Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 11735–11739, October 1996
Evolution

Evolution of alcohol dehydrogenase genes in the Palm and Grass families

(*Washingtonia*/gene duplication/gene family)

BRIAN R. MORTON*†‡, BRANDON S. GAUT§, AND MICHAEL T. CLEGG†

*Department of Biological Sciences, Barnard College, Columbia University, 3009 Broadway, New York, NY 10027; †Department of Botany and Plant Sciences, University of California, Riverside, CA 92521-0124; and §Department of Plant Sciences, Rutgers University, New Brunswick, NJ 08903

Contributed by Michael T. Clegg, June 12, 1996

ABSTRACT **The alcohol dehydrogenase (Adh; alcohol:NAD+ oxidoreductase, EC 1.1.1.1) gene family has two or three loci in a broad array of angiosperm species. The relative stability in the number of Adh loci led Gottlieb [Gottlieb, L. D. (1982) *Science* 216, 373–380] to propose that the Adh gene family arose from an ancient gene duplication. In this study, the isolation of three loci from the California fan palm (*Washingtonia robusta*) is reported. The three loci from palm are highly diverged. One palm Adh gene, referred to here as *adh*B, has been completely sequenced, including 950 nucleotides of the upstream regulatory region. For the second locus, *adh*A, 81% of the exon sequence is complete. Both show the same basic structure as grass Adh genes in terms of intron number and intron location. The third locus, *adh*C, for which only a small amount of sequence is available (12% of exon sequence) appears to be more highly diverged. Comparison of the Adh gene families from palms and grasses shows that the *adh*1 and *adh*2 genes of grasses, and the *adh*A and *adh*B genes of palms, arose by duplication following the divergence of the two families. This finding suggests that the multiple Adh loci in different monocot lineages are not the result of a single ancestral duplication but, rather, of multiple duplication events.**

A number of studies have been conducted on the evolutionary dynamics of plant gene families, including the gene families coding for the R and MADS-box regulatory proteins (1, 2), the small heat-shock proteins (3), chalcone synthase (4), and the chlorophyll a/b binding proteins (5). Most of these gene families consist of numerous loci and have a great deal of variation in copy number between species. Phylogenetic analyses indicate that much of this variation can be attributed to recent duplications. For example, there are at least nine chalcone synthase loci in alfalfa, all of them originating well after the divergence of the Legumes (4). The evolutionary picture emerging for these gene families is one of dynamic fluctuations of copy number through multiple gene duplication/deletion events.

The glycolytic proteins in plants are coded by small multigene families, which provide an interesting contrast to the high copy number gene families studied to date. Isozyme surveys covering an array of dicot and monocot species have revealed that most glycolytic enzymes have two forms in all species (6), probably reflecting a small, and stable, number of loci. The apparent stability of these gene families raises important questions regarding evolutionary dynamics. One issue is whether any given gene family emerged once by duplication and then differentiated, as suggested by Gottlieb (6). An alternative view posits a continuous, albeit slow, flux of gene duplication and loss that leads to an approximate dynamic equilibrium in copy number. The narrow range of gene family size for glycolytic enzymes suggests that additional constraints may act to determine copy number for this important class of genes. Alcohol dehydrogenase (Adh) genes encode glycolytic enzymes that have been characterized in several members of the grass family, but these genes have not been characterized in other monocot families. To further investigate the pattern of gene family evolution for glycolytic enzymes, it is crucial to investigate the dynamics of duplication by comparing gene family organization in a broader range of species.

Adh is an essential enzyme in anaerobic metabolism (7, 8). Transcription from Adh promoters increases under oxygen stress, as well as in response to cold stress in both maize and *Arabidopsis* and to dehydration in *Arabidopsis* (8). Two or three isozymes are observed in all flowering plant species, dicot and monocot (6), with the exception of *Arabidopsis* which has been shown to have a single Adh locus (9).

At the molecular level, the gene family has been most thoroughly studied in the grass family (10–12). Two Adh loci, *adh*1 and *adh*2, have been sequenced for maize (13–15) and rice (16) as well as barley from which a third locus, *adh3*, a recent duplication of *adh*2, has been isolated (10). In addition, the *adh*1 gene has been duplicated in some accessions of maize (15).

An analysis of animal and plant Adh genes indicated that the grass *adh*1 and *adh*2 genes diverged following the divergence of monocots and dicots (17). This result provides evidence that the gene family did not emerge from a single duplication event early in angiosperm evolution. Additionally, the isolation of a recent duplication product in barley (10), as well as the duplication of *adh*1 in other species, suggests that the gene family undergoes some copy number fluctuation. However, there has been little study of the Adh gene family in plants other than members of the grass family, and the issues raised concerning its evolution have yet to be addressed in other monocot lineages.

To further study the evolution of the Adh gene family in the flowering plants, we have characterized Adh loci from a representative of the palm family (Arecaceae), *Washingtonia robusta*. This species has three isozymes suggesting that there may be three functional Adh loci in this species (unpublished data). In the present study, three Adh loci have been isolated from *Washingtonia*. These loci do not correspond to the *adh*1 and *adh*2 genes of the grasses. Rather, the palm loci represent unique duplication events suggesting that the evolutionary history of the Adh gene family is dynamic. Moreover, it appears that none of the three palm loci are the result of a recent duplication event. Despite the apparent stability of the gene family copy number, Adh has undergone parallel duplications in different monocot lineages. This finding provides further evidence against the hypothesis that the Adh gene family emerged from a single duplication event. It also indi-

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: Adh, alcohol dehydrogenase; ARE, anaerobic response element.

‡To whom reprint requests should be addressed.

Downloaded by guest on December 1, 2021

cates that stability of isozyme number does not necessarily indicate stability of a gene family.

MATERIALS AND METHODS

A genomic library from *Washingtonia robusta* (California fan palm) was generated by ClonTech in a λEMBL-3 library. The library was screened with ^{32}P-labeled *Zea mays adh*1 generated by PCR amplification following the protocol described previously (12). This fragment spans exons 3–10 and is ≈2200 nucleotides long. The PCR product was cut from a SeaPlaque agarose gel (FMC Bioproducts) and resuspended in water at a ratio of 1.5 ml/g of gel. This resuspended fragment was denatured 10 min prior to labeling with ^{32}P dCTP. The labeling reaction followed the protocol supplied with the Boehringer Mannheim Random Prime kit using 18 μl of the resuspended PCR product in a 40 μl reaction. Hybridization was performed for 72 h in 50% formamide at a temperature of 32°C, and washed twice for 10 min at room temperature in 0.1× standard saline citrate (SSC)/0.5% SDS, followed by two washes for 45 min at 42°C in 2× SSC/0.5% SDS. Hybridization was detected by autoradiographic exposure for 72 h at −80°C. Roughly 1.8×10^6 plaques were screened.

From the first positive clone, a 1.8-kb *Sph*I restriction fragment that hybridized *adh*1 was subcloned into M13 mp18 and sequenced, initially using the universal primers (forward and reverse). Further sequencing of this λ clone, as well as the second positive clone, was completed using primers designed from the palm Adh sequence generated, sequencing directly using the fmol kit from Promega following the protocol supplied.

A portion of the Adh gene from the third positive clone was amplified directly from the λ clone using primers designed from sequencing the two previous positives. Primers from exons 5 (GGGTGCTGTAGGCCTTGC) and 8 (GATATCTGCATTTGAATGCG) were used to amplify a fragment of ≈400 nucleotides using 35 cycles of 2 min at 94°C, 2 min at 54°C, and 3 min at 72°C. This product was sequenced directly as described above, following a GeneClean (Bio 101) of the PCR product.

Sequence divergence was estimated by using the Kimura two-parameter model (18) employed by PHYLIP (19) with a transition/transversion ratio of 2.0. Phylogenetic analysis of Adh sequences was performed using the neighbor-joining method (20) of PHYLIP (19). Exon sequence only was included. Sequences used for the phylogenetic reconstruction were the palm Adh sequences, maize *adh*1 and *adh*2 (GenBank accession nos. L23548 and X01965), barley *adh*1 and *adh*2 (X07774 and X12733), rice *adh*1 and *adh*2 (X16296 and X16297), millet *adh*1 (X16547), and *Arabidopsis* Adh (M12196). Bootstrapping was performed resampling from the data 100 times. Maximum likelihood topologies were also generated (21), and the method of Kishino and Hasegawa (22) was used to compare alternative topologies with the same transition/transversion ratio used in the phylogenetic analysis.

RESULTS

Three different Adh sequences, designated *adh*A, *adh*B, and *adh*C, based on order of isolation, were detected in the five positive clones. If these three sequences represent three functional loci then the copy number is consistent with isozyme analyses of *W. robusta* seeds which show three monomorphic Adh loci (data not shown). Of the three sequences, the *adh*B sequence was represented three times and the other two by a single λ clone each. The *adh*B sequence was complete while the clone containing *adh*A started within the third intron and spanned beyond the 3′ end of the gene. The extent of the gene sequence within the *adh*C clone has not yet been completely determined.

Structure of Palm Adh. The *adh*B gene, which consists of a total of 10 exons, is diagrammed in Fig. 1. The coding sequence, with stop codon, consists of 1146 nucleotides corresponding to a polypeptide 381 amino acids long. The regions sequenced for the other two palm loci are indicated below the diagram. Since the complete gene sequence is available for palm *adh*B discussion of gene structure, including the upstream region, will be based on this sequence.

The palm *adh*B gene has the same intron/exon structure as the *adh*1 gene from grasses. Where *adh*A and *adh*C have been sequenced they, too, show conserved intron positions. This conservation of intron position is consistent with the observation that dicot Adh sequences have the same intron position as maize *adh*1 and *adh*2 (10), with the exception of the Adh gene from *Arabidopsis* which has only six exons (9). The coding sequence itself is complete and shows all features of Adh. The conserved amino acids of the two zinc-binding domains from grass Adh, Cys-47/His-69/Cys-178 and Cys-99/Cys-102/Cys-105/Cys-113, are all conserved in *adh*B. Therefore, the gene appears to code for a functional Adh protein.

Potential TATA boxes for RNA polymerase binding are found 194 and 88 bases upstream of the start codon in the palm *adh*B sequence. The site 88 nucleotides upstream may be a better candidate as it is immediately downstream of a pyrimidine-rich sequence, something commonly observed in Adh promoter regions (23). Two upstream sequences have been suggested to have a regulatory role for maize Adh. The sequences AAATCCTA (−410) and ATCCGAGC (−300) are both found upstream of *adh*1 and *adh*2 in maize (24). Very similar sequences to each are found upstream of *adh*B (AAAGCCTA and TTCCAAGC) and may represent conserved control regions.

Along with the upstream regulatory elements, an anaerobic response element (ARE) has been identified by deletion analysis in maize *adh*1 (25). The ARE is a regulatory sequence which mediates the increase in transcription observed from Adh loci when oxygen levels are low (8). The ARE in grass is located between −140 and −99 and consists of two adjacent sections, AREI and AREII (8). Similar sequences to both AREI and AREII are found upstream of palm *adh*B but do not lie in similar relative positions. Instead, the putative

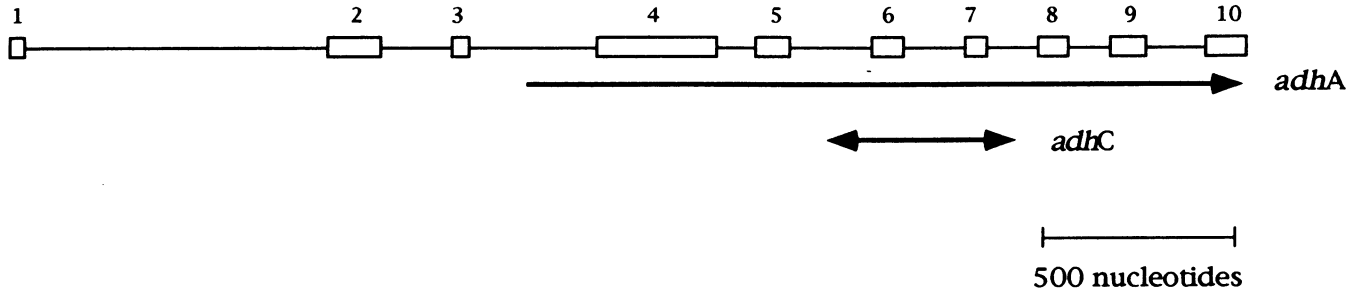

FIG. 1. Diagram of the *adh*B gene from *Washingtonia*. The open boxes represent exon sequences. The regions sequenced for the *adh*A and *adh*C loci are indicated under this (see text). An arrow indicates that the sequence of the clone continues in that direction.

Downloaded by guest on December 1, 2021

Table 1. Divergences, based on Kimura's two-parameter model, between grass and palm Adh genes

	*Zea adh*1*	*Zea adh*2	*adh*A	*adh*B	*adh*C
*Zea adh*1	—	0.203	0.279	0.261	0.472
*Zea adh*2	0.015	—	0.325	0.304	0.355
*adh*A	0.021	0.023	—	0.244	0.315
*adh*B	0.019	0.021	0.021	—	0.370
*adh*C	0.109	0.082	0.074	0.083	—

*Distances are given in the upper matrix, and standard errors in the lower matrix.

AREII is 131 bases upstream from the putative AREI. The lack of correct positioning raises doubts that the palm sequences are AREs, and any functional significance remains speculative.

In addition to the grass ARE, another regulatory element may be involved in anaerobic response. It has been suggested that, in barley Adh, a sequence within the transciption unit, upstream of the start codon, acts as a selective anaerobic translation control element (10). A sequence in the transcription unit of both maize *adh*1 and *adh*2 has been proposed to have a similar function (10). However, no sequence similar to this element is found in palm *adh*B within the probable nontranslated leader.

Divergence of the Adh Loci. Pairwise distances based on the Kimura two-parameter model for exon sequences of *Zea adh*1 and *adh*2 and the three palm loci are given in Table 1. Two points are apparent from this table. One is that the two maize loci are the most similar. The second is that the *adh*C gene from palms is highly diverged and must represent a duplication event predating the *adh*A/*adh*B and *adh*1/*adh*2 duplications. Comparing palm and grass Adh genes shows that the *adh*B gene from palms has an additional codon situated between what are codons 2 and 3 in both *Zea adh*1 and *adh*2. This indel is shared by dicot Adh genes, with the exception of *Arabidopsis*, indicating that it is likely that the grass *adh*1 and *adh*2 genes have lost the codon by deletion.

An unrooted phylogeny, based on the exon sequences available, is presented in Fig. 2. The *adh*C locus is omitted due

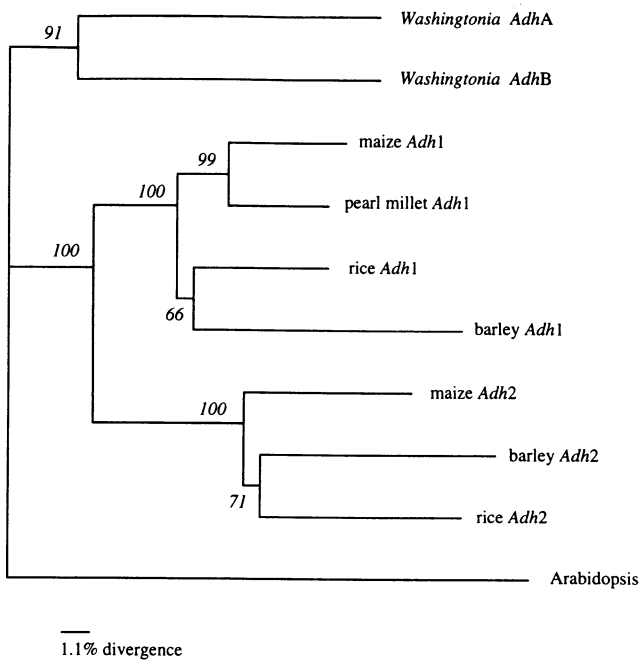

FIG. 2. Results of the neighbor-joining analysis of monocot Adh sequences with the dicot *Arabidopsis* as the outgroup. Numbers represent percentages from 100 bootstrap resamples.

Table 2. Results of likelihood tests of possible Adh orthologous pairs

Tree*	lnL	ΔlnL†	SD‡
(B,2), (A,1)	−6657.2	−76.7	18.4
(A,2), (B,1)	−6664.4	−83.8	16.7
(A,B), (1,2)	−6580.7	0	—

*Trees are given by indicating the orthologous pairings tested where A, B, 1, and 2 designate *adh*A, *adh*B, *adh*1, and *adh*2, respectively. All grass Adh sequences from Fig. 2 were included in each topology.
†Difference between lnL of the given topology and the most likely tree {(A,B), (1,2)}.
‡Number of standard deviations the difference represents.

to the short length of available sequence. It is clear from this phylogeny that the different loci from palms do not correspond to the two loci in grasses and that two duplication events have occurred, one in the palm lineage and one in the grasses. Topologies that require only a single duplication event are significantly less likely when the Kishino and Hasegawa (22) test is used (Table 2).

DISCUSSION

Palm Adh Genes. The palm family emerged ≈80 million years ago and as such it represents one of the lineages that radiated early in monocot evolution (26, 27). The grass family emerged and diversified later, beginning ≈60 million years ago and has been the most intensively studied monocot lineage because of its economic importance. The comparative analysis of these two monocot families presents an ideal opportunity to investigate the dynamics of angiosperm gene family evolution, and in particular, to expand our understanding of the evolution of the Adh gene family. Accordingly, we isolated three different Adh clones from a palm genomic library, and have denoted these as *adh*A, *adh*B, and *adh*C. Most of the *adh*A gene, and all of the *adh*B gene, have been sequenced. The sequences have conserved exon/intron junctions which are also consistent with intron/exon junctions of the grass Adh genes. The *adh*A and *adh*B sequences are most similar but are still separated by 0.244 substitutions per site. The *adh*C gene has only been partially sequenced it appears to be highly diverged from other palm and grass Adh genes (Table 1). Based on the levels of divergence observed between the three clones, we can conclude that they represent three distinct loci.

The pattern of substitution suggests that all three of the three isolated from the palm library code for functional Adh proteins. No premature stop codons are observed and substitution rates are greatest at third position, indicating that amino acid content is under selective constraint (data not shown). Further, the conservation of important residues supports the notion that the palm sequences code for functional enzymes. Although the expression of *adh*A, *adh*B, and *adh*C has not been demonstrated in palms, the isolation of three loci is consistent with isozyme studies which indicate three functional loci in *Washingtonia* (data not shown).

Several studies have examined the upstream region of grass *adh*1 and *adh*2 (10, 24, 25). Comparison of 950 bases upstream of the *adh*B coding region with grass upstream regions permits the detection of conserved sequences involved in regulating anaerobic metabolism. Upstream enhancers very similar to those found in grass Adh genes have been noted in the region upstream of palm *adh*B. However, no sequences similar to those proposed to be involved in anaerobic regulation (8, 10, 25) appear in the palm upstream region. This suggests that the expression in palms is either regulated in a different manner or that *adh*B is not regulated in response to anaerobic conditions. It is, of course, possible that the different loci in palm, for which upstream regions have not yet been isolated, have conserved anaerobic regulatory elements.

Downloaded by guest on December 1, 2021

Evolution of the Adh Multigene Family. A previous analysis of Adh has provided evidence against the *adh*1 and *adh*2 genes in grasses emerging from a single duplication event early in the evolutionary history of the angiosperms (17). The three palm Adh loci isolated in the present study provide further evidence against a single ancestral duplication event giving rise to the Adh gene family in flowering plants. When *adh*A and *adh*B from palms are compared with *adh*1 and *adh*2 from grasses, it is clear that two independent duplication events have occurred (Fig. 2). Forcing either *adh*A or *adh*B to be the result of a common duplication event with grass *adh*1 produces a significantly less likely tree (Table 2).

Analysis of the *adh*C sequence raises an interesting question concerning the evolution of the gene family. A phylogenetic analysis of the sequences available shows that the divergence of *adh*C predates the duplications that produced the other monocot Adh sequences presented in Fig. 2 (data not shown), a result consistent with the distances given in Table 1. If this section of the gene is representative of the evolution of the entire locus, then there must have been at least two loci at the time of the palm/grass divergence; the *adh*C progenitor and the progenitor of *adh*A/*adh*B and *adh*1/*adh*2. Therefore, *adh*C supports an ancestral duplication event. However, it also indicates that this ancestral duplication did not give rise to the two extant loci in grasses. Rather, the locus orthologous to *adh*C in the grasses must have been lost by deletion, or rendered nonfunctional, since isozyme analysis indicates only two functional Adhs (6). The *adh*C locus provides further evidence that, despite the stability of copy number for Adh, the copy number of this gene family undergoes some fluctuation.

Flowering Plant Multigene Families. The Adh gene family of flowering plants provides an interesting contrast to the angiosperm gene families studied to date. Many of these families have high copy numbers (1–4). The methods that have been used to isolate various members of the families differ from study to study, creating difficulties in making explicit comparisons. Despite this difficulty it is possible to divide the gene families studied to date into two main groups, a moderate copy number group with a high variance in copy number across lineages and a group with a low copy number and a low variance.

The first group, the high copy number gene families, is defined loosely as those with greater than three loci in at least one species. This includes the MADS-box genes, which has 14 members in *Arabidopsis* (2), the chalcone synthase family, which has 9 loci in alfalfa (4), the small heat shock proteins for which five loci have been isolated from *Glycine* (3) and the *cab* genes (5). In addition, the R protein gene family falls into this group, as four loci have been sequenced from *Pennisetum* (1), although in general it appears to be a low copy number family. One evolutionary feature of these gene families is that copy number fluctuates dramatically over time. Many of the extant loci have emerged from recent duplication events, suggesting that loci may be gained, and lost, continuously at a high rate.

The second group is composed of multigene families with three or fewer loci, as observed, to date, for the Adh gene family. Most glycolytic enzymes also appear to fall into this category, on the basis of isozyme studies, although for most of these families the actual number of loci has not yet been established. In contrast to the first group, the gene families of this group appear to have a relatively constant number of loci, suggesting that the evolutionary history, in terms of gene duplication, is much less dynamic. The individual members of the phytochrome gene family (28), which consists of functionally distinct loci that diverged prior to angiosperm radiation, may also fall into this group. No evidence yet exists for fluctuation in the copy number of any of these members. Although Adh from the flowering plants can be classified in this manner, the Adh gene family in pine has a significantly higher copy number (D. Perry and G. Furnier, personal communication) indicating that the gymnosperm gene family has substantially different evolutionary dynamics.

The analysis of Adh evolution presented here is a strong indication that stability of copy number does not necessarily indicate a a lack of copy number fluctuation at the molecular level. The dynamic evolutionary history of Adh in relation to the isozyme stability raises the question of what evolutionary process is keeping the isozyme and copy number of Adh, and, possibly, other glycolytic enzymes, relatively constant in different lineages.

Upstream regions must often play a role in the evolution of new gene functions. New duplications may be an advantage if the regulatory sequences are easily adapted to provide novel expression patterns. In this context, the evolution of Adh may not be representative of other glycolytic enzymes coded by small multigene families. The tissue specificity and regulated stress response of Adh may result in a unique pattern of evolution for this gene family relative to other glycolytic enzymes.

Low copy number gene families also have potential as markers for plant phylogenetic analyses to complement existing results from chloroplast encoded genes. The appealing features of these small families are that locus-specific amplification is easier to accomplish, and, if the copy number is stable, that it is easier to study an orthologous set of sequences. The Adh example shows that it will be important to establish orthologous patterns of descent as a precondition to phylogenetic analysis.

We are extremely grateful to Janet Clegg for performing isozyme analyses on *W. robusta*. We also owe Mary Durbin many thanks for excellent technical assistance and thank Tony Brown, Glen Furnier, and Spencer Muse for thoughts concerning the project. This work was supported in part by National Institute of Technology Grant GM 45144 to M.T.C.

1. Purugganan, M. D. & Wessler, S. (1994) *Genetics* **138,** 849–854.
2. Purugganan, M. D., Rounsley, S. D., Schmidt, R. J. & Yanofsky, M. F. (1995) *Genetics* **140,** 345–356.
3. Waters, E. R. (1995) *Genetics* **141,** 785–795.
4. Durbin, M. L., Learn, G. H., Huttley, G. A. & Clegg, M. T. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3338–3342.
5. Demmin, D. S., Stockinger, E. J., Chang, Y. C. & Walling, L. L. (1989) *J. Mol. Evol.* **29,** 266–279.
6. Gottlieb, L. D. (1982) *Science* **216,** 373–380.
7. Freeling, M. & Bennett, D. C. (1985) *Annu. Rev. Genet.* **19,** 297–323.
8. Dolferus, R., de Bruxelles, G., Dennis, E. S. & Peacock, W. J. (1994) *Ann. Bot. (London)* **74,** 301–308.
9. Chang, C. & Meyerowitz, E. M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 1408–1412.
10. Trick, M., Dennis, E. S., Edwards, K. J. R. & Peacock, W. J. (1988) *Plant Mol. Biol.* **11,** 147–160.
11. Gaut, B. S. & Clegg, M. T. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 2060–2064.
12. Gaut, B. S. & Clegg, M. T. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 5095–5099.
13. Dennis, E. S., Gerlach, W. L., Pryor, A. J., Bennetzen, J. L., Inglis, A., Llewellyn, D., Sachs, M. M., Ferl, R. J. & Peacock, W. J. (1984) *Nucleic Acids Res.* **12,** 3983–4000.
14. Sachs, M. M., Dennis, E. S., Gerlach, W. L. & Peacock, W. J. (1986) *Genetics* **113,** 449–467.
15. Osterman, J. C. & Dennis, E. S. (1989) *Plant Mol. Biol.* **13,** 203–212.
16. Xie, Y. & Wu, R. (1989) *Plant Mol. Biol.* **13,** 53–68.
17. Yokoyama, S. & Harry, D. E. (1993) *Mol. Biol. Evol.* **10,** 1215–1226.
18. Kimura, M. (1980) *J. Mol. Evol.* **16,** 111–120.
19. Felsenstein, J. (1992) PHYLIP (Univ. of Washington, Seattle), Version 3.5.
20. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.

Downloaded by guest on December 1, 2021

21. Felsenstein, J. (1981) *J. Mol. Evol.* **17,** 368–376.
22. Kishino, H. & Hasegawa, M. (1989) *J. Mol. Evol.* **29,** 170–179.
23. Llewellyn, D. J., Finnigan, E. J., Ellis, J. G., Dennis, E. S. & Peacock, W. J. (1987) *J. Mol. Biol.* **195,** 115–123.
24. Dennis, E. S., Sachs, M. M., Gerlach, W. L., Finnegan, E. J. & Peacock, W. J. (1985) *Nucleic Acids Res.* **13,** 727–743.
25. Walker, J. C., Howard, E. A., Dennis, E. S. & Peacock, W. J. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6624–6628.
26. Wilson, M. A., Gaut, B. S. & Clegg, M. T. (1990) *Mol. Biol. Evol.* **7,** 303–314.
27. Duvall, M. R., Clegg, M. T., Chase, M. W., Clark, W. D., Kress, J. W., Zimmer, E. A., Hills, H. G., Eguiarte, L. E., Smith, J. F., Gaut, B. S. & Learn, G. H. (1993) *Ann. Mo. Bot. Gard.* **80,** 607–619.
28. Mathews, S., Lavin, M. & Sharrock, R. A. (1995) *Ann. Mo. Bot. Gard.* **82,** 296–321.

Downloaded by guest on December 1, 2021