Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 2647–2651, March 1995
Evolution

Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms

DOUGLAS E. SOLTIS*, PAMELA S. SOLTIS*, DAVID R. MORGAN†, SUSAN M. SWENSEN‡, BETH C. MULLIN§, JULIE M. DOWD¶, AND PETER G. MARTIN¶‖

*Department of Botany, Washington State University, Pullman, WA 99164-4238; †Department of Biology, Western Washington University, Bellingham, WA 98225; ‡Department of Biology, Indiana University, Bloomington, IN 47405; §Department of Botany, Center for Legume Research, University of Tennessee, Knoxville, TN 37996; and ¶Department of Botany, University of Adelaide, Adelaide, South Australia 5005, Australia

Communicated by Michael T. Clegg, University of California, Riverside, CA, November 14, 1994

ABSTRACT **Of the approximately 380 families of angiosperms, representatives of only 10 are known to form symbiotic associations with nitrogen-fixing bacteria in root nodules. The morphologically based classification schemes proposed by taxonomists suggest that many of these 10 families of plants are only distantly related, engendering the hypothesis that the capacity to fix nitrogen evolved independently several, if not many, times. This has in turn influenced attitudes toward the likelihood of transferring genes responsible for symbiotic nitrogen fixation to crop species lacking this ability. Phylogenetic analysis of DNA sequences for the chloroplast gene *rbcL* indicates, however, that representatives of all 10 families with nitrogen-fixing symbioses occur together, with several families lacking this association, in a single clade. This study therefore indicates that only one lineage of closely related taxa achieved the underlying genetic architecture necessary for symbiotic nitrogen fixation in root nodules.**

Nitrogen-fixing symbioses in root nodules are known in only 10 of the approximately 380 families of angiosperms. Although only a fraction of legumes (Fabaceae) have been studied, nodulation has been found in more than 90% of the plants examined in subfamilies Mimosoideae and Papilionoideae and in 30% of Caesalpinioideae (1, 2). In addition to occurring in the legumes, nitrogen-fixing symbioses involving root nodules also occur in some members of Betulaceae, Casuarinaceae, Coriariaceae, Datiscaceae, Elaeagnaceae, Myricaceae, Rhamnaceae, Rosaceae, and Ulmaceae (Table 1) (4, 5, 7, 8). Nodules are induced and inhabited by either of two very different genera of bacteria. Species of Rhizobiaceae (Gram-negative motile rods) nodulate the legumes and *Parasponia* (subfamily Celtoideae, Ulmaceae) (9). Actinomycetes of the genus *Frankia* (Gram-positive, non-endospore-forming, mycelial bacteria) nodulate hosts in the remaining eight families (Table 1), plants referred to as actinorhizal (4, 5, 10).

The 10 families with nitrogen-fixing symbioses are distributed among four of Cronquist's (3) six major subgroups (subclasses) of dicotyledons: Magnoliidae, Dilleniidae, Rosidae, and Hamamelidae. Many of these families have been considered to be only distantly related (3, 11–14), engendering the hypothesis that nitrogen fixation evolved independently several, if not many, times (7, 8). This view has in turn influenced attitudes toward the likelihood of transferring genes responsible for symbiotic nitrogen fixation to crop species that do not possess this ability (15, 16). That is, the apparently great phylogenetic distance between some host plants suggests that the bacterial component can adapt to a wide range of genetic backgrounds. In the hope of elucidating phylogenetic relationships among the 10 families of angiosperms engaged in nitrogen-fixing symbioses, we employed comparative gene sequencing of *rbcL*.**

Table 1. Angiosperm families that participate in nodular nitrogen-fixing symbioses and the frequency of this association in each family

Prokaryote	Family	Total no. of genera*/genera having root nodules†
Rhizobium	Fabaceae	640/most
	Ulmaceae	18/1
Frankia	Betulaceae	6/1
	Casuarinaceae	4/4
	Elaeagnaceae	3/3
	Myricaceae	3/2
	Rhamnaceae	55/7
	Rosaceae	100/5
	Datiscaceae	3/1
	Coriariaceae	1/1

*From Cronquist (3).
†From Akkermans and van Dijk (4), Bond (5), and Torrey and Berg (6).

MATERIALS AND METHODS

Recently, phylogenetic trees were presented for angiosperms based on 499 sequences of the chloroplast gene *rbcL*, which encodes the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (EC 4.1.1.39), the chief carbon-fixing enzyme in photosynthesis (17). Included in that analysis were *rbcL* sequences representing the three subfamilies of legumes, as well as representatives of eight other families with nitrogen-fixing symbioses: Betulaceae, Casuarinaceae, Coriariaceae, Datiscaceae, Myricaceae, Rhamnaceae, Rosaceae, and Ulmaceae. Contrary to traditional systematic treatments, the phylogenetic trees based on *rbcL* sequences (17) suggested that angiosperm families that participate in nodular symbioses occur together, interspersed with families that lack these symbioses, in one clade corresponding generally to part of the dicot subclass Rosidae (called "Rosid-I"; Fig. 1). However, for most of the families with nitrogen-fixing symbioses the actual genera involved in root nodule symbiosis were not included in the analysis (17). Furthermore, the analysis of 499 *rbcL* sequences (17) resulted in 3900 trees. Because of the large number of sequences involved and the numerous most parsimonious trees obtained, Chase *et al.* (17) admit that their analysis suffers from uncertainties, although it provides a general guide for more detailed studies.

To elucidate the evolutionary origin of nodular nitrogen-fixing symbiosis in angiosperms, we therefore conducted a broad phylogenetic analysis of *rbcL* sequences, including sev-

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.

‖Deceased December 15, 1994.
**Sequences reported in this paper have been deposited in the GenBank data base (accession nos. V17038, V17039, and U20805).

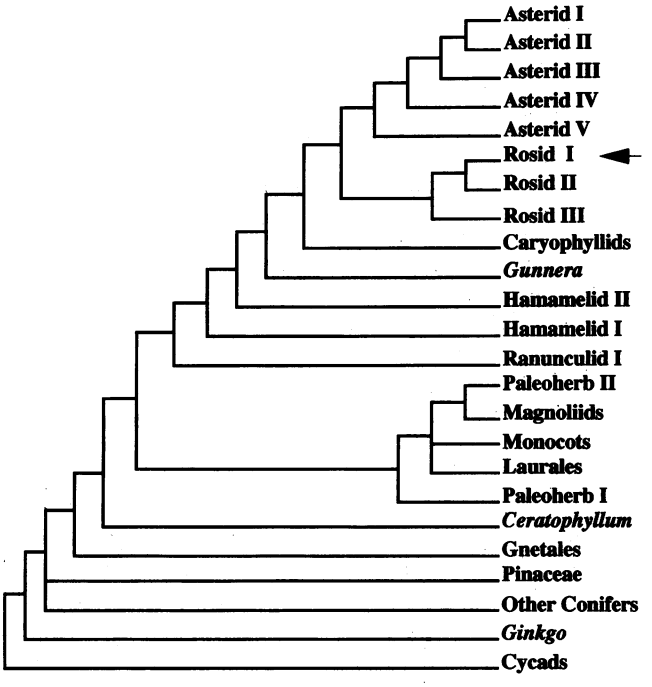

FIG. 1. Summary of the major clades identified in the strict consensus of 3900 equally parsimonious trees based on *rbcL* sequences for 499 taxa (17).

eral sequences for actinorhizal taxa (Table 2) not included previously by Chase *et al.* (17), to ensure that all angiosperm families known to participate in nodular symbioses were represented by taxa actually involved in nitrogen-fixing symbioses. Sequences were determined for members of Casuarinaceae (*Allocasuarina*) and Elaeagnaceae (*Elaeagnus* and *Shepherdia*), and recently published sequences for Rhamnaceae (*Ceanothus*) (22), Rosaceae (*Purshia* and *Cercocarpus*) (22), and Betulaceae (*Alnus*) (21) were also included (Table 2). As a result, all families of angiosperms (except Ulmaceae) known to host nitrogen-fixing bacteria in root nodules are represented in this analysis by *rbcL* sequences for genera that truly are hosts. In Ulmaceae, nitrogen fixation is known to occur only in *Parasponia*, a member of subfamily Celtoideae (9). Celtoideae are represented herein by sequences for the related genera *Celtis* and *Trema*. Also included were 78 additional *rbcL* sequences from genera lacking nitrogen-fixing symbioses in both nitrogen-fixing and non-nitrogen-fixing families of the Rosid I clade, chosen using the *rbcL* trees of Chase *et al.* (17) as a general guide, but with additional sequences added.

Five genera (*Trochodendron, Tetracentron, Platanus, Sabia,* and *Lambertia*) previously shown to be basal to the Rosid I lineage (17) were used as outgroups to root the phylogenetic trees. These 99 *rbcL* sequences were subjected to phylogenetic analysis using PAUP version 3.1.1 (24), which attempts to find the shortest (most parsimonious) trees. Because of the smaller scope of this project, we were able to conduct a more detailed phylogenetic analysis of the Rosid I clade than was possible in the larger analysis of Chase *et al.* (17). The heuristic search involved 800 replications with random taxon addition, using tree bisection-reconnection (TBR) branch swapping and saving a single shortest tree from each replicate. The shortest of these trees were then used as starting trees for further analyses,

Table 2. Representatives of angiosperm families with nitrogen-fixing symbioses for which *rbcL* sequences were used in the current study

Species	Family	Ref.
Species with nodular nitrogen-fixing symbioses and the families they represent		
Pisum sativum L.	Fabaceae (Papilionoideae)	18
Medicago sativa L.	Fabaceae (Papilionoideae)	19
Albizia julibrissin Durazz.	Fabaceae (Mimosoideae)	17
Bauhinia sp.	Fabaceae (Caesalpinioideae)	17
Coriaria myrtifolia L.	Coriariaceae	17
Datisca cannabina L.	Datiscaceae	20
Datisca glomerata L.	Datiscaceae	20
Myrica cerifera L.	Myricaceae	17
Casuarina litorea L.	Casuarinaceae	17
Allocasuarina muelleriana Miq.	Casuarinaceae	GenBank U20805
Alnus crispa (Ait.) Pursh	Betulaceae	21
Purshia tridentata (Pursh) D.C.	Rosaceae	22
Cercocarpus ledifolius Nutt.	Rosaceae	22
Ceanothus sanguineus Pursh.	Rhamnaceae	22
Shepherdia canadensis (L.) Nutt.	Elaeagnaceae	GenBank V17039
Elaeagnus angustifolia L.	Elaeagnaceae	GenBank V17038
Species lacking nodular nitrogen-fixing symbioses and the families they represent		
Octomeles sumatrana Miq.	Datiscaceae	20
Tetrameles nudiflora R. Br.	Datiscaceae	20
Betula nigra L.	Betulaceae	17
Rhamnus cartharticus L.	Rhamnaceae	17
Celtis yunnanensis C. K. Schneid.	Ulmaceae	17
Trema micrantha Blume	Ulmaceae	17
Geum chiloense Balb. ex Ser.	Rosaceae	23
Photinia fraseri Dress	Rosaceae	23
Prunus virginiana L.	Rosaceae	23
Rubus idaeus L.	Rosaceae	22
Spiraea vanhouttei (Briot) Zabel	Rosaceae	23

Species with and without nodular nitrogen-fixing symbioses were included. References indicate the source for each sequence. The remaining 72 sequences included in the analysis, from families not known to possess nodular nitrogen-fixing symbioses, were taken from Chase *et al.* (17). Legume subfamilies are given in parentheses.

again with TBR branch swapping, but with all most parsimonious trees saved.

The size of the data set precluded internal assessments of support by using bootstrap (25) and standard decay (26, 27) analyses. To assess support for the "nitrogen-fixing clade" described below, we performed a decay analysis on this single clade. We constructed a constraint tree having a basal polytomy and a single branch corresponding to the "nitrogen-fixing clade" and then saved all trees one and two steps longer than the most parsimonious trees that were incompatible with this specified topology. Groups of genera that collapse (i.e., cease to be monophyletic or "decay") in trees two or more

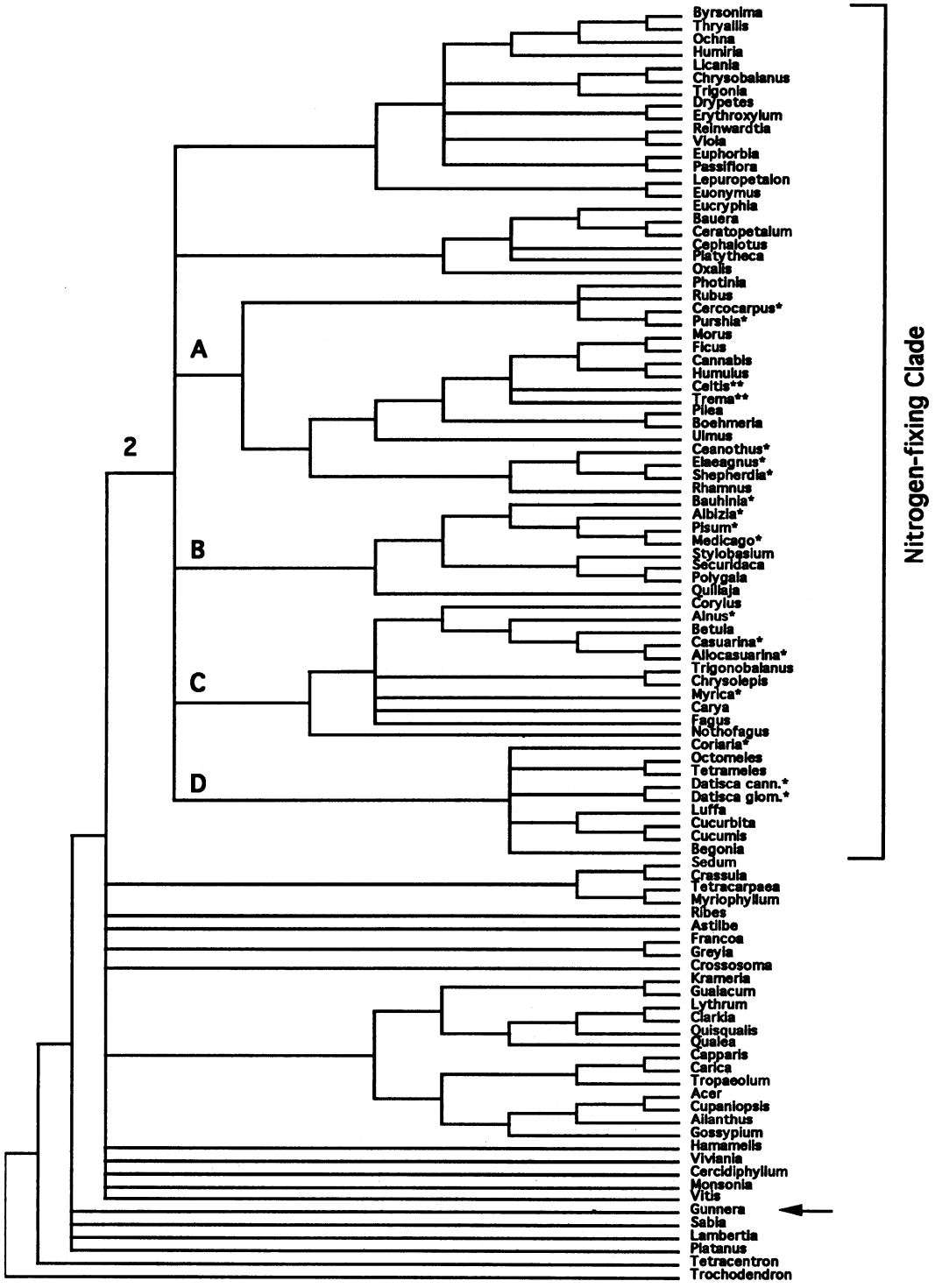

FIG. 2. Strict consensus of 558 shortest trees resulting from phylogenetic analysis of 99 *rbcL* sequences representing the Rosid I lineage (see Fig. 1) and additional taxa. All taxa engaged in nodular nitrogen-fixing symbioses are indicated by an asterisk. *Datisca cann., Datisca cannabina; Datisca glom., Datisca glomerata.* In Ulmaceae, nitrogen fixation is known to occur only in *Parasponia*, a member of subfamily Celtoideae. Celtoideae are represented herein by sequences of the related genera *Celtis* and *Trema*, each of which is marked by two asterisks. *Gunnera* (Gunneraceae), which hosts nitrogen-fixing cyanobacterial symbionts in leaf glands, rather than root nodules, is noted by an arrow. The "nitrogen-fixing clade" is supported by a decay value of 2, as indicated at the base of this clade (see text for further discussion). Letters (A–D) above lines designate the four subclades of the nitrogen-fixing clade that actually contain taxa involved in nitrogen-fixing symbioses. These four subclades form a monophyletic group in 74% of the 558 shortest trees.

steps longer than the shortest trees are considered better supported than those that decay at only a single step longer.

RESULTS AND DISCUSSION

The phylogenetic analysis produced 558 trees of 3552 steps distributed on three islands (28). The strict consensus of these trees (Fig. 2) indicates that all of the families with nitrogen-fixing symbioses occur together as part of the same large clade (labeled "nitrogen-fixing clade") with several families lacking these symbioses. This clade is maintained in all trees found that were a single step longer than the shortest trees, but it collapses in the strict consensus of all trees found that were two steps longer than the shortest trees. Although the nitrogen-fixing clade collapses in trees only two steps longer than the shortest trees, the maintenance of this clade through the first step of the decay analysis provides additional support for the monophyly of the nitrogen-fixing clade, especially given the large number and taxonomic breadth of taxa included in the analysis. Two of the six lineages of this clade do not contain any members known to host nitrogen-fixing bacteria. The four lineages of this clade that contain symbiotic nitrogen-fixing genera form a monophyletic group in 74% of the 558 trees, suggesting that all of the families with nodular nitrogen fixation may have indeed shared a common ancestor not shared by other families of angiosperms.

Phylogenetic analysis of *rbcL* sequences suggests a close relationship among the Rosaceae, the celtoid line (*Trema* and *Celtis*) of Ulmaceae, Rhamnaceae, and Elaeagnaceae, all with genera that host nitrogen-fixing bacteria, and the Moraceae, Cannabaceae, and Urticaceae, which do not. This relationship is supported not only by *rbcL* base substitutions but also by structural variation in the region flanking the 3′ end of *rbcL*. These families are characterized by a 4-bp duplication located 12 bp 3′ to the terminus of *rbcL*, although the sequences of *Humulus* (Cannabaceae) and *Pilea* and *Boehmeria* (Urticaceae) were not analyzed for this feature (Fig. 3). This duplication is absent from all other species of Rosidae examined (26 taxa in 19 families; Fig. 3) (22).

The presence in a single clade of all families that engage in symbiotic nodular nitrogen fixation indicates a much closer relationship among these taxa than is suggested by current classification schemes. Thus, molecular systematic data challenge the morphologically based inference that many families that host nitrogen-fixing bacteria are distantly related (3, 11–14). These findings, furthermore, suggest a single evolutionary origin of the underlying capacity for symbiotic nodular nitrogen fixation. Significantly, Gunneraceae, an angiosperm family that hosts nitrogen-fixing cyanobacterial symbionts in leaf glands (rather than in root nodules), are not a member of this nitrogen-fixing clade (Fig. 2). Gunneraceae, therefore, clearly represent an independent evolution of symbiotic nitrogen-fixing ability in angiosperms.

The clade containing symbiotic nitrogen-fixing taxa also includes nine families that do not participate in root nodule symbiosis (Begoniaceae, Cucurbitaceae, Fagaceae, Juglandaceae, Cannabaceae, Moraceae, Polygalaceae, Surianaceae, Urticaceae). Furthermore, with the exception of subfamily Papilionoideae and Mimosoideae of Fabaceae, the vast majority of species in families such as Rosaceae, Rhamnaceae, Ulmaceae, and the Caesalpinioideae (Fabaceae) are not known to nodulate and fix nitrogen (Table 1). For example, of the 18 genera in Ulmaceae, only *Parasponia* of subfamily Celtoideae has a symbiotic nitrogen-fixing association. Therefore, if the underlying predisposition to engage in symbiotic nitrogen fixation stems from a single, common origin, members of this clade that do not share this symbiosis must have lost the ability to form such associations. Alternatively, the ancestor of the nitrogen-fixing clade may have evolved the genetic components that would ultimately permit the evolution of symbiotic nodular nitrogen fixation. Following the establishment of these conditions, the necessary genetic background was present to allow parallel, recurrent evolution of symbiotic nitrogen fixation in the subsequent diversification of this clade. Both hypotheses call for mutations that established the symbiotic association with nitrogen-fixing bacteria, or at least the basis for this symbiosis, in the ancestor of the nitrogen-fixing clade.

Within the nitrogen-fixing clade, the legume/rhizobia symbioses and actinorhizal/frankiae symbioses fall into distinct lineages. One of the four subclades that contain members with symbiotic nitrogen-fixing associations consists solely of legumes and related nonsymbiotic relatives. Actinorhizals occur, along with nonsymbiotic taxa, in each of the other three clades. The celtoid line (*Trema* and *Celtis*) of Ulmaceae, representing the *Parasponia*/rhizobia symbiosis, occurs nested within a larger clade that includes actinorhizal/frankiae symbioses in Rosaceae, Rhamnaceae, and Elaeagnaceae. The distinctness of the legumes from all actinorhizals is further supported by aspects of the symbioses themselves. A major feature that distinguishes the legume/rhizobia symbioses from those involving actinorhizals and frankiae or *Parasponia* and rhizobia lies in the ontogeny of the nodule. Whether initiated by root hair infection, intercellular penetration, or infection at wound sites, all legume nodules result from cell divisions within the root cortex and have a stem-like anatomy with peripheral vascular tissue. All actinorhizal and *Parasponia* nodules, regardless of the mode of infection, form by modification of lateral roots and maintain a central vascular tissue. That organogenesis is directed by the host plant is confirmed by observations that *Parasponia* rhizobia induce typical legume nodules on compatible legumes and legume rhizobia induce root-like nodules on *Parasponia* (29). Also, some legumes may develop spontaneous nodules in the absence of a bacterial symbiont, confirming that the developmental pathway leading to the formation of nodules is encoded in the plant genome (30).

The fact that the Celtoid line of Ulmaceae and the legumes (families symbiotic with Rhizobiaceae) occur in two separate clades, with the former more closely allied with angiosperm families that are nodulated by actinomycetes (*Frankia*), may be relevant to the origin of root-nodule symbioses. The steps

```
                                  |- -|
Rubus          TAAtctagcaattacttactgttagttctcttaatt
Geum           TAAtccagcaattacttactcttagttcttttaatt
Elaeagnus      TAAtcttgtaattacttacagctcgttcttttaatt
Celtis         TAAtccagcaattacttcctgttccttttcttaatt
Morus          TAAtccagcaattacttactgttccttttcttaatt
Rhamnus        TAAtccagtaattacttactgttcgttcttgtaatt
Ceanothus      TAAtccagtaattacttactgttcgttctcttaatt
Chrysolepis    TAAtccagtaattac----cgctcgttctcttaatt
Myrica         TAAtccaataattac----ccctcgttcttttaatt
Casuarina      TAAtccagtaattat----cgctcgttctcttaatt
Ceratopetalum  TAAtccagtaattcc----tgttcgttctcttaatt
Bauera         TAAtccagtaattcc----tgttcgttctcttaatt
Cephalotus     TAAtccagtaattac----tgttcgttctcttaatt
```

FIG. 3. Duplicated sequence (in brackets) adjacent to 3′ end (bases in uppercase) of *rbcL* present in several families with nitrogen-fixing symbioses. *Chrysolepis, Myrica*, and *Casuarina*, representing families with nitrogen-fixing symbioses, lack the duplication. *Ceratopetalum, Bauera*, and *Cephalotus*, representing families without nitrogen-fixing symbioses, also lack the duplication.

leading to the formation of effective nodules in host plants involve a combination of plant and bacterial signals. The details of this "two-way molecular conversation" (31) vary depending on the host and bacteria involved. It seems unlikely that a new angiosperm host with the capacity for nodulation would successfully "converse" with two different bacterial species in separate recognition events. Rather, this angiosperm host more likely acquired the necessary cellular machinery that permitted symbiosis, of which one or the other type of bacteria could take advantage.

Literature in the area of symbiotic nitrogen fixation provides few clues as to the nature of the critical steps involved in the formation of the symbiotic association (reviews in refs. 32 and 33). Several lines of evidence actually suggest that the plant genes and proteins involved in nodulation are not unusual (reviewed in ref. 2). For example, recent studies suggest that the infection threads involved in nodulation may be related to normal cell functions that occur in all plants (34). Furthermore, the nodulation genes in rhizobia can be induced by flavonoids common to many angiosperms, as demonstrated by Peters *et al.* (35). In addition, one of the earliest appearing nodule-specific proteins (nodulins), ENOD12, is a cell-wall component also found in stems and flowers (36). Of the late nodulins, leghemoglobin may be present in all plants (37), and at least six of the enzymes involved in nodulation are normal "housekeeping" proteins (38). Evidence from legumes also indicates that particular steps essential to symbiotic establishment in one species may simply be by-passed or achieved differently in other species (39). Single-gene mutations can simultaneously prevent symbioses with nitrogen-fixing bacteria and association with vesicular-arbuscular mycorrhizae (40). Nodulation may therefore involve interactions among many different common genes, rather than unique genes ready to be switched on when appropriate microsymbionts are available (2).

The longstanding idea that symbiotic nitrogen-fixing species are taxonomically diverse (3, 11–14) has engendered the belief that nodulation could be extended to other plants by the molecular transfer of genes critical for nodular symbiosis. For example, rhizobial-induced galls were recently obtained on rice (41). Although the transfer of root nodulation and nitrogen fixation to plants such as rice may ultimately be possible, phylogenetic analyses of *rbcL* sequence data indicate clearly that of the approximately 250,000 to 300,000 species and 380 families of angiosperms, only one lineage of closely related taxa achieved the underlying genetic architecture necessary for root-nodule symbiosis. The likelihood of only a single origin of the predisposition for root-nodule symbioses in angiosperms therefore has implications regarding the crucial evolutionary events required in this process. Future efforts to unravel the process and evolution of nitrogen-fixing symbioses, and the transfer of this capacity to nonnodulating species, should first focus on related taxa from the nitrogen-fixing clade that possess and lack symbiotic nitrogen-fixing ability. Concomitantly, nodulating and nonnodulating members of this nitrogen-fixing clade should be examined to ascertain whether recurrent losses or recurrent gains of nitrogen-fixing ability have occurred.

The surviving authors dedicate this paper to our friend and colleague, Peter Martin, who died in December 1994. Peter provided inspiration, stimulus, and a wealth of ideas to a generation of botanists. This research was supported in part by National Science Foundation Grant BSR-9007614 to D.E.S., National Science Foundation Grant DEB-9306913 to S.M.S. and L. H. Rieseberg, and U.S. Department of Agriculture Grant 93-37305-9082 to B.C.M.

1. Allen, O. N. & Allen, E. K. (1976) *Symbiotic Nitrogen Fixation in Plants* (Cambridge Univ. Press, Cambridge, U.K.).
2. Sprent, J. I. & Sprent, P. (1990) *Nitrogen Fixing Organisms* (Chapman & Hall, New York).
3. Cronquist, A. (1981) *An Integrated System of Classification of Flowering Plants* (Columbia Univ. Press, New York).
4. Akkermans, A. D. L. & van Dijk, C. (1981) in *Nitrogen Fixation*, ed. Broughton, W. J. (Clarendon, Oxford), Vol. 1, pp. 57–103.
5. Bond, G. (1983) in *Biological Nitrogen Fixation in Forest Ecosystems*, eds. Gordon, J. C. & Wheeler, C. T. (Nijhoff, Dordrecht, The Netherlands), pp. 55–87.
6. Torrey, J. G. & Berg, R. H. (1988) *Am. J. Bot.* **75,** 864–874.
7. Baker, D. D. & Mullin, B. C. (1992) in *Biological Nitrogen Fixation*, eds. Stacey, G., Burris, R. H. & Evans, H. J. (Chapman & Hall, New York), pp. 259–291.
8. Mullin, B. C., Swensen, S. M. & Goetting-Minesky, P. (1990) in *Nitrogen Fixation Achievements and Objectives*, eds. Gresshoff, P. M., Roth, L. E., Stacey, G. & Newton, W. E. (Chapman & Hall, New York), pp. 781–787.
9. Trinick, M. J. & Galbraith, J. (1980) *New Phytol.* **86,** 17–26.
10. Torrey, J. G. & Tjepkema, J. D. (1979) *Bot. Gaz. Suppl.* **140,** i-ii.
11. Thorne, R. F. (1992) *Aliso* **13,** 365–389.
12. Thorne, R. F. (1992) *Bot. Rev.* **58,** 225–348.
13. Takhtajan, A. (1987) *System of Magnoliophyta* (U.S.S.R. Acad. Sci., Leningrad).
14. Dahlgren, R. M. T. (1980) *Bot. J. Linn. Soc.* **80,** 91–124.
15. Sprent, J. I. & Raven, J. R. (1992) in *Biological Nitrogen Fixation*, eds. Stacy, G., Burris, R. H. & Evans, H. J. (Chapman & Hall, New York), pp. 461–496.
16. Mullin, B. C. (1992) in *Nodulation and Nitrogen Fixation in Rice*, eds. Khush, G. S. & Bennett, J. (Int. Rice Res. Inst., Manila, The Philippines), pp. 67–75.
17. Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., *et al.* (1993) *Ann. Mo. Bot. Gard.* **80,** 528–580.
18. Zurawski, G., Whitfeld, P. R. & Bottomley, W. (1986) *Nucleic Acids Res.* **14,** 3975.
19. Aldrich, J., Cherney, B., Merlin, E. & Palmer, J. (1986) *Nucleic Acids Res.* **14,** 9535.
20. Swensen, S. M., Mullin, B. C. & Chase, M. W. (1994) *Syst. Bot.* **19,** 157–168.
21. Bousquet, J., Strauss, S. J. & Li, P. (1992) *Mol. Biol. Evol.* **9,** 1076–1088.
22. Morgan, D. R., Soltis, D. E. & Robertson, K. R. (1994) *Am. J. Bot.* **81,** 890–903.
23. Morgan, D. R. & Soltis, D. E. (1993) *Ann. Mo. Bot. Gard.* **80,** 631–660.
24. Swofford, D. L. (1993) PAUP: *Phylogenetic Analysis Using Parsimony* (Ill. Nat. Hist. Survey, Champaign, IL), Version 3.1.1.
25. Felsenstein, J. (1985) *Evolution* **39,** 783–791.
26. Bremer, K. (1988) *Evolution* **42,** 795–803.
27. Donoghue, M. J., Olmstead, R. G., Smith, J. F. & Palmer, J. D. (1992) *Ann. Mo. Bot. Gard.* **79,** 333–345.
28. Maddison, D. R. (1991) *Syst. Zool.* **40,** 315–328.
29. Becking, J.-H. (1992) in *Biological Nitrogen Fixation*, eds. Stacey, G., Evans, H. J. & Burris, R. H. (Chapman & Hall, New York), pp. 497–559.
30. Caetano-Anollés, G., Joshi, P. & Gresshoff, P. M. (1992) in *Plant Biotechnology and Development*, ed. Gresshoff, P. M. (CRC, Boca Raton, FL), pp. 61–70.
31. Fisher, R. F. & Long, S. R. (1992) *Nature (London)* **357,** 655–670.
32. Young, J. P. W. & Johnston, A. W. B. (1989) *Trends Ecol. Evol.* **4,** 341–349.
33. Dénarié, J., Debelle, F. & Rosenberg, C. (1992) *Annu. Rev. Microbiol.* **46,** 497–631.
34. VandenBosch, K. A., Bradley, D. J. & Knox, J. P. (1989) *EMBO J.* **8,** 335–342.
35. Peters, G. A., Frost, J. W. & Long, S. R. (1986) *Science* **223,** 977–980.
36. Scheres, B., Van de Wiel, C., Zalensky, A., Horvath, B., Spaink, H., Van Eck, H., Zwarthruis, F., Wolters, A. M., Gloudemans, T., Van Kammen, A. & Bisseling, T. (1990) *Cell* **60,** 281–294.
37. Landsmann, J., Dennis, E. S., Higgins, T. J. V., Appleby, C. A., Kortt, A. A. & Peacock, W. J. (1986) *Nature (London)* **324,** 166–168.
38. Sanchez, F., Padilla, J. E., Perez, H. & Lara, M. (1991) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **42,** 507–528.
39. Sprent, J. I. (1989) *New Phytol.* **111,** 129–153.
40. Duc, G., Trouvelot, A., Giannazzi-Pearson, V. & Giannazzi, S. (1989) *Plant Sci.* **60,** 215–222.
41. Al-Mallah, M. K., Davey, M. R. & Cocking, E. C. (1989) *J. Exp. Bot.* **40,** 473–477.