Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 7844–7848, August 1992
Evolution

Extensive variation in evolutionary rate of *rbcL* gene sequences among seed plants

(angiosperm/gymnosperm/large subunit of ribulose-1,5-bisphosphate carboxylase/variable molecular clock/phylogeny)

JEAN BOUSQUET*†‡, STEVEN H. STRAUSS*, ALLAN H. DOERKSEN*, AND ROBERT A. PRICE§

*Department of Forest Science, Oregon State University, Corvallis, OR 97331-5705; †Centre de Recherche en Biologie Forestière, Faculté de Foresterie et de Géomatique, Université Laval, Ste-Foy, Québec, Canada G1K 7P4; and §Department of Biology, Indiana University, Bloomington, IN 47405

Communicated by R. W. Allard, April 27, 1992

ABSTRACT **Extensive variation in synonymous and nonsynonymous rates of substitution was observed among 50 sequences of the gene coding for the large subunit of ribulose-1,5-bisphosphate carboxylase (*rbcL*) representing bryophyte, conifer, dicot, and monocot taxa. Relative rate tests revealed rate differences of up to 138% for nonsynonymous substitutions and up to 85% for synonymous ones. Within angiosperms, the annual forms evolved more rapidly, on average, than perennial forms. This rate heterogeneity was more extensive at nonsynonymous sites than synonymous ones, and it resulted primarily from a recent acceleration of substitution rate in many groups of angiosperms.**

Using the principle of the molecular clock (1), there have been a number of attempts to estimate divergence times of major plant taxa from chloroplast (2) and nuclear DNA sequences (2–4). An implicit assumption of the molecular clock, however, is an approximate constancy of evolutionary rate over time (5). Several recent studies have reported heterogeneity of substitution rates among plants in chloroplast DNA. Such heterogeneity has been observed for several chloroplast genes between lineages leading to pea and tobacco and spinach (6, 7), among cereals (8, 9), and among lineages leading to maize, tobacco, and liverwort (2). In a recent study involving *rbcL* (gene coding for the large subunit of ribulose-1,5-bisphosphate carboxylase) sequences, the rate of evolution appeared faster in annual angiosperms than in *Marchantia*, *Chlamydomonas*, and *Euglena* (10). For this gene, overall rates of substitution were also found heterogenous within the family Betulaceae (11) and within the Poaceae (12). When comparing overall rates of evolution of *rbcL* within the monocot family Arecaceae to those estimated in the Poaceae, an extensive slowdown was observed in the Arecaceae, the palm family (13). In the Saxifragaceae *sensu lato*, however, no significant heterogeneity was detected among *rbcL* sequences determined (14).

In this paper, we present evidence that greatly expands upon previous observations of rate heterogeneity for chloroplast genes. We report extensive variation in substitution rates among 50 *rbcL* gene sequences from *Marchantia* and from 18 families of seed plants representing the gymnosperms, the dicots, and the monocots. The observed patterns of variation were related to life-history factors, and rate variation was more extensive for nonsynonymous than synonymous rates of substitution.¶

MATERIALS AND METHODS

A total of 43 *rbcL* sequences was drawn from the literature (see Fig. 1, citations available from the authors upon request) and the complete *rbcL* nucleotide sequences were determined for eight additional taxa representing woody perennials: *Magnolia Soulangeana, Populus tremuloides, Pinus edulis, Pinus griffithii, Pinus longaeva, Pinus pinea, Pinus radiata*, and *Podocarpus gracilior*. Total genomic DNA was obtained from leaves (15) or from needles (16). Oligonucleotide primers for amplification and sequencing were kindly supplied by G. Zurawski, and additional primers were designed by identifying homologous regions among sequences already published (monocots, dicots, and *Marchantia*). The amplification of *rbcL* was conducted using the polymerase chain reaction with primers homologous to the beginning and the end of the coding region in conjunction with primers annealing to conserved regions of the flanking ORF512 and the gene *atpB* (17). Prior to amplification, one primer of each pair was phosphorylated while keeping the other unphosphorylated; by subsequently digesting the phosphorylated strand of amplified DNA with λ exonuclease (18), we obtained single-stranded DNA suitable for sequencing. This process was repeated for the alternate strand for each pair of primers. Amplification, DNA purification, and sequencing followed previously described procedures (11).

Pairwise synonymous (K_s) and nonsynonymous (K_a) numbers of substitutions corrected for multiple hits were calculated (19). Overall numbers of substitution (K_o) were calculated as weighted averages of K_s and K_a. The distance matrices obtained were analyzed with the neighbor-joining method of phylogenetic tree construction (20). Kimura's two-parameter method (21) was also used to estimate overall rates of substitution where rates of transitions and transversions are modeled separately. To assess the reliability of the topologies obtained from the neighbor-joining method, standard parsimony as well as various weighted parsimony strategies were conducted using PAUP (phylogenetic analysis using parsimony; ref. 22): (*i*) transitions and transversions received weights (1 vs. 2, respectively) proportional to the reciprocal of their frequency; (*ii*) first, second, and third codon positions received weights (4 vs. 7 vs. 1, respectively) proportional to the reciprocal of substitution frequency at each position; (*iii*) a combination of the two last weighting schemes. Bootstrap confidence intervals were estimated under the assumption of character independence from 100 replicates for standard parsimony.

Relying solely on the tree obtained from the neighbor-joining analysis of overall substitution rates (Fig. 1), a set of

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: K_a, nonsynonymous rate of substitution; K_s, synonymous rate of substitution; K_o, overall rate of substitution; *rbcL*, gene coding for the large subunit of ribulose-1,5-bisphosphate carboxylase.

‡To whom reprint requests should be addressed at †.

¶The sequences reported in this paper have been deposited in the GenBank data base [accession nos. M58393 (*Magnolia Soulangeana*), M58392 (*Populus tremuloides*), X58137 (*Pinus edulis*), X58131 (*Pinus griffithii*), X58132 (*Pinus longaeva*), X58133 (*Pinus pinea*), X58134 (*Pinus radiata*), and X58135 (*Podocarpus gracilior*)].

taxa representative of the major clades observed was chosen to conduct a series of relative rate tests (23) for K_s and K_a. When a clade was composed of more than two taxa, such as for the Pinaceae or the Poaceae, a representative taxon was selected that had roughly an average distance from the root of the tree for the clade. The reference taxa for comparison purposes were chosen in remote sister groups to avoid illegitimate choices based on local topological errors (details given in Table 2). The relative rate tests rely on pairwise rates of substitution and are therefore independent of the topology used to select the taxa to be compared. Rate differences between pairs of taxa relative to their common reference taxon are expressed as their pairwise differentiation: $[(K_{13} - K_{23})/K_{12}] \times 100$, where K is the fraction of nucleotides substituted among taxa, 1 and 2 are the taxa being compared, and 3 is the reference taxon.

Wilcoxon's signed-rank tests (24) were used to assess heterogeneity patterns observed among groups of taxa differing in life-history traits (e.g., annuals vs. perennials). The tests were conducted separately for K_s and K_a. Pairwise rate differences among all possible combinations of members of two contrasted life-history categories that were defined *a priori* were used for analysis. For instance, when testing synonymous rate differences between annuals (*Spinacea, Neurachne, Nicotiana, Gossypium, Pisum, Flaveria*, Table 2) and perennials (*Pinus, Penthorum, Quercus, Magnolia, Serenoa, Itea, Carpentaria, Francoa*, Table 2), we used the rates given in the lower left side of Table 2 (6 × 8 matrix).

RESULTS

Estimated Phylogenies. At higher levels of organization, the topology of the tree estimated from the neighbor-joining analysis of K_o (Fig. 1) was congruent with those obtained from the neighbor-joining analysis of K_a and from parsimony analysis using a simultaneous weighting of transition/transversion and first/second/third codon positions (not shown). This topology is also in general agreement with traditional views regarding the phylogeny of seed plants at higher levels of organization (25): conifers as members of the gymnosperms would have diverged first, followed by the diversification of angiosperms in their two basal classes, the monocots and the dicots, which formed two monophyletic groups. The topology of the Coniferales is also in agreement with existing views, with the Pinaceae forming a natural group distinct from *Podocarpus*, and the genus *Pinus* forming a distinct group split into the subgenera *Haploxylon* and *Diploxylon*. Notably, the neighbor-joining analysis of K_s and standard parsimony appeared unreliable as they grouped the Arecaceae with the dicots away from the Poaceae. The diversification of the dicots into their six subclasses (25) appears rapid from all the analyses: besides the constant earlier divergence of the Magnolidae here represented by the Magnoliaceae, the remaining five subclasses did not form natural groups; several of these subclasses appeared polyphyletic and showed topological variation from one method to the next, notably for the subclasses Hamamelidae, Rosidae, and Dilleniidae. In these cases, bootstrap estimates were generally lower than 50% for standard Wagner parsimony. Such inconsistent topologies, presumably a consequence of rapid radiation of the subclasses of the dicots, were also obtained from the analysis of sequences of the nuclear gene encoding glyceraldehyde-3-phosphate dehydrogenase (3), from partial sequences of 18S and 26S rRNAs (26–28), and from smaller subsets of dicot *rbcL* sequences (14, 28, 29).

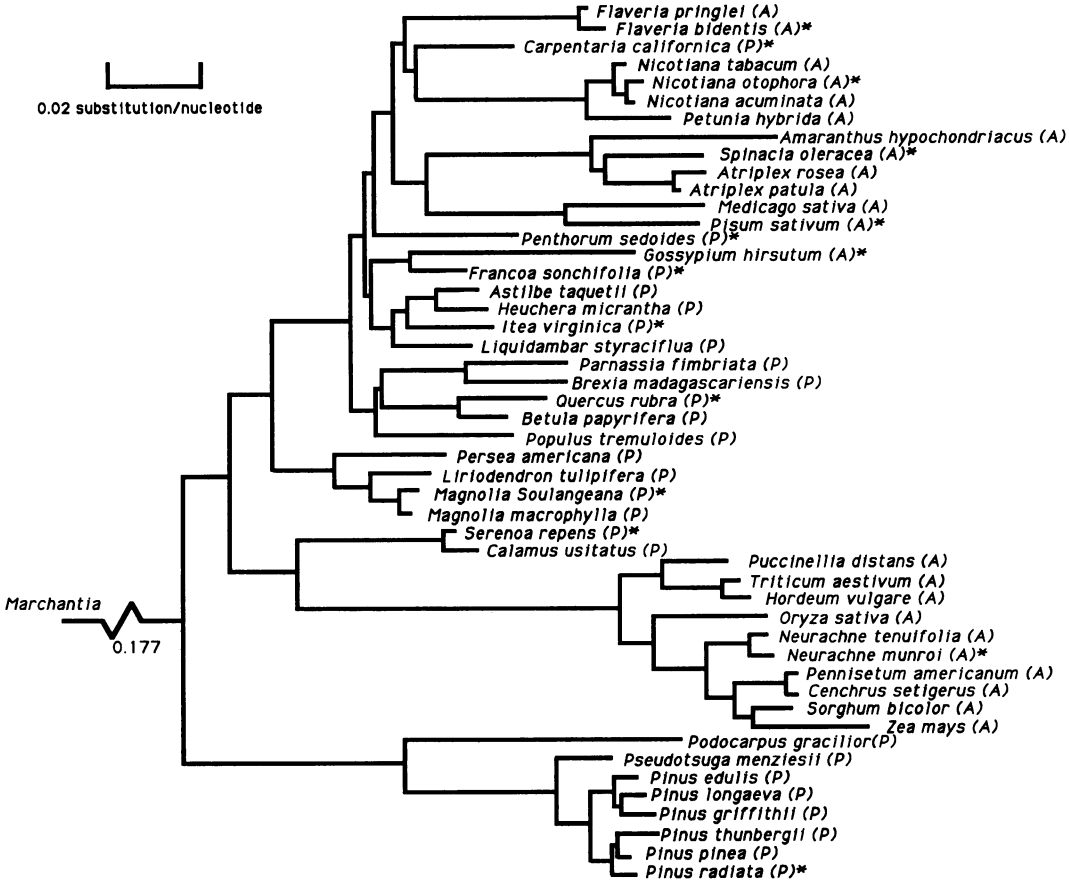

FIG. 1. Phylogenetic tree of 49 seed plant taxa based on neighbor-joining analysis of overall pairwise substitution rates (K_o) of *rbcL* sequences. P, perennials; A, annuals. Taxa followed by an asterisk were used in relative rate tests (Table 2).

Table 1. Comparisons of substitution rates per site among selected contrasting angiosperm taxa relative to *P. menziesii*

Taxon	No. of nucleotides compared	Observed no. of substitutions	Transitions (tn)	Transversions (tv)	tn/tv ratio	Kimura's two-parameter rate	K_a	K_s	K_o
Z. mays	1422	232	142	90	1.6	0.186	0.053	0.848	0.233
T. aestivum	1425	227	133	94	1.4	0.181	0.040	0.752	0.201
S. repens	985	122	84	38	2.2	0.137	0.023	0.675	0.167
M. macrophylla	1425	165	122	43	2.8	0.128	0.027	0.580	0.152
S. oleracea	1425	214	147	67	2.2	0.171	0.042	0.786	0.212
Q. rubra	1425	173	121	52	2.3	0.134	0.028	0.615	0.161
I. virginica	1407	166	123	43	2.9	0.129	0.021	0.631	0.154
P. hybrida	1425	209	144	65	2.2	0.166	0.043	0.740	0.202

Heterogeneity of Substitution Rates. There are striking differences in apparent rates of evolution (Fig. 1). Comparisons of substitution rates among a sample of diverse angiosperm taxa showed that this heterogeneity is made up of several components (Table 1). In some taxa such as *Magnolia* and *Itea*, the observed transition/transversion ratio is highly skewed and nearly double that of *Zea* and *Triticum*. K_a were roughly 20-fold lower in magnitude than K_s but appeared to vary substantially more; for example, *Zea* and *Itea* varied 2.5-fold at K_a but only 1.3-fold at K_s. Consequently, relative differences in branch length such as those observed on the neighbor-joining tree based on K_o (Fig. 1) appeared substantially larger when K_a were analyzed or when weighted parsimony strategies were used (not shown).

Relative rate tests (Table 2) indicated a large number of statistically significant ($P < 0.05$) instances of rate heterogeneity. The number (65 of 196) far exceeds that expected due to chance alone (about 10), given the number of tests conducted. The perennial seed plants tested, whether herbaceous or woody, appeared to evolve more slowly than the annual herbaceous plants. Wilcoxon's signed-rank tests between perennial and annual seed plant taxa showed that annuals are evolving significantly faster than perennials for K_s and K_a (Wilcoxon's T_s values of 98.5 and 12, respectively, both with $n = 48$ and significant at $P < 0.01$). Furthermore, the degree of heterogeneity for K_a substantially exceeded that for K_s: by comparing the 48 synonymous rate differences estimated among annuals and perennials tested to the 48 nonsynonymous rate differences, annuals were found to have evolved an average 25% (±26%) more rapidly than perennial seed plants for synonymous substitutions and 46% (±32%) more rapidly for nonsynonymous ones. This difference between synonymous and nonsynonymous substitutions was highly significant (parametric Student's t value of 3.51 and non-parametric Wilcoxon-derived Student's t value of 3.31, $P < 0.01$). When testing herbaceous perennials vs. woody taxa (*Penthorum* and *Francoa* vs. the six other woody perennials), no significant rate differences were observed for synonymous or nonsynonymous substitutions (both Wilcoxon's T_s values of 29 with $n = 12$, $P > 0.05$).

Variable Molecular Clock. The heterogeneity of substitution rates we observed could be the result of either a slowdown of evolutionary rate in the slower lineages or an acceleration of evolutionary rate in the faster lineages, or both. To analyze this problem, and because of the observed heterogeneity of substitution rates, a variable molecular clock strategy was developed where absolute rates of substitution per year were determined separately for each lineage (local clocks). Using *Chlamydomonas* as the outgroup, we determined the average branch lengths of a tree with six terminal groups (Fig. 2) that appear to be natural from our distance-matrix analyses of K_a and K_o, from our weighted parsimony analyses, and from current views on plant evolution. Only nonsynonymous rates of evolution could be used to construct this tree because of saturation in the overall and synonymous rates involving *Chlamydomonas*. Average nonsynonymous rates were calculated between groups, and with the topology fixed, terminal and internal branch lengths were estimated using a procedure that does not assume homogeneity of substitution rates (30). The tree obtained (Fig. 2)

Table 2. Relative rate tests among 15 selected taxa

Taxa	Spin	Neur	Nico	Goss	Pisu	Flav	Pinu	Pent	Quer	Magn	Sere	Itea	Carp	Fran	*Marchantia*
S. oleracea (Spin)	—	3	8	−2	9	**37***	**34***	**38***	**33***	**43***	**53†**	**58†**	**79†**	**86†**	**44†**
N. munroi (Neur)	−6	—	0	−5	12	20	17	**32***	**44†**	**46†**	**42***	**71†**	**55†**	**46†**	**27‡**
N. otophora (Nico)	−19	−6	—	−10	2	26	**31***	25	22	35	**61†**	**61†**	**72†**	**98†**	**37***
G. hirsutum (Goss)	−21	0	−7	—	12	**36***	26	**31***	**31***	5	**98†**	**49†**	**59†**	**138†**	**36***
P. sativum (Pisu)	5	−1	19	**28***	—	**40***	−3	32	36	15	**59†**	**60†**	**79†**	**125†**	18
F. bidentis (Flav)	−16	−14	4	10	−18	—	−3	−6	−10	28	**48***	22	38	**89†**	7
P. radiata (Pinu)	10	0	20	3	25	−6	—	20	13	−4	24	29	**22***	**67†**	4
P. sedoides (Pent)	**−38†**	**−33†**	−26	−16	**−43†**	**−28***	−25	—	−4	44	29	34	**52***	**81†**	7
Q. rubra (Quer)	−18	−18	−2	6	−22	5	1	22	—	−30	33	**60†**	**65***	**105†**	24
M. Soulangeana (Magn)	**−85†**	−25	**−77†**	**−66†**	**−67†**	**−73†**	12	**−62†**	**−58†**	—	14	20	6	23	18
S. repens (Sere)	−18	21	−7	−15	−11	4	29	21	13	32	—	7	−17	−14	17
I. virginica (Itea)	**−58†**	−17	**−43†**	**−37†**	**−52†**	**−47†**	−27	−2	**−40†**	**62†**	−11	—	23	**57***	8
C. californica (Carp)	**−55†**	−19	**−48†**	**−27***	**−55†**	**−50†**	−18	0	**−34†**	**75†**	−13	6	—	52	4
F. sonchifolia (Fran)	**−30***	−13	−15	−9	**−34†**	−20	7	13	−15	**41***	−5	**44***	32	—	−9

Numbers indicate rate differences in percent of pairwise substitution rates between taxa on the left and taxa at the top. K_a are above diagonal; K_s are below diagonal. For example, K_a in *S. oleracea* is 3% faster than *N. munroi* with reference to common reference taxon; K_s in *S. oleracea* is 6% faster than *N. munroi*. The choice of taxa compared is based on Fig. 1 (taxa followed by an asterisk in Fig. 1) except for *Marchantia* (selected based on Fig. 2). The choice of reference taxa is based on Fig. 1 (except for *Marchantia*, Fig. 2): reference taxon for comparisons involving dicots excluding *M. Soulangeana*: *M. macrophylla* (closest sister taxon); reference taxon for comparisons between *Magnolia* and other dicots: *S. repens* (closest sister taxon); reference taxon for comparisons between monocots and dicots: *P. menziesii* (closest sister taxon); reference taxon for comparisons between *P. radiata* and angiosperms: *Marchantia polymorpha*; reference for comparisons between *M. polymorpha* and seed plants: *Chlamydomonas reinhardtii* (based on Fig. 2). Bold character values are significant at $P < 0.05$ (*, test values exceeding 1.96) or $P < 0.01$ (†, test values exceeding 2.58). ‡, $P = 0.06$.

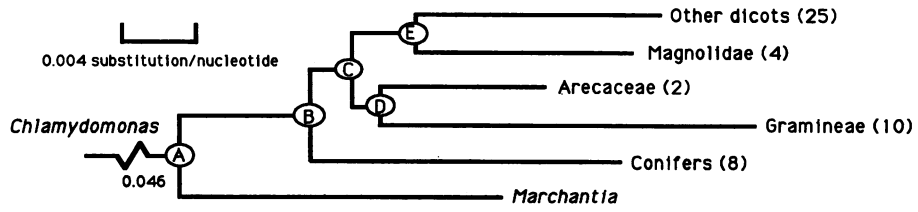

FIG. 2. Phylogenetic tree of six groups of land plants based on Li's method (30) using K_a of *rbcL* sequences. Numbers in parentheses indicate the numbers of taxa.

shows that the Poaceae are the fastest taxa, and the Arecaceae and *Marchantia* are the slowest ones. Using *Chlamydomonas* as the reference taxon, relative rate tests involving *Marchantia* showed its K_a to be significantly slower than several annuals tested; no significant differences, however, were detected relative to the perennials (Table 2).

We estimated substitution rates separately for several lineages based on calibrations using 400 million years (Myr) for the branching date of *Marchantia* (2) (node A in Fig. 2) and 200 Myr for the monocot–dicot split (2, 4) (node C in Fig. 2). We could therefore estimate rates per site per year separately for the branch leading to *Marchantia*, for the internal path between nodes A and C, and for each path emerging from node C. From the length of the branch leading to *Marchantia* (0.0178), the K_a is estimated at 4.4×10^{-11} substitutions per site per year. From the length of the internal path between nodes A and C (adding internal branch lengths AB + BC = 0.00695 + 0.00185) and using 200 Myr as the difference in time between nodes A and C, the K_a on this path alone is also estimated at 4.4×10^{-11} substitutions per site per year. From the length of each path emerging from node C (0.0009 + 0.0079 for the Arecaceae, 0.0009 + 0.217 for the Poaceae, 0.0035 + 0.0111 for the Magnoliaceae, and 0.0035 + 0.135 for the other dicots), the K_a per site per year are 4.4×10^{-11} for the Arecaceae, 11.3×10^{-11} for the Poaceae, 7.3×10^{-11} for the Magnoliaceae, and 8.5×10^{-11} for the other dicots. Therefore, when using *Chlamydomonas* as the outgroup, it appears that the nonsynonymous rate of evolution has been maintained relatively constant on the path linking *Marchantia* to the Arecaceae and that acceleration has taken place in the dicots (2-fold increase on average) and the Poaceae (close to 3-fold increase on average). However, for the other dicots, the average rapid rate of substitution masks a great deal of variability between perennial and annual taxa (Fig. 1).

DISCUSSION

rbcL evolved more rapidly in several modern lineages and, particularly, in annual plants. The so-called "most morphogically advanced forms" represented by the Asteridae and the Poaceae were the groups for which the fastest rates of substitution were observed. The slowest rates were observed for *Marchantia* and the perennial dicots and monocots that diverged much earlier than the Asteridae and Poaceae (25). This trend was particularly clear at nonsynonymous sites, and it was shown that these differences were caused by a rate acceleration. Rate heterogeneity among annual plants has also been observed at nonsynonymous sites for other chloroplast genes (2, 7), suggesting that the trend observed in this study could be representative of the chloroplast genome generally. No rate heterogeneity was reported among the same annual plants in the nuclear 18S and 26S rRNA genes (2) nor for the nuclear gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase (*rbcS*) among several annual dicots and monocots, one fern, and one pine (4). Thus, it appears that less rate heterogeneity exists for nuclear compared to chloroplast genes.

Several factors could be invoked to explain the observed heterogeneity of substitution rates in *rbcL*. Generation time (31), taken as the time to reach sexual maturity, is likely to be rejected because, at least in our limited sample, herbaceous and woody perennials had similarly slow rates. Moreover, in perennial plants where "germ-line" mutations can occur throughout their lifetime, the importance of generation time as an evolutionary factor may be much less than for animals. Furthermore, if generation time was the sole factor, the pattern of heterogeneity observed between perennial and annual forms would be expected to appear at least equally obvious at synonymous and nonsynonymous sites. A quite different pattern was observed, with larger rate differences at nonsynonymous sites between perennial and annual forms.

More recently, it has been suggested that efficiency of DNA replication or repair mechanisms (32), and selection against heterozygosity associated with mismatch-inhibited repair, could be a major force in shaping genetic variation. It would cause the rate of selectively neutral molecular evolution to decrease asymptotically with increasing population size (33). Although chloroplasts are usually uniparentally inherited, biparental inheritance (16, 34–36) and putative chloroplast recombination (37) have been observed. Although rare, these factors could be a significant evolutionary force when considering long time spans. Thus, assuming perennials maintain larger effective population sizes than annuals (38, 39), mismatch-inhibited repair may contribute to their slower rate of *rbcL* evolution.

Taxa with small population sizes may speciate more readily, and if genetic bottlenecks or genomic reorganization accompanies speciation, it may accelerate fixation of new variants. Speciation rate has been shown to vary extensively among seed plants, with long-lived perennial herbs and woody angiosperm taxa, as well as the gymnosperms, showing a slower rate of speciation (40). Although the genetic consequences of speciation are poorly understood in plants (41), annuals appear to more frequently develop strong postzygotic isolating mechanisms than perennials (42) and are more likely to undergo ecological shifts and founder events during speciation (43). In contrast, speciation in perennials appears to more often occur via long-term isolation of large, allopatric populations, with slow-developing isolating mechanisms (44) allowing hybridization to commonly occur in zones of secondary contact (16, 45, 46). Genetic bottlenecks would accelerate fixation of neutral mutations, which presumably comprise the majority of synonymous substitutions. Adaptation to new ecological niches may require fixation of functionally important mutations, which presumably comprise nonsynonymous substitutions. A lower frequency of speciation events in groups of perennials compared to annuals would result in fewer opportunities for such bottleneck and adaptation events. The net result may be an acceleration of nonsynonymous over synonymous rates in annuals, as the former is driven by natural selection and genetic drift, but the latter is driven only by drift and periodic selection (47). The effects of reduced population size would be augmented in organelle genes compared to nuclear genes because of their haploid nature.

Besides the rate heterogeneity that we observed among distantly related species, instances of heterogeneity of substitution rate within families have also been reported for *rbcL*. It has been recently shown that *Z. mays*, a predominantly outbreeding species, evolves faster than other Poaceae to which it was compared, although significant

differences were only found with *Puccinellia* and *Sorghum* (ref. 12; see Fig. 1). In the slow-evolving Betulaceae (see *Betula*, Fig. 1), an old family of perennial angiosperms, the two major clades delineated by the phylogenetic analysis of *rbcL* sequences were shown to evolve at distinct speeds, and this was correlated at the morphological level (11).

The extensive rate heterogeneity we observed indicates that extreme care must be taken when estimating phylogenies and branching dates using *rbcL* sequence data and perhaps other chloroplast genes. If topological errors are to be avoided, methods that assume homogeneity of substitution rates such as the unweighted pair-group method should be avoided or used with caution. Topologies obtained from unweighted standard parsimony must also be corroborated by other methods, as it is also sensitive to rate heterogeneity (48, 49).

We thank G. Zurawski (DNAX Institute of Genetics) for several of the primers used in this study, W.-H. Li (University of Texas), K. H. Wolfe (Trinity College, Dublin), and M. Nei (Pennsylvania State University) for making their computer programs available to us, the Institute of Forest Genetics, U.S. Forest Service, and the University of California Botanical Garden (both of Berkeley, CA), for use of their conifer collections, and L. Savard, M. Michaud, and P. Li (Laval University) for help in retrieving sequences and suggestions on earlier drafts of this manuscript. This research was supported by Grants NC-0642 and ER-0693 from Fonds pour la Formation des Chercheurs et l'Aide à la Recherche du Québec, Canadian National Sciences and Engineering Research Council Grant OGP0046273 and Visiting Research Fellowship to J.B., and National Science Foundation Grants BSR 895723 and 18819953 to S.H.S. and R.A.P.

1. Zuckerkandl, E. & Pauling, L. (1965) in *Evolving Genes and Proteins*, eds. Bryson, V. & Vogel, H. G. (Academic, New York), pp. 97–166.
2. Wolfe, K. H., Gouy, M., Yang, Y.-W., Sharp, P. M. & Li, W.-H. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6201–6205.
3. Martin, W., Gierl, A. & Saedler, H. (1989) *Nature (London)* **339**, 46–48.
4. Murakami, N., Tanaka, Y., Takishima, K., Minobe, Y., Matsuoka, M., Kiyota, S.-I., Hatanaka, S.-I. & Sakano, K. (1990) *Bot. Mag.* **103**, 419–434.
5. Kimura, M. (1968) *Nature (London)* **217**, 624–626.
6. Ritland, K. & Clegg, M. T. (1987) *Am. Nat.* **130**, S74–S100.
7. Wolfe, K. H., Li, W.-H. & Sharp, P. M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 9054–9058.
8. Palmer, J. D. (1985) in *Monographs in Evolutionary Biology: Molecular Evolutionary Genetics*, ed. MacIntyre, R. J. (Plenum, New York), pp. 131–240.
9. Rodermel, S. R. & Bogorad, L. (1987) *Genetics* **116**, 127–139.
10. Runnegar, B. (1991) *Philos. Trans. R. Soc. London B* **333**, 391–397.
11. Bousquet, J., Strauss, S. H. & Li, P. (1992) *Mol. Biol. Evol.* **9**, in press.
12. Doebley, J., Durbin, M., Golenberg, E. M., Clegg, M. T. & Ma, D.-P. (1990) *Evolution* **44**, 1097–1108.
13. Wilson, M. A., Gaut, B. & Clegg, M. T. (1990) *Mol. Biol. Evol.* **7**, 303–314.
14. Soltis, D. E., Soltis, P. S., Clegg, M. T. & Durbin, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4640–4644.
15. Bousquet, J., Simon, L. & Lalonde, M. (1990) *Can. J. For. Res.* **20**, 254–257.
16. Wagner, D. B., Furnier, G. R., Saghai-Maroof, M. A., Williams, S. M., Dancik, B. P. & Allard, R. W. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2097–2100.
17. Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, Y. B., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, K., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H. & Sugiura, M. (1986) *EMBO J.* **5**, 2043–2050.
18. Higuchi, R. G. & Ochman, H. (1989) *Nucleic Acids Res.* **17**, 5865.
19. Li, W.-H., Wu, C.-I. & Kuo, C.-C. (1985) *Mol. Biol. Evol.* **2**, 150–174.
20. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
21. Kimura, M. (1980) *J. Mol. Evol.* **16**, 111–120.
22. Swofford, D. L. (1991) PAUP, *Phylogenetic Analysis Using Parsimony* (Illinois Natural History Survey, Champaign), Version 3.0.
23. Li, W.-H. & Tanimura, M. (1987) *Nature (London)* **326**, 93–96.
24. Sokal, R. R. & Rohlf, F. J. (1981) *Biometry* (Freeman, New York), 2nd Ed.
25. Cronquist, A. (1988) *The Evolution and Classification of Flowering Plants* (Columbia Univ. Press, Irvington-on-Hudson, NY), 2nd Ed.
26. Zimmer, E. A., Hamby, R. K., Arnold, M. L., Leblanc, D. A. & Theriot, E. L. (1989) in *Nobel Symposium 70: The Hierarchy of Life*, eds. Fernholm, B., Bremer, K. & Jornvall, H. (Elsevier, Amsterdam), pp. 205–214.
27. Knaak, C., Hamby, R. K., Arnold, M. L., Leblanc, M. D., Chapman, R. L. & Zimmer, E. A. (1990) in *Biological Approaches and Evolutionary Trends in Plants*, ed. Kawano, S. (Academic, New York), pp. 135–158.
28. Martin, P. G. & Dowd, J. M. (1991) *J. Mol. Evol.* **33**, 274–282.
29. Les, D. H., Garvin, D. K. & Wimpee, C. F. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10119–10123.
30. Li, W.-H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1085–1089.
31. Kohne, D. E. (1970) *Quart. Rev. Biophys.* **33**, 327–375.
32. Britten, R. J. (1986) *Science* **231**, 1393–1398.
33. Stephan, W. & Langley, C. H. (1992) *Genetics*, in press.
34. Govindaraju, D. R., Dancik, B. P. & Wagner, D. B. (1988) *J. Evol. Biol.* **2**, 49–59.
35. White, E. E. (1990) *Theor. Appl. Genet.* **79**, 251–255.
36. Hong, Y.-P. (1991) Ph.D. dissertation (Oregon State University, Corvallis).
37. Govindaraju, D. R., Wagner, D. B., Smith, G. P. & Dancik, B. P. (1988) *Can. J. For. Res.* **18**, 1347–1350.
38. Levin, D. A. (1988) *Am. Nat.* **132**, 309–317.
39. Govindaraju, D. R. (1989) *Biol. J. Linn. Soc.* **37**, 345–357.
40. Levin, D. A. & Wilson, A. C. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 2086–2090.
41. Barrett, S. C. H. (1989) in *Speciation and Its Consequences*, eds. Otte, D. & Endler, J. A. (Sinauer, Sunderland, MA), pp. 257–283.
42. Grant, V. (1981) *Plant Speciation* (Columbia Univ. Press, Irvington-on-Hudson, NY), 2nd Ed.
43. Templeton, A. R. (1981) *Annu. Rev. Ecol. Syst.* **12**, 23–48.
44. Barton, N. H. & Charlesworth, B. (1984) *Annu. Rev. Ecol. Syst.* **15**, 133–164.
45. Keim, P., Paige, K. N., Whitham, T. G. & Lark, K. G. (1989) *Genetics* **123**, 557–565.
46. Whittemore, A. T. & Schaal, B. A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 2540–2544.
47. Maruyama, T. & Birky, C. W. (1991) *Genetics* **127**, 449–451.
48. Sourdis, J. & Nei, M. (1988) *Mol. Biol. Evol.* **5**, 298–311.
49. Saitou, N. & Imanishi, T. (1989) *Mol. Biol. Evol.* **6**, 514–525.