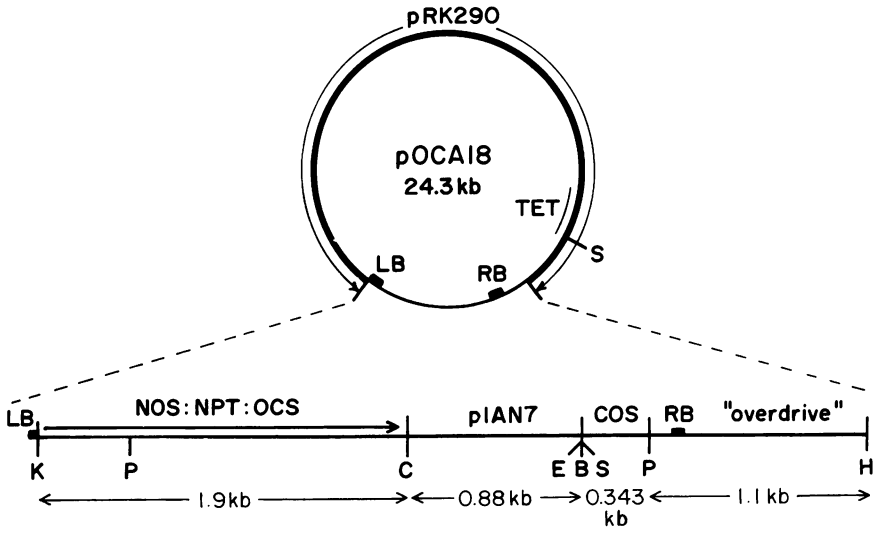

Figure 1. Map of pOCA18. The genetic loci indicated are; LB, left border repeat; NOS-NPT-OCS, a chimeric gene which is expressed in plant cells and confers resistance to kanamycin; πAN7 contains a ColE1 plasmid origin of replication and a bacterial supF tRNA gene (27); COS, the cos region from phage lambda; RB, right border repeat; TET, a bacterial tetracycline resistance gene. The direction of transcription for the NOS-NPT-OCS gene is indicated by the arrow. The T-DNA region contains the cos site, πAN7 and the NOS-NPT-OCS gene is delimited by the right and left border sequences. The restriction sites indicated are: B, BamHI; C, ClaI; E, EcoRI; H, HindIII; K, KpnI; P, PstI; and S, SalI.

should produce only border fragments. However, tandem arrays of the pOCA18 T-DNA (delimited by the left and right borders) should contain internal 3.2 kb EcoRI fragments (see Fig. 1). The occurrence of T-DNA arranged in tandem arrays has been reported previously (14,35). Longer exposures of the Southern blot shown in Fig. 2A revealed weaker bands of different size in the different transgenic plants; these probably represent border fragments between the pOCA18 T-DNA tandem arrays and the tobacco genomic DNA (not shown). The ratio of hybridization intensities between these presumptive border fragments and the 3.2 kb internal fragments indicated that 5-10 copies of the 3.2 kb fragment were present per haploid genomic equivalent.

The presence of the 3.2 kb EcoRI fragment with homology to pOCA18 in the kanamycin resistant plants is consistent with the interpretation that the border sequences functioned properly to delimit the pOCA18 DNA transferred. This interpretation is supported by the observation that hybridization to regions outside of the borders was not detected when a duplicate filter to the one shown in Fig. 2A was probed with pTJS133, a plasmid largely co-linear with pRK290 (36) (data not shown).

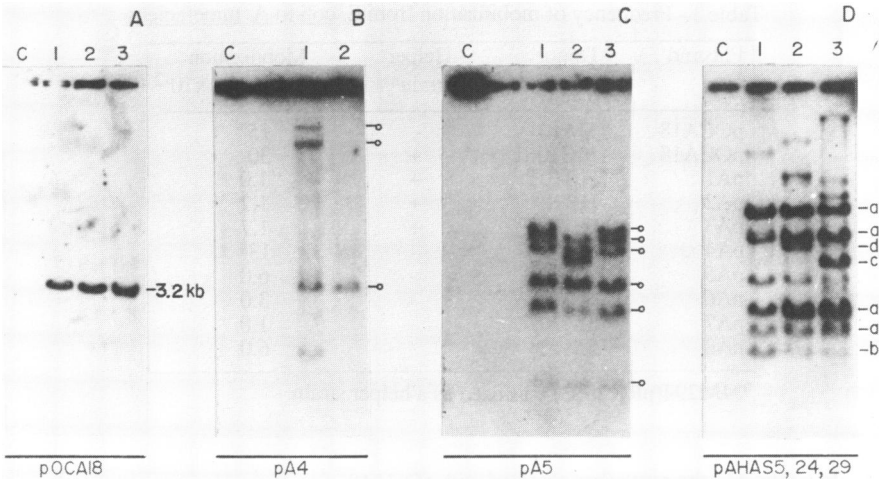

Figure 2. Genomic blots of transgenic tobacco plants. Genomic blots were prepared and hybridized as described in Methods. The plasmid which was used both to construct the transgenic plants and then as a hybridization probe is indicated below the panel. Numbered lanes contain genomic DNA from different transgenic plants and lanes labelled C contain DNA from control (untransformed) plants. Plasmids pA4 and pA5 are A. thaliana genomic clones. Plasmids pAHAS5, 24 and 29 are three independent A. thaliana genomic clones that contain the AHAS gene. -o marks the predicted location of the EcoRI fragments that are internal to the genomic clone analyzed in the panel. -a, indicates the predicted locations of T-DNA fragments common to all three AHAS genomic clones. -b and -c, indicate the predicted location of EcoRI fragments which are specific to pAHAS24 and pAHAS29, respectively.

Construction of an A. thaliana genomic library

An A. thaliana genomic library was constructed in pOCA18 by cloning size-selected TaqI partial digestion products of A. thaliana genomic DNA in the ClaI site of pOCA18. The average size of the A. thaliana inserts was approximately 20 kb.

To determine whether A. thaliana DNA inserts in pOCA18 were stably maintained in E. coli and A. tumefaciens, 21 random cosmid clones were conjugated from E. coli to A. tumefaciens and then back into E. coli. Following these conjugations, cosmid DNA was isolated from the resulting 21 E. coli strains and the EcoRI digestion patterns of the cosmid DNAs were compared to the digestion patterns of the cosmid DNAs prior to the conjugations. For each of the 21 cosmids tested, a single A. tumefaciens transconjugant was chosen as donor to re-conjugate the plasmid back to E. coli. Twenty clones did not undergo detectable changes during these conjugations; however, the majority of the insert in one clone was deleted (data not shown). Interestingly, intact copies of this apparently unstable clone must have been present in the A. tumefaciens transconjugant; when the A. tumefaciens transconjugant containing this clone was used to transform N. tabacum leaf discs, a plant containing an intact copy of the T-DNA was obtained (Fig. 2B, lane 1).

Table 1. Frequency of mobilization from E. coli to A. tumefaciens

Plasmid	Donor strain	Helper strain[a]	Mobilization frequency $x10^{-2}$
pOCA18	SM10	-	15
pOCA18	MC1061	+	30
pA1	HB101	+	1.0
pA2	HB101	+	1.2
pA3	HB101	+	1.7
pA4	SM10	-	13
pA5	SM10	-	6.0
pA6	SM10	-	3.0
pA7	SM10	-	1.0
pA8	SM10	-	6.0

[a]MM294(pRK2013) was used as a helper strain.

In contrast to the results we obtained with pOCA18, we previously reported that A. thaliana libraries constructed in another binary cosmid vector (pOCA7.9) were unstable in A. tumefaciens, with 70% of the clones suffering deletions of the insert DNA (37). pOCA18 and pOCA7.9 are based on the RK2 derived plasmids pRK290 and pTJS75, respectively. pRK290 contains an additional 13.3 kb of RK2 compared to pTJS75 and the additional RK2 sequences in pRK290 have been shown to confer greater stability of maintenance in a number of Gram negative bacteria (36). It is possible that these sequences also confer stability to the genomic inserts in the pOCA18 vector. Another possibility is that the stability of A. thaliana clones in RK2 replicons is proportional to the amount of A. thaliana DNA they carry and that the additional A. thaliana DNA carried in pOCA7.9 genomic clones compared to pOCA18 (approximately 13 kb) destabilizes them. Klee et al. (1) have also observed instability of genomic clones in a binary vector based on pTJS75 and stability of genomic clones in vectors based on the larger RK2 replicon, pRK292.

Genomic libraries constructed in pOCA18 must be transferred from E. coli to A. tumefaciens for use in phenotypic complementation experiments. In order to maintain a similar representation of individual genomic clones in E. coli and A. tumefaciens, all clones must transfer from E. coli to A. tumefaciens at approximately the same frequency. To investigate this matter, we measured the frequency of mobilization of pOCA18 and eight pOCA18 clones containing 20 kb A. thaliana DNA inserts from E. coli into A. tumefaciens (Table 1). The mobilization frequencies of genomic clones observed in triparental and biparental matings were similar and were all within 13 fold of each other, suggesting that most genomic clones in the library mobilize at similar frequencies; thus, the representation of clones in the library should be minimally affected by the transfer process, itself. However, the representation of the clones will be affected by the statistical dispersion that occurs during this process (see below).

Although pOCA18 itself mobilized at a higher frequency than several pOCA18 cosmids

containing inserts, the frequency in the library of clones containing no inserts was less than 1%, making it unlikely that the representation of pOCA18 would be greatly amplified during mobilization of the library to A. tumefaciens.

Transformation of N. tabacum by pOCA18 A. thaliana genomic clones

Seven randomly chosen pOCA18 A. thaliana genomic clones were tested to determine whether they could transform tobacco leaf discs. In addition, three different genomic clones which contained a mutant A. thaliana acetohydroxy acid synthase (AHAS) gene were pooled together and used to transform tobacco leaf discs (see below). All seven pOCA18 genomic clones and the pooled AHAS clones appeared to transform N. tabacum leaf discs; shoots resistant to 300 μg/ml kanamycin were obtained at frequencies ranging from 0.2 to 0.5 shoots per disc following co-cultivation with A. tumefaciens strain EHA101A containing the pOCA18 clones (data not shown).

The EcoRI digestion patterns of the T-DNA in 18 transgenic plants (two or three from each of the eight transformation experiments) were analyzed to determine if rearrangements in the T-DNA had occurred; representative data is shown in Fig. 2B-D. We were unable to determine the structure of the transferred A. thaliana DNA from two of the transgenic tobacco plants because the pOCA18 clone used in these cases contained A. thaliana T-DNA sequences homologous to repetitive tobacco sequences (data not shown). The other pOCA18 A. thaliana clones had only weak or no homology to DNA from untransformed tobacco plants. Ten of the transgenic tobacco plants contained A. thaliana DNA sequences which had the same EcoRI digestion pattern as the original clone in E. coli. In two transgenic plants, most of the A. thaliana DNA was deleted. In the other four cases where deletions occurred, only one or two of the EcoRI bands were altered. The occurrence of deletions did not appear to be specific to any particular clones. In the case of the AHAS clones and in five out of the six random genomic clones which could be analyzed, at least one of the analyzed transgenic plants had no deletions in the A. thaliana DNA.

When duplicate Southern blots of the ones described in Fig. 2 were hybridized with labeled pTJS133 DNA, no hybridization was detected (data not shown), indicating that pRK290 vector sequences were not stably transferred and that both the left- and right-border sequences were probably functioning to delimit the DNA stably incorporated into the tobacco genome.

The A. thaliana AHAS gene is expressed in N. tabacum

The A. thaliana library described in this paper was constructed using genomic DNA from an A. thaliana mutant (mutant line GH50, mutant csr-1; 28) that is resistant to the sulfonylurea herbicide chlorsulfuron. Chlorsulfuron inhibits acetohydroxy acid synthase (AHAS; E.C.4.1.3.18), the first enzymatic step specific to the branched-chain amino acid biosynthesis pathway. The csr-1 mutation is due to a mutation in the AHAS gene (37).

To determine whether the mutant A. thaliana AHAS gene could be expressed and whether its product could function in transgenic N. tabacum plants, we isolated pOCA18 clones from the

Table 2. Chlorsulfuron-resistant shoot and callus growth

	Number of shoots[a]				Callus number[b]	
	10 days post-cocultivation		4 weeks post-cocultivation		4 weeks post-cocultivation	
Chlorsulfuron (nM)	AHAS	pA10	pAHAS	pA10	pAHAS	pA10
0	++++	++++	++++	++++	++++	++++
1	++++	++++	++++	++++	++++	++++
10	++++	++++	++++	++++	+++	+
25	++	+	++	++	++	-
50	+	-	+	+	++	-
100	+	-	+	-	++	-
1000	-	-	-	-	-	-

[a]Shoot numbers ranged from: ++++, multiple shoots per leaf disc; to +, 2-4 shoots per 16 leaf discs. - indicates that no shoots were present.

[b]Only growing callus was scored. Some leaf discs initially produced a small amount of callus which did not grow. Callus numbers ranged from: ++++, 1 per leaf disc; to +, 4-5 per 16 leaf discs. - indicated that no callus was present.

A. thaliana library containing the mutated AHAS gene, used these clones to construct transgenic N. tabacum plants, and tested the transgenic plants for chlorsulfuron resistance. Three different pOCA18 clones containing the A. thaliana AHAS gene were identified by the colony hybridization procedure using a wild-type A. thaliana AHAS probe provided by B. Lalonde and G. Fink. These clones were mobilized into A. tumefaciens strain EHA101A and a mixture of the three A. tumefaciens strains containing the AHAS clones were co-cultivated with N. tabacum leaf discs. Table 2 shows that when inoculated leaf discs were placed on 50 and 100 nM chlorsulfuron, shoots were obtained within eight days whereas no shoots were observed on the control discs; however, with time, shoots also appeared on the non-transformed leaf discs. The kanamycin sensitivity of 7 shoots that had been transformed with the AHAS clones was determined by transferring them to rooting media (see Methods) containing 100 µg/ml of kanamycin; only one shoot formed roots and the remaining shoots became bleached and died. Authentic transformants do not die and generally form roots on this media (data not shown); suggesting that the majority of the shoots obtained in the presence of 50 or 100 nM chlorsulfuron were escapes.

In addition to shoot growth, in the case of the leaf discs that had been transformed with the AHAS clones, callus growth was also observed after about three weeks. Callus growth appeared to be more sensitive to chlorsulfuron than shoot growth; no growing callus was observed on control discs incubated with 25 nM chlorsulfuron (Table 2). In contrast to the shoots that formed in the presence of chlorsulfuron, the callus that formed under this selection appeared to all be transformed because it was possible to generate shoots that were resistant to

300 μg/ml of kanamycin from each of the 7 separate pieces that were tested. In addition, these shoots all formed roots on rooting media containing 100 μg/ml of kanamycin (data not shown).

It is unlikely that the shoots initially selected on chlorsulfuron were really transformed and were expressing chlorsulfuron but not kanamycin resistance, since all of the callus selected on chlorsulfuron produced kanamycin-resistant shoots. A more likely explanation is that shoots are less sensitive to chlorsulfuron than callus and that shoot development in the presence of chlorsulfuron is a mechanism for escaping the chlorsulfuron selection. In support of this explanation, we have observed that many shoots which form following kanamycin selection appear to derive from callus tissue, whereas shoots which arise following chlorsulfuron selection do not go through a callus intermediate (data not shown).

Three tobacco plants that presumptively contained the A. thaliana AHAS gene were used in further studies. These three plants were regenerated using different selection schemes. The first plant (AHAS-1) was selected as a shoot in the presence of 50 nM chlorsulfuron and then was rooted in the presence of 25 nM chlorsulfuron. The second plant (AHAS-2) was selected as a shoot in the presence of 300 μg/ml kanamycin and then rooted in the presence of 100 μg/ml kanamycin. The third plant (AHAS-3) was derived from callus selected in the presence of 50 nM chlorsulfuron that was induced to shoot and root in the presence of 300 and 100 μg/ml kanamycin respectively.

Southern blot analysis revealed that all three chlorsulfuron resistant tobacco plants contained complete and unrearranged copies of the T-DNA from the AHAS genomic clones (Fig. 2D lanes 1-3). These three plants were regenerated from leaf discs that had been co-cultivated with a mixture of three different genomic clones, each of which has EcoRI restriction fragments which are unique to, and diagnostic for, that clone. Because two of the AHAS-transformed plants contained EcoRI fragments corresponding to the unique EcoRI fragment(s) for more than one AHAS genomic clone (Fig. 2D lanes 2 and 3), we concluded that these plants had regenerated from a cell that had been transformed by more than one AHAS genomic clone or that these shoots had arisen from more than one transformed cell. Consistent with this interpretation, these unique EcoRI fragments were present in fewer copies than the fragments which are common to all of the AHAS clones. On the other hand, we have not ruled out the possibility that these presumptive diagnostic EcoRI fragments may have arisen by rearrangement of the A. thaliana DNA or that they represent the border fragments between the T-DNA and the tobacco genome.

The level of chlorsulfuron resistance conferred by the A. thaliana AHAS gene in the tobacco transformants was determined by assaying callus growth on different concentrations of chlorsulfuron (Fig. 3). Callus growth from control (untransformed) leaf discs was reduced by 5 nM and completely inhibited by 10 nM chlorsulfuron. Callus derived from AHAS-transformed leaves was resistant to at least a 20-fold higher concentration of chlorsulfuron than control callus. Vigorous callus growth was observed from AHAS leaf-discs at 100 nM, the highest concentration of chlorsulfuron tested (Fig. 3).

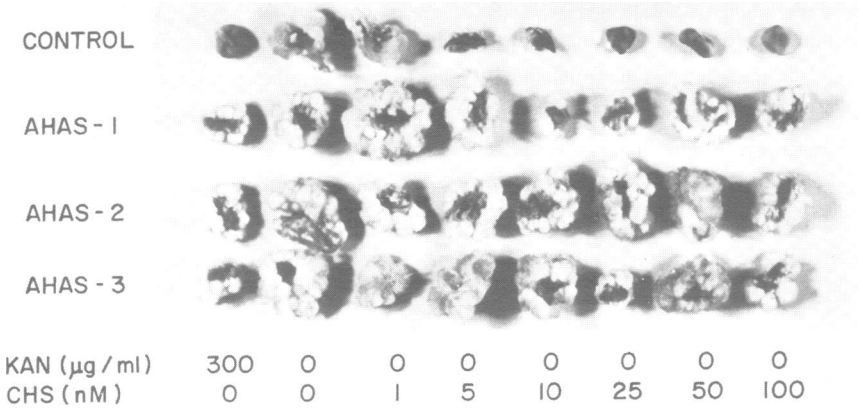

Figure 3. Chlorsulfuron sensitivity of transgenic plants. The chlorsulfuron sensitivity of callus growth on leaf discs derived from either plants transformed by genomic clones containing the AHAS gene (ALS1-ALS3) or genomic clone pA10 (a non-AHAS encoding clone) was assayed as described in Methods. The concentration of chlorsulfuron or kanamycin included in the media is indicated below the leaf discs.

Haughn et al. (38) have also identified genomic clones of the mutant AHAS gene which is responsible for the csr-1 mutation and used these clones to generate chlorsulfuron resistant N. tabacum plants.

Genetic segregation of kanamycin and chlorsulfuron resistance

Twenty-one transgenic tobacco plants were allowed to self-pollinate and set seed. Eighteen of these plants were transformed with A. thaliana genomic clones, and three were transformed with pOCA18 alone. These plants were the source of the genomic DNA that was used in the T-DNA structural analyses described above. Three of these plants were infertile and two had poor seed yields. For sixteen plants that had good seed yield, the ratio of kanamycin-resistant to kanamycin-sensitive seed progeny was scored by sowing the seeds on 200 µg/ml kanamycin (see Materials and Methods; Table 3). Kanamycin resistant plants grew at similar rates and appeared normal in both the presence and absence of kanamycin. Kanamycin sensitive progeny only grew to the expanded cotyledon stage and the cotyledons had a bleached appearance in the presence of kanamycin. Segregation ratios for the seed progeny from 10 plants were consistent with the hypothesis that kanamycin resistance was segregating as either a single (segregating 3:1), two independent (segregating 15:1) or four independent (segregating 255:1) Mendelian loci (Table 3).

The segregation of kanamycin resistant and sensitive seedlings from six plants was not consistent with simple Mendelian inheritance. All of the progeny from three plants (4-2, 10-1, and 18-3) were sensitive to kanamycin (Table 3). In two of these cases, analysis of genomic DNA from the parent plant indicated that the parent plant was transformed (Fig. 2A, lane 3 and Fig. 2B, lane 2). It is possible that the selectable marker was present in the progeny but not

Table 3. Segregation of resistance markers in transgenic progeny

Parent	Kanamycin selection			Chlorsulfuron selection		
	Resistant	sensitive	Predicted Ratio[a]	Resistant	Sensitive	Predicted Ratio
2-1[b]	187	60	3:1			
2-2	188	58	3:1			
3-2	151	37	3:1			
4-1	114	33	3:1			
4-2	1	487	---			
5-2	193	62	3:1	0	182	---
6-1	171	69	3:1			
6-2	114	9	15:1			
10-1	2	475	---			
10-2	200	60	3:1			
AHAS1	231	82	3:1	114	16	*
AHAS2	223	38	*	82	12	*
AHAS3	383	0	---	260	0	---
181	465	2	255:1			
182	140	31	*			
183	1	328	---			

[a]The chi-square value was calculated for the ratio predicted and the ratio is indicated only if P>0.05. * indicates that the ratio falls between 3:1 and 15:1.

[b]Parents 2-1 thru AHAS3 were all transformed using genomic clones. 181, 182 and 183 were transformed using pOCA18.

expressed or that the original transformed plants contained both transformed and nontransformed cells and that the germ line was derived from nontransformed cells. We have not tried to resolve this question. Loss of kanamycin resistance in progeny of transformed plants has also been observed and discussed by Budar et al. (39).

The segregation ratios for the progeny from two plants (AHAS-2 and 18-2) were intermediate to those predicted for segregation at a single locus or at two unlinked Mendelian loci. Segregation from multiple linked loci could explain these ratios. However, the reason for this segregation pattern is unclear. Finally, all of the seed progeny from the AHAS-3 plant were resistant to kanamycin. This could have been caused by transformation at multiple loci, or by gene conversion. No simple correlation is apparent between the integrity of the T-DNA and the inheritance pattern of the kanamycin resistance marker. Interestingly, all of the parents which had poor seed yield or were infertile contained an intact copy of the T-DNA (data not shown). It is not clear if the presence of intact A. thaliana DNA affected the fertility of these plants or if the infertility was caused by some other factor such as passage through tissue culture.

Segregation of chlorsulfuron resistance was scored in the seed progeny of AHAS transformed plants (see Materials and Methods section; Table 3). Under the selection conditions employed (500 nM chlorsulfuron), germination of control (transformed with a non-AHAS clone)

Table 4. Statistical dispersion of 4 kb genes in an A. thaliana genomic library during transfer

Library size in E. coli (N_1)	Gene frequency lower 99.5% confidence bound (f_1)[a]	Number of transconjugates (N_2)	Gene frequency lower 99% confidence bound (f_2)[b]	Plants required for 95% probability of observing at least one complementation (N_3)[c]
5×10^4	2.88×10^{-5}	1×10^6	1.50×10^{-5}	2.15×10^5
		1×10^7	2.44×10^{-5}	1.32×10^5
1×10^5	7.46×10^{-5}	1×10^6	5.24×10^{-5}	6.14×10^4
		1×10^7	6.76×10^{-5}	4.76×10^4
5×10^5	1.36×10^{-4}	1×10^6	1.06×10^{-4}	3.04×10^4
		1×10^7	1.26×10^{-3}	2.55×10^4
1×10^6	1.50×10^{-4}	1×10^6	1.19×10^{-4}	2.71×10^4
		1×10^7	1.40×10^{-4}	2.29×10^4
5×10^6	1.70×10^{-4}	1×10^6	$1.36\ 10^{-4}$	2.37×10^4
		1×10^7	1.59×10^{-4}	2.02×10^4
1×10^7	1.74×10^{-4}	1×10^6	1.40×10^{-4}	2.30×10^4
		1×10^7	1.63×10^{-4}	1.97×10^4

a) $f_1 = P - [2.57\sqrt{P(1-P)}/\sqrt{N_1}]$; Where P is the frequency of fragments containing an intact copy of a unique 4 kb gene in an infinitely large population of randomly generated 17 kb A. thaliana genomic DNA fragments and 2.57 is the 99.5 percentile point of the standard normal distribution. $P = 1/5400 = 13{,}000 / [7 \times 10^7 - 5\ (17{,}000)]$ where $7 \times 10^7 - 5\ (17{,}000)$ is the number of different possible 17 kb fragments which can be generated by randomly shearing the 7×10^7 A. thaliana genome and 13,000 is the number of different possible 17 kb fragments which can contain intact copies of the 4 kb gene. 5 (17,000) is a correction for size bounded end fragments.

b) $f_2 = f_1 - [2.57\sqrt{f_1(1-f_1)}/\sqrt{N_2}]$; since f_1 is calculated using the 99.5 percentile point of the standard normal distribution and f_2 is calculated using the 99.5 percentile point of the standard normal distribution, f_2 represents the product of these percentile points, or the lower 99% confidence bound.

c) N_3 is the smallest solution to $N_3 \geq \ln(1-.96)/\ln(1-f_2)$ so that .95 confidence is given by the product of 0.96 and 0.99 from the previous column.

seed progeny was delayed 4-5 days. Following germination the cotyledons were only partially green and did not expand. Root growth was also inhibited. Germination of chlorsulfuron-resistant seeds was similarly delayed. However, following germination, resistant seedlings developed normally. The segregation ratios observed for both chlorsulfuron resistance and kanamycin resistance were similar in seeds from AHAS-2 and AHAS-3 (Table 3). Segregation frequencies for chlorsulfuron resistance in AHAS-1 progeny were intermediate between those predicted for segregation from either a single locus or two independently segregating loci. In AHAS-1 progeny, segregation of kanamycin resistance was consistent with segregation from a single locus.

One possible explanation for the discrepancies in the segregation frequencies of the kanamycin and chlorsulfuron resistance markers in AHAS-1 is that chlorsulfuron inhibited the germination of a fraction of the sensitive but not resistant seeds; seeds which did not germinate were not scored. Since the segregation frequencies for both kanamycin and chlorsulfuron resistance in the progeny of the different transgenic plants varies while the segregation frequency for both kanamycin and chlorsulfuron resistance are similar within the progeny of each transgenic plant (Table 3), it seems likely that the chlorsulfuron resistance observed in AHAS-transformed tobacco plants was due to the expression of the A. thaliana AHAS gene.

Statistical dispersion of the library during transfer

Simoens et al. (16) calculated the statistical dispersion that will occur when an A. thaliana genomic library containing 15,000 members is transferred from E. coli to A. tumefaciens and then to plants. They calculated that 5.4×10^5 independently transformed plants would have to be generated to have a 95% probability of recovering a specific clone.

Because it is easier to increase either the size of the A. thaliana genomic library in E. coli or the number of A. tumefaciens transconjugants than it is to generate additional transformed plants, we calculated the effects of library size and A. tumefaciens transconjugant population size on the number of transformed plants which must be generated in order to have a 95% probability of observing a complementation event (Table 4). These calculations were made assuming that cloning, mating and transformation are all random events, that the average gene size is 4 kb, the average insert size is 17 kb and that the A. thaliana genome is 7×10^7 bp (2). As shown in Table 4, increasing either the library size or the number of A. tumefaciens transconjugants reduces the number of transgenic plants that must be generated (Table 4). Starting with a genomic library containing 1×10^7 members and an A. tumefaciens transconjugant population of 1×10^7, the lower confidence bound on the occurrence of each unique A. thaliana gene is 1.66×10^{-4}. To have a 95% probability of observing complementation with a clone represented at this frequency, it is necessary to generate 1.97×10^4 transgenic plants. For comparison, intact copies of a unique 4 kb gene are represented in an infinite population of randomly generated 17 kb fragments of A. thaliana genomic DNA at a frequency of 1.85×10^{-4} (Table 4). If this representation could be maintained in the A. tumefaciens population, 1.62×10^4 transgenic plants would be required

in order to have a 95% probability of observing complementation for any 4 kb gene.

It should be noted that the number of transgenic plants that must be generated to observe complementation is affected by the frequency of transformants containing intact T-DNA inserts. We have not tried to take into account the frequency of intact T-DNA inserts that is observed with the A. thaliana genomic library and calculate the number of transformants that must be obtained to observe complementation; although 38% of the transgenic plants obtained with this library lacked intact T-DNAs all contained some T-DNA (see above) and it is not clear how to deal with them in these calculations. However, the number of transformants that must be generated with this genomic library may be close to the values presented in Table 4 because we have calculated these values using 17 kb as the value for the average insert size and the average insert size in the library is actually 20 kb.

The data presented in Table 4 clearly indicate that effort spent generating a large library in E. coli and large numbers of A. tumefaciens transconjugants is well spent in view of the savings in the number of transgenic plants required. Generating an A. thaliana genomic library containing 1 x 10^7 members in pOCA18 should not be a problem since we generally obtain 2 x 10^6 library members/μg of size selected A. thaliana DNA in pOCA18 or closely related derivatives (37). An additional advantage of keeping the number of transgenic plants low is that as more transgenic plants are required to observe complementation it becomes more likely that spontaneous reversion of the mutation will interfere with detection of complementation events.

Conclusions

We have constructed a binary cosmid cloning vector, pOCA18, designed to transfer plant genomic libraries from A. tumefaciens to plants in phenotypic complementation experiments. We constructed an A. thaliana genomic library in pOCA18 that is stable in both E. coli and A. tumefaciens. When A. thaliana genomic clones were used to construct transgenic tobacco plants, 62% of the transgenic plants contained intact copies of the A. thaliana genomic DNA. We also used pOCA18 to transfer a mutant A. thaliana AHAS gene that encodes chlorsulfuron resistance into tobacco; the resulting transgenic tobacco plants were chlorsulfuron resistant and 62% of the progeny of the transgenic plants segregated the kanamycin resistance marker in a simple Mendelian fashion.

Simoens et al. (16) have calculated that to clone a gene by phenotypic complementation using an A. thaliana genomic library containing 15,000 members, 5.4 x 10^5 transgenic plants must be generated to have 95% chance of success. However, we present calculations indicating that by increasing the size of the E. coli library to 1 x 10^7 members and by using an inoculum containing 1 x 10^7 A. tumefaciens transconjugants, only 1.97 x 10^4 transgenic plants need to be regenerated.

Concurrent with this work, Klee et al. (1) rescued a model gene encoding kanamycin resistance from the A. thaliana genome by phenotypic complementation in callus following "shotgun transfer" of a genomic library into Petunia cells. The gene which was rescued was a

chimeric gene consisting of the bacterial neomycin phosphotransferase gene that had been introduced into the A. thaliana genome by A. tumefaciens mediated transformation. It remains to be demonstrated that this approach will work with endogenous genes.

+Present address: Department of Botany, University of Minnesota, St Paul, MN 55108, USA

ACKNOWLEDGMENTS

We thank B. Lalonde and G. Fink for providing pBL104 containing the wild-type A. thaliana AHAS gene prior to publication, G. Haughn and C. Somerville for providing A. thaliana line GH-50, M. Stempowski and J. Smith for synthesizing the left border oligonucleotide, R. Hyde for assistance in the preparation of this manuscript, and J. Chory for helpful discussions and donating the C to pOCA. This work was supported by a grant from Hoechst AG to Massachusetts General Hospital.

REFERENCES

1. Klee, H.J., Hayford, M.B. and Rogers, S.G. (1987) Mol. Gen. Genet. **210**, 282-287.
2. Leutwiler, L.S., Hough-Evans, B.R. and Meyerowitz, E.M. (1984) Mol. Gen. Genet. **194**, 15-23.
3. Lloyd, A.M., Barnason, A.R., Rogers, S.G., Byrne, M.C., Fraley, R.T. and Horsch, R.B. (1986) Science **234**, 464-466.
4. Hoekema, A., Hirsch, P.R., Hooykaas, P.J.J. and Schilperoort, R.A. (1983) Nature **303**, 179-181.
5. Klee, H.J., White, F.F., Iyer, V.N., Goldon, M.P. and Nester, E.W. (1983) J. Bacteriol. **153**, 878-883
6. Stachel, S.E. and Nester, E.W. (1986) EMBO J. **5**, 1445-1454.
7. Caplan, A., Herrera-Estrella, L., Inze, D., van Haute, E., van Montagu, M., Schell, J. and Zambryski, P. (1983) Science **222**, 815-821.
8. Shaw, C.H., Watson, M.D., Carter, G.H. and Shaw, C.H. (1984) Nucl. Acids Res. **12**, 6031-6041.
9. Wang, K., Herrera-Estrella, L., van Montagu, M. and Zambryski, P. (1984) Cell **38**, 455-462.
10. Jen, G.C. and Chilton, M.-D. (1986) Proc. Nat. Acad. Sci. USA **83**, 3895-3899.
11. Peralta, E.G., Hellmiss, R. and Ream, W. (1986) EMBO J. **5**, 1137-1142.
12. Rubin, R.A. (1986) Mol. Gen. Genet. **202**, 312-320.
13. Gheysen, G., Dhaese, P., van Montagu, M. and Schell, J. (1985) In Hohn, B. and Dennis, E.S. (eds.), Advances in Plant Gene Research, Vol. 2, pp. 11-47. Springer-Verlag, Berlin.
14. Bevan, M. (1984) Nucl. Acids Res. **12**, 8711-8721.
15. An, G., Watson, B.D., Stachel, S., Gordon, M.P. and Nester, E.W. (1985) EMBO J. **4**, 277-284.
16. Simoens, C., Alliotte, Th., Mendel, R., Muller, A., Schiemann, J., van Lijsebettens, M., Schell, J., van Montagu, M. and Inze, D. (1986) Nucl. Acids Res. **14**, 8073-8090.
17. Boyer, H.W. and Roulland-Dussoix, D. (1969) J. Mol. Biol. **41**, 459-472.
18. Simon, R., Priefer, U. and Puhler, A. (1983) In Puhler, A. (ed), Molecular Genetics of the Bacteria-Plant Interaction, pp. 98-106. Springer-Verlag, Berlin.
19. Casadaban, M.J., Chou, J. and Cohen, S.N. (1980) J. Bacteriol. **143**, 971-980.
20. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor.

21. Ooms, G., Hooykaas, P.J.J., van Veen, R.J.M., van Beelen, P., Regensburg-Tuink, T.J.G. and Schilperoort, R.A. (1982) Plasmid **7**, 15-29.
22. Miller, J.H. (1972) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor.
23. Hood E.E., Helmer, G.L., Fraley, R.T., and Chilton, M.-D. (1986) J. Bacteriol. **168**, 1291-1301.
24. Ditta, G., Stanfield, S., Corbin, D. and Helinski, D. (1980) Proc. Nat. Acad. Sci. USA **77**, 7347-7351.
25. Herrera-Estrella, L., de Block, M., Messens, E., Hernalsteens, J.P., van Montagu, M. and Schell J. (1983) EMBO J. **2**, 987-995.
26. Hain, R., Stebel, P., Czernilofsky, A.P., Steinbib, H.H., Herrera-Estrella, L. and Schell J. (1985) Mol. Gen. Genet. **199**, 161-168.
27. Lutz, C.T., Hollifield, W.C., Seed, B., Davie, J.M. and Huang, H.V. (1987) Proc. Nat. Acad. Sci. USA **84**, 4379-4383.
28. Haughn, G.W., and Somerville, C. (1986) Mol. Gen. Genet. **204**, 430-434.
29. Olszewski, N.E., Hagen, G., and Guilfoyle, T.J. (1982) Cell **29**, 395-402.
30. Kislev, N. and Rubenstein, I. (1980) Plant Physiol. **66**, 1140-1143.
31. Ruvkun, G.B. and Ausubel, F.M. (1981) Nature **289**, 85-88.
32. Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G. and Fraley, R.T. (1985) Science **227**, 1229-1231.
33. Horsch, R.B. and Klee, H.J. (1986) Proc. Nat. Acad. Sci. USA **83**, 4428-4432.
34. Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) Plant Mol. Biol. Rep. **1**, 19-21.
35. Zambryski, P., Holsters, M., Kruger, K., Depicker, A., Schell, J., van Montagu, M. and Goodman, H.M. (1980) Science **209**, 1385-1391.
36. Schmidhauser, T.J. and Helinski, D. (1985) J. Bacteriol. **164**, 446-455.
37. Olszewski, N., Voytas, D., Hu, J., Chory, J. and Ausubel, F. (1987) In Fox,J.E. and Jacobs, M. (eds), Molecular Biology of Plant Growth Control, pp. 145-155. Alan R. Liss, New York.
38. Haughn, G.W., Smith, J., Mazur, B. and Somerville, C. (1988) Mol. Gen. Genet. **211**, 266-271.
39. Budar, F., Thia-Toong, L., van Montagu, M. and Hernalsteens, J.-P. (1986) Genetics **114**, 303-313.