Directed Tagging of the Arabidopsis FATTY ACID ELONGATION1 (FAE1) Gene with the Maize Transposon Activator

Douglas W. James, Jr., Eda Lim, Janis Keller, Ingrid Plooy, Ed Ralston, and Hugo K. Dooner
DNA Plant Technology Corporation, 6701 San Pablo Avenue, Oakland, California 94608

The FATTY ACID ELONGATION1 (FAE1) gene of Arabidopsis is required for the synthesis of very long chain fatty acids in the seed. The product of the FAE1 gene is presumed to be a condensing enzyme that extends the chain length of fatty acids from C18 to C20 and C22. We report here the cloning of FAE1 by directed transposon tagging with the maize element Activator (Ac). An unstable fael mutant was isolated in a line carrying Ac linked to the FAE1 locus on chromosome 4. Cosegregation and reversion analyses established that the new mutant was tagged by Ac. A DNA fragment flanking Ac was cloned by inverse polymerase chain reaction and used to isolate FAE1 genomic clones and a cDNA clone from a library made from immature siliques. The predicted amino acid sequence of the FAE1 protein shares homology with those of other condensing enzymes (chalcone synthase, stilbene synthases, and β-ketoacyl-acyl carrier protein synthase III), supporting the notion that FAE1 is the structural gene for a synthase or condensing enzyme. FAE1 is expressed in developing seed, but not in leaves, as expected from the effect of the fael mutation on the fatty acid compositions of those tissues.

INTRODUCTION

The major components of seed storage lipids in higher plants are 16- and 18-carbon fatty acids. Oilseeds of the Cruciferae and a few other plants also accumulate C20 and C22 fatty acids, collectively referred to as very long chain fatty acids (VLCFAs) because of their relatively longer chain length compared with the more common fatty acids found in plants. The presence of VLCFAs in vegetable oils markedly affects their use. For example, erucic acid (22:1) has detrimental nutritional effects and is undesirable in edible oils. Rapeseed oil is naturally high in erucic acid, but through a concerted breeding effort, canola lines that are almost devoid of erucic acid have been developed (Stefansson et al., 1961; Loof and Appleqvist, 1972). On the other hand, vegetable oils high in erucic acid have found many industrial uses: directly as diesel fuel and as a raw material for an array of products, including paints, corrosion inhibitors, cosmetics, plastics, pharmaceuticals, and lubricants (Johnson and Fritz, 1989).

Elongation of the fatty acid carbon chain from C18 to C22 occurs by the sequential addition of two C2 moieties from malonyl coenzyme A (CoA) to a C18 carbon skeleton. This reaction is catalyzed by a particulate acyl CoA elongase complex (Downey and Craig, 1964; Stumpf and Pollard, 1983). Whether the two elongation reactions are performed by one or two different enzyme complexes is not clear at this time (Agrawal and Stumpf, 1985; Lessire et al., 1985; Taylor et al., 1992). Four different reactions are involved in the elongation system of plants: (1) condensation of 18:1 CoA with malonyl CoA to form a β-ketoacyl derivative, (2) reduction and (3) dehydration of the β-ketoacyl derivative, and (4) reduction of the double bond (Fehling and Mukherjee, 1991). However, because of the difficulties in solubilizing membrane-bound enzymes, the elongase complex has not been well characterized. Elongases have been partially purified from several plants, including rapeseed (Bessoule et al., 1989; Fehling et al., 1992; Creach and Lessire, 1993). This purification has resulted in an enrichment of three to four protein bands, an observation favoring the hypothesis that the elongase complex consists of several protein components and not of a large, single, multifunctional polypeptide (Bessoule et al., 1989).

In Arabidopsis, mutations in the FATTY ACID ELONGATION1 (FAE1) gene result in highly reduced levels of seed VLCFAs (James and Dooner, 1990; Lemieux et al., 1990) and in a deficiency in acyl chain elongation activities from C18 to C20 and C20 to C22, suggesting that the product of the FAE1 gene is required for both elongation steps (Kunst et al., 1992). Based on the observations that the other elongation systems for fatty acid biosynthesis (C4+C16 and C16+C18) share several enzyme activities, but differ in their condensing enzymes or ketoacyl synthases (Shimakata and Stumpf, 1982; Stumpf, 1984), and that fael mutants do not cause deleterious...
pleiotropic effects, Kunst et al. (1992) have argued that the FAE1 gene may encode a condensing enzyme unique to VLCFA synthesis.

The application of genetic approaches in Arabidopsis has proven particularly successful in the isolation of genes encoding membrane-bound fatty acid biosynthetic enzymes. Map-based cloning from a yeast artificial chromosome genomic library allowed the isolation of the FATTY ACID DESATURATION3 (FAD3) gene encoding an endoplasmic reticulum 18:2 desaturase (Arondel et al., 1992). T-DNA tagging was used as an alternative approach to isolate FAD3 (Yadav et al., 1993) and, recently, to isolate the FAD2 gene, which encodes another endoplasmic reticulum enzyme, an 18:1 desaturase (Okuley et al., 1994). In this study, we report the isolation of the Arabidopsis FAE1 gene by a third genetic approach, heterologous transposon tagging with the maize element Activator (Ac). The FAE1 protein predicted from the corresponding cDNA sequence shares homology with chalcone synthase (CHS) and stilbene synthase (STS), two condensing enzymes that also catalyze the addition of malonyl CoA molecules to a starter CoA ester, and with β-ketoacyl-acyl carrier protein synthase III (KASIIL) from Escherichia coli (Tsay et al., 1992) and spinach (Tai and Jaworski, 1993). This suggests that the product of the FAE1 gene is in fact a synthase or condensing enzyme. The FAE1 gene is expressed preferentially in seed, as could be anticipated from the specific effect of the fae1 mutation on seed fatty acid composition (James and Dooner, 1990; Lemieux et al., 1990).

RESULTS

Gene Isolation Approach

The strategy that we adopted to try to isolate fatty acid biosynthetic genes from Arabidopsis is illustrated in Figure 1. Because many of the fatty acid biosynthetic enzymes are membrane bound and difficult to solubilize and purify, we took a genetic approach. Specifically, we used transposon tagging with the autonomous maize element Ac. First, we induced and characterized a collection of mutants to define genetically loci that affect fatty acid composition in Arabidopsis seed (James and Dooner, 1990). We uncovered mutations at several loci; some of these were phenotypically similar to mutations isolated concurrently by Lemieux et al. (1990). Subsequent tests have shown that multiple alleles for at least three loci (FAD2, FAD3, and FAE1) were recovered independently by the two groups (Kunst et al., 1992; Miquel et al., 1993; D.W. James and H.K. Dooner, unpublished results).

Second, we generated 68 independent single-locus Arabidopsis transformants carrying Ac in the T-DNA (Keller et al., 1992). This number gives us >95% confidence that any gene in the Arabidopsis genome will be no farther than 15 centimorgans (cM) from an introduced Ac, assuming random integration of the T-DNA in the genome and a <600-cM genetic map (Koornneef et al., 1992). In Arabidopsis as well as in maize and other organisms, transposons of the Ac family transpose preferentially to sites linked to the donor site (Bancroft and Dean, 1993; Keller et al., 1993b). Therefore, to tag a specific gene, it is preferable to initiate the tagging experiment with an Ac (or Dissociation) element linked to the gene of interest. This approach, referred to as directed tagging, is laborious in that it requires mapping of the T-DNAs relative to the target loci; however, scoring for the presence or absence of the T-DNA is facilitated by the Hyg* (hygromycin resistance) transformation marker. Ultimately, the number of plants with transposed Ac (trAc) elements required to isolate a specific fatty acid mutation is greatly reduced relative to that required in a random transposon tagging approach.

We localized 24 T-DNAs to one of the five Arabidopsis chromosomes and found three significant linkages to FAE1. In transformants B805, B116, and C231, the T-DNAs were located, respectively, 15, 22, and 40 cM from FAE1. FAE1 is loosely linked to the restriction fragment length polymorphism (RFLP) marker 518 on chromosome 4 (D.W. James, J. Keller, E. Lim, and H.K. Dooner, unpublished results). We confirmed linkage of the B116 T-DNA to marker 518 (Chang et al., 1988) by a bulked segregant analysis procedure (Michelmore et al., 1991) in which paired DNA samples obtained from pooled homozygous Hygr/Hygr and +/+ segregants from a Hygr/+ heterozygote (Arabidopsis ecotypes Wassilewskija [WS] × Columbia) were scored for segregating RFLPs. If the T-DNA is linked to a RFLP marker, the hybridization intensity of the RFLP band contributed by the WS parent will be darker than that of the Columbia parent in the Hygr/Hygr DNA sample and lighter in the +/+ DNA sample because the WS ecotype is transformed with the Hygr gene. The converse will hold for the RFLP band contributed by the Columbia parent. The combination of two-point crosses and bulked segregant analyses has enabled us to assign a chromosomal location to 24 single-locus, Ac-containing T-DNA insertions to date.

In our constructs, Ac has been inserted in the 5′ untranslated region of the streptomycin phosphotransferase (SPT) gene (Jones et al., 1989; Keller et al., 1993a). Somatic excisions of Ac during the development of the cotyledons can be detected as green sectors on a white background in Arabidopsis seedlings germinated in streptomycin. Germinal excisions of Ac give rise to fully green seedlings, approximately half of which carry a trAc element somewhere in the genome (Dean et al., 1992; Keller et al., 1992). Therefore, SPT::Ac constitutes an efficient marker for selecting plants that have undergone transposition. We refer to the fully green derivatives as SPT*, following the convention used to designate phenotypic revertants from mutable alleles.

The transposition activity of Ac is generally low in Arabidopsis and results in a low frequency of germinal revertants (~1% on average) (Schmidt and Willmitzer, 1989; Dean et al., 1992; Keller et al., 1992). It is possible to trans activate low-activity Ac or Dissociation (Ds) elements with constructs in which the Ac transposase function is expressed behind a strong promoter (Bancroft and Dean, 1992; Grevelding et al., 1992; Swinburne
et al., 1992; Fedoroff and Smith, 1993; Honma et al., 1993; Keller et al., 1993b), but transposition events in Arabidopsis tend to be clonal whether the transposition function is supplied autonomously or by a second construct. Thus, germinal selections recovered from one plant will often be derived from a common premeiotic event and carry the same transposed element. To avoid extensive sampling of duplicates for the same transposition event, we usually transferred to the greenhouse no more than four green seedling selections from any one plant.

Identification of a New fae1 Mutation

We selected green SPT' seedlings on streptomycin, grew them to maturity in the greenhouse, and analyzed the fatty acid composition of their seed by gas chromatography. Because fae1 and several other mutations affecting seed fatty acid composition are codominant (James and Dooner, 1990; Lemieux et al., 1990), they can be identified in the heterozygous condition, a clear advantage when dealing with a phenotype that is difficult to score. Selections were made from lines in which Ac was either linked or unlinked to FAE7. Table 1 gives the number of SPT' selections analyzed, the number of plants that produced them, and the location of Ac relative to FAE7 in each line. Also given are estimates of the minimum and maximum number of independent Ac reinsertions screened, assuming a 50% Ac reinsertion frequency (Altmann et al., 1992; Dean et al., 1992; Keller et al., 1992). The minimum number, that is, the number screened if all the siblings resulted from independent transposition events, corresponds to half the number of selections analyzed. The actual number of independent Ac reinsertions assayed lies somewhere between the two values because both clonal and single Ac transposition events can be recovered from the same plant (Keller et al., 1992).

A total of 2243 SPT' selections from the two lines in which the T-DNAs were closest to FAE7 and over 3000 SPT' selections

<table>
<thead>
<tr>
<th>T-DNA Line</th>
<th>Ac-FAE7 Linkage</th>
<th>No. SPT' Selections Analyzed</th>
<th>No. Plants Producing Selections</th>
<th>Min.-Max. No. Independent Ac Reinsertionsa</th>
</tr>
</thead>
<tbody>
<tr>
<td>K805</td>
<td>15 cM</td>
<td>1522</td>
<td>660</td>
<td>330-761</td>
</tr>
<tr>
<td>B116</td>
<td>22 cM</td>
<td>721</td>
<td>272</td>
<td>136-361</td>
</tr>
<tr>
<td></td>
<td>Linkedb</td>
<td>2243</td>
<td>932</td>
<td>466-1122</td>
</tr>
<tr>
<td>C201</td>
<td>Unlinked</td>
<td>2107</td>
<td>324</td>
<td>162-1054</td>
</tr>
<tr>
<td>B246</td>
<td>Unlinked</td>
<td>680</td>
<td>227</td>
<td>114-340</td>
</tr>
<tr>
<td>A018</td>
<td>Unlinked</td>
<td>273</td>
<td>91</td>
<td>45-136</td>
</tr>
<tr>
<td></td>
<td>Unlinkedb</td>
<td>3060</td>
<td>642</td>
<td>321-1530</td>
</tr>
</tbody>
</table>

* Assumption: 50% reinsertions among SPT' selections. Minimum (Min.), all selections clonal; maximum (Max.), all selections independent.

Subtotal.
Table 3: all of the individuals carrying the new trAc were either Ac-Hybridizing Band selection, SPT/Ac). Resident site SPT/Ac and the latter, the Ac the new trAc at the progeny—a 3.3- and a 2.4-kb band; the former represents by gas chromatography analysis of their seed fatty acid composition of the pertinent six genotypic classes could be distinguished. DNA from 54 or +/+ by DNA gel blot analysis. Thus, Acl(-) individuals were scored as faeHfael or FaeHfael, FaeHfael, Fae1lfae1-G309, fael- (WS wild type), fael-2lfae1-G309, fael-2lfae1-2, and fael-2lfae1-G309.

Evidence That the New fael Mutation Is Tagged by Ac

Cosegregation of the fae Mutant Phenotype with an Ac-Hybridizing Band

DNA analysis of the fael-G309 derivative showed the presence of a new Ac-hybridizing band. We tested for the joint segregation of this new Ac band with the fael mutation in the self-progeny of a fael-G309/Fae1 heterozygote. Segregating individuals were scored as Fae1lFae1, Fae1lfael, or fael1fael by gas chromatography analysis of their seed fatty acid composition and as Ac(-) or +/- by DNA gel blot analysis. Thus, six genotypic classes could be distinguished. DNA from 54 individuals was digested with HindIII and hybridized with a probe from the 5' end of Ac (Kunze et al., 1987). As seen in Figure 2 (top), two Ac-hybridizing bands were segregating in the progeny—a 3.3- and a 2.4-kb band; the former represents the new trAc and the latter, the Ac at the SPT/Ac resident site in the T-DNA (the selfed parent was the original G309 green selection, SPT/SPT/Ac).

The results of the cosegregation analysis are presented in Table 3: all of the individuals carrying the new trAc were either homozygous or heterozygous for the new fael mutation. Conversely, all of the individuals that lacked the trAc were wild type. Therefore, no recombinants were recovered (linkage \( \chi^2 = 54, P < 0.001 \)).

A 1.5-kb fragment of DNA flanking the 5' end of the trAc was cloned by inverse polymerase chain reaction (IPCR) using Ac-specific primers on HindIII-digested genomic DNA from a fael-G309 homozygous individual that lacked the 2.4-kb SPT/Ac band. This flanking DNA was used to reprobe the same DNA gel blots. Figure 2 (bottom) shows the result of rehybridizing the blot shown in Figure 2 (top). As expected, the new probe detected the same 3.3-kb Ac-homologous fragment and, in addition, a new fragment of \( \sim 1.8 \) kb. If Ac had in fact tagged the FAE1 gene, all segregating fael mutants should be homozygous for the 3.3-kb band, all Fae1lfael plants should be heterozygous for the 3.3-kb and 1.8-kb bands, and all wild-type plants should be homozygous for the 1.8-kb band. The results of the analysis are presented in Table 4: 53 of the 54 segregants fit the expectation. The one exception (segregant 50) was fael1/ fael1, but heterozygous for the 3.3-kb trAc band. There are two

![Figure 2. DNA Gel Blot Analysis of Progeny from Green Selection G309 Carrying a trAc Element.](image)
Table 3. Segregation Data for the Self-Progeny of a Faellfael, trAc/+ Heterozygote: DNA from Segregants Scored with an Ac Probe

<table>
<thead>
<tr>
<th>FAE7 Genotype</th>
<th>trAc Genotype</th>
<th>+/+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faellfael</td>
<td>0</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Faellfael</td>
<td>29</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>fael1fael</td>
<td>16</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>9</td>
<td>54</td>
</tr>
</tbody>
</table>

Table 4. Segregation Data for the Self-Progeny of a Faellfael, trAc/+ Heterozygote: DNA from Segregants Scored with Probes for Ac and the Flanking DNA

<table>
<thead>
<tr>
<th>FAE7 Genotype</th>
<th>Presumed trAc Genotype</th>
<th>+/+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faellfael</td>
<td>trAcl/trAcl</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Faellfael</td>
<td>trAcl +</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>fael1fael</td>
<td>trAccl</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>15</td>
<td>30</td>
<td>54</td>
</tr>
</tbody>
</table>

plausible explanations for this exception: (1) The new fae mutation is not tagged by the trAc but is closely linked to it, and this individual is a recombinant between the fael mutation and the trAc. (2) The new fael mutant is tagged by the trAc, and this individual is the product of an Ac excision that did not restore gene function. Segregant 50 also had a new Ac-hybridizing band, which was absent in the other 53 siblings (data not shown), an observation suggesting that it might have originated by secondary transposition of Ac. The absence of additional bands hybridizing to the DNA adjacent to Ac (Figure 2, bottom) also indicates that Ac has inserted into a unique sequence in the Arabidopsis genome.

Reversion of the fael Mutant

To confirm that the fael-G309 mutation was tagged by Ac, we screened progeny of mutants for putative revertants, that is, individuals with an intermediate seed fatty acid composition, which could then be examined for excision of Ac from the cloned DNA. A total of 1052 offspring from four fael-G309 homozygous plants were screened, and three putative revertants to the wild type (lane 1) and the three putative revertants (lanes 2 to 4); hybridization was with a probe that spans the Ac insertion site (see the following discussion). As can be seen, all three exceptions were heterozygous for the 3.3-kb band in the original mutant allele and a 1.8-kb wild-type-sized band (the additional 1.4-kb band seen in lanes 2, 3, and 4 of Figure 3 is contributed by the mutant allele and corresponds to the other side of the Ac insertion). This is the result expected if these individuals originated from Ac excision events that restored gene function.

The IPCR-amplified, 1.5-kb fragment of genomic DNA was cloned into pBluescript KS– and used to isolate λ genomic clones of the wild-type WS progenitor allele. From a wild-type genomic clone, a 1.8-kb HindIII fragment, corresponding to the site of insertion of Ac in the mutant, was subcloned (clone 2203) and sequenced. The DNA around the trAc insertion site was then amplified by PCR (Saiki, 1990) and sequenced in the fael-G309 mutant and the three putative revertants. A comparison of the sequences is presented in Figure 4. Somewhat unexpectedly, the DNA sequences of the three putative revertants were identical to that of the wild-type progenitor, that is, they showed no evidence of the footprints that are typical of Ac excision events (Fedoroff, 1989). The absence of footprints in the putative revertants led us to consider other alternative origins for the exceptions. Wild-type sequence derivatives arising in a screen for revertants could represent either seed or pollen contaminants. Pollen contamination in the greenhouse is very rare in Arabidopsis (our male-sterile plants never set seed unless hand pollinated), but seed contamination is a possibility. However, three lines of evidence argue against this: (1) the putative revertants have an intermediate, not a wild-type, seed fatty acid composition, (2) they are heterozygous for the Ac insertion allele present in the original mutant selection, and (3) they are T-DNA homozygotes, based on the absence of segregation for the Hyg' marker in the introduced T-DNA.

Possibly, the sequence of the putative FAE1 protein around the site of insertion of Ac is intolerant of amino acid changes, so that only the rare events that restore not just the correct reading frame but the original sequence are selected as revertants. If so, segregant 50, our exception from the cosegregation analysis (Table 4), would represent an Ac excision event that did not restore the original amino acid sequence and thus failed to restore gene function. The DNA sequence of segregant 50 (Figure 4) revealed, in fact, a footprint at the Ac excision site, indicating that this exception had originated by an Ac excision that created a frameshift mutation and knocked out gene function.

To conclude, the analysis of revertants and of the one null exception in the cosegregation test confirms that the fael-G309 mutant arose by insertion of Ac into the FAE1 gene, that it is unstable, and that it can give rise to new alleles at the FAE1 locus. Hence, we have given it the official designation fael-m1(Ac) to denote that it is the first mutable allele of the FAE1 locus isolated and that it arose by insertion of the transposon Ac.
Figure 3. DNA Gel Blot Analysis of Revertants.
Genomic DNA (3 μg) was digested with HindIII, separated by electrophoresis on a 1% agarose gel, transferred to a nylon membrane, and probed with a 1.8-kb HindIII fragment that spans the Ac insertion site. Molecular length markers are given at right in kilobases. Lane 1, wild type (Fae1/Fae1); lanes 2 to 4, Fae' revertants (Fae/fae1-m1[Ac]).

Function of the FAE1 Gene in Arabidopsis

Because VLCFAs accumulate in seed, but not in leaves of Arabidopsis (Lemieux et al., 1990), FAE1 should be expressed preferentially in seed. Expression of the FAE1 gene was assayed by gel blot analysis of RNA from several tissues: leaf, flower, and immature siliques plus seed (pools of ~1-week-old and of 2- to 3-week-old siliques). Figure 5 shows the corresponding RNA blot, in which hybridization was with a FAE1-specific probe. As expected, a 1.7-kb transcript was found to accumulate in siliques containing developing seed, but not in leaves.

The 1.8-kb HindIII fragment (clone 2203) and a 700-bp BstXI-EcoRI fragment of the FAE1 gene (Figure 6B) were used as probes to screen a cDNA library made from 3-week-old developing siliques. A cDNA clone containing a 1.64-kb insert, roughly the size of the FAE1 transcript (Figure 5), was isolated and sequenced. Its nucleotide sequence is shown in Figure 6A: it matches the sequence of the genomic clone throughout its length, indicating that there are no introns in this segment of the FAE1 gene. The cDNA contains a long open reading frame but appears to be short of full length. If the ATG located in the corresponding FAE1 genomic sequence 20 nucleotides upstream from the 5' end of the cDNA is the initiation codon, the extended open reading frame would encode a protein of 507 amino acids with a predicted sequence as shown in Figure 6A. A diagram of the structure of the FAE1 genomic region, including the location of the cDNA and of the Ac insertion site in fae1-m1(Ac), is shown in Figure 6B.

DISCUSSION

The Transposon Tagging Strategy

We have isolated the FAE1 gene, which is required for the synthesis of VLCFAs in Arabidopsis seed, by directed tagging with the maize transposon Ac. The first reported examples of heterologous transposon tagging involved genes affecting phenotypes that were easily scored; they include flower color (Chuck et al., 1993), fruit development (Aarts et al., 1993), plant morphology (Bancroft et al., 1993), and chlorophyll production (Long et al., 1993). Those genes were isolated from random tagging experiments in which the maize transposons Ac, Ds, or Suppressor-mutator (Spm) did not target specific genes, but rather any gene in the genome. However, genes affecting phenotypes that are difficult to score are best targeted by a directed tagging approach in which the transposon is linked to the gene of interest to exploit one of the biological features of Ac and Spm, namely, their propensity to transpose to linked sites. This is precisely the approach taken by us to isolate FAE1 and by others to isolate disease resistance genes (Jones et al., 1994; Whitham et al., 1994).

Many mutations affecting seed fatty acid composition in Arabidopsis are codominant, that is, they can be identified in the heterozygous condition. Thus, our experimental design would have enabled us to identify mutations at five unlinked loci, FAD2, FAD3, FAE1, FATTY ACID BIOSYNTHESIS1 (FAB1), and FAB2 (James and Dooner, 1990; Lemieux et al., 1990). Yet, the only bona fide mutant that we obtained was fae1-m1(Ac), which originated in the Arabidopsis line B116, in which Ac was linked to FAE1. No mutations were recovered from three other lines in which Ac was unlinked to FAE1. Because only one mutation was isolated in the screen, it is not possible to make a meaningful comparison between mutation frequencies in the various lines carrying Ac. Nevertheless, the fact that the only mutant recovered arose in an Ac-linked line—when mutations at any one of five different loci could have been identified—shows the importance of linkage between the transposon and the target locus in transposon tagging experiments.

| Wild-type Fae1 | GTTGACTACTCGTGT TT |
| fae1-m1::Ac | GTTGACTACGTGACTACTCGTT |
| Fae1 revertants | GTTGACTACGTGACTACTCGTT |
| fae1 exception #50 | GTTGACTACGTGACTACGTACTCGTT |

Figure 4. Nucleotide Sequence of FAE1 around the Site of Insertion of Ac.

The eight bases duplicated upon insertion of Ac in the FAE1 gene are shaded. The boxed T residue indicates a transversion produced by excision of Ac in exception 50.
Total RNA (10 ng) was separated on a 1.1% agarose gel, transferred to a nylon membrane, and hybridized sequentially to two probes. (A) Hybridization to the 1.8-kb HindIII fragment of the FAE1 gene. (B) Hybridization to a wheat rDNA probe. Lane 1, leaf; lane 2, immature seed; lanes 3 and 4, respectively, 3- and 2-week-old siliques with seed. The top arrowhead indicates the position of the 28S rRNA, and the bottom arrowhead indicates that of the 18S rRNA.

**Figure 5. RNA Gel Blot Analysis of Several Arabidopsis Tissues.**

Total RNA (10 μg) was separated on a 1.1% agarose gel, transferred to a nylon membrane, and hybridized sequentially to two probes. (A) Hybridization to the 1.8-kb HindIII fragment of the FAE1 gene. (B) Hybridization to a wheat rDNA probe. Lane 1, leaf; lane 2, immature seed; lanes 3 and 4, respectively, 3- and 2-week-old siliques with seed. The top arrowhead indicates the position of the 28S rRNA, and the bottom arrowhead indicates that of the 18S rRNA.

**Reversion of fael-m1(Ac)**

Confirmation that a mutation has been tagged by a transposon is usually obtained from an analysis of revertants. We recovered revertants to the wild type from fael-m1(Ac) but at a very low frequency (three of 2104 chromosomes examined). The DNA sequence of the revertant alleles was identical to that of the wild type, that is, no transposon footprints were evident. When Ac excises, it usually leaves behind part of the target site repeat that it generates upon insertion; this excision footprint provides additional evidence that a transposon has visited the gene in question. Apparently, in fael-m1(Ac), reversion is rare because Ac is inserted in an exon of the FAE1 gene, and only those excision events that restore the original amino acid sequence of the FAE1 protein can generate revertants with a wild-type phenotype.

In the cosegregation analysis between the new fael mutation and the trAc band, an exceptional individual that had a homozygous mutant phenotype but a heterozygous trAc genotype proved to be derived from an Ac excision event that did not restore the original FAE1 reading frame. Thus, all the evidence from Ac excision events suggests that the FAE1 protein is highly intolerant of amino acid changes around the site of integration of Ac. The amino acid sequence of the FAE1 protein around that site is DYSC. Possibly, the relationship of the cysteine residue to neighboring amino acids is critical to FAE1 protein function because cysteine residues are known to be essential components of the active sites of enzymes that catalyze the condensation of acyl thioesters (Lanz et al., 1991; Bairoch, 1992).

The **FAE1 Gene and Its Product**

Kunst et al. (1992) have shown that the fael mutation results in a deficiency of the enzyme that catalyzes the condensation of malonyl CoA and oleoyl (18:1) CoA to form eicosenoyl (20:1) CoA. Evidence from this work supports the notion that the FAE1 gene encodes a condensing enzyme. A search of the combined GenBank and EMBL DNA sequence data bases with the TBLASTN program (Altschul et al., 1990) revealed that the predicted amino acid sequence of the FAE1 protein shares homology with three other condensing enzymes: CHS, STS, and KASIII. CHS and STS are condensing enzymes that use malonyl CoA in the synthesis of flavonoids and stilbenes, respectively, whereas KASIII is the condensing enzyme that initiates fatty acid biosynthesis in bacteria and plants by coupling acetyl CoA to the malonyl acyl carrier protein. The amino acid sequences of the four enzymes were aligned with the programs Pileup and Pretty from the Genetics Computer Group (Devereux et al., 1984) to produce the consensus sequence shown in Figure 7. The predicted sequence of FAE1 matches the consensus sequence at 17 positions in a 50-amino acid region close to the C terminus of the proteins (starting at position 391 in FAE1). This region is just upstream of the CHS–STS 12–amino acid "signature sequence" (Fiebigmann et al., 1992), which does not occur in the FAE1 protein. The FAE1 protein also shares homology with a conserved region in CHS and STS that is close to the active site cysteine identified by Lanz et al. (1991). However, this region of homology (L-A-K-D-L-X[9]-L-VV, residues 232 to 248) does not overlap the consensus sequence for the CHS/STS active site (G-C-[FY]-[GA]-G-G-TX[2]-R), but lies immediately downstream of it.

The nucleotide sequence of the FAE1 gene is highly homologous to Arabidopsis cDNA clone 1282 (Ats1282), which has been entered into GenBank as an expressed sequence tag. A comparison of the predicted amino acid sequences of the proteins encoded by FAE1 and Ats1282 reveals that the Ats1282 sequence may contain a reading frame error. If the entire Ats1282 sequence is placed in the same reading frame as FAE1 (by removing a T residue at position 303), the two predicted amino acid sequences are 58% identical and 76% similar over a 167–amino acid overlap. Thus, in all likelihood, the Ats1282 cDNA encodes another condensing enzyme that is also possibly involved in fatty acid biosynthesis. Figure 8 presents a dendrogram illustrating the relatedness of the FAE1 protein to the EST1282 product and to various CHS, STS, and KASIII enzymes. The predicted molecular mass of the FAE1 protein is 56 kD, which places it in the size range of the proteins enriched during the partial purification of the leek elongase complex (Bessoule et al., 1989).

The FAE1 gene is preferentially expressed in the developing seed of Arabidopsis. This was expected on the basis that the fael mutation results in the almost complete elimination of eicosenoic acid from the seed. No evidence of FAE1 expression was detected in the leaves, which do not accumulate VLCFAs (Lemieux et al., 1990). This pattern of expression contrasts with that of FAD2, which is a fatty acid biosynthetic gene.
that affects fatty acid desaturation in seed and leaves (Miquel and Browse, 1992) and is expressed in both tissues (Okuley et al., 1994). Thus, although the regulation of FAE1 and FAD2 may overlap, it is clearly not coextensive. Mutants that affect the seed 20:1 content but that are not allelic with FAE1 have been recovered by Kunst et al. (1992) and by us (D.W. James and H.K. Dooner, unpublished results). It will be interesting to determine whether those mutations have any effect on FAE1 expression.

METHODS

Arabidopsis Lines

The generation and characterization of the Arabidopsis thaliana lines C201, C231, B16, and B246, which carry a streptomycin phosphotransferase::Activator (SPT:Ac) marker, have been described previously (Keller et al., 1992). Line A018 contains two copies of the pJJ4075 construct (Keller et al., 1993a). Plant K0b5 contains one copy of the construct pJJ6224, which was made following the steps described for the construction of pJJ4141 (Keller et al., 1993a). However, the Nael fragment (positions 428 to 964 in Ac) was deleted from Ac in pJJ4368 prior to excision of Ac as an Sstl-Sal1 fragment and ligation of the cauliflower mosaic virus 35s promoter to the excision of Ac as an Sstl-Sal1 fragment and ligation of the cauliflower mosaic virus 35s promoter to the

Nucleic Acid Extraction and Analysis

DNA extraction and DNA blot analysis were performed as previously described (Keller et al., 1992). DNA from the region flanking the trans-posed Ac (trAc) in the fae1-ml(Ac) mutant was generated for cloning by inverse polymerase chain reaction (Ochman et al., 1988; Triglia et

Figure 6. Sequence of the FAE1 Gene and Structure of the FAE1 Genomic Region.

(A) Nucleotide sequence of the FAE1 gene and predicted amino acid sequence of the corresponding protein. The beginning and end of the FAE1 cDNA, which is slightly short of full length, are indicated by arrowheads.

(B) Structure of the FAE1 genomic region showing the location of the cDNA, the Ac insertion site in fae1-ml(Ac), and various probes. The IPCR-amplified fragment extends from the site of insertion of Ac to the nearest upstream HindIII site. Subclone 2203 corresponds to a 1.8-kb HindIII fragment that spans the site of insertion of Ac. The probe used for RNA gel blot analysis is a BstXI-EcoRI subclone from the end of one of the two overlapping λ genomic clones isolated. The EcoRI site designated R(λ) is a cloning site from the vector Lambda DASH and does not occur in the FAE1 gene.

METHODS

Arabidopsis Lines

The generation and characterization of the Arabidopsis thaliana lines C201, C231, B16, and B246, which carry a streptomycin phosphotransferase::Activator (SPT:Ac) marker, have been described previously (Keller et al., 1992). Line A018 contains two copies of the pJJ4075 construct (Keller et al., 1993a). Plant K0b5 contains one copy of the construct pJJ6224, which was made following the steps described for the construction of pJJ4141 (Keller et al., 1993a). However, the Nael fragment (positions 428 to 964 in Ac) was deleted from Ac in pJJ4368 prior to excision of Ac as an Sstl-Sal1 fragment and ligation of the cauliflower mosaic virus 35s promoter to the

Nucleic Acid Extraction and Analysis

DNA extraction and DNA blot analysis were performed as previously described (Keller et al., 1992). DNA from the region flanking the trans-posed Ac (trAc) in the fae1-ml(Ac) mutant was generated for cloning by inverse polymerase chain reaction (Ochman et al., 1988; Triglia et

Figure 6. Sequence of the FAE1 Gene and Structure of the FAE1 Genomic Region.

(A) Nucleotide sequence of the FAE1 gene and predicted amino acid sequence of the corresponding protein. The beginning and end of the FAE1 cDNA, which is slightly short of full length, are indicated by arrowheads.

(B) Structure of the FAE1 genomic region showing the location of the cDNA, the Ac insertion site in fae1-ml(Ac), and various probes. The IPCR-amplified fragment extends from the site of insertion of Ac to the nearest upstream HindIII site. Subclone 2203 corresponds to a 1.8-kb HindIII fragment that spans the site of insertion of Ac. The probe used for RNA gel blot analysis is a BstXI-EcoRI subclone from the end of one of the two overlapping λ genomic clones isolated. The EcoRI site designated R(λ) is a cloning site from the vector Lambda DASH and does not occur in the FAE1 gene.
ECOFABH 238 LdwLvpqan trtatsak xgsmdbnvv
SOKASIII 325 idwLlqan grtisaml renpsrlx
PSPCHSI 305 F...W...gsag PALLQEQAE LsklPeka
VUCHS 299 F...W...gsag PALLQEQAE LsklPeka
SOYCHSIV 298 F...W...gsag PALLQEQAE LsklPeka
AHGSNG11 176 F...W...gsag PALLQEQAE LsklPeka
PSSTS 302 F...W...gsag PALLQEQAE LsklPeka
ATEST1282 57 Fehfci RGAbV...e kLd...wmeb
FAEI 385 vd...fci RGAbV...e kLd...wmeb
CONSENSUS F...W...gs... ATEL... K L...K...

ECOFABH 268 T...nrngh tS...gvoeg..... sseavrdg...
SOKASIII 355 n...lhDNS tS...spLae..... seevrg...
PSPCHSI 333 T...hlseGN mS...cveFIL qer...kstd
VUCHS 327 T...hlseGN mS...cveFIL qer...kstd
SOYCHSIV 326 T...hlseGN mS...cveFIL qer...kstd
AHGSNG11 204 T...hlseGN mS...cveFIL qer...kstd
PSSTS 330 T...hlseGN mS...cveFIL qer...kstd
ATEST1282 87 s...rLrF...G tS...slwkk ayt...egk...
FAEI 41S s...rLrF...G tS...slwkk ayt...egk...
CONSENSUS T...rL...e...ILLA ATEL... K L...K...

Figure 7. Sequence Comparison of the FAE1 Protein and Related Proteins Starting at Amino Acid Residue 365 of FAE1.

The multiple alignment was obtained with the Genetics Computer Group multiple sequence comparison programs Pileup and Pretty (Devereux et al., 1984). GenBank accession numbers are given in parentheses after each abbreviation. From top to bottom: KASIII from E. coli (ECOFABH, M77744) and spinach (SOKAS III, Z22771), CHS from pea (PSPCHSI, X65333), Vigna unguiculata (VUCHS, X74821), and soybean (SOYCHSIV, L07647), STS from peanut (AHGSNG11, X62299) and Pinus sylvestris (PSSTS, X0753), the putative protein encoded by Arabidopsis ATEST1282 (Z26005), and FAE1. Amino acid residues that are identical or similar in at least six of the sequences are shown in uppercase letters and were used to derive the consensus sequence shown at the bottom. Identical and similar residues in the 50-amino acid conserved region starting from amino acid residue 391 in the FAE1 protein are shown shaded and dotted, respectively. Less conserved regions are represented with dashes in the consensus sequence. The 11-amino acid box demarcated with dashes corresponds to the CHS and STS "signature sequence" (Fliegmann et al., 1992). Numbers at left correspond to the number of the amino acid residue in each protein. Dots in each sequence correspond to gaps in the amino acid sequence introduced by the computer program to improve alignments.

al., 1988). DNA was obtained from progeny of fael-G309 that contained the 3.3-kb Ac-hybridizing HindIII fragment that cosegregated with the mutant fael phenotype and lacked the smaller 2.4-kb Ac-hybridizing band (e.g., DNA samples shown in lanes 1, 2, 3, 7, and 12 in Figure 2, top). Approximately 0.5 µg of genomic DNA was digested with HindIII and ligated overnight at 18°C under dilute conditions (200-µL reaction volume) to favor circularization of the HindIII fragments. Primer FL125, oriented outward from the 5'end of Ac (CGCTTATACGATAACGGTCG), and primer JK30, just 5' from the first HindIII site in Ac (GTCGCTAGTACGGTG), were used to amplify the 1.5 kb of genomic DNA flanking the 5'end of the trAc. The resulting DNA fragment was cloned into pBluescript KS- (Stratagene) and sequenced. It was also used as a probe to isolate two λ clones from an Arabidopsis ecotype WS library of partial Sau3A fragments cloned into Lambda DASH (Stratagene). The clones were 135 and 105 bp in length and overlapped by ∼9 kb. DNA was sequenced with either the Sequenase kit (U.S. Biochemical Corp.) or the T7 oligo (Promega) following the recommendations of the manufacturers.

DNA from various Arabidopsis tissues (developing siliques, leaves, and flowers) was isolated by a phenol–SDS method (Napoli et al., 1990), separated by formaldehyde–agarose gel electrophoresis, and blotted to Duralon-UV membranes (Stratagene). Poly(A) RNA was isolated from 1 g of 2- to 5-week-old green siliques using a PolyATtract system 1000 kit (Promega). An immature silique cDNA library was made from poly(A) RNA using the Lambda ZAP cDNA synthesis kit (Stratagene).

Fatty Acid Analysis

Fatty acid composition was determined by gas chromatography as described previously (James and Dooner, 1990).

ACKNOWLEDGMENTS

We thank Quincy Huoh, Josh Levine, Kevin Clark, and Florin Georgescu for technical assistance; Joyce Hayashi for figure preparation; Rob Narberes for photography; Cara Velton and Sarah Glenn for plant care; and Elliot Meyerowitz for the RFLP probes. The project was supported in part by National Science Foundation Small Business Innovation Research Phase II Grant No. 9122871. This is DNA Plant Technology Research Paper No. 5-27.
REFERENCES


Tagging of \textit{FAE7} in Arabidopsis


Schmidt, R., and Willmitzer, L. (1989). The maize autonomous element \textit{Activator (Ac)} shows a minimal germinal excision frequency of 0.2%–0.5% in transgenic \textit{Arabidopsis thaliana} plants. Mol. Gen. Genet. 220, 17–24.


Directed tagging of the Arabidopsis FATTY ACID ELONGATION1 (FAE1) gene with the maize transposon activator.
D W James, Jr, E Lim, J Keller, I Plooy, E Ralston and H K Dooner

*Plant Cell* 1995;7;309-319
DOI 10.1105/tpc.7.3.309

This information is current as of June 29, 2017