Molecular evolutionary history of ancient aquatic angiosperms

(rbcL sequencing/systematics/Nymphaeales/phylogeny)

DONALD H. LES, DENISE K. GARVIN, AND CHARLES F. WIMPEE

Department of Biological Sciences, University of Wisconsin, Milwaukee, WI 53201

Communicated by David L. Dilcher, August 19, 1991 (received for review May 29, 1991)

ABSTRACT Aquatic plants are notoriously difficult to study systematically due to convergent evolution and reductionary processes that result in confusing arrays of morphological features. Plant systematists have frequently focused their attention on the "water lilies," putative descendants of the most archaic angiosperms. Classification of these 10 plant genera varies from recognition of one to three orders containing three to six families. We have used DNA sequence analysis as a means of overcoming many problems inherent in morphologically based studies of the group. Phylogenetic analyses of sequence data obtained from a 1.2-kilobase portion of the chloroplast gene rbcL provide compelling evidence for the recognition of three distinct lineages of "water lily" plants. Molecular phylogenies including woody Magnoliidae sequences and sequences of these aquatic plants depict Ceratophyllum as an early diverging genus. Our results support hypotheses that most taxonomic concepts of the order Nymphaeales reflect polyphyletic groups and that the unusual genus Ceratophyllum represents descendants of some of the earliest angiosperms.

Precise clues to Darwin's "abominable mystery," the origin of flowering plants, have eluded systematists for more than a century. Fossils document the rapid diversification of angiosperm subclasses by the lower Cretaceous (1), yet the scarcity of flowering plant fossils prior to this period has impeded attempts to reconstruct patterns of divergence among ancestral angiosperm lineages (2). Instead, systematists have conceptualized "primitive" angiosperms by comparing anatomical, biochemical, embryological, and morphological data derived from among extant taxa. This approach has yielded a number of modern classifications, each striving to depict hierarchical taxonomic groups in a fashion that putatively parallels their phylogeny.

Most systematists assign extant flowering plants to either monocots or dicots, but debate continues over which group is more ancient and over which lineage within each group is most primitive. Of particular relevance to questions of early angiosperm evolution is the phylogenetic position of the aquatic order Nymphaeales (water lilies). Most authors treat the water lilies as dicots although some place them with monocots (3). Regardless, many contemporary taxonomists view the Nymphaeales as primitive angiosperms somehow implicated with monocot origins (4-6). At the turn of the century, Shaffner (7) advocated the recognition of more than two major angiosperm groups, arguing that the water lilies represent a common stock from which all flowering plants originated. Although Shaffner's perspective (7) was never widely adopted, his observations of monocotyledon features in the Nymphaeales were eventually confirmed (8). A synthesis of available evidence expanded this concept to hypothesize that the Nymphaeales represent descendants of angiosperm lineages existing prior to the divergence of monocots and dicots (9), an arrangement supported by rRNA

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.

sequence analysis (10). The great age of these aquatic plants has been verified by fossils of the genus Ceratophyllum among the oldest known reproductive angiosperm remains (11). These observations emphasize the importance of reconciling phylogenetic relationships of the Nymphaeales, a "pivotal" group in questions of early angiosperm relationships. Such clarification should enhance the understanding of early angiosperm evolution and lead to improvements in existing classifications.

Disarray in classifications of the Nymphaeales impairs the testing of evolutionary hypotheses. The 10 genera included in the broadest ordinal concept (Barclaya, Brasenia, Cabomba, Ceratophyllum, Euryale, Nelumbo, Nuphar, Nymphaea, Ondinea, and Victoria) have been aligned variously among one to three orders including from three to six families (Fig. 1). Although certain tendencies emerge from this synopsis, it is difficult to fully accept any existing classification of the Nymphaeales. Furthermore, emphatic arguments both for (12) and against (9) the monophyly of the order itself make it difficult to pursue any meaningful discussion of relationships to other angiosperms.

Here we report on the use of *rbcL* sequence data* to construct estimates of phylogeny for aquatic plant genera allied previously with the Nymphaeales. We relate our results to past efforts aimed at deducing phylogenetic relationships in this unusual group of plants.

MATERIALS AND METHODS

A 1183-base-pair (bp) portion of rbcL was sequenced for Barclaya longifolia, Brasenia schreberi, Cabomba caroliniana, Ceratophyllum demersum, Euryale ferox, Nelumbo lutea, Nelumbo nucifera, Nuphar variegata, Nymphaea odorata, and Victoria cruziana. Euryale plants were grown from seeds donated by Longwood Gardens (Kennett Square, PA); Barclaya and Victoria plants were acquired from Suwanee Laboratories (Lake City, FL); Cabomba plants were obtained from Carolina Biological Supply. Remaining plants were collected from field localities (or for Nelumbo nucifera, were in cultivation) in Wisconsin.

Total genomic DNA was isolated (19) and a 1.2-kilobase fragment of rbcL was amplified (20) using the PCR and internal primers corresponding to positions 5–26 and 1210–1231 of the coding region. PCR products were purified by electroelution from 1% agarose gels and (Cabomba and Ceratophyllum) blunt-end ligated into the bacteriophage M13 vector mp19 (Bethesda Research Laboratories) or (remaining genera) ligated into M13 Phagescript (Stratagene) using Not I restriction sites attached to the primer ends. Single-stranded phage DNA was sequenced (21) using modified T7 DNA polymerase (Sequenase; United States Biochemical) and a set of synthetic primers based on the rbcL sequence of corn (provided by G. Zurawski, DNAX). Sequences were verified for separate clones of both strands of amplified DNA. The

^{*}The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M77027-M77036).

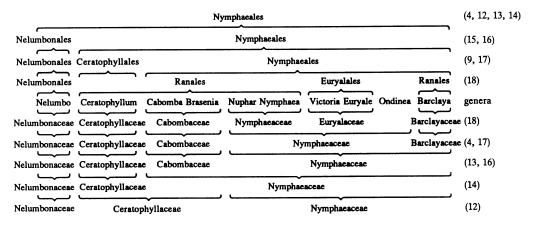


Fig. 1. Contrasting classifications (references at right) of "water lily" genera (center) represented as one to three orders (upper portion of figure) and three to six families (lower portion of figure). Discovery of the genus *Ondinea* postdates one treatment.

few observed discrepancies were resolved by sequencing additional clones.

The rbcL sequences were compared to published sequences of the genera Magnolia, Liriodendron, Persea (subclass Magnoliidae), and the gymnosperm Pseudotsuga (22, 23). Parsimony analyses with character state changes weighted equally and Pseudotsuga specified as the outgroup were carried out using the DNAPARS program of PHYLIP Version 3.3 (provided by J. Felsenstein, University of Washington, Seattle) and PAUP Version 2.4 (24). Most parsimonious trees were found using the "branch and bound" algorithm. A bootstrap analysis of the data was conducted using 1000 replicate computer runs (200 replicates each from five different random number seeds) with the DNABOOT program of Phylip Version 3.3 (25). A strict consensus tree of all most-parsimonious trees was constructed.

A maximum likelihood estimation (MLE) of the phylogeny was made with the DNAML program in PHYLIP Version 3.3 to perform global searches using empirically determined base frequencies and a transition/transversion ratio of 2.0. Empirically determined weights for each codon position were used initially but did not affect resulting topologies and were later eliminated to minimize run times. The MLE approach was optimized by testing log-likelihood values obtained from these runs with those generated by "user trees" representing all topologies obtained by parsimony methods. When user trees provided better likelihoods, the input order of genera was rearranged until the lower values were duplicated. The "jumble" option was then used with 10 different random number seeds to determine whether further manipulation of input order would produce lower log-likelihood values. These "best" MLE topologies were input as user trees and tested for significant differences (26).

RESULTS AND DISCUSSION

Understandably, traditional approaches that ascertain phylogeny using comparative morphological data have reached inconsistent conclusions of evolutionary relationships both within and among these aquatic genera (9). Problems plaguing the systematic study of groups of such ancient divergence arise from the length of time over which evolutionary changes can occur. The lapse of time surpassing 120 million years since the origin of the angiosperms (27) has allowed for considerable modification of characteristics originally existing in flowering plant progenitors. Consequently, it is difficult to ascertain homologies among spurious patterns due to convergence, parallelism, and character/state losses. These problems are perhaps most severe in aquatic plants that experience reduction in features during adaptation to aquatic environments (28) and show widespread convergence and

parallelism for traits linked to survival in water (6). Although morphology has been investigated intensively in the Nymphaeales (29–31), contradictory recommendations for classification have been made from phenetic studies (9, 15) and cladistic analysis (12) of principally morphological data.

A search for greater consistency has inspired testing of water lily relationships using chemosystematic approaches. Because homology of biochemical characters can be less ambiguous than for morphological traits (32), their use in systematic studies has been advocated. Biochemical studies of the Nymphaeales have been informative but inconclusive. Comparative serology has provided convincing evidence for the removal of Nelumbo from the Nymphaeales and its recognition as a distinct order (33). This disposition is supported by the contrasting alkaloid chemistry of Nelumbo and members of the Nymphaeaceae (34). Properties of plastocyanin and ferredoxin are vastly different in Ceratophyllum compared to other plants (35). Starch grain structure indicates an affinity between water lilies and "true" monocotyledons (36); however, the presence of ellagic acid in some Nymphaeaceae cautions against their inclusion within monocots, which lack the compound (37). Water lilies also lack raphides crystals, which are widely distributed among monocots (6). Overall, chemosystematic studies have failed to resolve many of the same questions left unanswered by morphological studies.

Molecular systematics (the use of nucleic acid data for phylogenetic study) has been promoted as a superior technique for surmounting problems associated with the evolutionary study of groups whose origins span long time intervals (38, 39). For "high-level" questions in plant systematics (concerning suprageneric relationships), the study of the chloroplast encoded gene rbcL has been encouraged because of its appropriate degree of divergence (38, 39). Long species divergence times relative to gene coalescent times (40) are theorized to minimize distinctions between "gene trees" and "species trees" (41).

Our sequencing of 1183 bp of rbcL (approximately 85% of the gene) for 10 species in 9 genera is summarized in Fig. 2. In this data set, substitutions occur in 12.6% of the total sequence mainly at third codon positions (74.5%) and less at first (15.4%) and second (10.1%) codon positions. Pairwise comparisons of transition to transversion ratios varied from 0.8 to 4.0, with an overall average of 1.9. For all sequences analyzed, 20.9% of the characters were variable and 9.8% represented informative sites. Because the rbcL sequence of both Nelumbo species differed at only one site, we limited our analyses to one of them (Nelumbo lutea).

Four equally shortest trees (length = 357; consistency index = 0.782) were found by parsimony analysis. The topology of each tree was identical except for placement of

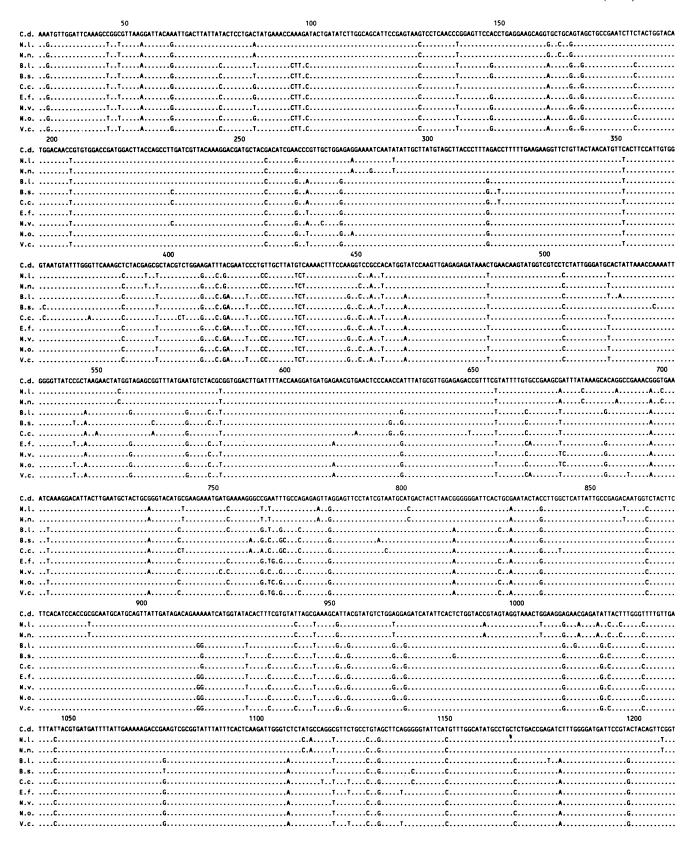


FIG. 2. An 1183-bp region of rbcL comparing Ceratophyllum demersum (C.d.) to Nelumbo lutea (N.l.), Nelumbo nucifera (N.n.), Barclaya longifolia (B.l.), Brasenia schreberi (B.s.), Cabomba caroliniana (C.c.), Euryale ferox (E.f.), Nuphar variegata (N.v.), Nymphaea odorata (N.o.), and Victoria cruziana (V.c.).

Barclaya and Nuphar. In every case, Ceratophyllum was basal to the other genera of Magnoliidae, followed by Nelumbo, which occupied an isolated position. Above these genera, a dichotomy separated two monophyletic groups,

one consisting of the three genera representing Magnoliales (Magnolia, Liriodendron) and Laurales (Persea) and the other including all remaining genera (Barclaya, Brasenia, Cabomba, Euryale, Nuphar, Nymphaea, and Victoria). A

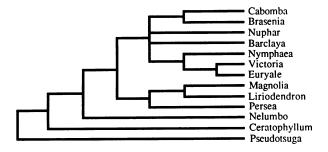


Fig. 3. Strict consensus tree of four equally shortest trees found by parsimony analysis of *rbcL* data in "water lily" genera and woody Magnoliidae rooted using a gymnosperm outgroup.

strict consensus tree (Fig. 3) depicted consistent sister groups of [Persea (Magnolia, Liriodendron)], (Cabomba, Brasenia), and [Nymphaea (Victoria, Euryale)]. In three cladograms, Nuphar and Barclaya interchange or share the position basal to a branch including Nymphaea, Victoria, and Euryale. In the fourth cladogram, Nuphar was a sister group to Cabomba and Brasenia, and Barclaya was the sister group to the six remaining "water lily" genera. A majority-rule consensus tree of the 1000 bootstrap replicates and suggested confidence intervals for nodes of the tree are shown in Fig. 4.

The four tree topologies generated by parsimony methods were identical to those of four "best" maximum likelihood trees with the lowest observed log-likelihood values (Fig. 5). The best log-likelihood values ranged from -3569.61 to -3571.66 but did not differ significantly.

Our analysis of rbcL sequence data warrants that the broadest concept of the order Nymphaeales suggested by Thorne (13), Cronquist (4), and Tamura (14) should be abandoned because inclusion of either Ceratophyllum or Nelumbo with the water lilies represents a polyphyletic group. Forcing the topology of our parsimony analysis to create one monophyletic group of Ceratophyllum, Nelumbo, and the water lilies adds an additional 9 steps to the tree and increases to 43 steps the branch length separating these two genera from the water lilies. Therefore, we conclude that true water lilies are a well-defined clade that includes Barclaya, Brasenia, Cabomba, Euryale, Nuphar, Nymphaea, and Victoria. This arrangement is consistent with opinions of Takhtajan (17) and Les (9) who recognize these genera (along with Ondinea—not surveyed here) to comprise the order Nymphaeales. Classifications including either Nelumbonaceae (42) or Ceratophyllaceae (15, 16, 18) in the Nymphaeales are now disputed by our rbcL phylogenies, by rRNA phylogenies (10), and by recent morphologically based cladistic analyses (43). The lack of close relationship among these genera agrees with phenetic analyses (9) indicating that the Nymphaeales sensu lato embodies three lineages: Nymphaeales, Nelumbonales, and Ceratophyllales.

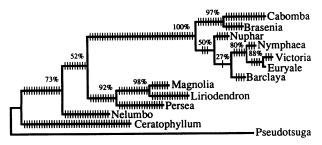


Fig. 4. Majority-rule bootstrap consensus tree from parsimony analysis of genera in Fig. 3. Short vertical bars represent substitutions supporting portions of the topology. Percentages give confidence intervals for nodes based upon 1000 bootstrap samples. Branch lengths are proportional to the number of inferred nucleotide substitutions.

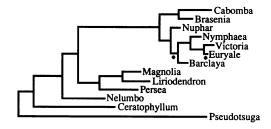


Fig. 5. One of "best" maximum-likelihood topologies (showing relative branch lengths) agrees with parsimony analysis of genera analyzed in Fig. 4. No significantly better tree could be found using this algorithm. *, Branches whose lengths are not significantly positive.

The question of an appropriate familial classification for these aquatic genera (Fig. 1) can also be largely resolved by rbcL data. We advocate continued recognition of the Nelumbonaceae and Ceratophyllaceae as distinct families, remote phylogenetically from true water lilies. This arrangement is supported by most contemporary classifications (Fig. 1). The suggestion to include Cabomba and Brasenia in the Ceratophyllaceae (12), however, is unacceptable. Sequence data from rbcL support recognition of the family Cabombaceae (Cabomba and Brasenia), which consistently appears as a monophyletic group within the Nymphaeales. Except for Ito (12) and Tamura (14), this disposition reflects the opinion of other contemporary taxonomists (Fig. 1). The recognition of a monotypic family Barclayaceae (4, 17, 18) is not supported by rbcL data. The association of Barclaya with Nuphar. Nymphaea, Victoria, and Euryale in our analyses indicates a close relationship among these genera. Our analyses provide similar arguments against Nakai's segregation (42) of Nuphaceae (Nuphar) from the Nymphaeaceae. Li's concept (18) of Nymphaeaceae (Nymphaea, Nuphar) is also rejected given that rbcL data place Nymphaea closer phylogenetically to Victoria and Euryale than to Nuphar (Figs. 3-5). Most contemporary authors (Fig. 1) recognize Victoria and Euryale as members of the Nymphaeaceae. Phylogenetic analysis of rbcL data supports the recognition of two families in the Nymphaeales, Cabombaceae (Cabomba, Brasenia) and Nymphaeaceae (Barclaya, Euryale, Nuphar, Nymphaea, and Victoria), but provides insufficient grounds for further taxonomic subdivision.

A high degree of sequence similarity (99.9%) exists in the 1183-bp region of rbcL sequenced in Nelumbo lutea and Nelumbo nucifera, the two extant species of the Nelumbonaceae. These species are extremely similar morphologically, with notable differences occurring only in their petal color and fruit shapes (44). High rbcL sequence similarity is consistent with the complete serological identity observed for these species (33). Although Nelumbo lutea (North America) and Nelumbo nucifera (Asia and Australia) are presently disjunct, fossil evidence indicates that their geographic isolation may have occurred as recently as the early Pleistocene (45). These observations verify expected correspondence among recent divergence time, slight morphological differentiation, serological identity, and low level of nucleotide divergence in these species.

In summary, the study of *rbcL* divergence in ancient aquatic plants has provided important insights. The hypothesis (based on a morphological reduction series) that the Ceratophyllaceae represent a highly specialized offshoot of the Nymphaeales (4) is rejected by *rbcL* phylogenies. Instead, the *rbcL* phylogeny is consistent with a hypothesis of relationships based upon a comprehensive study of nonmolecular data (9). Such concordance lends increased confidence in the capability of *rbcL* sequences to test phylogenetic hypotheses involving ancient angiosperms. Erroneous con-

cepts of relationships in the Nymphaeales have been due largely to convergent evolution (e.g., dissected foliage in Ceratophyllum and Cabomba or floating leaves in Nelumbo and true water lilies). By overcoming intrinsic difficulties of interpreting convergent morphological evolution, DNA sequence analysis is proving to be of great utility in the study of aquatic plant phylogenies.

Correlation of high rbcL homology with other measures of evolutionary divergence supports inferences of recent divergence between the two surviving Nelumbo species.

Analysis of rbcL data provides additional evidence that the modern aquatic Ceratophyllales represent descendants of some very early flowering plants (9). In our phylogenetic analyses, the Ceratophyllales precede the woody Magnoliidae (e.g., Magnoliales), which have long been assumed to represent the "basal" angiosperm lineage (4, 43), and the Nymphaeales, which have also been implicated as ancestral angiosperms (9, 10). Naturally, the full resolution of phylogenetic relationships among early angiosperm lineages ultimately awaits global analyses that include a large number of additional angiosperm sequences from rbcL and other genes. Sequencing additional rbcL genes in genera within the dicot subclass Magnoliidae and members of the primitive aquatic monocot subclass Alismatidae should provide further data pertinent to resolving this critical question. A preliminary result maintains the same maximum likelihood topology of these aquatic genera as additional genera from Magnoliidae and Alismatidae are added to our analysis.

Are the angiosperms of an aquatic ancestry? Although evidence is mounting that implicates aquatic plants as descendants of early flowering plants, it is highly unlikely that the group evolved from aquatic ancestors. Aquatic plants typically exhibit stasis and extremely long species durations (46, 47). Their ability to survive over long time periods has probably biased the sample of surviving archaic angiosperms to overly represent hydrophytes.

We thank N. Harriman, E. Schneider, D. Bryne, and Mr. P. Nutt, for plant material, T. Schuck for care of plants, and S. Duncan for computer assistance. This project was supported by National Science Foundation Grant BSR-8817992.

- Raven, P. & Axlerod, D. I. (1974) Ann. Mo. Bot. Gard. 61, 539-673.
- Axlerod, D. I. (1952) Evolution 6, 29-60.
- Haines, R. W. & Lye, K. A. (1975) Bot. J. Linn. Soc. 70, 255-265.
- Cronquist, A. (1988) The Evolution and Classification of Flowering Plants (New York Botanical Garden, Bronx, NY).
- Takhtajan, A. L. (1980) Bot. Rev. 46, 225-359.
- Dahlgren, R. M. T. & Rasmussen, F. N. (1983) Evol. Biol. 16, 255-395.
- Schaffner, J. H. (1904) Ohio Nat. 4, 83-92.
- Dahlgren, R. M. T., Clifford, H. T. & Yeo, P. F. (1985) The Families of Monocotyledons (Springer, Berlin).
- Les, D. H. (1988) Taxon 37, 326-345.
- Hamby, K. & Zimmer, E. A. (1991) in Molecular Systematics

- of Plants, eds. Soltis, P., Soltis, D. & Doyle, J. J. (Chapman & Hall, London), pp. 50-91.
- Dilcher, D. L. (1989) Am. J. Bot. 76, 162 (abstr.). Ito, M. (1987) Bot. Mag. 100, 17-35.
- 12.
- 13. Thorne, R. F. (1976) Evol. Biol. 9, 35-106.
- Tamura, M. (1982) Acta Phytotaxon. Geobot. 33, 336-345.
- Bukowiecki, H., Furmanowa, M. & Oledzka, H. (1972) Acta Pol. Pharm. 29, 319-327.
- Dahlgren, R. M. T. (1980) Bot. J. Linn. Soc. 80, 91-124.
- Takhtajan, A. (1987) Systema Magnoliophytorum (Leninopoli).
- Li, H. (1955) Am. Midl. Nat. 54, 33-41. 18.
- 19 Doyle, J. J. & Doyle, J. L. (1987) Phytochem. Bull. 19, 11-15.
- Saiki, R., Gelfund, D., Stoffel, S., Scharf, S., Higuchi, R., Horn, G., Mullis, K. & Erlich, H. (1988) Science 239, 487-491.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- Golenberg, E. M., Gianassi, D. E., Clegg, M. T., Smiley, C. J., Durbin, M., Henderson, D. & Zurawski, G. (1990) Nature (London) 344, 656-658.
- Hipkins, V., Tsai, C. H. & Strauss, S. H. (1990) Plant Mol. Biol. 15, 505-507.
- Swofford, D. L. (1985) PAUP: Phylogenetic Analysis Using Parsimony, Version 2-4 (Illinois Nat. Hist. Surv., Champaign).
- Felsenstein, J. (1985) Evolution 39, 783-791
- Kishino, H. & Hasegawa, M. (1989) J. Mol. Evol. 29, 170-179.
- Beck, C. B. (1980) Origin and Early Evolution of Angiosperms (Columbia Univ. Press, New York).
- Sculthorpe, C. D. (1967) Biology of Aquatic Vascular Plants (Edward Arnold, London).
- Osborn, J. M. & Schneider, E. L. (1988) Ann. Mo. Bot. Gard. **75**, 778–794.
- Ito, M. (1986) Bot. Mag. 99, 169-184.
- Khanna, P. (1965) Aust. J. Bot. 13, 379-387.
- Young, D. A. & Siegler, D. S. (1981) Phytochemistry and Angiosperm Phylogeny (Praeger, New York).
- 33. Simon, J. P. (1970) Aliso 7, 243-261.
- 34. Gibbs, R. D. (1974) Chemotaxonomy of Flowering Plants (McGill-Queen's Univ. Press, Montreal).
- Oganezova, E. P. & Nalbandyan, R. M. (1976) Biokhimiya (Moscow) 41, 794-800.
- Czaja, A. T. (1978) Taxon 27, 463-470.
- Dahlgren, R. (1977) Plant Syst. Evol. (Supp.) 1, 253-283.
- 38. Ritland, K. & Clegg, M. T. (1987) Am. Nat. 130, S74-S100.
- Zurawski, G. & Clegg, M. T. (1987) Annu. Rev. Plant Physiol. **38**, 391–418.
- 40. Weir, B. S. (1990) Genetic Data Analysis (Sinauer, Sunderland, MA).
- Tateno, Y., Nei, M. & Tajima, F. (1982) J. Mol. Evol. 18,
- Nakai, T. (1943) Ordines, Familiae, Tribi, Genera, Sectiones, Species, Varietates, Formae et Combinationes Novae a Prof. Nakai-Takenoschin Adhuc ut Novis Edita (Imperial University, Tokyo).
- Donoghue, M. J. & Doyle, J. A. (1989) in The Hierarchy of Life, eds. Fernholm, B., Bremer, K. & Jornvall, H. (Elsevier, Amsterdam), pp. 181-193.
- Cook, C. D. K. (1970) Water Plants of the World (Junk, The Hague, The Netherlands).
- Good, R. (1964) The Geography of Flowering Plants (Wiley, New York).
- Les, D. H. (1988) Ann. Mo. Bot. Gard. 75, 819-835.
- Herendeen, P. S., Les, D. H. & Dilcher, D. L. (1990) Am. J. Bot. 77, 7-16.