Complete Congruence between Morphological and rbcL-based Molecular Phylogenies in Birches and Related Species (Betulaceae)

Jean Bousquet,* † Steven H. Strauss,† and Peng Li*

*Centre de Research en Biologie Forestière, Faculté de Foresterie et de Géomatique, Université Laval; and †Department of Forest Science, Oregon State University

Estimations of phylogenies from morphological and molecular data often show contrasting results. We compared morphological and molecular phylogenies in an ancient family of woody dicots, the Betulaceae (birch family). The phylogeny of the family was estimated from parsimony analysis of morphological characters in the genera Alnus, Betula, Carpinus, Corylus, Ostrya, and Ostryopsis and from parsimony and distance-matrix analyses of DNA sequences of the chloroplast gene encoding the large subunit of ribulose-1,5-biphosphate carboxylase (rbcL) in the genera Alnus, Betula, Carpinus, Corylus, and Ostrya and in two outgroups, Quercus and Liquidambar. The topologies obtained by the different methods were completely congruent, and bootstrapping strongly supported the division of the family Betulaceae into two major clades, Betuleae (Alnus and Betula) and Coryleae (other members). Only slightly more homoplasy was present in the rbcL sequence data set than in the morphological set. Relative-rate tests indicated that the Coryleae clade had a faster rate of rbcL evolution than did the Betuleae clade. Heterogeneity of rates of morphological evolution also paralleled those for rbcL.

Introduction

The Betulaceae (the birch family) is composed of six genera and 120–150 species of anemophilous shrubs and trees, many of which are ecologically and economically important (Abbe 1974; Furlow 1990). The family is widely distributed in the northern hemisphere (Asia, Europe, and North America), and some species occur as far south as northern South America. All species of the genus Alnus so far investigated are actinorhizal—bearing nitrogen-fixing nodules on their root system, resulting from infection by the actinomycete Frankia (Normand and Bousquet 1989). However, this has not been reported in any of the remaining five genera of the family.

There is no controversy about the division of the Betulaceae into six genera (Alnus, Betula, Carpinus, Corylus, Ostrya, and Ostryopsis). Opinions differ, however, regarding both the circumscription of genera into higher taxonomic ranks and the level of these ranks. Most early taxonomists (see Winkler 1904), as well as a recent cladistic analysis of reproductive characters (Crane 1989), supported the division of the six genera into two tribes: Betuleae (Alnus and Betula), and Coryleae (Carpinus, Corylus, Ostrya, and Ostryopsis). On the other hand, the basic two tribes defined by

1. Key words: Actinorhizal plants, Betulaceae, large subunit of ribulose-1,5-biphosphate carboxylase, Liquidambar, morphology, phylogeny, Quercus, rate of evolution. Abbreviations: $K_s$ = nonsynonymous rate of substitution; $K_a$ = synonymous rate of substitution; $K_\text{overall}$ = overall rate of substitution; rbcL = gene coding for the large subunit of ribulose-1,5-bisphosphate carboxylase.

Address for correspondence and reprints: Jean Bousquet, C.R.B.F, Faculté de foresterie, Université Laval, Ste-Foy, Québec, Canada G1K 7P4.

© 1992 by The University of Chicago. All rights reserved.
0737-4038/92/0906-0006$02.00
Winkler were elevated to the rank of subfamilies (Furlow 1990) or families (Hutchinson 1964, p. 27; Dahlgren 1980). On the basis of floral morphology, *Carpinus* and *Ostrya* were separated from the Corylaceae, to form a third tribe, Carpineae (Abbe 1974; Furlow 1990). Kuprianova (1963) emphasized the importance of some characters—especially the compound pollen apertures—shared by *Carpinus*, *Ostrya*, and *Ostryopsis*, and he placed these three genera in a third family Carpinaceae, leaving only *Corylus* in the Corylaceae.

DNA sequences from the chloroplast gene encoding the large subunit of the enzyme ribulose-1,5-biphosphate carboxylase (*rbcL*) have been fruitfully used for estimating phylogeny in angiosperms, at the class and subclass levels (Ritland and Clegg 1987). More recently, sequences of this gene have been used to elucidate evolutionary relationships of plant taxa at the family level, and the results generally have been congruent with traditional views regarding the Gramineae (Doebley et al. 1990), while quite divergent from traditional views regarding the Saxifragaceae sensu lato (Soltis et al. 1990). Because the Betulaceae has been well characterized at the morphological level (Crane 1989), it presents an ideal opportunity to contrast phylogenies estimated from morphological and molecular data.

**Material and Methods**

**Morphological Data**

Thirty-five binary characters derived from study of palynogy, morphology, wood anatomy, and chemical constituents were assembled (table 1), from previous studies, for the genera *Alnus*, *Betula*, *Carpinus*, *Corylus*, *Ostrya*, and *Ostryopsis* (Woodworth 1930; Abbe 1935; Tippe 1938; Metcalfe and Chalk 1950, vol. 1 pp. 587–593, vol. 2 pp. 1302–1312; Hall 1952; Hardin and Bell 1986; Meurer et al. 1988; Crane 1989; Morawetz and Samuel 1989; Wolfe 1989). If a character was not available or not applicable in one taxon, it was designated as missing. If the two states of a character occurred in the same taxon, the character was declared polymorphic.

Wagner parsimony (Farris 1970) was used to find the most parsimonious unrooted trees. They were rooted using two outgroups. The first one, the Fagaceae, is thought to represent a closely related sister group of the family Betulaceae, and both are placed in the order Fagales within the subclass Hamamelidae (Dahlgren 1980; Takhtajan 1980; Cronquist 1988, pp. 292–298; Thorne 1989). Although the origin of the Fagales is not resolved, they are considered to be related to the Hamamelidales (including the family Hamamelidaceae) (Takhtajan 1980; Cronquist 1988, pp. 292–298). Therefore, the family Hamamelidaceae was chosen as the more ancient outgroup. The suitability of using both the Fagaceae and Hamamelidaceae as outgroups has been shown in previous cladistic analysis of the Hamamelidae (Zavada and Dilcher 1986; Barabe et al. 1987; Crane 1989; Nixon 1989). The parsimony analyses were conducted using the branch-and-bound algorithm of PAUP V3.0 (Swofford 1991). Bootstrap analyses (Felsenstein 1985) were used to test the statistical reliability of the estimated phylogenetic trees, under the assumption of character independence. Each of these analyses was based on 1,000 replications. The consistency index excluding uninformative sites was used to measure the level of homoplasy (Kluge and Farris 1969).

**Molecular Data**

We determined the complete *rbcL* nucleotide sequence (1,428 bp) for one species in each of five genera of the family Betulaceae: *Alnus incana* (L.) Moench., *Betula papyrifera* Marsh., *Carpinus caroliniana* Walt., *Corylus cornuta* Marsh., and *Ostrya virginiana* (Mill.) Koch. To root the phylogenetic trees, the complete *rbcL* sequence
Table 1
Morphological Data Matrix and Definition of Characters

| Character Status* | TAXON  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 |
|------------------|--------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| *0 = Absence of character; 1 = presence of character; 2 = polymorphic state; and – = information not available. Definitions of characters and character states (i.e., numbers shown in col. heads) are as follows: wood anatomy (Tippo 1938; Metcalfe and Chalk 1950, vol. 1 pp. 587-593, vol. 2 pp. 1302-1315; Hall 1952)—1 = tracheid present (0) or absent (1); 2 = vessels not in dentritic pattern (0) or in dentritic pattern (1); 3 = spiral thickening in vessels, absent (0) or present (1); 4 = vessel perforation plates exclusively scalariform (0) or at least partially simple (1); 5 = bars on scalariform plates, many (0) or intermediate or few (1); 6 = scalariform plates bordered (0) or nonbordered (1); / = intervessel pitting at least partially opposite (0) or exclusively alternate (1); 8 = vascular rays heterogenous (0) or homogenous (1); Y = aggregate rays absent (0) or present (1); and 10 = xylem parenchyma at least partially diffuse (0) or exclusively metatracheal (1), pollen (Crane 1989)—1 = pollen tricolpate or tricolporate (0) or triporate (1); 12 = pollen sculpture various (0) or weakly rugulate with fine spinules (1); 13 = endosperm of the pollen absent (0) or present (1), and 14 = pollen shape various (0) or oblong (1); 15 = ovules tenuinucellate unitegmic (0) or crassinucellate unitegmic (1); 16 = baccate ovary diagonal (0) or transverse (1); 17 = perianth on bisexual or pistillate flowers, present (0) or absent (1); 18 = bract and bracteoles associated with dichasia with bisexual or pistillate flowers, various (0) or fused into a scale (1); 19 = bracteoles associated with pistillate flowers, various (0) or fused into an involute (1); 20 = central flower of pistillate dichasia, present (0) or absent (1), and 21 = perianth on bisexual or staminate flowers, present (0) or absent (1); 22 = receptacle-torus of staminate flowers, glabrous (0) or pilose (1); 23 = stamens not inserted into a torus (0) or inserted into a torus (1); 24 = staminate flowers with collateral bracteoles (0) or fused with primary bract (1); 25 = anthers not pilose at apex (0) or pilose (1); 26 = mature fruits various (0) or small and strongly dorsiventrally compressed (1); and 27 = surface of nutlet smooth (0) or with prominent longitudinal ribs (1), leaf structures (Harland and Bell 1986; Crane 1989; Wolfe 1989)—28 = leaf vernation conuplicate (0) or plicate (1); 29 = leaf areoles with branches (0) or simple free ending veins (1); 30 = secondary veins extending directly into leaf teeth (0) or curvingly toward leaf teeth (1); 31 = peltate scale absent (0) or present (1); and 32 = stalk of stipitate gland uniseriate (0) or multisieriate (1), chemistry (Meurer et al. 1988)—33 = N'N'-caffeoyl-L-lysine absent (0) or present (1); and 34 = N',N'-difuroyl-L-lysine absent (0) or present (1); and chromosome number (Woodworth 1931; Morawetz and Samuel 1989)—35 = base chromosome number 8 (0) or 12-14 (1).
was also determined in the family Fagaceae, for *Quercus rubra* L., and in the family Hamamelidaceae, for *Liquidambar styraciflua* L. In addition, the first 100 nucleotides of the 3' flanking region of *rbcL* ORF were determined for these seven species. Leaf material was collected, from natural stands in Québec, for *C. cornuta*, *O. virginiana*, and *Q. rubra*. For *C. carolina* and *L. styraciflua*, leaf material was collected from one individual located on the Campus of Oregon State University. For *A. incana* and *B. papyrifera*, leaf tissues were collected from clones AI-4 and BP-34, respectively, described elsewhere (Bousquet et al. 1989) and maintained at Laval University.

Total genomic DNA was extracted from lyophylized leaf tissues according to the procedures of Bousquet et al. (1990). Amplification and sequencing primers were supplied by Dr. G. Zurawski, and additional primers were designed by identifying homologous regions among published sequences (monocots, dicots, and liverwort). Polymerase chain reaction was conducted by using primers homologous to the beginning and the end of the *rbcL* coding region, in conjunction with primers annealing to conserved regions of the flanking ORF512 and *atpB* genes (Shinozaki et al. 1986). The typical volume of amplification reaction was 100 μl, containing 100 ng of genomic DNA, 50 pmol of each primer, 0.2 mM of each dNTP, 50 mM KCl, 10 mM Tris HCl pH 9.0, 1.5 mM MgCl₂, 0.01% gelatin, 0.1% Triton X-100, and 2.5 units *Taq* polymerase (Promega). Amplification was conducted in a DNA thermal cycler (Perkin Elmer Cetus) for 35 cycles, each consisting of a denaturing step of 1 min at 94°C, an annealing step of 1 min at 50°C, and an extension step of 2.5 min at 72°C. The last 20 extension steps were progressively extended by 5 s/cycle. The last cycle was followed by a period of 10 min at 72°C to ensure that primer extension reactions proceeded to completion. Amplified fragments were purified by electrophoresis through 1.5% agarose, electroelution using a microelectroelutor (Amicon), and ultrafiltration with Centricon-30 (Amicon). DNA was ethanol precipitated before being resuspended in 10 μl T.E (10 mM Tris-HCl pH 8.0, and 0.1 mM ethylenediaminetetraacetate).

From a typical 100-μl amplification reaction, enough DNA was obtained for four different labeling reactions. These were directly conducted according to the dideoxy protocol of Winship (1989), with 5 μCi ³⁵S-dATP (USB), 3 units of Sequenase 2.0 (USB), and 10 pmol of primer per reaction. This method uses denaturing conditions (dimethylsulfoxide) to reduce renaturation of double-stranded DNA template. For sequencing the 5'-3' strand, 26-34mer internal primers annealing to positions 1, 103, 234, 427, 674, 895, 1020, 1204, and 1264 were used. Primers annealing to positions 153, 358, 674, 895, 1020, 1204, and 1375 were used for sequencing the opposite strand. Sequencing was conducted using 6% polyacrylamide-7M uracil gels according to the manufacturer's recommendations (Bio-Rad). X-ray films were exposed for 24 h to several days.

Pairwise synonymous and nonsynonymous numbers of substitutions (Ks and Ka, respectively) corrected for multiple hits were calculated, along with their standard errors, according to the methods of Li et al. (1985). Overall number of substitutions (K0) was calculated as a weighted average of Ks and Ka. The different distance matrices obtained were analyzed with the neighbor-joining method of phylogenetic tree construction (Saitou and Nei 1987). This method does not assume homogeneity of substitution rates and has been shown to be robust under a variety of conditions (Sourdis and Nei 1988; Saitou and Imanishi 1989). In addition, parsimony analyses (Fitch 1977) were conducted using the branch-and-bound algorithm of PAUP 3.0 (Swofford 1991). We examined standard parsimony as well as transition/transversion parsimony. With the latter strategy, character-state differences were weighted such that transitions
and transversions received weights proportional to the reciprocal of their frequencies. The consistency index excluding uninformative sites was also calculated to measure the level of homoplasy in the standard parsimony approach. Bootstrap confidence intervals (Felsenstein 1985) were calculated from 1,000 replications for both types of parsimony, by using PAUP 3.0 and for the neighbor-joining tree, by using the program NJBOOT with the method of Jukes and Cantor (1969) (T. S. Whittam, personal communication).

Relative-rate tests were used to assess heterogeneity of substitution rates within the Betulaceae family by comparing two taxa of interest with reference taxon (Wu and Li 1985). If $K_r$ is to be used in the relative-rate tests, covariances between taxa need to be estimated for both synonymous and nonsynonymous substitutions. However, the small number of nonsynonymous substitutions (only nine) within the Betulaceae family might result in large errors in estimating between-taxon covariances in nonsynonymous substitutions (Wu and Li 1985). Therefore, for the relative-rate tests we used the number of nucleotide substitutions estimated from Kimura's (1980) two-parameter method.

**Results**

**Morphology**

For the six genera of Betulaceae, with the Fagaceae and the Hamamelidaceae as the outgroup, our analyses resulted in one most parsimonious tree with a consistency index of 0.853. Thirty-seven character-state changes were detected within the Betulaceae, of which 31 were informative (fig. 1A). When *Ostryopsis* was excluded from the analysis (to enable a comparison with results from *rbcL*), one most parsimonious tree was obtained with a consistency index of 0.93, and the number of changes among the Betulaceae was 34, of which 27 were informative (fig. 1B). The use of Fagaceae as the only outgroup did not change these topologies. The family was divided into two clades, corresponding to the tribes Betuleae and Coryleae (fig. 1B). The number of synapomorphies was much larger for the Coryleae clade (15) than for the Betuleae clade (4–6). The Betuleae have the following synapomorphies: reduced pistillate flowers, small compressed fruits, and peltate scales on the leaf surface. All Coryleae share the following synapomorphies: reduced staminate flowers, advanced wood anatomy features, and the presence of spermidines in pollens. Within Coryleae, *Corylus* is the sister taxon to the *Carpinus-Ostrya-Ostryopsis* subclade. This subclade has compound pollen, a base chromosome number of 8, and a plicate leaf vernation pattern. Within this subclade, *Ostryopsis* is the sister group of *Carpinus* and *Ostrya*, with the later two sharing seven synapomorphies. Bootstrap estimates were relatively large, with higher values for the Coryleae clade (fig. 1A and B).

**rbcL Sequences**

Of the 1,425 sites (excluding the stop codon) compared among the five Betulaceae *rbcL* sequences, 33 sites were found to be variable. Of these, 24 were synonymous substitutions and 9 were nonsynonymous ones. The number of informative sites was 15, with 12 at synonymous positions and 3 at nonsynonymous ones. Within the family, the average synonymous, nonsynonymous and overall numbers of substitution per nucleotide were, respectively, 0.038 ± 0.015, 0.004 ± 0.001, and 0.012 ± 0.004, while the average number of substitutions estimated from Kimura's two-parameter method was 0.012 ± 0.004. The average transitions/transversions ratio was 2.2 ± 1.1.

With *Liquidambar* (Hamamelidaceae) or *Quercus* (Fagaceae) as the outgroup,
FIG. 1.—A and B, Trees obtained from Wagner parsimony analysis of 35 morphological characters, with Fagaceae and Hamamelidaceae as outgroups. C, Tree obtained from standard parsimony analysis of rbcL sequences, with Quercus (Fagaceae) and Liquidambar (Hamamelidaceae) as outgroups. D, Tree obtained from neighbor-joining analysis of overall numbers of substitutions/site of rbcL sequences, with Quercus (Fagaceae) and Liquidambar (Hamamelidaceae) as outgroups [the bootstrap analysis was based on the number of substitutions reported by Jukes and Cantor (1969)]. Numbers on branch lengths (in A–C) indicate numbers of steps. Circled numbers on nodes (in A–D) indicate bootstrap estimates from 1,000 replications. c.i. = consistency index.
the same topology was obtained from both parsimony (fig. 1C) and neighbor-joining analyses using the overall number of substitutions (fig. 1D) as well as the numbers of substitutions estimated according to Jukes and Cantor (1969) and Kimura (1980) (not shown), and it was in complete agreement with the topology derived from morphological data. (The overall number of substitutions and the numbers of substitutions estimated by the Jukes-Cantor and Kimura methods were virtually identical with each other.) Two major clades were obtained—Betuleae, containing *Alnus* and *Betula*, and Coryleae, containing *Carpinus*, *Corylus* and *Ostrya*. Standard parsimony (fig. 1C) led to a single most parsimonious tree with 36 changes within the Betulaceae, of which 15 were informative. The consistency index was 0.83 when *Quercus* was included and was 0.875 when it was excluded. One single most parsimonious tree with the same topology was obtained with transition/transversion parsimony by using a weighting scheme of 2:1 for transversions/translations. Within the Coryleae clade, *Corylus* was the sister taxon to *Carpinus* and *Ostrya*. The bootstraps obtained were generally high (>80%) and essentially the same for the three methods used, with estimates of 93%, 83%, 100%, and 90% for standard parsimony (fig. 1C), 92%, 83%, 100%, and 90% for the neighbor-joining tree using Jukes and Cantor's (1969) method (fig. 1D), and 97%, 87%, 99%, and 93% for weighted parsimony (not shown).

Standard parsimony as well as the neighbor-joining method (fig. 1C and D) indicated a faster speed of evolution in the Coryleae clade. The relative-rate tests for all 10 pairwise comparisons within the Betulaceae family when *Quercus* was used as a reference taxa showed that *Carpinus* and *Ostrya* evolved significantly faster than *Betula* (at the 5% probability level) or *Alnus* (at the 10% probability level) (table 2). However, a relative-rate test for examining the rate constancy between two lineages indicated that the Coryleae clade overall evolved significantly faster than the Betulaceae clade (at the 5% probability level) (Li and Bousquet 1992).

When the first 100 nucleotides downstream from *rbcL* were compared, four positions were found to be variable in the Betulaceae, of which three were informative (fig. 2, positions 32, 39, and 99). All the observed differences are congruent with the family topology estimated from *rbcL* sequences and morphology data. In the com-

<table>
<thead>
<tr>
<th>TAXA COMPARED</th>
<th>REFERENCE TAXON: 3</th>
<th>NO. OF SUBSTITUTIONS/NUCLEOTIDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betula</td>
<td>Quercus</td>
<td>$K_{(12)}$ 0.0064 0.0280 0.0294 0.0014 -23 0.509</td>
</tr>
<tr>
<td>Betula</td>
<td>Quercus</td>
<td>$K_{(13)}$ 0.0128 0.0280 0.0339 0.0059 -46 0.057</td>
</tr>
<tr>
<td>Betula</td>
<td>Quercus</td>
<td>$K_{(23)}$ 0.0149 0.0280 0.0353 -0.0073 -49 0.029</td>
</tr>
<tr>
<td>Betula</td>
<td>Quercus</td>
<td>$K_{(3-23)}$ 0.0156 0.0280 0.0361 -0.0081 -52 0.019</td>
</tr>
<tr>
<td>Alnus</td>
<td>Quercus</td>
<td>$K_{(12)}$ 0.0128 0.0294 0.0339 -0.0045 -35 0.150</td>
</tr>
<tr>
<td>Alnus</td>
<td>Quercus</td>
<td>$K_{(13)}$ 0.0149 0.0294 0.0353 -0.0059 40 0.078</td>
</tr>
<tr>
<td>Alnus</td>
<td>Quercus</td>
<td>$K_{(23)}$ 0.0156 0.0294 0.0361 -0.0067 43 0.052</td>
</tr>
<tr>
<td>Alnus</td>
<td>Quercus</td>
<td>$K_{(3-23)}$ 0.0092 0.0339 0.0353 -0.0014 -16 0.542</td>
</tr>
<tr>
<td>Alnus</td>
<td>Quercus</td>
<td>$K_{(12)}$ 0.0071 0.0339 0.0361 -0.0022 -31 0.342</td>
</tr>
<tr>
<td>Alnus</td>
<td>Quercus</td>
<td>$K_{(13)}$ 0.0063 0.0353 0.0361 -0.0008 12 0.726</td>
</tr>
</tbody>
</table>

*($K_{(3-23)}/K_{(12)}) \times 100$. 

Table 2
Relative-Rate Tests Used to Assess Heterogeneity of Numbers of Substitution per Nucleotide, Estimated from Kimura's (1980) Two-Parameter Method
FIG. 2.—Alignment of the first 100 nucleotides of the 3' flanking region of the gene rbcL including the rbcL stop codon TAA. Only positions different from the consensus Alnus sequence are shown. A dash (-) denotes a deletion/insertion.

parison with outgroups, these differences appeared as synapomorphies in the Coryleae clade, again suggesting a faster speed of evolution in this group.

Discussion

The phylogenetic analysis of morphological characters and rbcL sequences showed the existence of two distinct clades within the family Betulaceae, Betuleae and Coryleae. Even though Ostryopsis was not sequenced, the rbcL trees clearly showed that Betuleae is a natural group while all the other genera undoubtedly belong to the Coryleae. Our results conform well both to those derived from preliminary cladistic analyses of the family that are based on reproductive characters (Crane 1989) and to traditional taxonomic views of relationships among genera (Winkler 1904; Anderson and Abbe 1938; Hall 1952; Kikuzawa 1982). Preliminary analysis of isozymes and restriction-fragment-length polymorphisms of ribosomal DNA in the family also indicated considerable similarity between the genera Alnus and Betula of the clade Betuleae (Bouquet and Lalonde 1990). Thus, these results indicate that the family is composed of two major clades, which could be elevated to the tribal level or subfamily level. Our results, however, do not support the split of Betuleae into two tribes as suggested by serological studies (Brunner and Fairbrothers 1979), nor do they fully support the division of Coryleae into either two tribes, Coryleae (Corylus, Ostryopsis) and Carpinaceae (Carpinus, Ostrya) (Abbe 1974; Furlow 1990), or two families, Corylaceae (Corylus) and Carpinaceae (others) (Kuprianova 1963).

The fossil records of the Betulaceae are extensive and conform well to the phylogenies obtained. Pollen records showed that Alnus and Betula were differentiated as early as in the Santonian (80 Mya). Pollens of the Coryleae appeared later: Corylus in the Maestrichtian (67 Mya) and Carpinus and Ostrya in the Paleocene (60 Mya B.P.) and Eocene (41 Mya) (Muller 1984). On the basis of reproductive structures, Crane (1989) established the unequivocal generic differentiation of the Betuleae at least by the middle of the Eocene (45 Mya). In the Coryleae, Corylus could be differentiated by the Paleocene (60 Mya) while Carpinus, Ostrya, and Ostryopsis were differentiated later, in the later Eocene and Oligocene (30–40 Mya) (Crane 1989).

Phylogenetic trees obtained from morphological characters and rbcL sequences all indicated a faster rate of evolution in the Coryleae clade compared with the Betuleae
clade. This trend was shown to be statistically significant by relative-rate tests. Therefore, absolute rates of evolution for the family could be obtained from sequence data if the principle of a variable molecular clock is adopted. Because the Betuleae clade shows a distinct slowdown of evolutionary rate, compared with the Coryleae clade, separate rates of evolution should be estimated for each. If 80 Myr and 45 Myr are used as boundaries for the divergence time (T) between *Alnus* and *Betula*, the overall rate of substitution per site per year (*Ko*/2T) for the Betuleae clade is 3.7 × 10⁻¹¹–6.7 × 10⁻¹¹. Similarly, for the divergence of *Corylus* from *Carpinus* and *Ostrya*, taking 67 Myr and 60 Myr as boundaries, we calculated rates of 6.0 × 10⁻¹¹–6.7 × 10⁻¹¹. When 60 Myr and 40 Myr are used for the divergence of *Carpinus* from *Ostrya*, rates of 5.0 × 10⁻¹¹–7.5 × 10⁻¹¹ were obtained. Overall, these estimates are well within the range of each other, indicating that most of the heterogeneity in rate of evolution occurs along the branches leading to each clade, rather than within clades. This is also evident at the morphological level, for which rates of evolution are surprisingly correlated with those for rbcL. There do not seem to be any obvious population or life-history differences between contemporary taxa of the two clades. However, because much of the rate heterogeneity could be accounted for by internal branches, and because little information is available from the ancestors of contemporary members of each clade, potential causal factors such as mutation repair mechanism (Britten 1986) and generation time (Kohne 1970) are difficult to assess. Speciation rate and related historical influences involving bottleneck effects were recently suggested as a possible mechanism that could account for the important heterogeneity, in rbcL substitution rates, observed among widely divergent seed plant taxa (Bousquet et al. 1992).

In rbcL rate of evolution, within-family differences such as those observed in our study have not been observed in the Saxifragaceae sensu lato (Soltis et al. 1990) but were found in the Gramineae, with an accelerated speed of evolution in the lineage leading to *Zea* (Doebely et al. 1990). Moreover, the estimates of substitution per site per year in the Betulaceae are considerably lower than those obtained from rbcL sequences in the Gramineae (50 × 10⁻¹⁰–67 × 10⁻¹⁰ substitutions per site per year; Zurawski et al. 1984; Doebely et al. 1990) and, to a lesser extent, than those obtained for the Arecaceae (13 × 10⁻¹¹; Wilson et al. 1990). This suggests both an extreme diversity of substitution rate for rbcL, at least among angiosperms, and caution in utilizing this gene as a molecular clock (Bousquet et al. 1992). Because of this heterogeneity, methods of phylogenetic tree construction that assume homogeneity of substitution rates, such as the unweighted pair-group method (UPGMA; Sokal and Michener 1958), should be used with caution (Nei 1987, pp. 309–313).

The association with *Frankia*, which is only observed in *Alnus*, does not appear on the basis of parsimony, to be ancestral to the Betulaceae family. If our estimated topologies are correct, an ancestral state of association would necessitate two character changes. If the association is considered as a derived state, only one change on the branch leading to *Alnus* is necessary to explain the pattern of variation among taxa. It is perhaps surprising that *Betula* does not associate with *Frankia*, considering the close relationship between *Alnus* and *Betula*. This lack of association could also be considered as an acquired resistance to *Frankia* as a root invader, because it has been shown that *Frankia* can be abundant under *Betula* stands (Smolander 1990). This suggests the possibility of interaction between *Betula* roots and *Frankia* populations in the rhizosphere. Given both that *Frankia* is only found associated with phylogenetically ancient woody dicots dispersed in eight families (Bousquet and Lalonde 1990) and that, with reference to host specificity, it started diversifying ~100 Mya
In the present study, there was complete agreement among analyses of different types of characters: coding versus noncoding sequences, DNA versus morphological characters. The complete congruence in topology between DNA and morphological characters could not arise by chance alone, because the probability that two five-member cladograms completely match is only 0.012 (Simberloff 1987). There was also complete agreement when different methods of phylogenetic tree construction were used: the parsimony and distance-matrix (neighbor-joining) analysis of rbcL sequences led to the same topology and similar patterns of relative rates of evolution. rbcL sequence data also showed slightly more homoplasy than did morphological data, which supports the hypothesis that, in studies including small numbers of taxa, molecular data are more homoplastic than are morphological data (Sanderson and Donoghue 1989). However, the recorded levels of homoplasy correspond well to those obtained from the estimation of Gramineae phylogeny when rbcL and morphology are used (Doebley et al. 1990). Our results indicate that congruent topologies and correlated rates of evolution can be obtained with widely different phylogenetic data sets.

Sequence Availability

The rbcL nucleotide sequences reported in this paper are available from EMBL: Alnus incana (L.) Moench. (X56618), Betula papyrifera Marsh. (X56617), Carpinus caroliniana Walt. (X56621), Corylus cornuta Marsh. (X56619), Ostrya virginiana (Mill.) Koch. (X56620), Quercus rubra L. (M58391), and Liquidambar styraciflua L. (M58394).

Acknowledgments

We would like to thank G. Zurawski (DNAX Institute of Genetics, California) for several of the primers used in this study, M. Nei and T. S. Whittam (Institute of Molecular Evolutionary Genetics, Pennsylvania State University), W.-H. Li (Center for Demographic and Population Genetics, University of Texas), and K. Wolfe (Trinity College, Dublin) for making computer programs available to us; and L. Savard (CRBF, University of Laval) for suggestions on an earlier draft of the manuscript. This research was supported by Canadian NSERC grant OGP0046273 and Québec FCAR grants ER-0693 and NC-0642, all to J.B.; by NSF grants BSR 895723 and 8817953, both to S.H.S.; and by Canadian NSERC fellowship to P.L.

LITERATURE CITED


MASATOSHI NEI, reviewing editor

Received January 2, 1992; revision received April 27, 1992

Accepted April 27, 1992