*Proc. Natl. Acad. Sci. USA*
Vol. 95, pp. 9019–9024, July 1998
Plant Biology

# Cloning of the *Arabidopsis ent*-kaurene oxidase gene *GA3*

CHRIS A. HELLIWELL*, CANDICE C. SHELDON*, MARK R. OLIVE*, AMANDA R. WALKER†, JAN A. D. ZEEVAART‡, W. JAMES PEACOCK*, AND ELIZABETH S. DENNIS*§

*Commonwealth Scientific and Industrial Research Organization, Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia; †Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, United Kingdom CB2 3EA; ‡Michigan State University–Department of Energy Plant Research Laboratory, Michigan State University, East Lansing, MI 48824-1312

*Contributed by W. James Peacock, May 14, 1998*

**ABSTRACT** **The *ga3* mutant of *Arabidopsis* is a gibberellin-responsive dwarf. We present data showing that the *ga3-1* mutant is deficient in *ent*-kaurene oxidase activity, the first cytochrome P450-mediated step in the gibberellin biosynthetic pathway. By using a combination of conventional map-based cloning and random sequencing we identified a putative cytochrome P450 gene mapping to the same location as *GA3*. Relative to the progenitor line, two *ga3* mutant alleles contained single base changes generating in-frame stop codons in the predicted amino acid sequence of the P450. A genomic clone spanning the P450 locus complemented the *ga3-2* mutant. The deduced GA3 protein defines an additional class of cytochrome P450 enzymes. The *GA3* gene was expressed in all tissues examined, RNA abundance being highest in inflorescence tissue.**

Gibberellins (GAs) are a large group of diterpenoid compounds, some of which are active plant growth regulators. GAs are involved in the regulation of a number of plant growth and developmental processes including germination, stem elongation, and reproductive development (1). The GA biosynthetic pathway can be conveniently divided into three parts: (*i*) the synthesis of *ent*-kaurene from geranyl geranyl diphosphate by copalyl diphosphate synthase and *ent*-kaurene synthase activities; (*ii*) the conversion of *ent*-kaurene to $GA_{53}$ by cytochrome P450 enzymes; and (*iii*) the further metabolism to biologically active GAs by dioxygenases.

Substantial progress has been made in isolating the genes encoding the enzymes of the GA biosynthetic pathway from a variety of species (2). Copalyl diphosphate synthase and *ent*-kaurene synthase genes have been isolated (3–6) and genes encoding the soluble dioxygenases of several of the later steps in GA biosynthesis have also been isolated—20-oxidases (7–9), a 7-oxidase (10), 3$\beta$-hydroxylases (11, 12), and a 2$\beta$,3$\beta$-hydroxylase (13). The genes encoding the microsomal steps of the GA biosynthetic pathway (Fig. 1) have proved more difficult to isolate. The maize *Dwarf3* gene encodes a cytochrome P450 (14) and mutations in this gene produce a GA-responsive dwarf. The available data suggest that *Dwarf3* may encode a 13-hydroxylase activity, converting $GA_{12}$ to $GA_{53}$. The *dwarf3* mutant has low levels of $GA_{53}$ (15) consistent with a block before $GA_{53}$ synthesis. In feeding experiments, the *dwarf3* mutant responded to $GA_{53}$ but not to *ent*-kaurene (16). Therefore the block in *dwarf3* is between *ent*-kaurene and $GA_{53}$, but the exact point of the lesion in GA biosynthesis has not been defined (14).

The first microsomal step of the GA biosynthetic pathway is catalyzed by *ent*-kaurene oxidase. Pea *lh* mutants are blocked in the oxidation of *ent*-kaurene, *ent*-kaurenol, and *ent*-kaurenal (17), suggesting that *ent*-kaurene oxidase catalyzes three steps in the GA biosynthetic pathway (Fig. 1). *ent*-Kaurene oxidase is a cytochrome P450 enzyme (18, 19).

The *ga3* mutant was identified as a nongerminating, GA-responsive dwarf after an ethyl methanesulfonate mutagenesis of *Arabidopsis* ecotype Landsberg *erecta* (L.*er*) (20). Two alleles (*ga3-1* and *ga3-2*) have been identified, both being recessive. *ga3-1* is a more severe dwarf and requires exogenous GA for good germination rates, whereas *ga3-2* will germinate in the absence of GA (C.A.H. and C.C.S., unpublished results).

We show that the *GA3* gene encodes a cytochrome P450 enzyme which we identify as *ent*-kaurene oxidase on the basis of data which show that *ga3-1* is deficient in *ent*-kaurene oxidase activity.

## MATERIALS AND METHODS

**Plant Material.** A mapping population was generated by crossing the triple mutant line *ttg ga3-2 ch5* (stock no. NW134, Nottingham *Arabidopsis* Stock Centre, Nottingham, U.K.) with either the Niederzenz or RLD ecotypes. $F_2$ plants were selected that showed recombination between either the *TTG* or *CH5* and *GA3* loci. The mapping population generated contained 150 recombinants between *TTG* and *GA3* and 57 recombinants between *GA3* and *CH5*. The *ga3-1* mutant was obtained from the Nottingham *Arabidopsis* Stock Centre, line no. CS60.

Plant transformation was by vacuum infiltration (21). Seeds from vacuum-infiltrated plants were grown on Murashige and Skoog agar plates, supplemented with 50 mg/liter kanamycin and $10^{-5}$ M $GA_3$ where appropriate.

***ent*-Kaurene Measurements and Feeding Experiments.** Plants for *ent*-kaurene measurements were grown in short day conditions initially, then transferred to long days (9). After transfer to long days plants were watered daily with water or $2 \times 10^{-5}$ M tetcyclacis. Rosettes (10–15) were harvested for each treatment (≈1 g dry weight). *ent*-Kaurene levels were measured as described (22) by isotope dilution (23) using combined GC-selected ion monitoring.

Plants used for the feeding experiments were grown initially in short-day conditions. After transfer to long days the GA precursors (*ent*-kaurene, *ent*-kaurenol, and *ent*-kaurenoic acid) or $GA_1$ were applied. The applications were 2.5 $\mu$g per plant for the precursors and 1.0 $\mu$g per plant of $GA_1$. The applications were repeated after 4 days. The stem height was measured after 25 days.

**DNA Methods.** Bacterial artificial chromosomes (BACs) were subcloned by partially digesting with *Sau*3A1 and size-selecting fractions of 2–4 kbp using agarose gel electrophoresis followed by agarase digestion of the gel. Fragments were

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.



Abbreviations: GA, gibberellin; BAC, bacterial artificial chromosome; L.*er*, Landsberg *erecta*.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF047719–AF047721).

§To whom reprint requests should be addressed. e-mail: liz.dennis@pican.pi.csiro.au.

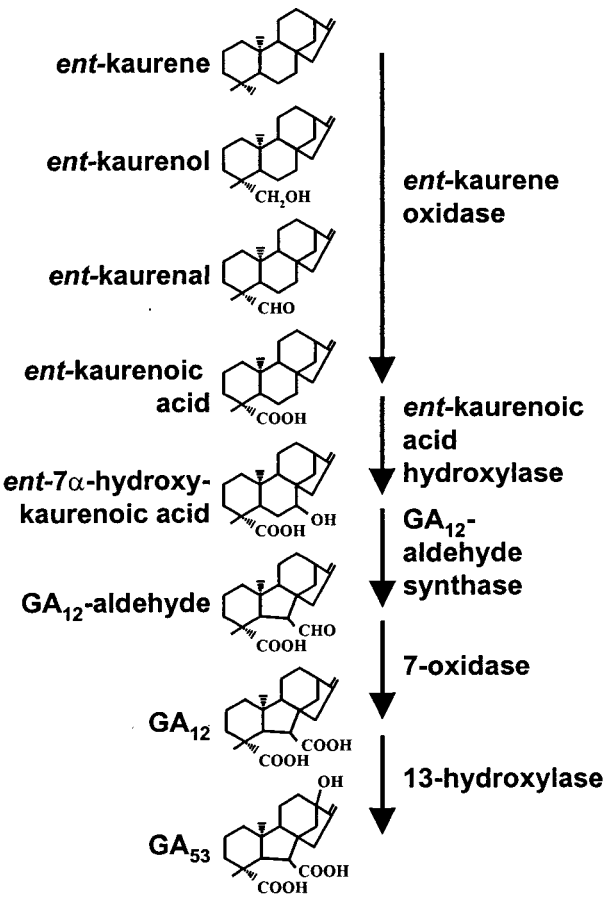


FIG. 1. Outline of the gibberellin biosynthetic pathway from *ent*-kaurene to $GA_{53}$. The reactions catalyzed by *ent*-kaurene oxidase, *ent*-kaurenoic acid hydroxylase, $GA_{12}$-aldehyde synthase, 7-oxidase and 13-hydroxylase are shown together with the structures of the intermediates.

cloned in to pBluescript $KS^+$ (Stratagene). Larger fractions of 6–10 and 10–15 kbp were also isolated and cloned into pBin19 (24).

Large-scale plant genomic DNA extraction was by the method of Dellaporta (25). Small-scale extraction of PCR-grade plant genomic DNA was by the method of Edwards (26). PCR was carried out by using 1 unit/50 $\mu$l reaction of Amplitaq DNA polymerase (Perkin–Elmer) using the manufacturer's buffer, 1.5 mM $MgCl_2$, 250 $\mu$M dNTPs, 1 $\mu$M oligonucleotide primers (shown in Fig. 3, also Columbia ecotype (Col)-specific primer 5′-CTTGAGAGTAATAGGTAGG-3′ and L.*er*-specific primer 5′-AATGTGAATTGACTCGGTG-3′) and $\approx$0.1 $\mu$g genomic DNA. Cycling program; 95°C, 5 min, 95°C, 20 s; 50°C, 15 s; 72°C, 30 s; 35 cycles, 72°C, 5 min. Products were assessed by electrophoresis of a 5 $\mu$l sample on 1% agarose gels. PCR products were prepared for sequencing by using Wizard PCR preps (Promega). Sequencing reactions were carried out by using the Applied Biosystems dRhodamine Dye Terminator sequencing mix according to the manufacturer's instructions and analyzed using an Applied Biosystems 377 sequencing machine. DNA sequences were analyzed by using the GCG suite of programs and BLAST searches were carried out via the Australian National Genome Information Service.

**RNA Methods.** Total RNA was extracted from $\approx$1 g of plant material (27). RNA probes were synthesized by using the Promega Riboprobe system according to the manufacturer's protocol. RNase protection assays were carried out using the Hybspeed RPA kit (Ambion, Austin, TX) with $10^5$ cpm of the *GA3* sense-specific probe and 10 $\mu$g of total plant RNA (determined spectrophotometrically and by ethidium bromide-stained agarose gels) in each reaction, following the manufacturer's protocol. Protected fragments were precipitated and separated on 5% acrylamide/8 M urea gels. Position and intensity of radioactive bands was determined by PhosphorImager.

## RESULTS

**The *ga3-1* Mutant Is Deficient in *ent*-Kaurene Oxidase Activity.** The *ent*-kaurene content of four GA-deficient mutants of *Arabidopsis* together with wild-type L.*er* was measured by isotope dilution by using GC-selected ion monitoring (22, 23). *ent*-Kaurene content was determined at 3 and 5 days after transfer to long days, both with and without the application of the P450 inhibitor tetcyclacis (Table 1). The *ga1-2* and *ga2-1* mutants (deficient in copalyl diphosphate synthase and *ent*-kaurene synthase, respectively) did not show any *ent*-kaurene accumulation under any of the conditions, consistent with them being blocked in *ent*-kaurene synthesis. The wild-type L.*er* had low levels of *ent*-kaurene in the untreated control plants; the level increased $\approx$50-fold with the application of tetcyclacis. The *ga4-1* mutant (deficient in 3$\beta$-hydroxylase activity) also accumulated *ent*-kaurene after tetcyclacis treatment. These results show that the tetcyclacis application was inhibiting the oxidation of *ent*-kaurene. The untreated *ga3-1* mutants had levels of *ent*-kaurene similar to those measured in the tetcyclacis-treated wild-type plants and did not accumulate *ent*-kaurene to higher levels after application of tetcyclacis. Therefore in *ga3-1* the oxidation of *ent*-kaurene must be blocked.

The growth response of the *ga1-2*, *ga2-1,* and *ga3-1* mutants to early intermediates of the GA biosynthetic pathway was tested (Table 2). The *ga1-2* and *ga2-1* mutants both respond to applications of *ent*-kaurene, *ent*-kaurenol, *ent*-kaurenoic acid, and $GA_1$ as expected for mutants blocked in kaurene synthesis. The *ga3-1* mutant did not respond to *ent*-kaurene application, showed a slight response to *ent*-kaurenol application, and a larger response to both *ent*-kaurenoic acid and $GA_1$. The *ga1-2* and *ga2-1* plants treated with *ent*-kaurenol flowered and produced fruits, whereas the *ga3-1* plants showed slight stem elongation, but did not flower. These feeding data confirm the block in *ent*-kaurene oxidation and suggest that the three steps between *ent*-kaurene and *ent*-kaurenoic acid may be blocked in the *ga3-1* mutant.

**Identification of a Cytochrome P450 Gene Mapping to the *GA3* Locus.** Restriction fragment length polymorphism analysis was used in conjunction with a mapping population of recombinants between *ttg* and *ga3-2* or *ga3-2* and *ch5* to place the *GA3* locus between the markers *TINY* (28) and the yeast artificial chromosome end CIC2E3R on chromosome 5 (Fig. 2). A BAC contig was constructed between *TINY* and CIC2E3R. The region (estimated to be $\approx$200 kbp) was spanned by two BACS, T12M10, and F8E12 and a third BAC T1N24 placed between *TINY* and CIC2E3L.

Randomly selected subclones of the two BACS spanning the *GA3* locus (T12M10 and F8E12) were mapped on the contig

Table 1. *ent*-Kaurene content of *Arabidopsis* L.*er* and GA-deficient mutants with and without tetcyclacis treatment

| | *ent*-Kaurene, per pmol/g of dry weight | | | | |
|---|---|---|---|---|---|
| Treatment | *ga1-2* | *ga2-1* | *ga3-1* | *ga4-1* | L. *er* |
| 3 days | | | | | |
| Control | 0 | 0 | 2,879 | 69 | 59 |
| Tetcyclacis | 0 | 0 | 2,750 | 3,470 | 2,266 |
| 5 days | | | | | |
| Control | 0 | 0 | 3,393 | 200 | 57 |
| Tetcyclacis | 0 | 0 | 3,374 | 7,092 | 2,811 |

Numbers are the average of two separate measurements.

Table 2. Effects of *ent*-kaurenoids and $GA_1$ on stem growth of three GA-deficient mutants

| Treatment | Stem height, cm | | |
|---|---|---|---|
| | *ga1-2* | *ga2-1* | *ga3-1* |
| Control | 1.1 ± 0.11 | 2.6 ± 0.12 | 1.0 ± 0.08 |
| *ent*-Kaurene | 7.8 ± 0.47 | 6.3 ± 0.34 | 1.0 ± 0.09 |
| *ent*-Kaurenol | 16.9 ± 1.34 | 14.0 ± 0.90 | 2.4 ± 0.15 |
| *ent*-Kaurenoic acid | 11.6 ± 0.44 | 10.1 ± 0.87 | 7.3 ± 0.67 |
| $GA_1$ | 18.3 ± 0.64 | 11.8 ± 1.23 | 10.1 ± 1.14 |

Numbers are average stem heights ± SEM of 7–13 plants for each treatment taken 25 days after the initial application of the precursors or $GA_1$.

and sequenced from each end. These end sequences were amplified from the parental ecotypes of the mapping population and examined for sequence polymorphisms. A total of eight sequences were tested in this manner, six of them contained at least one sequence polymorphism between the parental ecotypes. The polymorphisms were used to fine map the *GA3* locus to the F8E12 BAC by sequencing the corresponding regions of relevant plants from the mapping population.

The final part of the strategy was to search for a P450-encoding gene in the F8E12 BAC. None of the P450-like cDNAs in the *Arabidopsis* expressed-sequence tag collection hybridized to F8E12 (data not shown) so we initiated random sequencing of the BAC. We generated 92 random sequences of F8E12 subclones, each at least 500 bp in length. These sequences, along with preliminary sequence of T1N24 (http://genome.wustl.edu/gsc/) were used to search the GenBank database for putative P450 genes using the BLAST algorithm. This search identified a putative P450 gene located on both the F8E12 and T1N24 BACs. The putative P450 gene was in a region that had been demonstrated to cosegregate with *GA3* by using sequence polymorphism markers (Fig. 2).

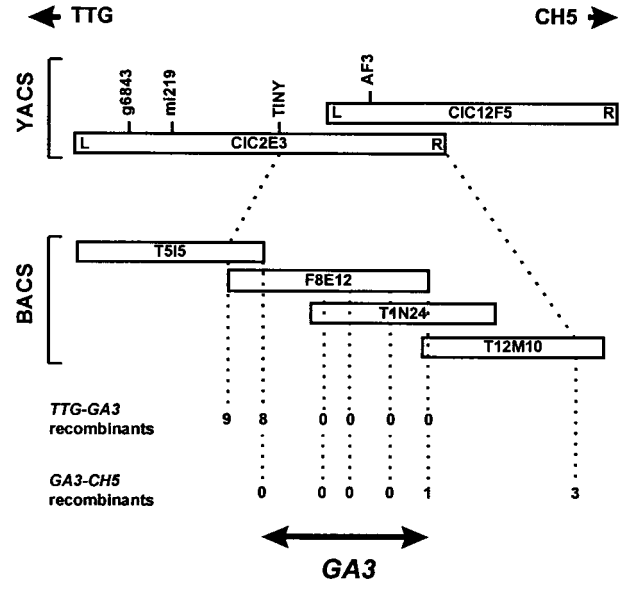


FIG. 2. Yeast artificial chromosome and BAC contig across the *GA3* locus. The yeast artificial chromosomes CIC2E3 and CIC12F5 were isolated from the CIC yeast artificial chromosome library (29). The BACS were either from the TAMU (T prefix) or IGF (F prefix) libraries (30). Dotted lines indicate molecular markers used in the analysis of the recombinant mapping population. The recombinant scores given for the markers are the numbers of plants that did not show segregation of the marker with *GA3*. The data for all recombination events between the *TTG* and *GA3* loci and for the *GA3* and *CH5* loci are summed.

**The *ga3* Alleles Contain Point Mutations.** The region containing the putative *GA3* gene was amplified as four overlapping fragments from wild-type L.*er* and from the mutant lines *ga3-1* and *ga3-2*. These fragments were directly sequenced from both ends. Fig. 3 shows the nucleotide sequence of the

```
   1 TGACATTCGCTGGACCCGTTAATCCAACTTGGAAAATGGTCCGATTACCTCTATTAATTG
  61 ATGATCCAAAGAAAAACTCAATTATGTCGAGATAGGTAAAAAAACATTATTTAAAAGAA
 121 ATAATGAATTTAAAATTAGTAAATAAAATTTAAAATTTAAAAATCCTCCTATAAATTTTG
 181 TGTCTCACAACTTTCTTTCAAACTTTTTTTCTTCAAAAATATACCGTATGCTGATGAAAT
 241 TTTTATCTTTACCAATTAAGAATATAATCTTTTGTTTCCATCTCAAATCATCATCATCAC
                          primer1 >
 301 CTCCAAAGCGTTTAGCAACCCCAAAACCAACCACTTGGCGACTACACAGTATTTTACTCA
 361 AAACATCACAAGTTTCTATGAAACGTAAAACGATGAGTTTAGTAAAAAAAAACTAGAGAT
 421 TATAGTCAGAAAAACTATAATAATAAGAAGATAACATTAAACGGATAGATTCATGTCTCT
 481 TTTTTGAAATCCAAAATTGGGTAAAATCCAAAAACCAAAATCTTGCCAAATGCCAATAAG
 541 AGCAAAAATCCAAACTTTCCTTTCTCTTCACTATTACACAATGAAATCAACACTACATTT
 601 TCACATCAAAATCACATTCCTTCCATGGCCTTCTTCTCCATGATCTCCATTCTCCTTGGC
                                  M  A  F  F  S  M  I  S  I  L  L  G   12
 661 TTTGTTATCTCCTCCTTCATCTTCATCTTCTTCTTCAAGAAACTTCTCTCCTTCTCCAGA
      F  V  I  S  S  F  I  F  I  F  F  F  K  K  L  L  S  F  S  R   32
 721 AAGAACATGTCTGAAGTCTCCACTCTCCCCTCTGTTCCAGGTTTTTTTTTTTTTTTTCC
      K  N  M  S  E  V  S  T  L  P  S  V  P  V                     46
 781 GATTTTGTAATCTGGGTTTTGCGATTCATAGGTTTGGTGACTACTTCTGTTTCAAAGTTT
 841 GAAACTTTCTTTGATTCTAGTGATTTATGGAACTAAAAATTTGATCTTTGAGAACTTTAA
                                                 primer2 >
 901 AAGTTGGGTTTTTGATGTAGTGGTACCAGGGTTTCCTGTTATTGGGAACTTGCTGCAACT
                                  V  P  G  F  P  V  I  G  N  L  L  Q   59
 961 AAAAGAGAAGAAACCTCACAAGACTTTCACTAGATGGTCAGAGATTTATGGTCCTATTTA
      K  E  K  K  P  H  K  T  F  T  R  W  S  E  I  Y  G  P  I  Y   79
                                        < primer 3
1021 CTCTATAAAGATGGGTTCTTCTTCTCTTATTGTCCTCAATTCTACTGAGACTGCCAAAGA
      S  I  K  M  G  S  S  S  L  I  V  L  N  S  T  E  T  A  K  E   99
1081 GGTTAGCTTTCTAGGTTTGATTGATATGTTTGGAATCGAAAAATCGTTCCTTTTTCGATC
1141 TTGAACACATGACTAAAGCCTTACACAACTTGATGATTGCTCTTGTTTATATGATTGTAG
1201 TCTGAACTTATGTAGAAAAGTTGAGTTGATGTCTTCTCTTATCGTTTAGGTTACTCTTTA
1261 TGTGTTTGTTTTCAGGCCATGGTGACGCGGTTTTCGTCTATCTCAACGAGGAAGTTGTCA
                                  A  M  V  T  R  F  S  S  I  S  T  R  K  L  S  114
1321 AATGCGTTGACAGTCCTTACTTGTGACAAATCTATGGTTGCTACTAGTGATTATGATGAT
      N  A  L  T  V  L  T  C  D  K  S  M  V  A  T  S  D  Y  D  D  134
1381 TTCCACAAGTTGGTGAAACGGTGTCTCTTGAACGGTCTTTTGGGTGCTAATGCACAGGTG
      F  H  K  L  V  K  R  C  L  L  N  G  L  L  G  A  N  A  Q     153
1441 AATTCATTGTTCGTTTATCCCGCAGCTTCTCCTTGGTATACATAATGGCAACTGAATCAA
1501 TCTTTTGTGTTGTGTTGGTTTTTTTTGCAGAAACGAAAAAGACATTACAGAGATGCACTC
                                  K  R  K  R  H  Y  R  D  A  L  163
                                                    TAA(ga3-1)
1561 ATTGAAAATGTGTCTTCCAAGTTGCATGCCCATGCTAGGGACCATCCACAAGAACCTGTA
      I  E  N  V  S  S  K  L  H  A  H  A  R  D  H  P  Q  E  P  V  183
      primer4 >
1621 AACTTCAGAGCTATATTTGAGCATGAGCTTTTCGGTGTAGCATTGAAGCAAGTGAGTGTT
      N  F  R  A  I  F  E  H  E  L  F  G  V  A  L  K  Q           200
                                        < primer5
1681 TCGCTTATCCTATACATCATTTCTTCGTCAACATAAAGTATGAAAATGGATTTTTGGTTT
1741 TGTTTTCATAGGCTTTTGGGAAAGATGTGGAATCCATTTATGTTAAAGAACTCGGTGTGA
                  A  F  G  K  D  V  E  S  I  Y  V  K  E  L  G  V  T  217
1801 CTTTGTCGAAAGACGAGATCTTCAAGGTTTTAGTACATGACATGATGGAAGGTGCAATTG
      L  S  K  D  E  I  F  K  V  L  V  H  D  M  M  E  G  A  I  D  237
                                        TGA(ga3-2)
1861 ATGTTGATTGGAGAGACTTCTTCCCATACTTGAAATGGATTCCAAATAAAAGTTTTGAAG
      V  D  W  R  D  F  F  P  Y  L  K  W  I  P  N  K  S  F  E  A  257
1921 CAAGAATCCAGCAAAAGCATAAACGTAGACTCGCGGTGATGAATGCTCTGATTCAAGATC
      R  I  Q  Q  K  H  K  R  R  L  A  V  M  N  A  L  I  Q  D  R  277
1981 GACTGAAGCAGAATGGTTCAGAATCGGTTAGTTTTGTTTGAATAGTCTTTCATATATTGC
      L  K  Q  N  G  S  E  S                                       285
2041 ATTACACTGCTGATCTAATCTCTACTTATAAACAGGATGATGATTGCTATCTCAACTTCT
                              D  D  D  C  Y  L  N  F  L  294
2101 TGATGTCGGAAGCGAAAACACTAACCAAGGAGCAAATTGCTATCTTGGTTTGGGAGACGA
      M  S  E  A  K  T  L  T  K  E  Q  I  A  I  L  V  W  E  T  I  314
                                                 primer6 >
2161 TTATCGAGACAGCTGACACTACTTTGGTTACAACTGAATGGGCCATCTATGAGCTCGCTA
      I  E  T  A  D  T  T  L  V  T  T  E  W  A  I  Y  E  L  A  K  334
2221 AGCATCCAAGTGTCCAAGATCGTCTGTGTAAAGAAATCCAAAATGTCTGCGGAGGAGAAA
      H  P  S  V  Q  D  R  L  C  K  E  I  Q  N  V  C  G  G  E  K  354
                                        < primer7
2281 AGTTCAAAGAAGAGCAATTGTCTCAAGTTCCTTATCTCAATGGAGTATTCCATGAAACGC
      F  K  E  E  Q  L  S  Q  V  P  Y  L  N  G  V  F  H  E  T  L  374
2341 TTAGGAAATACAGTCCTGCTCCTCTAGTCCCCATTCGCTATGCCCACGAAGATACGCAAA
      R  K  Y  S  P  A  P  L  V  P  I  R  Y  A  H  E  D  T  Q  I  394
2401 TCGGAGGCTATCATGTCCCTGCAGGAAGTGAGGTACACCGAGTCAATTCACATTTTGAAG
      G  G  Y  H  V  P  A  G  S  E                                 404
                                        < alternately spliced
2461 TTCATTTACAAAGTCTAGTTTCAAAATCTTGATGTGTAGATTGCAATAAACATCTATGGA
                                        I  A  I  N  I  Y  G  411
   region of intron 6 >
2521 TGCAACATGGATAAGAAGCGTTGGGAGAGACCAGAGGACTGGTGGCCGGAGCGGTTTCTT
      C  N  M  D  K  K  R  W  E  R  P  E  D  W  W  P  E  R  F  L  431
2581 GATGATGGCAAATACGAAACGTCGGATCTTCACAAGACAATGGCGTTTGGAGCGGGAAAG
      D  D  G  K  Y  E  T  S  D  L  H  K  T  M  A  F  G  A  G  K  451
2641 AGGGTTTGTGCTGGTGCTCTTCAAGCATCTCTCATGGCAGGCATTGCCATTGGGAGGTTA
      R  V  C  A  G  A  L  Q  A  S  L  M  A  G  I  A  I  G  R  L  471
2701 GTGCAAGAATTCGAGTGGAAGCTTAGAGACGGTGAAGAAGAGAATGTGGATACATATGGC
      V  Q  E  F  E  W  K  L  R  D  G  E  E  E  N  V  D  T  Y  G  491
2761 TTGACCTCTCAGAAGCTTTATCCTCTTATGGCTATTATCAATCCAAGGCGTTCTTAAGAC
      L  T  S  Q  K  L  Y  P  L  M  A  I  I  N  P  R  R  S  *     509
                                        < primer8
2821 AGAAATCTCTTCTTTCTTTTGTTAGTTTTTAACCTTGTATCAACTTTTATTTGTAACAAA
2881 TGCTGTTTTGTAGTACTATTTATGTTTTACTGTGTTATATTTTACTATCCTCTGAGATTG
```

FIG. 3. Nucleotide sequence of *GA3*. The nucleotide sequence of the *GA3* gene from the L.*er* ecotype of *Arabidopsis* is shown with the sequence of the L.*er* cDNA clone underlined. The predicted amino acid sequence is shown under the nucleotide sequence. The sequence of the eight oligonucleotide primers used in the sequencing of the *ga3* mutant alleles is shown italicized with the direction indicated by < or >. The nucleotide changes in the *ga3-1* and *ga3-2* mutants are shown above the sequence with the changed nucleotide underlined. The alternatively spliced region 3′ of intron 6 is shown italicized and indicated above the sequence. Putative TATA boxes are shown underlined at bases 438 and 535. Nucleotides are numbered on the left, amino acids on the right.

P450 gene from L.*er*, together with the primers used to amplify the gene. Comparison of the L.*er* and mutant sequences identified a single base change in each of the mutants. In *ga3-1* there was a change of C to T (nucleotide 1,609), leading to the introduction of an in-frame stop codon in place of a glutamine residue in the predicted ORF of the gene (Fig. 3). Similarly a change of G to A (nucleotide 1898) in the *ga3-2* mutant introduced a stop codon in place of tryptophan in the predicted reading frame. The base change in *ga3-2* also generated an *Eco*RI site not present in wild-type L.*er*. This site was used to generate a cleaved amplified polymorphic sequence marker for the *ga3-2* allele, which confirmed that the P450 gene cosegregated with the *ga3-2* dwarf phenotype in the mapping population.

The gene sequence was also determined from the Columbia ecotype. Comparison of the Col and L.*er* sequences identified 15 single base differences between the two ecotypes within the putative protein-coding region. None of these differences changed the putative amino acid sequence. The Col allele of the gene also contained an extra 12 bp within the putative intron 6. DNA gel blot analysis at high stringency showed that there is a single copy of the *GA3* gene in *Arabidopsis* (data not shown).

**The Isolated P450 Gene Complements the *ga3* Mutation.** A complementation test confirmed that the P450 gene is the *GA3* gene. A genomic clone from the Columbia ecotype containing the whole coding region together with 2.4 kbp of 5′ flanking region and 3.0 kbp of 3′ flanking region was constructed. This fragment did not contain any other P450 sequences The *Arabidopsis* line used for transformation was homozygous for the *ga3-2* and *ttg* (transparent testa glabra, distinguishable by lack of trichomes and pale seed coats) mutations.

Among 8,000 seeds obtained after the transformation treatment, 19 kanamycin resistant plants were identified and transferred to soil. The *ttg ga3-2* mutant line and wild-type L.*er* were grown under the same conditions without kanamycin selection. The *ttg ga3-2* mutant line showed typical *ga3-2* phenotypes, including dwarfism, reduced leaf elongation, dark green leaves, poor flower formation, and short siliques. These plants also lacked trichomes and had yellow seeds due to the presence of the *ttg* mutation. The 19 kanamycin resistant plants were of wild-type height, had more elongated leaves, were the same shade of green as the wild type, had normal flower formation, and elongated siliques. They all lacked trichomes and had yellow seed, indicating that they were homozygous for the *ttg* mutation, showing they were from the *ttg ga3-2* line used for the transformation process. Three of the kanamycin resistant plants were selected for further analysis; these plants are shown together with L.*er* and *ttg ga3-2* in Fig. 4*A*. A PCR method showed that the kanamycin resistant plants contained the Columbia *GA3* transgene. The extra 12 bp in intron 6 of the Col copy of the gene enabled oligonucleotide primers to be designed to differentiate between the Col and L.*er* genes. These were used in conjunction with primer 4 (Fig. 3) to amplify a DNA fragment from the region containing the point mutation in *ga3-2*. The results of this analysis are shown in Fig. 4*B*. PCR using Col specific primer shows that the Col gene is present in the kanamycin resistant plants. PCR using the L.*er* specific primer shows that the L.*er* copy of the gene is also present in the resistant plants as expected. Digestion of the L.*er* specific PCR products with *Eco*RI shows that the L.*er* copies of the gene have the *ga3-2* point mutation. This confirms that the kanamycin resistant plants were mutant at the *GA3* locus and were restored to wild-type phenotype when the Col transgene is present.

**The *GA3* cDNA Encodes an Additional Class of Cytochrome P450 Protein.** The PCR products generated to sequence the L.*er GA3* gene were used as probes against a cDNA library constructed from mRNA prepared from developing siliques of L.*er* with embryos at the 8- to 16-cell stage (A. Koltunow, CSIRO Plant Industry, Adelaide, Australia). This library was chosen because the *ent*-kaurene biosynthesis rate is high early in seed development in *Arabidopsis* (31, 32). A cDNA was isolated that was 1,678 bp in length (Fig. 3) and encoded a putative polypeptide of 509 amino acids. Comparison of the cDNA sequence to the L.*er* genomic sequence showed six introns in the gene. An in-frame start codon is present in a sequence corresponding to a good translation start consensus, with putative TATA box motifs upstream. The amino acid sequence contains conserved P450 motifs (33) and has a hydrophobic N terminus, consistent with it being localized in the endoplasmic reticulum.

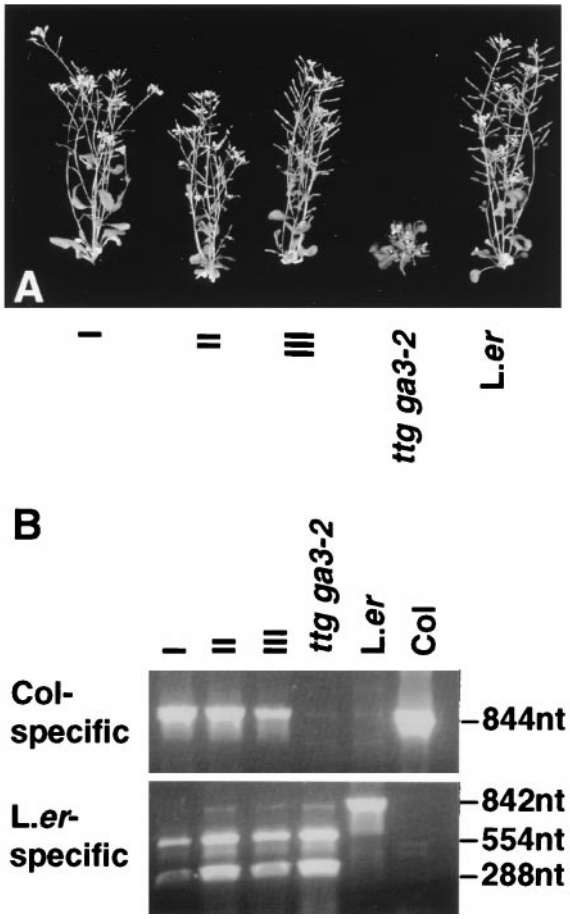


FIG. 4. Complementation of the *ga3-2* mutant. (*A*) Photographs of three kanamycin resistant plants from the transformation of *ga3-2* with a genomic clone containing the P450 gene (I-III), the *ga3-2 ttg* line and wild-type L.*er*. All plants were germinated on medium containing $GA_3$ (the seeds from the transformation were germinated on medium that also contained kanamycin), transferred to medium without GA after cotyledons were expanded (11 days) and transferred to soil after a further 13 days growth. Photographs were taken 10 days after transfer to soil. (*B*) PCR analysis of kanamycin resistant plants. PCR was carried out using primers specific to the L.*er* and Col copies of the gene. L.*er* specific products were digested with *Eco*RI to distinguish between *ga3-2* and *GA3* copies of the gene.

The putative amino acid sequence was used to search the protein database. The strongest similarities were with other cytochrome P450s. The highest amino acid identity was 28% with CYP77A14 from eggplant (34), a cytochrome P450 of unknown function. There were no P450 amino acid sequences with >40% amino acid identity to GA3 and therefore it falls into a new class of cytochrome P450 proteins (35). It is noteworthy that the identity to the maize Dwarf3 protein, also implicated as being a GA biosynthetic P450, is only 20%.

A number of truncated *GA3* cDNAs were also isolated and sequenced. Of five cDNA clones analyzed, two had differential

splicing of intron 6. The presumed aberrant splice results in the removal of an extra 39 bp at the 3′ end of the intron. The location of the second 3′ splice site is shown in Fig. 3. The aberrant splicing maintains the ORF of *GA3*, however the 13 amino acids lost in the putative polypeptide produced from this aberrant splice are conserved in other P450s. This suggests that the polypeptide produced from the aberrantly spliced mRNA may have reduced or no activity.

**Expression of the *GA3* mRNA Is Developmentally Regulated.** RNase protection was used to analyze the expression pattern of the *GA3* mRNA. The probe extended from 14 bp 5′ of intron 6 to 225 bp 3′ of the intron. This probe detects both splicing products from the intron, giving protected fragments of 225 bp for the shorter splice or 186 bp for the longer splice. Fig. 5*A* shows expression in young seedlings; the transcript abundance decreased ≈50% as the seedlings aged from 5 to 19 days. The transcript abundance was relatively low in both rosette and cauline leaves (Fig. 5*B*), but increased ≈4-fold in elongating stems, and 8-fold in inflorescence above the rosette leaf measurement. Expression was also examined in inflorescences of L.*er* with and without GA treatment to investigate whether there was evidence of negative feedback by GA on the expression of the *GA3* mRNA as has been reported for the mRNAs of two C-20 oxidases operating later in the pathway (8). No significant difference between expression levels was observed in the RNA from GA-treated plants compared with the untreated L.*er* plants (Fig. 5*C*). The expression of *GA3* mRNA was also measured in the *ga1-3*, *ga3-1*, *ga4-1,* and *ga5-1* mutants (Fig. 5*C*). The abundance of the message was reduced to <50% of that in L.*er* in the *ga3-1* mutant. This may be due to an effect on the stability of the *GA3* mRNA. No significant changes in *GA3* transcript abundance were detected in the other three mutants. These experiments were also carried out for leaf tissue of L.*er* and the mutants (data not shown), again showing no effect of GA treatment or of the presence of the *ga* mutations on the *GA3* mRNA level.

In all cases the relative abundance of the two differently spliced mRNAs was the same, the presumed correctly spliced message being present at ≈1.5-fold the abundance of the aberrantly spliced message (taking into account the difference in specific activity of the two protected fragments).

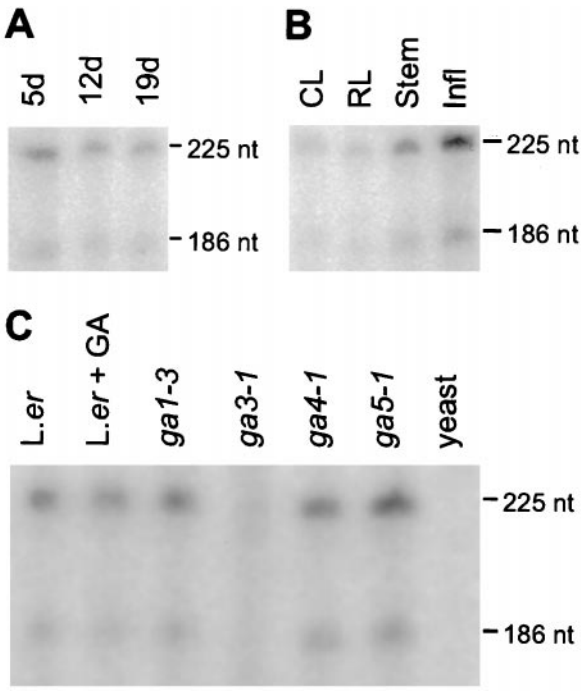


FIG. 5. Expression of the *GA3* mRNA. The abundance of the *GA3* transcript was determined by RNase protection. For each panel the amount of input RNA for each lane was shown to be equal by spectrophotometer and ethidium stained agarose gel analysis. The sizes of the protected fragments from the full-length mRNA (225 nt) and the product of aberrant splicing of intron 6 (186 nt) are shown. Sizes of protected fragments were determined by comparison to end-labeled single stranded DNA fragments from an *Msp*I digest of pBR322. (*A*) Expression of the GA3 mRNA in *Arabidopsis* (L.*er*) seedlings at 5, 12, and 19 days old. RNA was prepared from the aerial parts of the plants. (*B*) Expression of the GA3 mRNA in tissues of flowering *Arabidopsis* (L.*er*) plants. RL; rosette leaf, CL; cauline leaf, Stem; elongating stems stem, Infl; inflorescence. (*C*) Effect of gibberellic acid ($GA_3$) and *ga* mutations on the expression of the GA3 mRNA. RNA was prepared from the inflorescence of flowering plants. Yeast control reaction contained 50 μg of yeast RNA hybridized to the GA3 probe and treated with RNase.

## DISCUSSION

We have demonstrated that the *ga3-1* mutant of *Arabidopsis* is deficient in *ent*-kaurene oxidase activity. The mutant accumulates *ent*-kaurene to the same level as tetcyclacis-treated wild-type plants and shows a growth response to *ent*-kaurenoic acid but not to *ent*-kaurene. It is therefore likely that *GA3* encodes *ent*-kaurene oxidase, the first cytochrome P450-mediated step of the GA biosynthetic pathway (18). We isolated the *GA3* gene by mapping the gene to a single BAC and used DNA sequence data to identify a single putative cytochrome P450 gene located on that BAC. We sequenced the P450 gene from wild-type L.*er* and from the two mutant alleles of *ga3*. Each of the mutant alleles contained a single base change in the P450 gene. In both cases the mutation introduced a stop codon into the predicted ORF of the P450 gene. To confirm that this P450 was the *GA3* gene, we complemented the *ga3-2* mutant with a wild-type genomic clone. The transformed plants containing the transgene were restored to the wild-type phenotype.

Our data suggest that there is a single *ent*-kaurene oxidase activity in *Arabidopsis*. DNA gel blots show that *GA3* is a single copy gene and application of tetcyclacis to the *ga3-1* mutant does not result in increased *ent*-kaurene accumulation. This contrasts with the GA 20-oxidase, catalyzing later reactions in the pathway, for which there are three genes in *Arabidopsis* (8). If there is an alternative *ent*-kaurene oxidase gene it is not closely related to *GA3* and the gene product may not be a cytochrome P450 as its activity is not inhibited by tetcyclacis.

Swain *et al.* (17) showed that the pea *lh-2* mutant is blocked in the three steps from *ent*-kaurene to *ent*-kaurenoic acid. Extracts from embryos of the *lh-2* mutant were unable to oxidize *ent*-kaurene, *ent*-kaurenol, and *ent*-kaurenal but were able to synthesize *ent*-kaurene and also to oxidize *ent*-kaurenoic acid to *ent*-7α-hydroxy kaurenoic acid, suggesting that *ent*-kaurene oxidase catalyzes all three reactions. Our feeding data show that *ga3-1* responds to *ent*-kaurenoic acid but not to *ent*-kaurene. A slight growth response to *ent*-kaurenol was observed, but notably did not lead to flowering as was the case when *ga1-2* and *ga2-1* were treated in the same way. This slight response to *ent*-kaurenol may be due to spontaneous oxidation or another oxidase activity that can act on *ent*-kaurenol but not *ent*-kaurene.

We have observed that the *ga3-1* mutant has a more severe phenotype than *ga3-2*. The putative polypeptides encoded by both alleles are truncated before the heme-binding domain and would therefore not be expected to have *ent*-kaurene oxidase activity. However the less severe *ga3-2* phenotype suggests that the *ga3-2* gene product is capable of conferring some *ent*-kaurene oxidase activity. This could be by alternate splicing out of exon 5, which would restore the heme-binding domain or be due to an interaction of the truncated polypeptide with other proteins.

The expression pattern of *GA3* mRNA is consistent with its role in GA biosynthesis. The transcript is high in young

seedlings where there is a high proportion of dividing and expanding cells and declines with increased age to a low level in mature leaves. The *GA3* transcript level is also high in the elongating stems and further increased in the inflorescence. Our cDNA library screening also indicates that *GA3* mRNA is expressed in developing seeds. GAs are particularly important in germination, stem growth, and flowering.

A detailed study of expression of the *GA1* gene (36) has shown that the *GA1* promoter directs highest GUS expression in the rapidly growing tissues of transgenic *Arabidopsis* plants including embryos, meristems, anthers, and vascular tissue. *GA1* encodes copalyl diphosphate synthase, a chloroplast-located enzyme catalyzing the first step of *ent*-kaurene synthesis. Our expression data suggest that the expression patterns of *GA1* and *GA3* may be closely related, as might be expected of genes encoding enzymes catalyzing successive steps in the GA biosynthetic pathway.

The three GA 20-oxidases of *Arabidopsis* show differential tissue-specific expression; mRNA levels of at least one of the genes are high in stems, inflorescences, and siliques (8). Expression levels of the mRNAs of two of the 20-oxidase genes are increased in the *ga1-2* mutant, and decreased in these plants by the application of endogenous GA, suggesting that there is feedback regulation in the pathway. We found no evidence of this type of regulation of the *GA3* gene.

Comparison of the derived amino acid sequence of the *GA3* gene to other P450 enzymes shows that it encodes the first member of a new class of cytochrome P450 enzymes. The GA3 protein is not closely related to the maize Dwarf3 protein, also thought to be a GA biosynthetic P450. The number and positions of the introns in the *GA3* gene differ from those in other known plant P450 genes. We plan to use the *GA3* gene as a probe at reduced stringency to isolate related P450s. *ent*-Kaurenoic acid hydroxylase, the enzyme catalyzing the conversion of *ent*-kaurenoic acid to *ent*-7$\alpha$-hydroxy kaurenoic acid is also a cytochrome P450 mediated reaction. There are biosynthetic pathways in which adjacent steps are catalyzed by related P450s (37, 38), so it is possible that the *ent*-kaurenoic acid hydroxylase amino acid sequence may be related to that of *ent*-kaurene oxidase.

Data from *Thlaspi arvense*, a species closely related to *Arabidopsis*, demonstrate that the activities of *ent*-kaurene oxidase and *ent*-kaurenoic acid hydroxylase increase specifically in the shoot tip after vernalization (39). These changes in enzyme activities are coupled with a reduction in the pools of *ent*-kaurene and *ent*-kaurenoic acid in the shoot tip (40). This suggests that *ent*-kaurenoic acid hydroxylase and possibly *ent*-kaurene oxidase have a controlling role in the GA biosynthetic pathway and that the enzyme activities are under control of low temperature.

We thank Ying Luo and Janice Norman for technical assistance. C.A.H. was supported in part by a postdoctoral fellowship from the Royal Society, United Kingdom. J.A.D.Z. was supported by U.S. Department of Energy Grant DE-FG02–91ER20021.

1. Hooley, R. (1994) *Plant Mol. Biol.* **26,** 1529–1555.
2. Hedden, P. & Kamiya, Y. (1997) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48,** 431–460.
3. Sun, T.-p., Goodman, H. M. & Ausubel, F. M. (1992) *Plant Cell* **4,** 119–128.
4. Sun, T.-p. & Kamiya, Y. (1994) *Plant Cell* **6,** 1509–1518.
5. Yamaguchi, S., Saito, T., Abe, H., Yamane, H., Murofushi, N. & Kamiya, Y. (1996) *Plant J.* **10,** 203–213.
6. Yamaguchi, S., Sun, T.-p., Kawaide, H. & Kamiya, Y. (1998) *Plant Physiol.* **116,** 1271–1278.
7. Lange, T., Hedden, P. & Graebe, J. E. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 8552–8556.
8. Phillips, A. L., Ward, D. A., Uknes, S., Appleford, N. E. J., Lange, T., Huttly, A. K., Gaskin, P., Graebe, J. E. & Hedden, P. (1995) *Plant Physiol.* **108,** 1049–1057.
9. Xu, Y.-L., Li, L., Wu, K., Peeters, A. J. M., Gage, D. A. & Zeevaart, J. A. D. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 6640–6644.
10. Lange, T. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 6553–6558.
11. Chiang, H.-H., Hwang, I. & Goodman, H. (1995) *Plant Cell* **7,** 195–201.
12. Lester, D. R., Ross, J. J., Davies, P. J. & Reid, J. B. (1997) *Plant Cell* **9,** 1435–1443.
13. Lange, T., Robatzek, S. & Frisse, A. (1997) *Plant Cell* **9,** 1459–1467.
14. Winkler R. G. & Helentjaris, T. (1995) *Plant Cell* **7,** 1307–1317.
15. Fujioka, S., Hisakazu, Y., Spray, C. R., Gaskin, P., MacMillan, J., Phinney, B. O. & Takahashi, N. (1988) *Plant Physiol.* **88,** 1367–1372.
16. Phinney, B. O. & Spray, C. (1982) in *Plant Growth Substances*, ed. Wareing, P. F. (Academic, New York), pp. 101–110.
17. Swain, S. M., Reid, J. B. & Kamiya, Y. (1997) *Plant J.* **12,** 1329–1338.
18. Murphy, P. J. & West, C. A. (1969) *Arch. Biochem. Biophys.* **133,** 395–407.
19. Coolbaugh, R. C. & Moore, T. C. (1971) *Phytochemistry* **10,** 2401–2412.
20. Koornneef, M. & van der Veen, J. H. (1980) *Theor. Appl. Genet.* **58,** 257–263.
21. Bechtold, N., Ellis, J. & Pelletier, G. (1982) *C. R. Acad. Sci.* **316,** 1194–1199.
22. Zeevaart, J. A. D. & Gage, D. A. (1993) *Plant Physiol.* **101,** 25–29.
23. Grosselindemann, E., Graebe, J. E., Stöckl, D. & Hedden, P. (1991) *Plant Physiol.* **96,** 1099–1104.
24. Bevan, M. (1984) *Nucleic Acids Res.* **12,** 8711–8721.
25. Dellaporta, S. L., Wood, J. & Hicks, J. B. (1983) *Plant Mol. Biol. Rep.* **1,** 19–21.
26. Edwards, K., Johnstone, C. & Thompson, C. (1991) *Nucleic Acids Res.* **19,** 1349.
27. Longemann, J., Schell, J. & Willmitzer, L. (1987) *Anal. Biochem.* **163,** 16–20.
28. Wilson, K., Long, D., Swinburne, J. & Coupland, G. (1996) *Plant Cell* **8,** 659–671.
29. Creusot, F., Fouilloux, E., Dron, M., Lafleuriel, J., Picard, G., Billault, A., Le Paslier, D., Cohen, D., Chabouté, M.-E., Durr, A., *et al.* (1995) *Plant J.* **8,** 763–770.
30. Choi, S., Creelman, R. A., Mullet, J. E. & Wing, R. A. (1995) *Weeds World* **2,** 17–20.
31. Barendse, G. W. M. & Koornneef, M. (1982) *Arabidopsis Information Service* **19,** 25–28.
32. Barendse, G. W. M., Kepczynski, J., Karssen, C. M. & Koornneef, M. (1986) *Physiol Plant.* **67,** 315–319.
33. Kalb, V. F. & Loper, J. C. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7221–7225.
34. Toguri, T. & Tokugawa, K. (1994) *FEBS Lett.* **338,** 290–294.
35. Nebert, D. W. & Nelson, D. R. (1991) *Methods Enzymol.* **206,** 3–11.
36. Silverstone, A. L., Chiang, C., Krol, E. & Sun, T. (1997) *Plant J.* **12,** 9–19.
37. Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grün, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R. B., Briggs, S. P., Simcox, K. & Gierl, A. (1997) *Science* **277,** 696–699.
38. Choe, S., Dilkes, B. P., Fujioka, S., Takatsuto, S., Sakurai, A. & Feldmann, K. A. (1998) *Plant Cell* **10,** 231–243.
39. Hazebroek, J. P., Metzger, J. D. & Mansager, E. R. (1993) *Plant Physiol.* **102,** 547–552.
40. Hazebroek, J. P. & Metzger, J. D. (1990) *Plant Physiol.* **94,** 157–165.