Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 3338–3342, April 1995
Evolution

Evolution of the chalcone synthase gene family in the genus *Ipomoea*

(gene phylogenies/morning glory/anthocyanin/flavonoids)

MARY L. DURBIN, GERALD H. LEARN, JR.*, GAVIN A. HUTTLEY†, AND MICHAEL T. CLEGG‡

Department of Botany and Plant Sciences, University of California, Riverside, CA 92521-0124

Contributed by Michael T. Clegg, December 23, 1994

ABSTRACT **The evolution of the chalcone synthase [CHS; malonyl-CoA:4-coumaroyl-CoA malonyltransferase (cyclizing), EC 2.3.1.74] multigene family in the genus *Ipomoea* is explored. Thirteen CHS genes from seven *Ipomoea* species (family Convolvulaceae) were sequenced—three from genomic clones and the remainder from PCR amplification with primers designed from the 5′ flanking region and the end of the 3′ coding region of *Ipomoea purpurea* Roth. Analysis of the data indicates a duplication of CHS that predates the divergence of the *Ipomoea* species in this study. The *Ipomoea* CHS genes are among the most rapidly evolving of the CHS genes sequenced to date. The CHS genes in this study are most closely related to the *Petunia* CHS-B gene, which is also rapidly evolving and highly divergent from the rest of the *Petunia* CHS sequences.**

Chalcone synthase [CHS; malonyl-CoA:4-coumaroyl-CoA malonyltransferase (cyclizing), EC 2.3.1.74] is a key enzyme in the biosynthesis of flavonoids. CHS catalyzes the condensation of three acetate residues from malonyl-CoA with *p*-coumaroyl-CoA to form naringenin chalcone (1). This is the initial step of the phenylpropanoid pathway that leads to the production of flavonoids. Flavonoids have many functions in plants. They are important for the pigmentation of flowers and, hence, act as attractants to pollinators (2). Flavonoids also play an important role in protection against UV light (3), plant pathogen defense (4, 5), induction of nodulation (6), auxin transport (7), pollen function (8), and insect resistance (9).

CHS has been shown to be under a complex system of regulation, and a number of potential regulatory elements have been identified (10). To date, there are CHS sequences from 16 species in the GenBank data base. Another closely related gene, stilbene synthase (SS), has also been sequenced but from only three species. The SSs are thought to have evolved independently from CHS at least three times in seed plant evolution (11). The two enzymes use identical substrates but yield different products. Only a small number of amino acid changes are necessary to convert CHS to SS function (11). The two enzymes are results of a process whereby a member of a gene family diverges and evolves a new function.

Common morning glory (*Ipomoea purpurea* Roth) exhibits a number of flower color polymorphisms in populations of the southeastern United States (12). These range from fully pigmented pink or purple flowers to white and variegated phenotypes. In contrast, *I. purpurea* populations in Mexico are typically monomorphic for the purple color phenotype (unpublished data). Much work has been done on the ecological genetics of these flower color polymorphisms, including characterizing the effect of flower color on pollinator preference and the selective advantages or disadvantages of the genes that determine flower color (13, 14). White flowers have reduced pollinator visitation when the white phenotype is a minority in the population (15–17). Self-fertilization rates increase in plants with white flowers relative to those with darkly pigmented flowers under conditions of reduced pollinator visitation (18). Hence, pollinator preference for certain flower color phenotypes can influence the transmission of genes within a population.

The diversity in flower color is almost certainly due to differences in either the structural or the regulatory genes of the flavonoid biosynthetic pathway. The genetics of the variegated phenotype, for example, has been shown to be consistent with the presence of an active transposable element (19, 20). The white phenotypes presumably contain blockages at some point in the flavonoid biosynthetic pathway. As a first step in linking ecological information with molecular changes, we examine in this report the evolution of CHS in six *Ipomoea* species (Convolvulaceae). We have sequenced three CHS genes and a fourth, probable pseudogene, from *I. purpurea*. Other CHS sequences were also obtained from *Ipomoea nil* (also known as *Pharbitis nil*), *Ipomoea cordatotriloba*, *Ipomoea triloba*, *Ipomoea trifida*, and *Ipomoea platensis*.§ We address the following questions based on comparison of these gene sequences. What factors modulate the evolution of the CHS gene family? Are there conserved domains in CHS between these species that may indicate regions of functional importance? How do the CHS genes in *Ipomoea* compare with other CHS and SS genes previously studied?

MATERIALS AND METHODS

All *I. purpurea* plant materials used were derived from isogenic lines described elsewhere (14) and grown under greenhouse conditions. *I. platensis* was obtained from the teaching collection at the University of California, Riverside. Genomic DNA from *I. nil* was provided by Sharman O'Neill (University of California, Davis) and DNA from *I. triloba*, *I. trifida*, and *I. cordatotriloba* was provided by Deborah Glover (University of California, Riverside).

Three *I. purpurea* CHS clones were isolated from a λEMBL3 genomic library by plaque hybridization with a tomato CHS clone kindly provided by Sharman O'Neill and colleagues (21). The λEMBL clones were subcloned into pUC119 and sequenced by the Sanger dideoxynucleotide chain-termination method (22). Primers designed from *I. purpurea* CHS-A were used to PCR amplify and sequence CHS from several other *Ipomoea* species. The numbering of the primers corresponds to the translation start site in CHS-A. The 5′ forward primers were −52F (5′-ACATTTCTCACTAAACCGTGAAGCC-3′) and −150F (5′-ATTTGAGCCCTTCCATGCCC-3′). The

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: CHS, chalcone synthase; SS, stilbene synthase.

*Present address: Department of Microbiology SC-42, University of Washington, Seattle, WA 98195.
†Present address: Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick Cancer Research and Development Center, Frederick, MD 21702.
‡To whom reprint requests should be addressed.

§The sequences reported in this paper have been deposited in the GenBank data base (accession nos. U15941–U15953).

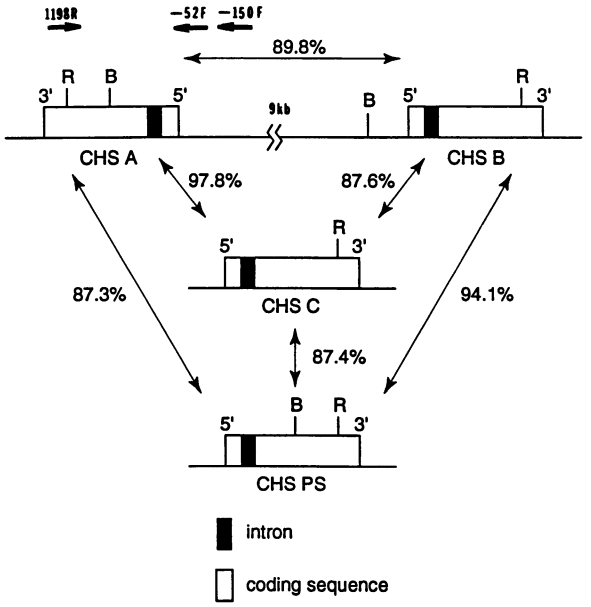

FIG. 1. Organization of the four CHS genes characterized from *I. purpurea*. R, *Eco*RI; B, *Bgl* II. % refers to percentage similarity at the nucleotide level including the intron. Positions of the primers used in PCR amplification are shown above CHS-A. Drawing is not to scale.

−52F primer is located at the transcription start site. The −150F primer is located near the putative CAAT site. Because the primers span regions of potential regulatory significance, there is likely to be sufficient conservation in this region to allow PCR amplification of CHS genes in other *Ipomoea* species. This would allow amplification of more full-length gene products. The 3′ reverse primer was 1198R (5′-CCCATTC-CAAGCCCTCCCCAGTAGT-3′). The reverse primer is located 71 bp 5′ from the end of the coding region. The positions of the primers are shown in Fig. 1.

PCR amplification conditions typically consisted of 30 cycles of 95°C for 30 sec, 50°C for 20 sec, and 72°C for 5 min. These conditions were designed to minimize the formation of chimeric sequences during amplification from a mixture of closely related gene family members (23). To verify that our sequences did not contain any chimeric artifacts, another amplification was performed with a different 5′ primer. Since such chimeric formations should be random events, it would be highly unlikely that two different amplifications would result in the same sequence unless it was correct. PCR products were cloned using the TA Cloning system (Invitrogen). Several clones were selected by colony hybridization to CHS-A from *I. purpurea* and sequenced by PCR-based sequencing using the *fmol* DNA sequencing system (Promega).

A cDNA library made from the mRNA of UV- and light-induced *I. purpurea* seedlings was screened using the *I. purpurea* CHS-A clone. Poly(A)$^+$ mRNA was isolated directly from plant tissue using the PolyATtract system (Promega). cDNA was made using SuperScript RNase H$^-$ reverse transcriptase according to the manufacturer's instructions (GIBCO/BRL).

Nonsynonymous and synonymous substitution distances for exons (Table 1) were estimated by the method of Nei and Gojobori (24) corrected for multiple substitutions using the Jukes–Cantor (25) formula as implemented in the MEGA package (26). Divergence times were estimated using a synonymous substitution rate of 8×10^{-9} substitution per site per yr based on the synonymous substitution rate of alcohol dehydrogenase 1 in grasses (27). Since the alignment of intron sequences was problematic, only exon sequences were used to estimate phylogenetic relationships. The sequences included in the analysis were from the start of exon 1 to 100 bp 5′ to the end of exon 2. Gene phylogenies were estimated by neighbor-joining (28) and maximum-likelihood methods (29). DNADIST using the two-parameter method of Kimura (30) and NEIGHBOR and DNAML programs from version 3.5c of the PHYLIP package (31) were used for phylogeny reconstruction. The MACCLADE program version 3.04 was used to analyze patterns of amino acid change on the phylogenetic tree (32).

RESULTS

Cloning of CHS Genes. Four CHS sequences were isolated from the dark pink phenotype of *I. purpurea* by cloning into λEMBL3 and by a PCR-based strategy. Repeated screenings of a λEMBL3 genomic library yielded three clones, designated CHS-A, CHS-B, and CHS-PS. CHS-A and CHS-B were obtained on the same λ genomic clone. Single- and double-restriction mapping showed that the two CHS genes were separated by ≈9 kbp and oriented 5′ end to 5′ end (Fig. 1). The cDNA sequence corresponding to CHS-A was obtained from a cDNA library made from mRNA of the hypocotyls of UV- and light-induced seedlings, from a PCR product of cDNA made from mRNA of the corolla of immature flower buds, and from mRNA of the anthers and pollen of open flowers. A fourth CHS sequence, CHS-C, was obtained from genomic DNA by PCR amplification from cDNA made from mRNA from the throat tissue of mature flower buds and from anther and pollen cDNA from the open flower. The PCR survey of

Table 1. Evolutionary distances for synonymous and nonsynonymous sites for the 13 *Ipomoea* CHS sequences and *Petunia* CHS-B

	purpA	*purpC*	*nilA*	*platA*	*cordB*	*trifB*	*purpB*	*nilB*	*purpS*	*trilB*	*trilA*	*cordA*	*trifA*	*PetB*
purpA		0.062	0.101	0.149	0.245	0.250	0.302	0.349	0.367	0.311	0.371	0.401	0.402	*
purpC	0.023		0.073	0.136	0.232	0.240	0.330	0.350	0.357	0.324	0.399	0.409	0.404	*
nilA	0.030	0.023		0.141	0.259	0.238	0.301	0.321	0.381	0.292	0.359	0.387	0.381	*
platA	0.033	0.043	0.040		0.181	0.182	0.288	0.288	0.297	0.262	0.304	0.310	0.311	*
cordB	0.043	0.052	0.054	0.047		0.096	0.297	0.310	0.334	0.336	0.345	0.357	0.352	*
trifB	0.045	0.051	0.054	0.046	0.019		0.290	0.297	0.359	0.291	0.338	0.344	0.344	*
purpB	0.061	0.075	0.063	0.054	0.056	0.057		0.076	0.150	0.144	0.227	0.238	0.244	2.724
nilB	0.073	0.084	0.069	0.063	0.065	0.066	0.012		0.147	0.150	0.231	0.230	0.236	3.325
purpS	0.071	0.085	0.079	0.066	0.069	0.070	0.017	0.021		0.174	0.267	0.272	0.273	*
trilB	0.072	0.083	0.073	0.063	0.064	0.063	0.028	0.035	0.039		0.199	0.214	0.214	*
trilA	0.077	0.086	0.077	0.069	0.067	0.069	0.038	0.048	0.050	0.040		0.029	0.025	*
cordA	0.076	0.084	0.074	0.067	0.065	0.067	0.036	0.046	0.048	0.043	0.017		0.012	*
trifA	0.070	0.078	0.067	0.062	0.059	0.062	0.031	0.041	0.043	0.038	0.012	0.010		*
PetB	0.226	0.233	0.216	0.218	0.209	0.218	0.209	0.212	0.218	0.225	0.225	0.217	0.214	

Distances are adjusted for multiple substitutions using the Jukes and Cantor correction (25). Distances for synonymous sites are above and to the right of the diagonal while those for nonsynonymous sites are below and to the left of the diagonal.
*Comparisons that were not calculated.

I. purpurea genomic DNA yielded only the four CHS genes shown in Fig. 1.

The *I. platensis* sequence was obtained from PCR amplification of cDNA made from mRNA isolated from the throat of the mature flower. The rest of the CHS sequences were obtained by PCR amplification of genomic DNA using primers in the 5′ flanking regions and in a conserved region near the 3′ end of the coding sequence. The products were cloned, and several were completely sequenced from each individual.

Since the amplification primers were designed from the 5′ flanking sequence of *I. purpurea* CHS-A to yield as full-length sequences as possible, it is possible that there are CHS genes in the other *Ipomoea* species that were not recovered in our PCR survey.

Sequence Comparison. A single intron divides the gene into two exons. The size of exon 1 varies among *Ipomoea* species from 187 to 202 bp. Exon 2 was 995 bp in *I. purpurea* CHS-A, CHS-B, and CHS-PS. The remaining genes are PCR-derived sequences and are not full length because they were generated by a primer that was within the coding region at the 3′ end. In the first 80 bp of coding sequence of exon 1, there are a number of indels of multiples of 3 nucleotides, thus maintaining proper translational frame. Only *I. purpurea* CHS-A, CHS-B, and CHS-PS are full length. The position of the intron, splitting a cysteine codon (corresponding to codon 69 relative to the start of translation of *I. purpurea* CHS-A), is conserved in all CHS sequences published to date. The introns in the *Ipomoea* CHS sequences are all very small, ranging in size from 84 to 125 nucleotides. The introns are A+T-rich (74.7%) and are characterized by many addition/deletion events. The 5′ and 3′ splice junctions conform to the GT–AG rule for excision of the intron (33).

In one CHS sequence obtained from one of the genomic λ clones (CHS-PS of *I. purpurea*), the ATG start site has mutated to ATA. There is no alternative in-frame start site for this sequence, yet there are no other features, such as premature stop codons or frameshifts, that would mark it as a pseudogene.

Molecular Phylogenies. An unrooted maximum-likelihood phylogram based on exon sequences for the 13 CHS genes from *Ipomoea* species and CHS-B from *Petunia hybrida* is shown in Fig. 2. The sequences clearly fall into two groups. Each group contains sequences from each *Ipomoea* species studied with the exception of *I. triloba*, which has both sequences in group 2, and *I. platensis*, for which only a single sequence was obtained.

An amino acid alignment of the 13 *Ipomoea* CHS genes (available on request) reveals that nine amino acid changes unequivocably distinguish the sequences of these two groups consistent with the grouping based on exon sequences (Fig. 2). In addition, the sequences in group 2 all have an asparagine residue at codon position 258 (relative to the start of *I. purpurea* CHS-A), characteristic of the SS, while the sequences in group 1, except *I. cordatotriloba* CHS-B, have an aspartic acid residue characteristic of CHS sequences at position 258 (10). All sequences contain an asparagine at codon position 54, known to be essential for CHS activity in *Antirrhinum majus* (34), and a cysteine at codon position 173, which is located at the binding site of 4-coumaroyl CoA and is absolutely required for CHS function (35).

The time of divergence between the gene groups was estimated by determining the mean pairwise synonymous site distance between the two groups of sequences. For each sequence in group 1 (left side of phylogram in Fig. 2), a distance to each group 2 sequence was obtained. Similarly, for each group 2 sequence, a mean synonymous site distance to the group 1 sequences was calculated. The grand mean for all of these distances was 0.337 substitution per synonymous site. Assuming a synonymous substitution rate of 8×10^{-9} substitution per site per yr, the date of the group 1/group 2 divergence is ≈21 million yr ago. The divergence dates among the *Ipomoea* species in this analysis would appear to be more recent since their divergences are reflected within each group of CHS sequences. The mean for all of the nonsynonymous site substitutions between the two *Ipomoea* groups was 0.069 substitution per site. The number of synonymous site changes in *Ipomoea* was ≈4.9 times the number of nonsynonymous site changes.

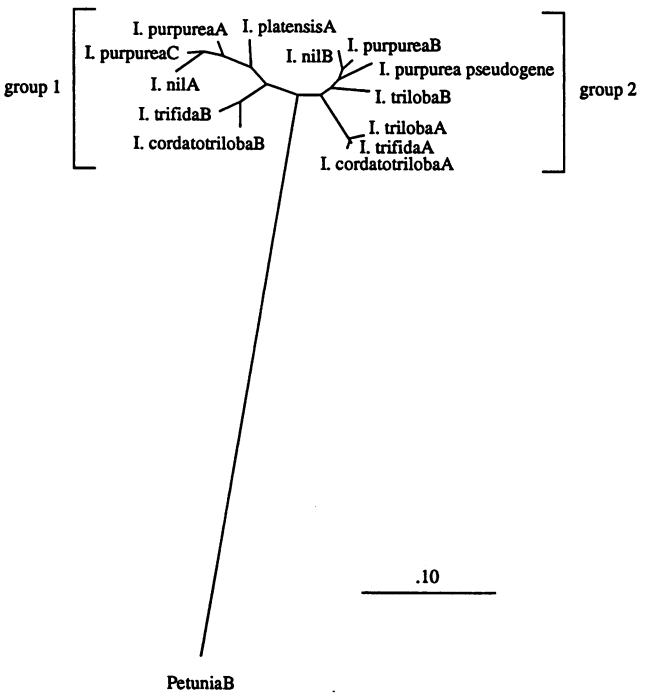

FIG. 2. Unrooted maximum-likelihood tree based on exon sequences from 13 *Ipomoea* CHS sequences and *Petunia* CHS-B. Group 1 sequences are on the left and group 2 sequences are on the right.

The time of divergence between the *Ipomoea* CHS genes and *Petunia* CHS-B was estimated based on nonsynonymous site changes. The mean for the nonsynonymous site changes between the *Ipomoea* CHS sequences and *Petunia* CHS-B was 0.218 substitution per nonsynonymous site. Using an adjusted nonsynonymous substitution rate of 1.6×10^{-9} substitution per site per yr (the synonymous rate calculated for *Pennisetum* alcohol dehydrogenase divided by the ratio of synonymous/nonsynonymous site changes for *Ipomoea*), the date of the *Ipomoea* CHS/*Petunia* CHS-B divergence is ≈68 million yr ago.

Fig. 3 shows the phylogeny from a neighbor-joining analysis showing the 13 *Ipomoea* sequences from this study along with the 52 sequences identified as encoding CHS and related proteins (SS and other phytoalexin synthases). The *Ipomoea* sequences are clearly most closely related to the anomalous CHS-B sequence from *P. hybrida* and these sequences form a distinct group of genes. The SS and bibenzyl and resveratrol synthases of *Pinus*, *Arachis*, and *Vitis* fall among the CHS.

DISCUSSION

CHS is encoded by a multigene family in most of the species examined to date. Only *Arabidopsis* and *Antirrhinum* appear to contain single gene copies (1). Gene duplication followed by differentiation can result in the production of proteins with new functions. The duplicated gene must, however, diverge fast enough to escape the homogenizing effects of gene conversion (36) or recombination (37). The SSs are an example of just such a duplication and differentiation event. They have apparently developed independently from CHS at least three

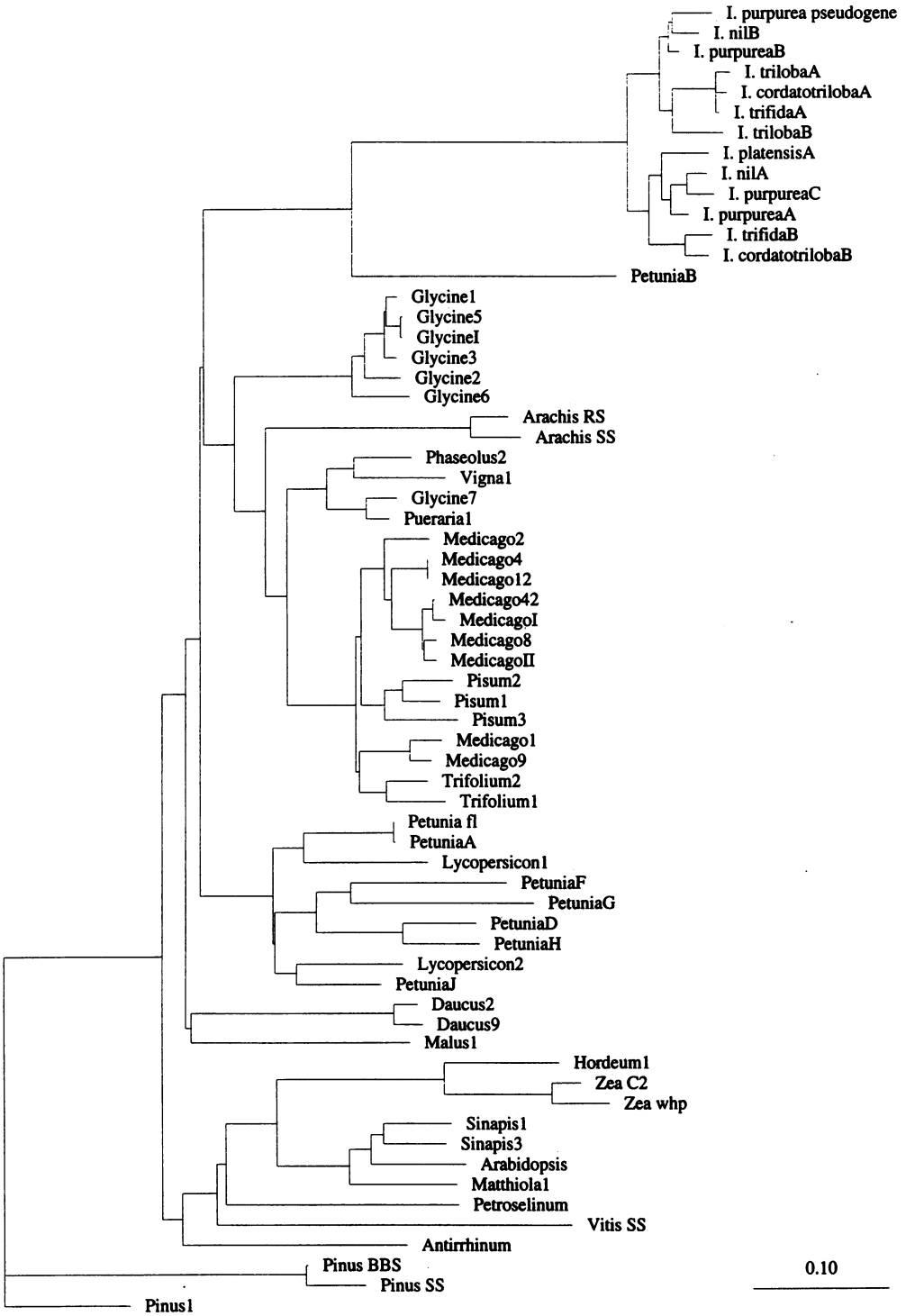

FIG. 3. Neighbor-joining phylogram based on exon sequences from CHS and SS using *Pinus* as an outgroup.

times in the course of seed plant evolutionary history (Fig. 3; ref. 11). In this study, we have evidence for divergence in identity of the CHS sequences in *Ipomoea*, perhaps analogous to the divergence of the stilbene synthases.

I. purpurea CHS-A, *I. purpurea* CHS-C, and *I. platensis* CHS-A are the only genes in this study so far confirmed to be expressed (unpublished data). These sequences are all found in group 1. It is not unusual for one or two family members to account for most of the CHS gene expression. In *Petunia*, there are about eight family members, but one gene, CHS-A, accounts for 90% of the gene expression, while CHS-J accounts for 10%, with CHS-B and CHS-G being expressed at very low levels (38). Of course CHS-B in *I. purpurea* may be expressed in a tissue or developmental stage not yet examined. There are no representatives of *I. triloba* in group 1, but there may be more family members in *I. triloba* that did not turn up in our PCR survey.

The introns in the *Ipomoea* sequences are very short (84–125 bp) in comparison to those found in most other CHS sequences. In *Petunia*, for example, the introns range in size from 0.695 to >15 kb (38). Very short introns are also reported in two clover CHS sequences (39). Aside from many addition/deletion events, the introns in all the *Ipomoea* CHS sequences are remarkably similar, sharing many sequence motifs.

The divergence of the *Ipomoea* CHS gene family into two groups was relatively recent (≈21 million yr ago), and considerable duplication has occurred during this period. It is noteworthy that the *Ipomoea* CHS family is most closely related to the *Petunia* CHS-B sequence. Based on nonsynonymous substitution rates, the *Ipomoea* CHS and *Petunia* CHS-B diverged ≈68 million yr ago. *Petunia* CHS-B itself is highly diverged from the rest of the *Petunia* sequences. It is possible that other CHS sequences exist in the *Ipomoea* genome that are more closely related to the *Petunia* A–H gene cluster.

Petunia CHS-B and the *Ipomoea* CHS genes appear to be among the most rapidly evolving CHS genes, based on relative branch lengths in Fig. 3. The nonsynonymous substitution rate for *Ipomoea* is 12.8 times the nonsynonymous substitution rate for *Pennisetum* alcohol dehydrogenase. The factors involved in this rapid rate of evolution are not clear, but it is possible that divergence of function or minimal selective constraints associated with varying functional demands may be a factor.

In many cases, CHS gene sequences form clusters consistent with taxonomic groupings (Fig. 3). For example, genes from the genera *Ipomoea*, *Glycine* (except for *Glycine* 7), *Zea*, *Pisum*, and *Trifolium* each cluster together. The representatives of the Solanaceae are more diverged with *Lycopersicon* and *Petunia* CHS sequences interspersed. *Petunia* CHS-B is highly diverged from the remaining *Petunia* sequences, indicating an ancient divergence in this gene family. This pattern suggests the frequent origin of new CHS genes through duplication with occasional highly diverged elements preserved over longer intervals of evolutionary time. As already noted, functional divergence from CHS to SS may occur repeatedly. In addition, loss of function must occur at a substantial rate. The overall picture of CHS evolution is that of a highly fluid gene family where copy number varies greatly among plant lineages and where functional divergence may occur repeatedly.

This work is dedicated to the memory of Deborah Elaine Glover (1959–1989). Deborah was a promising young scientist who died just as she was completing her Ph.D. Her love of morning glories and her enthusiasm for life and learning were an inspiration to all who knew her. We thank Dr. Brian Morton, Dr. Sharman O'Neill, Dr. Richard Meagher, and Dr. Jim Eckenwalder for helpful comments; Dr. Bryan Epperson for plant materials; and Ms. Patti Fagan for preparation of the manuscript. This work was supported by National Science Foundation Grant 9006984 to M.T.C.

1. Forkman, G. (1993) in *The Flavonoids: Advances in Research Since 1986*, ed. Harborne, J. B. (Chapman & Hall, London), pp. 537–564.
2. Brouillard, R. (1988) in *The Flavonoids*, ed. Harborne, J. B. (Chapman & Hall, London), pp. 525–538.
3. Schmelzer, E., Jahnen, W. & Hahlbrock, K. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 2989–2993.
4. Dixon, R. A., Dey, P. M. & Lamb, C. J. (1983) *Adv. Enzymol.* **55,** 1–136.
5. Lamb, C. J., Lawton, M. A., Dron, M. & Dixon, R. A. (1989) *Cell* **56,** 215–224.
6. Long, S. (1989) *Cell* **56,** 203–214.
7. Jacobs, M. & Rubery, P. H. (1988) *Science* **241,** 346–349.
8. Taylor, L. P. & Jorgensen, R. (1992) *J. Hered.* **83,** 11–17.
9. Hedin, P. A. & Waage, S. K. (1986) in *Plant Flavonoids in Biology and Medicine: Biochemical, Pharmacological and Structure-Activity Relationships*, eds. Cody, V., Middleton, E. & Harborne, J. B. (Liss, New York), pp. 87–103.
10. Martin, C. R. (1993) in *International Review of Cytology*, eds. Jeon, K. W. & Jarvik, J. (Academic, New York), pp. 233–284.
11. Tropf, T. L., Rensing, S. A., Schroder, J. & Schroder, F. (1994) *J. Mol. Evol.* **38,** 610–618.
12. Ennos, R. A. & Clegg, M. T. (1983) *J. Hered.* **74,** 247–250.
13. Brown, B. A. & Clegg, M. T. (1984) *Evolution* **38,** 796–803.
14. Epperson, B. K. & Clegg, M. T. (1988) *J. Hered.* **79,** 64–68.
15. Epperson, B. K. & Clegg, M. T. (1986) *Am. Nat.* **128,** 840–858.
16. Epperson, B. K. & Clegg, M. T. (1987) *Evolution* **41,** 1302–1311.
17. Epperson, B. K. (1990) *Genetics* **124,** 757–771.
18. Rausher, M. D. & Fry, J. D. (1993) *Genetics* **134,** 1237–1247.
19. Epperson, B. K. & Clegg, M. T. (1987) *J. Hered.* **78,** 346–352.
20. Epperson, B. K. & Clegg, M. T. (1992) *J. Hered.* **83,** 405–409.
21. O'Neill, S. D., Tong, Y., Sporlein, B., Forkmann, G. & Yoder, J. (1990) *Mol. Gen. Genet.* **224,** 279–288.
22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
23. Meyerhans, A., Vartanian, J. & Wain-Hobson, S. (1990) *Nucleic Acids Res.* **18,** 1687–1691.
24. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
25. Jukes, T. H. & Cantor, C. R. (1969) in *Mammalian Protein Metabolism*, ed. Munro, H. N. (Academic, New York), pp. 21–132.
26. Dumar, S., Tamura, K. & Nei, M. (1993) MEGA, Molecular Evolutionary Genetic Analysis (Pennsylvania State Univ., University Park), Version 1.0.
27. Gaut, B. S. & Clegg, M. T. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 2060–2064.
28. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
29. Felsenstein, J. (1981) *J. Mol. Evol.* **17,** 368–376.
30. Kimura, M. (1980) *J. Mol. Evol.* **16,** 111–120.
31. Felsenstein, J. (1993) PHYLIP, Phylogeny Inference Package (Univ. of Washington, Seattle), Version 3.5c.
32. Maddison, W. P. & Maddison, D. R. (1992) MACCLADE, Analysis of Phylogeny and Character Evolution (Sinauer, Sunderland, MA), Version 3.0.
33. Brown, J. W. S. (1986) *Nucleic Acids Res.* **14,** 9549–9559.
34. Luo, D., Coen, E. S., Doyle, S. & Carpenter, R. (1991) *Plant J.* **1,** 59–69.
35. Lanz, T., Tropf, S., Marner, F. J., Schröder, J. & Schröder, G. (1991) *J. Biol. Chem.* **266,** 9971–9976.
36. Walsh, J. B. (1987) *Genetics* **117,** 543–557.
37. Drouin, G. & Dover, G. A. (1990) *J. Mol. Evol.* **31,** 132–150.
38. Koes, R. E., Spelt, C. E., van den Elzen, P. J. M. & Mol, J. N. M. (1989) *Gene* **81,** 245–257.
39. Arioli, T., Howles, P. A., Weinman, J. J. & Rolfe, B. G. (1994) *Gene* **138,** 79–86.