

# Arabidopsis Genes Essential for Seedling Viability: Isolation of Insertional Mutants and Molecular Cloning

Gregory J. Budziszewski,\* Sharon Potter Lewis,\* Lyn Wegrich Glover,\* Jennifer Reineke,\*<sup>1</sup>  
Gary Jones,\*<sup>2</sup> Lisa Schlater Ziemnik,\*<sup>3</sup> Jennifer Lonowski,\* Beat Nyfeler,\* George Aux,\*  
Qing Zhou,\* John McElver,\* David A. Patton,\* Robert Martienssen,<sup>†</sup> Ueli Grossniklaus,<sup>†,4</sup>  
Hong Ma,<sup>†,5</sup> Marcus Law\* and Joshua Z. Levin\*

\*Syngenta Biotechnology, Inc., Research Triangle Park, North Carolina 27709 and <sup>†</sup>Cold Spring Harbor Laboratories,  
Cold Spring Harbor, New York 11724

Manuscript received April 24, 2001

Accepted for publication September 17, 2001

## ABSTRACT

We have undertaken a large-scale genetic screen to identify genes with a seedling-lethal mutant phenotype. From screening ~38,000 insertional mutant lines, we identified >500 seedling-lethal mutants, completed cosegregation analysis of the insertion and the lethal phenotype for >200 mutants, molecularly characterized 54 mutants, and provided a detailed description for 22 of them. Most of the seedling-lethal mutants seem to affect chloroplast function because they display altered pigmentation and affect genes encoding proteins predicted to have chloroplast localization. Although a high level of functional redundancy in *Arabidopsis* might be expected because 65% of genes are members of gene families, we found that 41% of the essential genes found in this study are members of *Arabidopsis* gene families. In addition, we isolated several interesting classes of mutants and genes. We found three mutants in the recently discovered nonmevalonate isoprenoid biosynthetic pathway and mutants disrupting genes similar to *Tic40* and *tatC*, which are likely to be involved in chloroplast protein translocation. Finally, we directly compared T-DNA and *Ac/Ds* transposon mutagenesis methods in *Arabidopsis* on a genome scale. In each population, we found only about one-third of the insertion mutations cosegregated with a mutant phenotype.

**W**HAT genes are essential for the viability of a plant? Because of the complexity of the multitude of biological processes required for a plant to grow and develop, a large and diverse set of genes are likely to be involved. A forward genetics approach to this question is a powerful method to identify the relevant genes. This approach involves the isolation of embryo-defective mutants and seedling-lethal mutants, which are likely to comprise the largest classes of visible mutants in *Arabidopsis*. There is often an overlap in the mutants identified in embryo and seedling screens because embryo-defective mutants that form seeds capable of germination may also be identified as seedling-lethal mutants. Genes with a seedling-lethal phenotype are likely to encode genes specifically required during early seedling development

as well as more generally functioning genes whose absence becomes critical during seedling development. Although a saturation ethyl methanesulfonate (EMS) mutagenesis has identified several *Arabidopsis* genes with a seedling-lethal mutant phenotype (JURGENS *et al.* 1991; MAYER *et al.* 1991), subsequent analysis has been limited to a small subset of genes with unusual body patterns. Similarly, >300 embryo-defective *Arabidopsis* mutants have been isolated (ERRAMPALLI *et al.* 1991; CASTLE *et al.* 1993; MEINKE 1994) and ~150 were mapped (FRANZMANN *et al.* 1995). Molecular cloning of genes in these two classes on a genome-wide scale has not been reported.

Previous studies provide conflicting estimates of how many genes are essential for embryogenesis in *Arabidopsis*. On the basis of the frequency of multiple alleles in genes with an embryo-defective phenotype, there are estimated to be only 500 genes essential for embryogenesis (FRANZMANN *et al.* 1995). By contrast, an estimated 3500–4000 genes are predicted to be essential for embryogenesis on the basis of the frequency of fusca mutants in large-scale seed color and seedling-lethal screens (JURGENS *et al.* 1991; MISERA *et al.* 1994). The number of genes essential for the seedling stage of growth has not been reported. About 1100 genes are estimated to be essential for gametophytic function on the basis of an analysis of transmissible deficiencies hav-

Corresponding author: Joshua Z. Levin, Syngenta Biotechnology, Inc., 3054 Cornwallis Rd., Research Triangle Park, NC 27709.  
E-mail: joshua.levin@syngenta.com

<sup>1</sup> Present address: Duke University, Durham, NC 27708.

<sup>2</sup> Present address: BASF Plant Science, Research Triangle Park, NC 27709.

<sup>3</sup> Present address: University of Cincinnati College of Medicine, Cincinnati, OH 45267.

<sup>4</sup> Present address: Institute for Plant Biology, University of Zürich, CH-8008 Zürich, Switzerland.

<sup>5</sup> Present address: Department of Biology, Penn State University, University Park, PA 16802.

ing average DNA losses of <90 kb (VIZIR *et al.* 1994). Estimates from genetic studies of other eukaryotes are relatively similar to each other with respect to the proportion of essential genes in a genome. In *Saccharomyces cerevisiae* 17% of the genes were shown to be essential for viability in rich medium by creating a collection of mutant lines, each with the precise deletion of one open reading frame (ORF; WINZELER *et al.* 1999). For *Caenorhabditis elegans*, estimates range from 15% (BRENNER 1974; HERMAN 1988) to 25% (STEWART *et al.* 1998). Recent studies with RNA interference suggest that ~10% of *C. elegans* genes are essential (FRASER *et al.* 2000). For *Drosophila melanogaster*, 28% of the genes have been estimated to be essential (BOSSY *et al.* 1984).

In this study, we have initiated a comprehensive functional genomics effort to identify the genes required for viability at the seedling stage of plant development. Parallel efforts to identify genes required at the embryo stage of development are described in MCELVER *et al.* (2001). Because these genes encode essential proteins in Arabidopsis, this research may also have applications to the identification of new herbicidal compounds (WARD and BERNASCONI 1999). From a screen of Arabidopsis T-DNA and *Ds* insertion lines, we isolated >500 mutants with a seedling-lethal phenotype as the first step in this process. Many of these mutants are likely to be required for chloroplast function on the basis of phenotype and sequence data. In addition, mutant phenotypes, such as elongated or reduced hypocotyls, may be due to the disruption of essential genes in signal transduction pathways. In light of the large-scale nature of the project, we took alternative high-throughput approaches, enabling us to focus on the identification of DNA sequences for a large number of these genes. From cosegregation analysis of >200 mutants, we were able to directly compare the frequency of "tagged" mutants in T-DNA and *Ac/Ds* mutageneses. Finally, we present sequence information for an initial set of genes that are essential for seedling growth and development.

## MATERIALS AND METHODS

**Arabidopsis insertional mutant collections:** All *Ds* lines were generated according to SUNDARESAN *et al.* (1995) and generously provided by R. Martienssen (Cold Spring Harbor Laboratories), H. Ma (Pennsylvania State University), and U. Grossniklaus (Zurich University). T-DNA lines were generated as described in MCELVER *et al.* (2001). The T-DNA vectors included pPCVICEn4HPT, pSKI015, pCSA104, and pDAP101.

**Arabidopsis seedling screening and growth conditions:** Between 50 and 75 seeds for each line were placed on MS media [4.3 g/liter Murashige and Skoog salts (Life Technologies, Rockville, MD), 8 g/liter Phytagar (Life Technologies)] containing the fungicides Benomyl (5 mg/liter; Sigma, St. Louis) and Maxim (1 mg/liter; Syngenta, Greensboro, NC). Top agar with fungicides, identical to MS media except with 6 g/liter Phytagar, was added to spread out the seeds on each plate. Plates were placed at 4° for 1–7 days to synchronize germination. Seedlings were germinated and grown at 19°–23° under

lights (80–100  $\mu\text{E}/\text{sec}/\text{m}^2$ ) with a 16-hr light and 8-hr dark photoperiod. Seven and 14 days after being moved from 4°, the plates were screened with a dissecting microscope for abnormal seedlings. For a line in which it was uncertain whether the mutant phenotype was lethal, three mutant seedlings were transplanted to soil and grown at 18°–23° with 14.5 hr of light per day. For a given line, the mutant phenotype was considered to be lethal only if all three mutant seedlings died. For a small portion of the lines, it was necessary to rescreen a line by replating sterilized seeds on MS + 2% sucrose media. This medium was used because it enabled us to distinguish more easily between seedling-lethal mutants and sick nonmutant seedlings. It is possible that a few of these lines had mutants that would have been inviable on MS media lacking sucrose, but were viable on media containing sucrose. Because such replated lines were screened only on media with sucrose, they would not have been identified as lethal in our screen and the total number of lethal mutants may be slightly underestimated.

**Cosegregation analysis:** For T-DNA lines, ~75–150 T<sub>2</sub> seeds were sterilized and grown on germination medium (GM; GUYER *et al.* 1998) containing 30 mg/liter hygromycin B for pPCVICEn4HPT lines and 15 mg/liter Basta for pSKI015, pCSA104, and pDAP101 lines. For a small number of lines, only ~40–75 plants were analyzed because there were not enough seeds or there was poor germination. The ratio of resistant to sensitive seedlings (*R:S* ratio) was used to determine the most likely number of insertion loci. If the *R:S* ratio was <6.0, the line most likely had a single insertion locus and 32 resistant seedlings were transplanted to soil. This cutoff was derived empirically and was based on chi-square analysis and a strategy to prevent lines with a single insertion locus being assigned to the two insertion loci category; this analysis may have led to a slight overestimate of the number of single insertion lines. In most cases, siliques were screened for a potential embryo-defective phenotype. If all the resistant plants segregated progeny with an embryo phenotype, the line was considered "tagged." If some, but not all, the resistant plants segregated progeny with an embryo-defective phenotype, the line was considered "not tagged." If no embryo phenotype was detected, seeds were collected from each resistant plant and plated on MS media with fungicides. If a seedling-lethal phenotype was detected among the progeny of each of the resistant plants, the line was considered tagged. For tagged mutants, the number of resistant plants checked for cosegregation of the selectable marker and the lethal phenotype was usually ~30 and ranged from 24 to 59. If the *R:S* ratio was >6.0 and <20, then 17 or more resistant seedlings were usually transplanted to allow the identification in the next generation of a "subfamily" that segregated a single insertion and the seedling lethal phenotype. If an appropriate subfamily was identified, cosegregation analysis proceeded as described for lines with a single insertion locus. If an appropriate subfamily was not identified, the line was not analyzed further (Table 3). If the *R:S* ratio was >20, the line was not usually analyzed because there were likely to be more than two insertion loci. Because of the possibility of having two insertion loci linked to each other, *R:S* ratio data could not definitively determine the number of insertion loci for every line analyzed. For a small number of lines, cosegregation analysis was done according to MCELVER *et al.* (2001).

Experiments were performed similarly for *Ds* lines, except 50 mg/liter kanamycin monosulfate was used. In addition, 20 *Ds* lines segregated only resistant progeny (Table 2), which indicated that each line was homozygous for a *Ds* element and heterozygous for the lethal mutation (Tables 4 and 5). This situation arose because F<sub>4</sub> seeds derived from homozygous F<sub>3</sub> plants were used for some *Ds* lines, while T<sub>2</sub> seeds

derived from hemizygous  $T_1$  plants were used for T-DNA lines (see RESULTS). Among these 20 lines, it is most likely that each has a single *Ds* element that did not cosegregate with the lethal phenotype. It remains possible that a very small number of these lines contained two *Ds* elements and were tagged.

The following calculations provide an estimate of the accuracy of the cosegregation analysis. Although most lines designated as tagged contained an insertion that caused the lethal phenotype, at a low frequency, it is possible that an apparently tagged line had an insertion that was tightly linked to a second mutation with a lethal phenotype but that no recombination was detected between the two mutations. In this case, the line is not truly tagged because the insertion did not disrupt the gene responsible for the lethal phenotype. If no recombinant plants were detected among 30 resistant plants analyzed, the recombination frequency ( $p$ ) between the insertion and a hypothetical, linked second mutation was  $\leq 0.034$  (or 3.4 cM). Thus, a hypothetical, linked mutation would have been within a 6.8-cM interval spanning the insertion. Each line used for cosegregation is estimated to harbor about three mutations on the basis of the cosegregation frequency of 29–33% (Table 3); this estimate implies that on average there are one insertion and two other mutations per line. With  $\sim 550$  cM in the Arabidopsis genome (SCHMIDT 1998), the frequency of a hypothetical, linked mutation being in that 6.8-cM region is 2.5% [ $2 \times (6.8/550)$ ]. Therefore, the designation of a line as tagged based on cosegregation of the lethal phenotype and the selectable marker is likely to be correct for 97.5% of the lines when 30 resistant plants were analyzed.

The cosegregation process, which consists of a single self-cross of a heterozygote in most cases, served as an opportunity for other mutations to segregate away from the seedling-lethal mutations in these lines. Because of the large scale of the experiments, these mutants were not subjected to backcrossing. Based on the cosegregation data (Table 3), there are likely to be only about three to five mutations in most lines, which is much less than in standard Arabidopsis EMS seed mutageneses based on the frequency of embryo-lethal mutants observed (REDEI and KONCZ 1992; McELVER *et al.* 2001). It seems likely that most background mutations would not prevent the detection of a seedling-lethal phenotype because they would be weaker or affect processes later in development.

**Molecular biology:** Arabidopsis genomic DNA was prepared according to REITER *et al.* (1992) or using the Nucleon Phyto-Pure Plant DNA isolation kit (Amersham International, Buckinghamshire, England) or the Puregene DNA isolation kit (Gentra Systems, Minneapolis) as modified by McELVER *et al.* (2001). Other procedures were carried out according to standard methods (AUSUBEL *et al.* 1998).

**Plasmid rescue:** For each pPCVICEn4HPT and pSKI015 T-DNA line with a tagged seedling-lethal mutation, genomic DNA was isolated from tissue collected from either heterozygotes or a mixture of homozygotes, heterozygotes, and wild-type plants. Following Southern blot analysis to determine appropriate restriction enzymes to use for plasmid rescue, genomic DNA was cut with an appropriate restriction enzyme to rescue the right or left border of the T-DNA. The ligated genomic DNA was transformed into *Escherichia coli* cells and ampicillin-resistant colonies were isolated. Plasmid clones from these colonies were analyzed by restriction enzyme digestion and sequenced to determine the location of the insertion in the Arabidopsis genome.

**Thermal asymmetric interlaced PCR:** Thermal asymmetric interlaced (TAIL)-PCR (LIU *et al.* 1995) was performed as modified by McELVER *et al.* (2001). The arbitrary degenerate primers and T-DNA primers used are described in McELVER *et al.* (2001). The *Ds* nested primers used were 5a (5' ACTAGC

TCTACCGTTTCCGTTTCCGTTTAC 3'), 5b (5' TTACCTCG GGTTTCGAAATCGATCGGGATAA 3'), 5c (5' AAATCGGTTA TACGATAACGGTCCGGTACGGGA 3'), 3a (5' GGGTCTTGC GGATCTGAATATATGTTTTTCATGTGTG 3'), 3b (5' TACCG AACAAAATACCGTTCCCGTCCGATTTCGAC 3'), and 3c (5' GGATCGTATCGGTTTTCGATTACCGTATTTATCC 3'). The DNA sequence of PCR products was determined as described by McELVER *et al.* (2001).

**Confirmation of sequences flanking insertions:** Results from plasmid rescue experiments were confirmed either by Southern blot with a probe derived from the flanking genomic DNA or by PCR with one primer in the insertion and the other in the flanking genomic DNA. Results from TAIL-PCR were confirmed by a second PCR reaction with a gene-specific primer and an insertion-specific primer. For four T-DNA lines, results were considered confirmed when the same border sequence was obtained from TAIL-PCR reactions with two or more different arbitrary degenerate primers. Both borders of each insertion were identified and confirmed for all but three of the lines.

**Photography and image processing:** Plants were photographed with a DEI-750 video camera (Optronics Engineering, Goleta, CA) and images were captured with Scion Image (Scion Corporation, Frederick, MD) software. Images were adjusted for brightness, contrast, and color and assembled for figures with Adobe (San Jose, CA) Photoshop (version 5.5).

## RESULTS

**Isolation of seedling-lethal mutants:** To identify mutants with a seedling-lethal phenotype, we screened both T-DNA and *Ds* transposon Arabidopsis mutant collections. The generation of the T-DNA collection has been described (McELVER *et al.* 2001). For the T-DNA lines, we screened  $T_2$  self progeny from a single  $T_1$  parent (McELVER *et al.* 2001), so that we expected to find one-quarter homozygous mutant progeny for a line segregating a recessive mutation. We generally performed our screening by examining the growth of seedlings on MS media without sugar to allow us to identify the greatest number of lethal mutants possible. Addition of sugar to the media might rescue the phenotype of some mutants that show lethality on media without sugar.

From 26,187 independent T-DNA lines, we isolated 407 lines segregating seedling-lethal mutants. Although we used some T-DNA lines generated with an activation tagging vector (pPCVICEn4HPT and pSKI015, GenBank accession no. AF187951; WALDEN *et al.* 1994; WEIGEL *et al.* 2000), we did not observe any dominant or semidominant lethal mutants in this screen. Dominant lethal mutations would presumably die in the  $T_1$  generation so that they would not be present in the  $T_2$  progeny. The frequency of seedling-lethal mutants identified from the activation tagging lines (1.52%; 55/3609) and the frequency from the other T-DNA lines (1.56%; 352/22,578) are essentially identical. This result suggests that most of the mutants identified in the activation tagging lines were the result of loss of gene function rather than altered gene function. In addition, molecular analysis of the tagged seedling-lethal mutants indicated that most of the lines have a T-DNA disrupting an ORF (see be-

low). When an ORF is disrupted by a T-DNA, it is most likely that the protein function is disrupted rather than being misexpressed.

To make a direct comparison between T-DNA lines and *Ds* transposon lines, we also screened a large collection of *Ds* lines (SUNDARESAN *et al.* 1995). This collection was generated by crossing parental *Ac* and *Ds* lines and selecting for F<sub>2</sub> progeny with transposition events unlinked to the original *Ac* and *Ds* elements. This approach provided the opportunity to identify insertions into most of the genome (PARINOV *et al.* 1999). F<sub>3</sub> seeds from each F<sub>2</sub> plant were collected individually and a stock of F<sub>4</sub> seeds was sometimes created by collecting seeds from kanamycin-resistant F<sub>3</sub> progeny. Using the same seedling screening protocol as for the T-DNA lines, we screened F<sub>3</sub> or F<sub>4</sub> progeny from 12,196 lines and identified 98 lines with a seedling-lethal mutant phenotype. Unlike the T-DNA lines, most of the *Ds* lines in the collection were not screened for an embryo-defective phenotype.

**Phenotypic classes:** The seedling-lethal mutants displayed a wide range of phenotypes, which were classified as affecting pigmentation and/or morphology (Figures 1 and 2). Among the T-DNA and *Ds* insertion mutants, the frequencies of pigmentation (81% *vs.* 79%), pigmentation and morphology (11% *vs.* 12%), and morphology (8% *vs.* 9%) mutants were nearly identical (Table 1). This distribution of mutants seems to differ from that obtained in another large-scale seedling mutant screen that found only 50% of seedling mutants had defects in pigmentation but not morphology (JURGENS *et al.* 1991). This difference might stem from the other screen having used a different classification scheme or not having been limited to lethal mutants. The frequency of pigmentation subclasses in our study was also comparable between the two mutant populations (Table 1). The albino, yellow, and pale green mutants (Table 1 and Figure 1, A, B, and E–H), which were the majority of the mutants, included a range of pigmentation phenotypes and the assignment of mutants to these subclasses was based on visual inspection. We isolated 12 mutants exhibiting an albino phenotype on media without sucrose and a striking purple-tinted (“fusca”) phenotype superimposed on the albino phenotype on media containing sucrose (Figure 1, C and D). Because only a subset of mutants were grown on both types of media, there are likely to be more mutants in this subclass that have been classified in this study as albino. About 1 week after germination on media containing sucrose, the purple coloration begins to fade and the seedlings gradually appear more like typical albinos. Wild-type *Arabidopsis* seedlings grown on media with sucrose exhibit a much milder version of this purple phenotype, which is due to anthocyanin accumulation, particularly in the hypocotyl at its junction with the cotyledons and along the edges of the cotyledons (KUBASEK *et al.* 1992). Two mutants had a distinctive phenotype

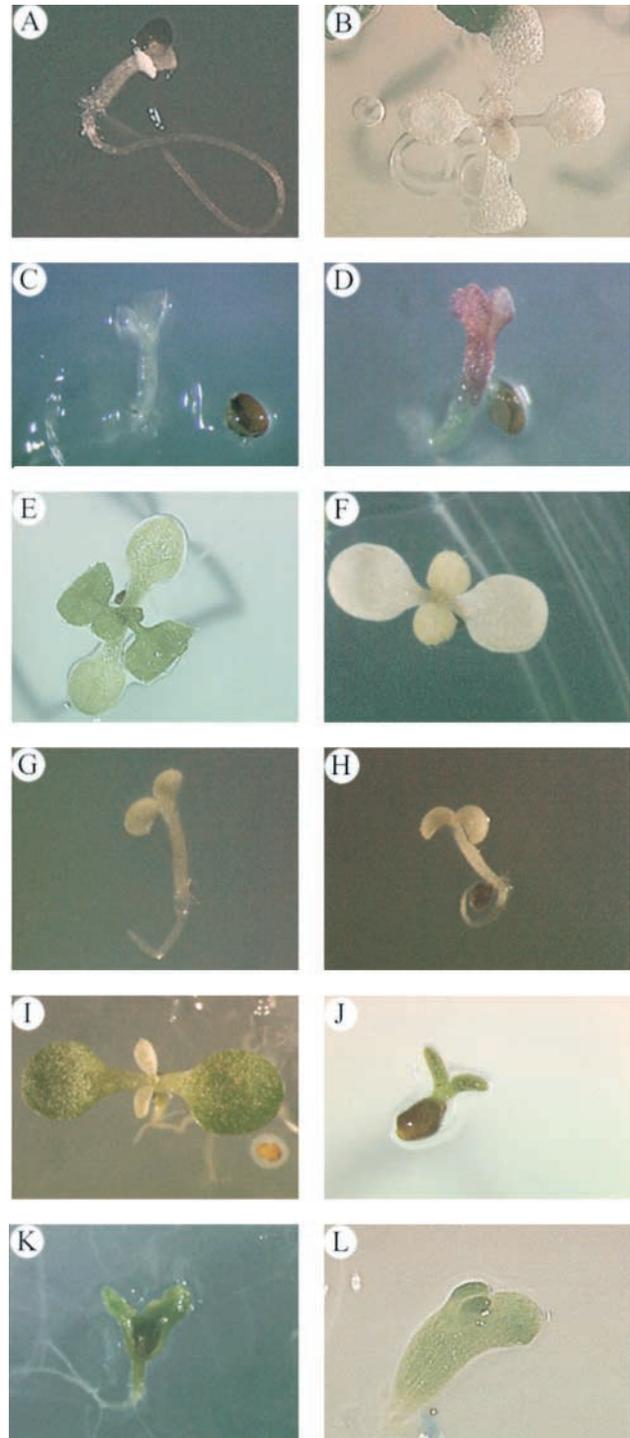


FIGURE 1.—Seedling-lethal pigmentation phenotypes. (A) GT0946, albino, no leaves (day 15). (B) 4036, albino, with leaves (day 14). (C) 4788, albino (day 7). (D) 4788, fusca (day 7). (E) 245, pale green (day 14). (F) GT6839, yellow leaves, albino cotyledons (day 12). (G) GT1209, yellow, small cotyledons, reduced root growth (day 14). (H) GT0992, yellow, small cotyledons, reduced root growth (day 14). (I) 5007, white leaves, green cotyledons (day 14). (J) 2973, fusca cotyledons with purple dots (day 18). (K) 22084, dark green, cotyledons not separated (day 14). (L) 22433, dark green, variable cotyledon number and size, short thick hypocotyl, little root growth (day 13). All seedlings were grown on MS media, except B and E were supplemented with 5% sucrose and D, F, and K were supplemented with 2% sucrose.

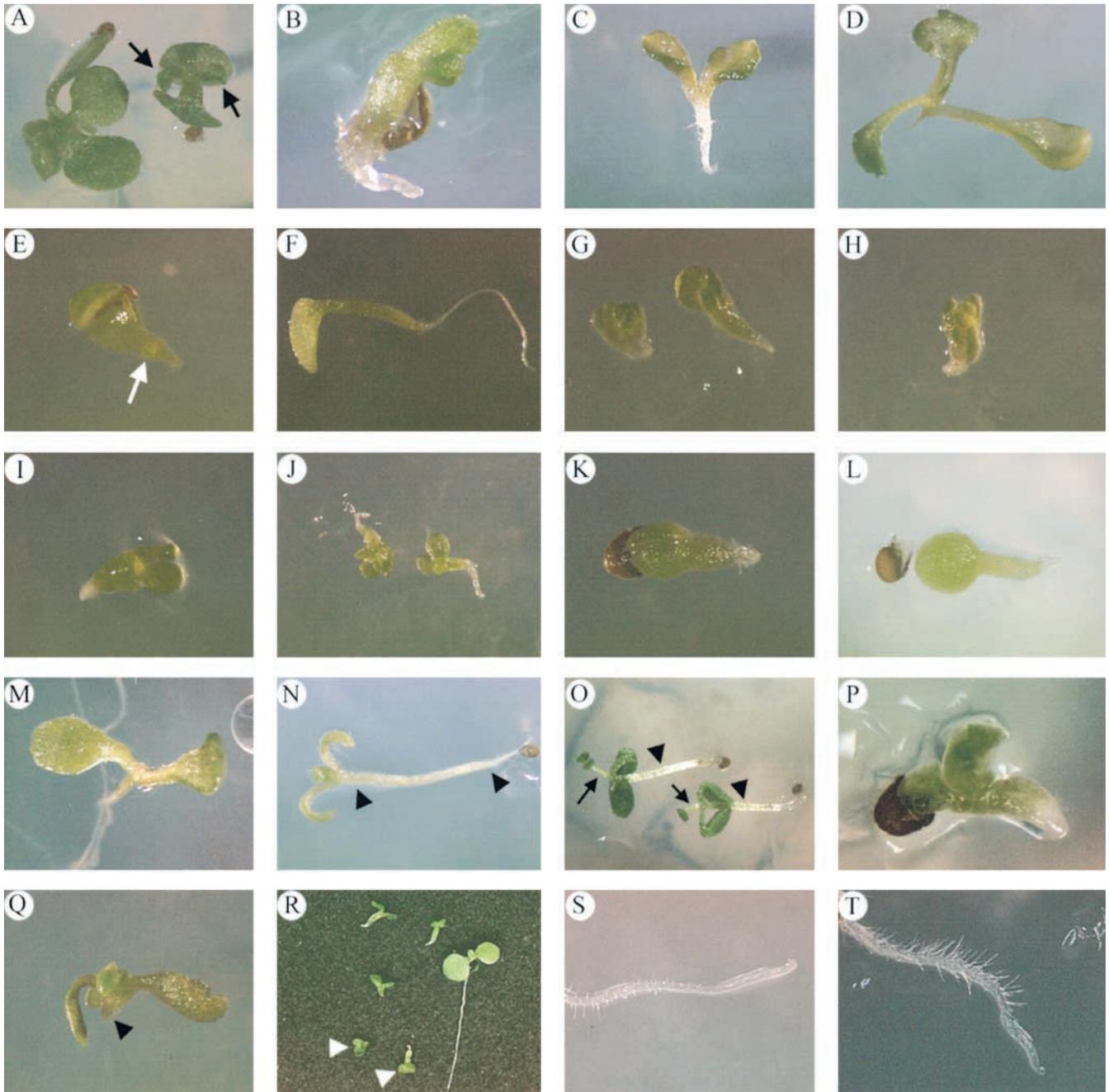


FIGURE 2.—Seedling-lethal morphological phenotypes. (A) 3963, (right) small leaves with irregular margins (arrows) and a wild-type sibling (left). (B) 44446, no leaves, three small cotyledons, thick hypocotyl, reduced root growth. (C) 59928, no leaves, short thick hypocotyl. (D) 55582, single concave cotyledon, reduced hypocotyl, very little root growth (not visible). (E) 59438, single cotyledon, two small leaves (arrow), reduced hypocotyl, no root. (F) 59930, single cotyledon. (G) ET4386, two seedlings with variable cotyledon number, disrupted phyllotaxy, little root growth. (H) 58972, stubby cotyledons, variable cotyledon number, very little root growth. (I) 47091, variably shaped cotyledons, possibly four cotyledons, reduced hypocotyl, very little root growth. (J) 59270, two seedlings with three or four cotyledons, short and thick hypocotyl, little root growth. (K) 59095, no leaves, small closed cotyledons, short and thick hypocotyl, very little root growth. (L) ET5262, no leaves, small closed cotyledons, little root growth. (M) 58424, no leaves. (N) 57348, elongated hypocotyl (arrowheads), small leaves, pale green. (O) 46153, two seedlings with an elongated hypocotyl (arrowheads), elongated leaf petioles (arrows). (P) GT5602, small cotyledons, reduced hypocotyl, very little root growth. (Q) 54196, no root (arrowhead), short hypocotyl. (R) 5283, seedlings with very little root growth compared with wild-type (Col-0) seedling, variable cotyledon defect (arrowheads). (S) Col-0, root with small root hairs. (T) 5283, root with decreased overall length and longer root hairs compared to S. All seedlings were grown on MS media, except C, D, M, N, and R were supplemented with 2% sucrose and S and T were grown on GM. Seedlings in R, S, and T were grown on plates placed at a slant so that the roots would grow on the agar surface.

**TABLE 1**  
Seedling-lethal phenotypic classes

Phenotypic class <sup>b</sup>	No. of lines (%) <sup>a</sup>	
	T-DNA	<i>Ds</i> transposon
Pigmentation <sup>c</sup>	375	89
Albino	131 (32)	29 (30)
Yellow or pale green <sup>d</sup>	216 (53)	52 (53)
Dark green	9 (2)	4 (4)
Fusca <sup>e</sup>	3 (1)	2 (2)
White true leaves <sup>f</sup>	2 (0.5)	0 (0)
Other <sup>g</sup>	14 (3)	2 (2)
Morphology only	32 (8)	9 (9)

<sup>a</sup> Percentages do not add to 100 because of rounding. Pigmentation mutant percentages are only shown for subclasses.

<sup>b</sup> Classification as a morphological mutant is limited to mutants with abnormalities other than smaller organs. Mutants with no roots or very little roots are included, but those with some root growth (greater than ~3 mm) are not.

<sup>c</sup> Subclasses of pigmentation mutants are listed in indented rows below. They include 44 T-DNA and 12 *Ds* mutants with both pigmentation and morphology defects.

<sup>d</sup> Includes pale green then albino, yellow/green, yellow/brown, pale yellow, and yellow then albino.

<sup>e</sup> Fusca (purple) includes dark green/purple.

<sup>f</sup> White true leaves and green cotyledons.

<sup>g</sup> Includes three other subclasses: albino without sucrose and fusca with sucrose in media, green without sucrose and fusca with sucrose, and albino cotyledons with green true leaves.

of green cotyledons and small white leaves (Figure 11). This phenotype has been observed previously in *Arabidopsis* for mutants in the *TZ*, *TH-1*, and *PY* genes, which are likely to encode thiamin biosynthetic genes (LI and REDEI 1969; KOORNNEEF and HANHART 1981). Both of these mutants seem to be thiamin auxotrophs because they appeared normal when grown on media supplemented with 0.1 mM thiamin (data not shown). Eleven of 13 dark green lethal mutants also displayed morphological defects (for example, Figure 1, K and L), which may imply that the dark green phenotype is indicative of a type of defect different from other pigmentation defects. A dark green leaf phenotype has been reported for dwarf mutants with a deficiency in either brassinosteroid or gibberellic acid hormonal pathways and it has been suggested that this defect may be due to smaller cell size (CLOUSE *et al.* 1996; BENNETT *et al.* 1998).

In addition to the pigmentation-defective lethal mutants, mutants were isolated with a wide array of morphological defects (Figure 2). We observed phenotypes ranging from those that seemed to affect only a single structure to others that seemed to affect all seedling structures (leaves, cotyledons, hypocotyl, and roots). Mutant 3963 appeared normal, except for small leaves with irregular margins (Figure 2A). For lethals with defects in cotyledon number, either a single cotyledon (Figure 2, D–F) or multiple cotyledons (Figure 2, B and G–J), the number of cotyledons often varied among the seed-

**TABLE 2**  
Number of insertion loci in seedling-lethal mutants

No. of insertion loci <sup>a</sup>	T-DNA (%) <sup>b</sup>	<i>Ds</i> transposon (%) <sup>b</sup>
Single locus	90 (53)	72 (94)
Two loci	35 (21)	5 (6)
More than two loci	44 (26)	0 (0)
Not determined	238 <sup>c</sup>	21 <sup>d</sup>

<sup>a</sup> Number of loci was determined from the ratio of resistant to sensitive (*R:S*) seedlings on appropriate selective media. Single locus, *R:S* < 6.0; two loci, 6.0 < *R:S* < 20; more than two loci, *R:S* > 20.

<sup>b</sup> Number of lines in each category is listed. Percentages do not include the “not determined” lines.

<sup>c</sup> Includes 2 lines with all sensitive progeny.

<sup>d</sup> Includes 20 lines with all resistant progeny that are likely to be homozygous for the *Ds* element.

lings of a given mutant line. More than 40 mutants exhibited a short, thick, or reduced hypocotyl (Figure 2, B–D and H). Four mutants displayed an elongated hypocotyl (Figure 2, N and O). Previously isolated elongated hypocotyl mutants, which affected gibberellic acid or light signal transduction, were viable (JACOBSEN and OLSZEWSKI 1993; BRIGGS and HUALA 1999; REED *et al.* 2000), so that these essential genes might define novel components in these pathways. Ethylene and auxin are also known to control hypocotyl elongation (COLLETT *et al.* 2000). More than 30 mutants had very little or no root growth (Figure 2, P–R). For mutant 5283, although we saw increased root growth on media with sucrose (Figure 2T) compared to media without sucrose (Figure 2R), there was still considerably less root elongation than in wild type (Figure 2S). Several genetic screens have identified mutants that lack roots or have reduced root systems, but the number of *Arabidopsis* genes with this mutant phenotype remains unclear (MAYER *et al.* 1991; CHENG *et al.* 1995; BERLETH *et al.* 1996; SCHERES *et al.* 1996).

Because many of the seedling-lethal mutants exhibited defects soon after germination, these mutants were also examined for phenotypes during embryogenesis. For the T-DNA lines, 191 of 407 seedling-lethal mutants had a detectable embryo-defective phenotype (data not shown; McELVER *et al.* 2001). For the *Ds* lines, 35 of 72 examined seedling-lethal mutants had a detectable embryo-defective phenotype (data not shown). Generally, these embryo phenotypes appeared late in embryogenesis, for example, as pale mature or albino embryos.

**Genetic analysis of mutants:** Because *Arabidopsis* insertional mutants also contain noninsertional mutations, it was necessary to determine whether an insertion cosegregated with the seedling-lethal phenotype. If the insertion cosegregated with the lethal phenotype, the mutant was considered tagged, but if the insertion did not cosegregate with the lethal phenotype, the line was

**TABLE 3**  
Cosegregation analysis

Tagged? <sup>a</sup>	T-DNA (%) <sup>b</sup>	<i>Ds</i> transposon (%) <sup>b</sup>
Yes	32 (29)	32 (33)
No	80 (71)	66 (67)
Not determined (two insert loci) <sup>c</sup>	7	0
Not determined (high insert no.) <sup>d</sup>	41	0

<sup>a</sup> "Tagged" implies that the lethal phenotype and the selectable marker always cosegregated together.

<sup>b</sup> Number of lines in each category is listed. Percentages do not include the "not determined" lines.

<sup>c</sup> No subfamily with a low *R:S* ratio and segregating the mutant phenotype was identified.

<sup>d</sup> Lines with *R:S* ratios >20 were usually not analyzed for cosegregation.

considered not tagged and a noninsertional mutation was likely to be the cause of the lethal phenotype. From the initial *R:S* ratio in cosegregation analysis, the number of insertion loci in a line was determined (see MATERIALS AND METHODS; Table 2). Based on the segregation of the selectable markers, 54% of the T-DNA lines had a single insertion locus, while 94% of the *Ds* lines had a single insertion locus. The higher number of insertion loci per line in the T-DNA lines probably explains the higher frequency of seedling-lethal mutants isolated from the T-DNA population compared to the *Ds* population. From the cosegregation analysis, we identified 32 T-DNA and 32 *Ds* lines as tagged (Table 3). The frequency of tagged lines in both populations was about one-third.

We refer to the mutants by line number and have not named the corresponding genes because the large number of mutants isolated in this study made the use of complementation tests or genetic mapping inefficient for this purpose. Instead, we used the molecular position of an insertion in the Arabidopsis genome to identify mutants with disruptions of the same gene.

**Molecular analysis of mutants:** For each tagged seedling-lethal mutant line identified, we attempted to identify the DNA sequence of the gene or genes disrupted by the insertion. Table 4 shows a summary of this molecular analysis. Initially to isolate Arabidopsis genomic DNA sequences adjacent to T-DNA insertions, we used a plasmid rescue approach (HOLSTERS *et al.* 1982) and obtained flanking sequences for 12 of 13 lines. Although plasmid rescue was an effective tool to identify genomic sequence flanking insertions, we subsequently switched to TAIL-PCR (LIU *et al.* 1995) because it allowed a higher throughput. We used TAIL-PCR experiments to identify flanking sequences for 15 of 19 T-DNA lines. We used only TAIL-PCR for *Ds* insertions, which lack the sequence elements necessary for plasmid rescue, and obtained flanking sequences for 31 of 32 lines. BLASTn (ALTSCHUL *et al.* 1997) was used to identify the insertion position in Arabidopsis genomic sequence entries in GenBank. Additional BLAST searches identified genes with sequence similarity from Arabidopsis and other species. Results from plasmid rescue and TAIL-PCR ex-

periments were confirmed either by Southern blot or by PCR (see MATERIALS AND METHODS). For *Ds* lines ET2614 and GT2929, we were unable to clearly identify the essential gene due to a *Ds* insertion into the *indole acetic acid hydrolase (iaaH)* negative selectable marker, which might have been linked to the lethal mutation (SUNDARESAN *et al.* 1995; PARINOV *et al.* 1999). For 10 of 64 mutants, no confirmed Arabidopsis genomic sequence flanking an insertion was recovered (Table 4). For 15 of the remaining 54 mutants, it was not possible to easily assign the gene that was responsible for the lethal phenotype because an insertion was between two genes or it was accompanied by a deletion or rearrangement affecting more than one gene (Table 4). Chromosomal rearrangements have been reported previously for T-DNA lines (CASTLE *et al.* 1993; NACRY *et al.* 1996). Additional analysis of 30 of 37 genes disrupted in these seedling-lethal mutants included the identification, either experimentally or from GenBank, of a cDNA clone that contained the entire protein-coding sequence (G. J. BUDZISZEWSKI and J. Z. LEVIN, unpublished results).

Molecular analysis of the tagged seedling-lethal mu-

**TABLE 4**  
Molecular analysis of tagged lines

Analysis category	T-DNA	<i>Ds</i> transposon
Tagged mutants analyzed	32	32
No flanking sequence obtained <sup>a</sup>	5	0
Insertion in <i>iaaH</i> gene	0	2 <sup>c</sup>
No flanking sequence confirmed <sup>b</sup>	2	1 <sup>c</sup>
Essential gene not identified <sup>d</sup>	11	4
Essential gene identified	14	25 <sup>e</sup>

<sup>a</sup> Results from plasmid rescue and/or TAIL-PCR.

<sup>b</sup> Lines were not analyzed further.

<sup>c</sup> Line ET2614 is counted only in the insertion in the *iaaH* gene category, but it also had another TAIL-PCR sequence that could not be confirmed by PCR.

<sup>d</sup> Lines with possible rearrangements, multigene deletions, and insertions between two genes in which the essential gene was not identified.

<sup>e</sup> Three of the 25 *Ds* mutant lines disrupt the same gene.

**TABLE 5**  
**Sequence data for lines with a seedling-lethal phenotype**

Line <sup>a</sup>	Seedling phenotype	Nucleotide accession no. <sup>b</sup>	Insertion sites <sup>c</sup>	Chr. <sup>d</sup>	Protein function (gene) <sup>e</sup>
245	Pale green	AL137080	6600	3	Peptide chain release factor 2 ( <i>E. coli</i> RF-2)
868	Albino	AL161515	93008, 134144 <sup>f</sup>	4	N/A
1055	Albino	AL161542	9579, 9637	4	<i>CLA1</i> , D-1-deoxyxylulose-5-phosphate synthase <sup>g</sup>
2490	Pale green	AB008270	37913	5	Cp protein import ( <i>B. napus</i> Toc36) <sup>h</sup>
3963	Abnormal leaves, small	AB010695	37129	5	DNA repair nuclease ( <i>S. cerevisiae</i> <i>MRE11</i> )
4036	Albino	AB009053	31380, 31412	5	D-1-Deoxyxylulose-5-phosphate reductoisomerase
4144	Pale green	AL161515	118366	4	Cp ATP synthase $\delta$ -subunit (tobacco <i>atpD</i> ) <sup>h</sup>
4788	Albino; fusca (sucrose) <sup>i</sup>	AC002505	60451	2	Pyrimidine permease (maize <i>Lpe1</i> )
5283	Reduced root	AC004473	32858	1	RNA splicing ( <i>S. cerevisiae</i> <i>PRP31</i> )
45928	Surface blotchy and rough	AC025416	33063, 33066	1	<i>KEULE</i> , vesicle trafficking <sup>g</sup>
ET1158	Albino	AC005936	5647, 5655	2	Zn finger (potato PCP1)
ET4401	Albino	AC002062	58754	1	Unknown function <sup>f</sup>
		AC010796	5642	1	Between two genes <sup>f</sup>
ET5262	No root or leaves	AC026238	1505, 1513	1	$\beta$ -1,3-Endoglucanase-like (wheat E13B)
ET5745	Fusca	AL161516	172038	4	<i>DET1</i> <sup>g</sup>
GT0946	Albino	AC004136	21164, 21173	2	4-Diphosphocytidyl-2C-methyl-D-erythritol synthase
GT1209	Pale yellow	AC007592	33100	1	Unknown function
GT1354	Albino	AC006533	54823, 54823	2	Unknown function
GT1802	Yellow	AL161496	91363, 91370	4	Cytochrome <i>b<sub>6</sub>-f</i> complex iron-sulfur subunit (spinach)
GT1998	Albino	AC006072	42110	2	<i>PAC</i> , Cp mRNA maturation? <sup>g,h</sup>
GT6839	Pale yellow, then albino	AC006837	84110, 84111	2	<i>tatC</i> , Cp protein import <sup>h</sup>
GT8096	Yellow, then albino	AC006837	83652, 83659	2	<i>tatC</i> , Cp protein import <sup>h</sup>
ET7536	Yellow	AC006837	84166, 84172	2	<i>tatC</i> , Cp protein import <sup>h</sup>

<sup>a</sup> *Ds* lines start with "ET" or "GT." Other lines are T-DNA.

<sup>b</sup> Arabidopsis genomic sequence (GenBank).

<sup>c</sup> Position of insertion within nucleotide accession. If only one site is listed, the sequence flanking the other border was not determined precisely. Small deletions often accompanied an insertion. For the *Ds* lines, an 8-bp duplication was often observed.

<sup>d</sup> The chromosome location based on sequence identity with a genomic sequence entry in GenBank.

<sup>e</sup> Protein functions and similar genes assigned on the basis of BLAST searches. Characterized genes not from Arabidopsis list the organism in parentheses. Expect values from BLASTP searches for protein reported were  $<10^{-17}$ .

<sup>f</sup> Line appears to have a deletion or rearrangement affecting more than one gene.

<sup>g</sup> Arabidopsis mutants are known for *CLA1* (MANDEL *et al.* 1996), *KEULE* (ASSAAD *et al.* 2001), *DET1* (PEPPER *et al.* 1994), and *PAC* (REITER *et al.* 1994).

<sup>h</sup> Cp, chloroplast.

<sup>i</sup> Albino phenotype without sucrose in media; fusca phenotype with sucrose in media.

tants revealed that a diverse set of genes is essential for seedling viability. Table 5 shows the locations of insertions within Arabidopsis genomic clones and the identities of genes disrupted in 20 of the 39 mutants for which the gene responsible could be deduced (Table 5). A detailed molecular characterization of the remaining 19 mutants will be presented in a future article. For two other lines, 868 and ET4401, the gene responsible for the seedling-lethal phenotype could not be identified (Table 5). Line 868 appears to contain a rearrangement or deletion that spans a region including the gene disrupted in line 4144, so that the lethal phenotypes of these two lines might be due to the inactivation of the same gene. Among those genes identified in Table 5 are four that were previously shown to have

seedling-lethal phenotypes: *DET1* (PEPPER *et al.* 1994), *CLA1* (MANDEL *et al.* 1996), *KEULE* (ASSAAD *et al.* 2001), and *PALE CRESS* (*PAC*; REITER *et al.* 1994), which may have a role in chloroplast mRNA maturation (MEURER *et al.* 1998). Disruption of the translational apparatus resulted in pigmentation-defective lethal phenotypes in line 245. Line 245 had a defect in a putative peptide chain release factor, which was predicted by TargetP (EMANUELSSON *et al.* 2000) to be localized in the mitochondria. Disruption of the photosynthetic apparatus also resulted in pigmentation-defective lethal phenotypes in lines 4144, with a defect in a putative chloroplast ATP synthase  $\delta$ -subunit, and GT1802, with a defect in a putative cytochrome *b<sub>6</sub>-f* complex iron-sulfur subunit. Antisense experiments with two similar genes in tobacco

resulted in plants with extremely slow growth due to decreases in photosynthesis (PRICE *et al.* 1995). In contrast to these four lines, the identification of a putative RNA splicing gene as defective in line 5283, which exhibited reduced root growth and variable defects in the number and shape of cotyledons and leaves (Figure 1, R and T), did not provide a clear explanation of the defect in this mutant; however, these data can be used as a starting point in future studies of this mutant. Consistent with the mutant phenotype, the mRNA for the gene disrupted in line 5283 is detected by Northern blot analysis in both roots and aboveground seedling tissues (data not shown). In summary, the predicted roles for the identified essential genes (Table 5) indicate that a diverse set of pathways and processes within the plant contain an essential component.

**Nonmevalonate isoprenoid pathway mutants:** Three of the genes identified in this study disrupt the recently discovered nonmevalonate isoprenoid pathway. In plants, two independent pathways are responsible for the synthesis of isoprenoids: a cytosolic acetate/mevalonate pathway and a plastidic nonmevalonate 1-deoxy-D-xylulose-5-phosphate pathway (reviewed in LICHTENTHALER 1999 and LICHTENTHALER *et al.* 2000). Initial studies in *Scenedesmus obliquus* suggested that such a pathway existed, but it was unclear whether there was redundancy between the two pathways (SCHWENDER *et al.* 1996). Subsequently, the genes involved in this pathway have been identified and characterized in *E. coli*. We identified albino seedling-lethal mutants disrupting genes encoding the first three enzymes in this pathway. This phenotype was likely due to a block in the formation of carotenoids, phytol side chains of chlorophylls, and plastoquinone-9, which are products of this pathway. The first enzyme, 1-deoxy-D-xylulose-5-phosphate synthase (DXS), converts pyruvate and glyceraldehyde 3-phosphate to 1-deoxy-D-xylulose-5-phosphate (LOIS *et al.* 1998). Line 1055 disrupts the Arabidopsis gene encoding DXS (Table 5), which has been previously identified as *CLA1* on the basis of its albino mutant phenotype (MANDEL *et al.* 1996). The second enzyme, 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR), converts 1-deoxy-D-xylulose-5-phosphate to 2-C-methyl-D-erythritol 4-phosphate (TAKAHASHI *et al.* 1998) and the Arabidopsis homolog has been cloned and characterized (SCHWENDER *et al.* 1999). Line 4036 (Figure 1B) shows that the Arabidopsis homolog is an essential gene. The third enzyme, 4-diphosphocytidyl-2C-methylerythritol synthase, converts 2-C-methyl-D-erythritol 4-phosphate and CTP to 4-diphosphocytidyl-2C-methyl-D-erythritol (ROHDICH *et al.* 1999) and the Arabidopsis homolog has been cloned and characterized (ROHDICH *et al.* 2000). Line GT0946 (Figure 1A) shows that the Arabidopsis homolog is also an essential gene. Interestingly, the enzymes in this pathway are not found in animals and have been proposed to be novel targets for herbicides

and antibacterial drugs that are based on the compound fosmidomycin (ROHMER 1998).

**Chloroplast protein translocation:** Two of the mutants identified in this study may disrupt the translocation of nuclear-encoded proteins into the chloroplast. Most of these proteins are imported by protein complexes composed of Toc (translocons of the outer envelope of chloroplasts) and Tic (translocons of the inner envelope of chloroplasts) proteins (SCHLEIFF and SOLL 2000). After import into the chloroplast, four pathways are proposed to be involved in the translocation of proteins into or across the thylakoid membrane (ROBINSON *et al.* 2001). Membrane proteins are translocated by an SRP-dependent or spontaneous pathway. Lumen proteins are translocated by a Sec-dependent or  $\Delta$ pH-dependent pathway. Line 2490, which had a pale green phenotype, contained a disruption in a gene similar to the pea Tic40 (STAHL *et al.* 1999) and *Brassica napus* Toc36 genes (KO *et al.* 1995; Table 5). Tic40 has been shown to be an inner chloroplast envelope-localized protein acting in protein translocation and displays some sequence similarity to Hsp70 interacting proteins, but its exact role remains unclear. Although some Arabidopsis chloroplast protein translocation mutants, *e.g.*, *ppi2* (BAUER *et al.* 2000) and 2490, display lethal phenotypes, others such as *ppi1* (JARVIS *et al.* 1998), *ffc*, and *chaos* (AMIN *et al.* 1999) have a pale nonlethal phenotype that might be a result of partial functional redundancy among import pathways. Lines GT6839, GT8096, and ET7536, which had a yellow phenotype (Figure 1F), contained a disruption in the Arabidopsis homolog of the *E. coli* *tatC* gene (Table 5; BOGSCH *et al.* 1998; MORI *et al.* 1999). The pea *tatC* protein has been shown to be required for the thylakoid  $\Delta$ pH-dependent pathway *in vitro* (MORI *et al.* 2001). Maize mutants in the *tha4* and *hcf106* genes, which are similar to *E. coli* *tatA* and *tatB*, disrupt this pathway and have a seedling-lethal phenotype (SETTLES *et al.* 1997; WALKER *et al.* 1999).

**Predicted localization of essential seedling proteins:**

In light of the large fraction of seedling-lethal mutants with pigmentation defects (Table 1), we attempted to determine whether these mutants had molecular defects in chloroplast function. We used the TargetP program (EMANUELSSON *et al.* 2000) to predict whether the genes identified in this study encode proteins containing chloroplast transit peptides (CTPs) at their N termini. TargetP is reported to be the most accurate method for predicting the presence of CTPs and is estimated to be correct for 85% of plant proteins. In most cases, the coding region for these proteins was derived from full-length cDNA sequences (data not shown). Nine of 13 genes from mutants with only a pigmentation phenotype in Table 5 were predicted to contain a CTP. Analysis of a larger set of 30 essential genes identified from mutants with only a pigmentation phenotype in this study indicated that 21 genes were predicted to contain a CTP (J. Z. LEVIN, unpublished results). None of the

four genes from mutants with morphological phenotypes in Table 5 was predicted to contain a CTP (data not shown). Analysis of three additional genes from mutants with morphological phenotypes in this study predicted these genes also would not contain a CTP (J. Z. LEVIN, unpublished results). Among all Arabidopsis proteins, ~14% are predicted by TargetP to have CTPs (ARABIDOPSIS GENOME INITIATIVE 2000). These results show a significant enrichment of chloroplast proteins in the pigmentation-defective mutant class.

**Gene family membership:** In light of the high frequency of genes belonging to gene families in Arabidopsis (ARABIDOPSIS GENOME INITIATIVE 2000), we analyzed the genes identified in this study as essential for seedling viability for their membership in gene families. In an analysis of the entire genome, 65% of Arabidopsis genes were considered to be members of gene families on the basis of BLAST (ALTSCHUL *et al.* 1997) and FASTA (PEARSON and LIPMAN 1988) analyses (ARABIDOPSIS GENOME INITIATIVE 2000). Using the same criteria, we determined that 44% of the 18 essential genes identified in Table 5 were members of gene families. Three of these essential genes are in gene families with two members; 1 is in a gene family with three members; 1 is in a gene family with four members; and 3 are in gene families with more than five members. This distribution is similar to that found for the entire genome (ARABIDOPSIS GENOME INITIATIVE 2000). Analysis of a larger set of 37 essential genes identified from mutants in this study indicated that 41% are members of gene families (J. Z. LEVIN, unpublished results).

## DISCUSSION

With the determination of the genome sequence of Arabidopsis complete, the challenge for plant biologists is to understand the function of every gene. It is estimated that there are ~25,500 genes in Arabidopsis (ARABIDOPSIS GENOME INITIATIVE 2000), but only a small fraction of them have been characterized. Determination of the function of all Arabidopsis genes is a goal of the Arabidopsis research community before the year 2010 (CHORY *et al.* 2000). One approach to this problem is to generate large insertional mutant collections and to use them to identify mutants in each of the predicted genes (reviewed in PARINOV and SUNDARESAN 2000). With the aim of identifying the genes necessary for seedling viability, we isolated and molecularly characterized Arabidopsis seedling-lethal mutants on a genome-wide scale. Using both T-DNA and *Ds* insertion mutant populations, we isolated >500 seedling-lethal mutants and molecularly characterized >50 genes disrupted in these mutants. Among these essential genes are some that might be considered "housekeeping genes" such as a peptide release factor in line 245 and an ATP synthase subunit in line 4144 and others that might be classified as "developmental regulatory genes" such as *DET1* (PEP-

PER *et al.* 1994) in line ET5745 (Table 5). A high proportion of seedling-lethal mutants with pigmentation defects are likely to affect nuclear-encoded chloroplast proteins. Seedling development seems to depend primarily on chloroplast function because of the need for energy. This hypothesis is supported by the ability of sucrose to bypass the need for light in dark-grown wild-type Arabidopsis seedlings that were able to flower when grown on vertical petri dishes with sucrose-containing media (ROLDAN *et al.* 1999). Although many complex processes involve the chloroplast during seedling development, such as coordination of plastid and nuclear gene expression (SUSEK *et al.* 1993), it is unclear whether these functions have an essential component beyond the need for energy.

Beyond the usefulness of the identification of a loss-of-function phenotype for a gene, each mutant can be a starting point for future detailed characterization of specific cellular and developmental processes. This collection of seedling-lethal mutants could be a resource for such experiments. The application of other functional genomic methods, such as mRNA or metabolite profiling, to these seedling-lethal mutants may yield a greater understanding of the roles these genes play in plant growth and development. In particular, mutants disrupting the nonmevalonate isoprenoid pathway (Table 5) could shed light on the regulation of these genes and metabolites within this newly discovered biosynthetic pathway.

Seedling-lethal mutants have been identified previously in other types of genetic screens, including those for pigmentation defects. High chlorophyll fluorescence (*hcf*) occurs in plants when there is a reduction in photosynthetic activity beyond photosystem II and it can be visually detected as red plants in response to UV irradiation. In a screen for *hcf* mutants, 23 of the 34 Arabidopsis mutants identified were also seedling lethals (MEURER *et al.* 1996). Although it is unresolved how many genes can mutate to a *hcf* phenotype, there is clearly an overlap between mutants that could be found in a *hcf* screen and those identified in this study as seedling lethal. Only 5 of the 130 maize *hcf* mutants identified were allelic with each other, indicating that this is a large class of genes (MILES 1994). Molecular analysis of a limited number of maize *hcf* genes has revealed chloroplast proteins acting in protein translocation (SETTLES *et al.* 1997), mRNA processing and translation (FISK *et al.* 1999), and translation (SCHULTES *et al.* 2000). Genes identified in our study play roles in these processes as well (Table 5). A screen for chlorophyll-deficient *xantha* Arabidopsis mutants identified many mutants but focused on seven genes specifically affecting chlorophyll synthesis or integration into the photosynthetic membrane (RUNGE *et al.* 1995).

Multiple large-scale duplication events during the last 200 million years have been proposed to explain the extensive duplication within the Arabidopsis genome

(VISION *et al.* 2000). As a result of these duplications, Arabidopsis is reported to have 65% of its genes within gene families compared to only 29% for *S. cerevisiae*, 28% for *D. melanogaster*, and 45% for *C. elegans* (ARABIDOPSIS GENOME INITIATIVE 2000). Members of gene families can exhibit functional redundancy, depending on the extent of divergence of function by changes in their coding sequences and/or expression patterns. It will be valuable to determine the extent of functional redundancy within the entire Arabidopsis genome. An interesting example of partial redundancy within gene families has been reported for the Arabidopsis *CAULIFLOWER* and *APET-ALA1* genes in which double mutants have a dramatic cauliflower-like floral meristem defect, while *cauliflower* single mutants have a wild-type phenotype and *apetala1* single mutants have a milder floral-defective phenotype (KEMPIN *et al.* 1995). For the large R2R3 MYB transcription factor gene family, it appears that there may be considerable functional redundancy as a significant number of genes have no visible single mutant phenotype, although further study may reveal subtle phenotypes (MEISSNER *et al.* 1999). Somewhat surprisingly, 41% of the genes found in this study to be essential for seedling viability are members of gene families. For these genes, no other gene family member can replace the function of the mutated gene. While it is possible that all the members of a gene family have the same function and the lethal phenotype is the result of a decrease in the aggregate level of gene function below a threshold, it seems more likely that there has been a divergence in function for these genes. One other possibility is that members of a gene family encode proteins that function nonredundantly in different cellular compartments, *e.g.*, cytoplasm and chloroplast. These results suggest that many of the members of gene families in Arabidopsis may have nonredundant functions.

Although we isolated >500 seedling lethals, our research program is ongoing and additional effort will be required to establish exactly how many genes can mutate to this phenotype. An estimate of how close we are to saturation in this screen can be made in several ways. First, for a phenotypic class with a known number of genes based on previous saturation mutageneses, we can compare the number of mutants isolated in this study. We detected five mutants with a fusca phenotype (Table 1) and 10 genes could have been detected in a seedling screen (MISERA *et al.* 1994). We detected two seedling-lethal mutants with white leaves and thiamin auxotrophy (Table 1) and 3 genes with this phenotype are known (KOORNNEEF and HANHART 1981). Both of these results suggest that there are >500 genes with a seedling-lethal phenotype. Second, an estimate can be made relative to the *emb* lethals isolated from the same T-DNA population (McELVER *et al.* 2001). McElver and co-workers found that 2.5% of T-DNA lines segregate an *emb* mutant and they estimate that there are 500–750 *emb* genes (McELVER *et al.* 2001). In comparison, 1.6%

of T-DNA lines segregated a seedling-lethal mutant, suggesting that 320–480 genes are in this class. Third, only 1 gene was mutated more than once among the 54 tagged lines characterized molecularly. Thus, the molecular analysis will require additional information on multiple alleles before we can use this criterion to determine the extent of saturation.

As part of our seedling-lethal screen, we performed a large-scale direct comparison of T-DNA and *Ds* transposon insertional mutagenesis methods in Arabidopsis. The spectrum of phenotypes obtained in each screen appears to be similar (Table 1). Many of the differences between the methods result in T-DNA lines being more complex to analyze than *Ds* lines. On average, T-DNA lines had more insertion loci per line than *Ds* lines (Table 2). At a given insertion locus, a T-DNA line often had more than one copy and the insertion frequently was partially rearranged, while the *Ds* elements showed no evidence of rearrangements (data not shown). T-DNA lines were more likely to affect multiple genes or to have an insertion between two predicted genes. We attempted to identify the gene disrupted by an insertion in 32 T-DNA lines and 32 *Ds* lines. Among the 15 tagged lines analyzed in which the essential gene could not be identified for these reasons, only 4 were *Ds* lines (Table 4). We obtained a similar frequency of tagged mutants in both populations (Table 3). This frequency reflects the number of mutations in a line other than insertions carrying the selectable marker. Most of these mutations are likely to be point mutations or partial insertion copies that might be caused by DNA-modifying enzymes involved in T-DNA insertion or transposition. The frequency of 29% for T-DNA mutants is similar to the 36% found by CASTLE *et al.* (1993) and the 34% reported by McELVER *et al.* (2001). Finding only 33% of the *Ds* mutants were tagged was a surprise to us. Previous reports of *Ds* transposon tagging frequency in Arabidopsis suggested a higher rate might be found. However, these estimates were based on analysis of fewer mutants. Because each *Ac/Ds* system was slightly different, it is not possible to definitively determine the causes of the reported tagged-mutant frequency. For the two largest studies reported, 15 of 29 (LONG *et al.* 1997) and 4 of 28 (ALTMANN *et al.* 1995) mutants were tagged. In three smaller studies, 3 of 5 (BANCROFT *et al.* 1993), 3 of 4 (LONG *et al.* 1993), and 2 of 6 (BHATT *et al.* 1996) mutants were tagged. What might be the reason for these differences? It is possible that the low frequency obtained by ALTMANN *et al.* (1995) was due to the absence of selection against the continued presence of the *Ac* element, resulting in additional excision and insertion events that lowered the frequency of tagged mutants. Because the lines in this study (SUNDARESAN *et al.* 1995) and 26 of the lines in LONG *et al.* (1997) were both generated from *Ac* lines expressing transposase under the control of the 35S promoter, the strength of the promoter used to express transposase does not

seem to account for these differences. It is possible that previous studies do not have sufficiently large sample sizes to statistically distinguish them from the results in this study.

To increase the confidence in the assignments of particular genes as responsible for the seedling-lethal phenotype in a given line, additional experiments will be necessary. These assignments are estimated to be correct in almost every case on the basis of the cosegregation results (see MATERIALS AND METHODS). Four of the genes identified here are previously known to be essential (Table 5). Additional efforts could include the isolation of additional alleles, complementation with a wild-type transgene, reversion of *Ds* mutants, and creation of transgenic plants with antisense or dsRNA constructs (WATERHOUSE *et al.* 1998; CHUANG and MEYEROWITZ 2000; LEVIN *et al.* 2000). The most efficient strategy for a large-scale effort would probably be the first alternative.

We acknowledge Joanna Barton, Paul Burt, Parna Chattaraj, Hodan Guled, Karen Maguylo, and Sarah Williams for technical assistance, and Bob Dietrich, Mark Johnson, David Meinke, and Cathy Frye for critical reading of the manuscript. We thank David Meinke, Mary Ann Cushman, Amy Schetter, and Kelsey Smith (all from Oklahoma State University) for assistance with cosegregation analysis. We are grateful to Joseph Simorowski for sending us additional seeds for the *Ds* lines on several occasions. We also thank the Syngenta Biotechnology, Inc., sequencing facility, greenhouse facility, and media kitchen for their excellent assistance.

*Note added in proof:* After the submission of the revised version of this article, a report describing the phenotype of Arabidopsis *tatC* mutants was published (R. MOTOHASHI, N. NAGATA, T. ITO, S. TAKAHASHI, T. HOB0 *et al.*, 2001, An essential role of a *TatC* homologue of a Delta pH-dependent protein transporter in thylakoid membrane formation during chloroplast development in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **98**: 10499–10504).

#### LITERATURE CITED

- ALTMANN, T., G. FELIX, A. JESSOP, A. KAUSCHMANN, U. UWER *et al.*, 1995 *Ac/Ds* transposon mutagenesis in *Arabidopsis thaliana*: mutant spectrum and frequency of *Ds* insertion mutants. Mol. Gen. Genet. **247**: 646–652.
- ALTSCHUL, S. F., T. L. MADDEN, A. A. SCHAFER, J. ZHANG, Z. ZHANG *et al.*, 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**: 3389–3402.
- AMIN, P., D. A. C. SY, M. L. PILGRIM, D. H. PARRY, L. NUSSAUME *et al.*, 1999 Arabidopsis mutants lacking the 43- and 54-kilodalton subunits of the chloroplast signal recognition particle have distinct phenotypes. Plant Physiol. **121**: 61–70.
- ARABIDOPSIS GENOME INITIATIVE, 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408**: 796–815.
- ASSAAD, F. F., Y. HUET, U. MAYER and G. JURGENS, 2001 The cytokinesis gene *KEULE* encodes a Sec1 protein that binds the syntaxin KNOLLE. J. Cell Biol. **152**: 531–543.
- AUSUBEL, F. M., R. BRENT, R. KINGSTON, D. D. MOORE, J. G. SEIDMAN *et al.*, (Editors), 1998 *Current Protocols in Molecular Biology*. John Wiley & Sons, New York.
- BANCROFT, I., J. D. JONES and C. DEAN, 1993 Heterologous transposon tagging of the *DRL1* locus in Arabidopsis. Plant Cell **5**: 631–638.
- BAUER, J., K. CHEN, A. HILTBUNNER, E. WEHRLI, M. EUGSTER *et al.*, 2000 The major protein import receptor of plastids is essential for chloroplast biogenesis. Nature **403**: 203–207.
- BENNETT, M., J. KIEBER, J. GIRAUDAT and P. MORRIS, 1998 Hormone regulated development, pp. 107–150 in *Arabidopsis*, edited by M. ANDERSON and J. ROBERTS. Sheffield Academic Press, Sheffield, England.
- BERLETH, T., C. S. HARDTKE, G. K. H. PRZEMECK and J. MULLER, 1996 Mutational analysis of root initiation in the *Arabidopsis* embryo. Plant Soil **187**: 1–9.
- BHATT, A. M., T. PAGE, E. J. LAWSON, C. LISTER and C. DEAN, 1996 Use of *Ac* as an insertional mutagen in Arabidopsis. Plant J. **9**: 935–945.
- BOGSCH, E. G., F. SARGENT, N. R. STANLEY, B. C. BERKS, C. ROBINSON *et al.*, 1998 An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. J. Biol. Chem. **273**: 18003–18006.
- BOSSY, B., L. M. C. HALL and P. SPIERER, 1984 Genetic activity along 315kb of the *Drosophila* chromosome. EMBO J. **3**: 2537–2541.
- BRENNER, S., 1974 The genetics of *Caenorhabditis elegans*. Genetics **77**: 71–94.
- BRIGGS, W. R., and E. HUALA, 1999 Blue-light photoreceptors in higher plants. Annu. Rev. Cell. Dev. Biol. **15**: 33–62.
- CASTLE, L. A., D. ERRAMPALLI, T. L. ATHERTON, L. H. FRANZMANN, E. S. YOON *et al.*, 1993 Genetic and molecular characterization of embryonic mutants identified following seed transformation in *Arabidopsis*. Mol. Gen. Genet. **241**: 504–514.
- CHENG, J. C., K. A. SEELEY and Z. R. SUNG, 1995 RML1 and RML2, Arabidopsis genes required for cell proliferation at the root tip. Plant Physiol. **107**: 365–376.
- CHORY, J., J. R. ECKER, S. BRIGGS, M. CABOCHE, G. M. CORUZZI *et al.*, 2000 National Science Foundation-sponsored workshop report: “The 2010 Project” functional genomics and the virtual plant. A blueprint for understanding how plants are built and how to improve them. Plant Physiol. **123**: 423–426.
- CHUANG, C. F., and E. M. MEYEROWITZ, 2000 Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA **97**: 4985–4990.
- CLOUSE, S. D., M. LANGFORD and T. C. McMORRIS, 1996 A brassinosteroid-insensitive mutant in *Arabidopsis thaliana* exhibits multiple defects in growth and development. Plant Physiol. **111**: 671–678.
- COLLETT, C. E., N. P. HARBERD and O. LEYSER, 2000 Hormonal interactions in the control of Arabidopsis hypocotyl elongation. Plant Physiol. **124**: 553–561.
- EMANUELSSON, O., H. NIELSEN, S. BRUNAK and G. VON HEIJNE, 2000 Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. **300**: 1005–1016.
- ERRAMPALLI, D., D. PATTON, L. CASTLE, L. MICKLESON, K. HANSEN *et al.*, 1991 Embryonic lethals and T-DNA insertional mutagenesis in Arabidopsis. Plant Cell **3**: 149–157.
- FISK, D. G., M. B. WALKER and A. BARKAN, 1999 Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondrial and chloroplast gene expression. EMBO J. **18**: 2621–2630.
- FRANZMANN, L. H., E. S. YOON and D. W. MEINKE, 1995 Saturating the genetic map of *Arabidopsis thaliana* with embryonic mutations. Plant J. **7**: 341–350.
- FRASER, A. G., R. S. KAMATH, P. ZIPPERLEN, M. MARTINEZ-CAMPOS, M. SOHRMANN *et al.*, 2000 Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. Nature **408**: 325–330.
- GUYER, D., A. TUTTLE, S. ROUSE, S. VOLRATH, M. JOHNSON *et al.*, 1998 Activation of latent transgenes in *Arabidopsis* using a hybrid transcription factor. Genetics **149**: 633–639.
- HERMAN, R. K., 1988 Genetics, pp. 17–46 in *The Nematode Caenorhabditis elegans*, edited by W. B. WOOD. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- HOLSTERS, M., R. VILLAROEI, M. VAN MONTAGU and J. SCHELL, 1982 The use of selectable markers for the isolation of plant DNA/T-DNA junction fragments in a cosmid vector. Mol. Gen. Genet. **185**: 283–289.
- JACOBSEN, S. E., and N. E. OLSZEWSKI, 1993 Mutations at the SPINDLY locus of Arabidopsis alter gibberellin signal transduction. Plant Cell **5**: 887–896.
- JARVIS, P., L. J. CHEN, H. LI, C. A. PETO, C. FANKHAUSER *et al.*, 1998 An Arabidopsis mutant defective in the plastid general protein import apparatus. Science **282**: 100–103.

- JURGENS, G., U. MAYER, R. A. TORRES RUIZ, T. BERLETH and S. MISERA, 1991 Genetic analysis of pattern formation in the Arabidopsis embryo. *Development* **91**(Suppl. 1): 27–38.
- KEMPIN, S. A., B. SAVIDGE and M. F. YANOFKY, 1995 Molecular basis of the cauliflower phenotype in Arabidopsis. *Science* **267**: 522–525.
- KO, K., D. BUDD, C. WU, F. SEIBERT, L. KOURTZ *et al.*, 1995 Isolation and characterization of a cDNA clone encoding a member of the Com44/Cim44 envelope components of the chloroplast protein import apparatus. *J. Biol. Chem.* **270**: 28601–28608.
- KOORNNEEF, M., and C. J. HANHART, 1981 A new thiamine locus in Arabidopsis. *Arabidopsis Inf. Serv.* **18**: 52–58.
- KUBASEK, W. L., B. W. SHIRLEY, A. MCKILLOP, H. M. GOODMAN, W. BRIGGS *et al.*, 1992 Regulation of flavanoid biosynthetic genes in germinating Arabidopsis seedlings. *Plant Cell* **4**: 1229–1236.
- LEVIN, J. Z., A. J. DE FRAMOND, A. TUTTLE, M. W. BAUER and P. B. HEIFETZ, 2000 Methods of double-stranded RNA-mediated gene inactivation in *Arabidopsis* and their use to define an essential gene in methionine biosynthesis. *Plant Mol. Biol.* **44**: 759–775.
- LI, S. L., and G. P. REDEL, 1969 Thiamine mutants of the crucifer. *Arabidopsis. Biochem. Genet.* **3**: 163–170.
- LICHTENTHALER, H. K., 1999 The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**: 47–65.
- LICHTENTHALER, H. K., J. ZEIDLER, J. SCHWENDER and C. MULLER, 2000 The non-mevalonate isoprenoid biosynthesis of plants as a test system for new herbicides and drugs against pathogenic bacteria and the malaria parasite. *Z. Naturforsch.* **55**: 305–313.
- LIU, Y. G., N. MITSUKAWA, T. OOSUMI and R. F. WHITTIER, 1995 Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* **8**: 457–463.
- LOIS, L. M., N. CAMPOS, S. R. PUTRA, K. DANIELSEN, M. ROHMER *et al.*, 1998 Cloning and characterization of a gene from *Escherichia coli* encoding a transketolase-like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. *Proc. Natl. Acad. Sci. USA* **95**: 2105–2110.
- LONG, D., M. MARTIN, E. SUNDBERG, J. SWINBURNE, P. PUANGSOMLEE *et al.*, 1993 The maize transposable element system Ac/Ds as a mutagen in Arabidopsis: identification of an albino mutation induced by Ds insertion. *Proc. Natl. Acad. Sci. USA* **90**: 10370–10374.
- LONG, D., J. GOODRICH, K. WILSON, E. SUNDBERG, M. MARTIN *et al.*, 1997 Ds elements on all five Arabidopsis chromosomes and assessment of their utility for transposon tagging. *Plant J.* **11**: 145–148.
- MANDEL, M. A., K. A. FELDMANN, L. HERRERA-ESTRELLA, M. ROCHA-SOSA and P. LEON, 1996 CLAI1, a novel gene required for chloroplast development, is highly conserved in evolution. *Plant J.* **9**: 649–658.
- MAYER, U., R. A. TORRES RUIZ, T. BERLETH, S. MISERA and G. JURGENS, 1991 Mutations affecting body organization in the Arabidopsis embryo. *Nature* **353**: 402–407.
- MCELVER, J., I. TZAFRIR, G. AUX, R. ROGERS, C. ASHBY *et al.*, 2001 Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* **159**: 1751–1763.
- MEINKE, D., 1994 Seed development in *Arabidopsis thaliana*, pp. 253–295 in *Arabidopsis*, edited by E. M. MEYEROWITZ and C. R. SOMERVILLE. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- MEISSNER, R. C., H. JIN, E. COMINELLI, M. DENEKAMP, A. FUERTES *et al.*, 1999 Function search in a large transcription factor gene family in Arabidopsis: assessing the potential of reverse genetics to identify insertional mutations in R2R3 MYB genes. *Plant Cell* **11**: 1827–1840.
- MEURER, J., K. MEIERHOFF and P. WESTHOFF, 1996 Isolation of high-chlorophyll-fluorescence mutants of *Arabidopsis thaliana* and their characterisation by spectroscopy, immunoblotting and northern hybridisation. *Planta* **198**: 385–396.
- MEURER, J., C. GREVELDING, P. WESTHOFF and B. REISS, 1998 The PAC protein affects the maturation of specific chloroplast mRNAs in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **258**: 342–351.
- MILES, D., 1994 The role of high chlorophyll fluorescence photosynthesis mutants in the analysis of chloroplast thylakoid membrane assembly and function. *Maydica* **39**: 35–45.
- MISERA, S., A. J. MULLER, U. WEILAND-HEIDECKER and G. JURGENS, 1994 The *FUSCA* genes of *Arabidopsis*: negative regulators of light responses. *Mol. Gen. Genet.* **244**: 242–252.
- MORI, H., E. J. SUMMER, X. MA and K. CLINE, 1999 Component specificity for the thylakoidal Sec and Delta pH-dependent protein transport pathways. *J. Cell Biol.* **146**: 45–55.
- MORI, H., E. J. SUMMER and K. CLINE, 2001 Chloroplast TatC plays a direct role in thylakoid (Delta)pH-dependent protein transport. *FEBS Lett.* **501**: 65–68.
- NACRY, P., C. CAMILLERI, B. COURTIAL, M. CABOCHE and D. BOUCHEZ, 1996 Major chromosomal rearrangements induced by T-DNA transformation in Arabidopsis. *Genetics* **149**: 641–650.
- PARINOV, S., and V. SUNDARESAN, 2000 Functional genomics in Arabidopsis: large-scale insertional mutagenesis complements the genome sequencing project. *Curr. Opin. Biotechnol.* **11**: 157–161.
- PARINOV, S., M. SEVUGAN, D. YE, W. C. YANG, M. KUMARAN *et al.*, 1999 Analysis of flanking sequences from dissociation insertion lines. A database for reverse genetics in Arabidopsis. *Plant Cell* **11**: 2263–2270.
- PEARSON, W. R., and D. J. LIPMAN, 1988 Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**: 2444–2448.
- PEPPER, A., T. DELANEY, T. WASHBURN, D. POOLE and J. CHORY, 1994 DET1, a negative regulator of light-mediated development and gene expression in Arabidopsis, encodes a novel nuclear-localized protein. *Cell* **78**: 109–116.
- PRICE, G. D., J.-W. YU, S. VON CAEMMERER, J. R. EVANS, W. S. CHOW *et al.*, 1995 Chloroplast cytochrome *b<sub>6</sub>/f* and ATP synthase complexes in tobacco: transformation with antisense RNA against nuclear-encoded transcripts for the Rieske FeS and ATP $\delta$  polypeptides. *Aust. J. Plant Physiol.* **22**: 285–297.
- REDEL, G. P., and C. KONCZ, 1992 Classical mutagenesis, pp. 16–82 in *Methods in Arabidopsis Research*, edited by C. KONCZ and N.-H. CHUA. World Scientific Publishing, Singapore.
- REED, J. W., P. NAGPAL, R. M. BASTOW, K. S. SOLOMON, M. J. DOWSON-DAY *et al.*, 2000 Independent action of ELF3 and phyB to control hypocotyl elongation and flowering time. *Plant Physiol.* **122**: 1149–1160.
- REITER, R. S., R. M. YOUNG and P. A. SCOLNIK, 1992 Genetic linkage of the Arabidopsis genome: methods for mapping with recombinant inbred and random amplified polymorphic DNAs (RAPDs), pp. 170–190 in *Methods in Arabidopsis Research*, edited by C. KONCZ, N.-H. CHUA and J. SCHELL. World Scientific Publishing, Singapore.
- REITER, R. S., S. A. COOMBER, T. M. BOURETT, G. E. BARTLEY and P. A. SCOLNIK, 1994 Control of leaf and chloroplast development by the Arabidopsis gene pale cress. *Plant Cell* **6**: 1253–1264.
- ROBINSON, C., S. J. THOMPSON and C. WOOLHEAD, 2001 Multiple pathways used for the targeting of thylakoid proteins in chloroplasts. *Traffic* **2**: 245–251.
- ROHDICH, F., J. WUNGSINTAWEEKUL, M. FELLERMEIER, S. SAGNER, S. HERZ *et al.*, 1999 Cytidine 5'-triphosphate-dependent biosynthesis of isoprenoids: YgbP protein of *Escherichia coli* catalyzes the formation of 4-diphosphocytidyl-2-C-methylerythritol. *Proc. Natl. Acad. Sci. USA* **96**: 11758–11763.
- ROHDICH, F., J. WUNGSINTAWEEKUL, W. EISENREICH, G. RICHTER, C. A. SCHUHR *et al.*, 2000 Biosynthesis of terpenoids: 4-diphosphocytidyl-2C-methyl-D-erythritol synthase of *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **97**: 6451–6456.
- ROHMER, M., 1998 Isoprenoid biosynthesis via the mevalonate-independent route, a novel target for antibacterial drugs? *Prog. Drug Res.* **50**: 135–154.
- ROLDAN, M., C. GOMEZ-MENA, L. RUIZ-GARCIA, J. SALINAS and J. M. MARTINEZ-ZAPATER, 1999 Sucrose availability on the aerial part of the plant promotes morphogenesis and flowering of *Arabidopsis* in the dark. *Plant J.* **20**: 581–590.
- RUNGE, S., B. VAN CLEVE, N. LEBEDEV, G. ARMSTRONG and K. APEL, 1995 Isolation and classification of chlorophyll-deficient *xantha* mutants of *Arabidopsis thaliana*. *Planta* **197**: 490–500.
- SCHERES, B., H. MCKHANN, C. VAN DER BERG, V. WILLEMSSEN, H. WOKENFELT *et al.*, 1996 Experimental and genetic analysis of root development in *Arabidopsis thaliana*. *Plant Soil* **187**: 97–105.
- SCHLEIFF, E., and J. SOLL, 2000 Travelling of proteins through membranes: translocation into chloroplasts. *Planta* **211**: 449–456.
- SCHMIDT, R., 1998 The *Arabidopsis thaliana* genome: towards a complete physical map, pp. 1–30 in *Arabidopsis*, edited by M. ANDER-

- SON and J. ROBERTS. Sheffield Academic Press, Sheffield, England.
- SCHULTES, N. P., R. J. SAWERS, T. P. BRUTNELL and R. W. KRUEGER, 2000 Maize high chlorophyll fluorescent 60 mutation is caused by an *Ac* disruption of the gene encoding the chloroplast ribosomal small subunit protein 17. *Plant J.* **21**: 317–327.
- SCHWENDER, J., M. SEEMAN, H. K. LICHTENTHALER and M. ROHMER, 1996 Biosynthesis of isoprenoids (carotenoids, sterols, prenyl side-chains of chlorophylls and plastoquinone) via a novel pyruvate/glyceraldehyde 3-phosphate non-mevalonate pathway in the green alga *Scenedesmus obliquus*. *Biochem. J.* **316**: 73–80.
- SCHWENDER, J., C. MULLER, J. ZEIDLER and H. K. LICHTENTHALER, 1999 Cloning and heterologous expression of a cDNA encoding 1-deoxy-D-xylulose-5-phosphate reductoisomerase of *Arabidopsis thaliana*. *FEBS Lett.* **455**: 140–144.
- SETTLES, A. M., A. YONETANI, A. BARON, D. R. BUSH, K. CLINE *et al.*, 1997 Sec-independent protein translocation by the maize Hcf106 protein. *Science* **278**: 1467–1470.
- STAHL, T., C. GLOCKMANN, J. SOLL and L. HEINS, 1999 Tic40, a new “old” subunit of the chloroplast protein import translocon. *J. Biol. Chem.* **274**: 37467–37472.
- STEWART, H. I., N. J. O’NEIL, D. L. JANKE, N. W. FRANZ, H. W. CHAMBERLIN *et al.*, 1998 Lethal mutations defining 112 complementation groups in a 4.5 Mb sequenced region of *Caenorhabditis elegans* chromosome III. *Mol. Gen. Genet.* **260**: 280–288.
- SUNDARESAN, V., P. SPRINGER, T. VOLPE, S. HAWARD, J. D. JONES *et al.*, 1995 Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes Dev.* **9**: 1797–1810.
- SUSEK, R. E., F. M. AUSUBEL and J. CHORY, 1993 Signal transduction mutants of *Arabidopsis* uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell* **74**: 787–799.
- TAKAHASHI, S., T. KUZUYAMA, H. WATANABE and H. SETO, 1998 A 1-deoxy-D-xylulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proc. Natl. Acad. Sci. USA* **95**: 9879–9884.
- VISION, T. J., D. G. BROWN and S. D. TANKSLEY, 2000 The origins of genomic duplications in *Arabidopsis*. *Science* **290**: 2114–2117.
- VIZIR, I. Y., M. L. ANDERSON, Z. A. WILSON and B. J. MULLIGAN, 1994 Isolation of deficiencies in the *Arabidopsis* genome by  $\gamma$ -irradiation of pollen. *Genetics* **137**: 1111–1119.
- WALDEN, R., K. FRITZE, H. HAYASHI, E. MIKLASHEVICH, H. HARLING *et al.*, 1994 Activation tagging: a means of isolating genes implicated as playing a role in plant growth and development. *Plant Mol. Biol.* **26**: 1521–1528.
- WALKER, M. B., L. M. ROY, E. COLEMAN, R. VOELKER and A. BARKAN, 1999 The maize *tha4* gene functions in sec-independent protein transport in chloroplasts and is related to *hcf106*, *tatA*, and *tatB*. *J. Cell Biol.* **147**: 267–275.
- WARD, E., and P. BERNASCONI, 1999 Target-based discovery of crop protection chemicals. *Nat. Biotechnol.* **17**: 618–619.
- WATERHOUSE, P. M., M. W. GRAHAM and M.-B. WANG, 1998 Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc. Natl. Acad. Sci. USA* **95**: 13959–13964.
- WEIGEL, D., J. H. AHN, M. A. BLAZQUEZ, J. O. BOREVITZ, S. K. CHRISTENSEN *et al.*, 2000 Activation tagging in *Arabidopsis*. *Plant Physiol.* **122**: 1003–1014.
- WINZELER, E. A., D. D. SHOEMAKER, A. ASTROMOFF, H. LIANG, K. ANDERSON *et al.*, 1999 Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**: 901–906.

Communicating editor: C. S. GASSER