J Mol Evol (1991) 32:253–261

Journal of
Molecular Evolution

Angiosperm Origin and Early Stages of Seed Plant Evolution Deduced from rRNA Sequence Comparisons

A.V. Troitsky, Yu.F. Melekhovets, G.M. Rakhimova, V.K. Bobrova, K.M. Valiejo-Roman, and A.S. Antonov

A.N. Belozersky Laboratory of Molecular Biology and Bioorganic Chemistry, Moscow State University, Moscow 119899, USSR

Summary. Complete or partial nucleotide sequences of five different rRNA species, coded by nuclear (18S, 5.8S, and 5S) or chloroplast genomes (5S, 4.5S) from a number of seed plants were determined. Based on the sequence data, the phylogenetic dendrograms were built by two methods, maximum parsimony and compatibility. The topologies of the trees for different rRNA species are not fully congruent, but they share some common features. It may be concluded that both gymnosperms and angiosperms are monophyletic groups. The data obtained suggest that the divergence of all the main groups of extant gymnosperms occurred after the branching off of the angiosperm lineage. As the time of divergence of at least some of these gymnosperm taxa is traceable back to the early Carboniferous, it may be concluded that the genealogical splitting of gymnosperm and angiosperm lineages occurred before this event, at least 360 million years ago, i.e., much earlier than the first angiosperm fossils were dated. Ancestral forms of angiosperms ought to be searched for among Progymnospermopsida. Genealogical relationships among gymnosperm taxa cannot be deduced unambiguously on the basis of rRNA data. The only inference may be that the taxon Gnetopsida is an artificial one, and *Gnetum* and *Ephedra* belong to quite different lineages of gymnosperms. As to the phylogenetic position of the two Angiospermae classes, extant monocotyledons seem to be a paraphyletic group located near the root of the angiosperm branch; it emerged at the earliest stages of angiosperm evolution. We may conclude that either monocotyledonous characters arose independently more than once in different groups of ancient Magnoliales or that monocotyledonous forms rather than dicotyledonous Magnoliales were the earliest angiosperms. Judging by the rRNA trees, Magnoliales are the most ancient group among dicotyledons. The most ancient lineage among monocotyledons leads to modern Liliaceae.

Key words: Chloroplast 4.5S rRNA — Cytosolic and chloroplast 5S rRNAs — 5.8S rRNA — 18S rRNA — Nucleotide sequences — Phylogenetic trees — Angiosperms — Gymnosperms — Monocotyledons — Dicotyledons

Introduction

Phylogenetic relationships among the main groups of seed plants remain obscure and are widely discussed in the botanical literature (Beck 1976, 1988; Doyle 1978; Meyen 1984; Crane 1985; Doyle and Donoghue 1987; Krassilov 1989; and papers cited therein). In this paper we take a molecular approach to the problem. The nucleotide sequences of plant cytosolic 5.8S and 5S rRNAs, chloroplast 4.5S and 5S rRNAs, and also the partial sequences of cytosolic 18S rRNA (totally, about 760 positions, approximately 11,800 nucleotide residues) have been used to construct phylogenetic trees by the compatibility and maximum parsimony methods.

Although it is widely accepted that rRNA is an appropriate molecule for inferring phylogenetic relations, until recently plant rRNAs were not the object of close scrutiny in this respect. Molecular phylogenetic studies of plants were performed most-

ly by sequencing a single molecular species, be it a protein or rRNA. There is only one paper in which the data obtained for different plant molecular species were analyzed together (Martin et al. 1985). Most of the macromolecules considered in that paper were proteins and the only rRNA species was cytosolic 5S rRNA. Only dicotyledonous angiosperms were studied. The authors confronted certain difficulties while trying to combine individual dendrograms into a global tree.

In this paper we present the results of phylogenetic reconstructions based on the sequences of five different rRNA species and obtained by the two methods, compatibility and maximum parsimony. Preliminary considerations were published earlier (Rakhimova et al. 1989; Troitsky et al. 1989b).

Materials and Methods

The low molecular weight rRNAs were isolated by the hot phenol extraction procedure at pH 5.1 and purified by ion-exchange chromatography on DEAE-Toyoperl 650M (Toyo Soda, Japan) and electrophoresis on 8% polyacrylamide gel with 7 M urea, pH 8.3 as described in Troitsky et al. (1984, 1989b). The sequencing was performed by the method of Peattie (1979).

The procedure of total high molecular RNA isolation and partial sequencing with the use of reverse transcriptase (Liang et al. 1983) were described earlier (Rakhimova et al. 1989). The primer for sequencing d(CTTGCTTTGAGCACTCTAATTT) specifically interacts with the nucleotides 1533–1549 of nuclear-encoded 18S rRNA [numeration according to Dams et al. (1988)].

Dendrograms were constructed by the compatibility method (Estabrook 1983; Le Quesne 1983) using the original algorithm (Omelyanchuk and Kolchanov 1985) briefly described in our earlier papers (Rakhimova et al. 1989; Troitsky et al. 1989b) and by the maximum parsimony method using a program from the PHYLIP package (Felsenstein 1989). The minimal numbers of fixed mutational events in the branches of the trees obtained were calculated by the Fitch procedure (Fitch 1971) using a program from the VOSTORG package (Zharkikh et al. 1990).

The algorithm of the compatibility method is based on the analysis of elementary trees for the quartets of sequences from the alignment. Three topologies for the elementary tree are possible for each quartet. Each of the three topologies may be characterized by a number of compatible sites n_1, n_2, and n_3. According to the main theorem of the compatibility method (Estabrook and McMorris 1980; Omelyanchuk and Kolchanov 1985), each set of compatible sites may be used to construct an additive tree. It is obvious that only one of these trees reflects the real process of divergence of the four sequences analyzed. It is suggested that this is the tree with the maximal number of compatible sites $N_{max} = \max\{n_1, n_2, n_3\}$. To evaluate the validity of prevalence of one topology over the other two by the number of compatible sites, the following criterion was used. Let us suppose the total number of compatible sites in three elementary trees is $M = n_1 + n_2 + n_3$, and these sites are uniformly distributed among the trees. Then the probability of obtaining, by chance, in one of the three elementary trees the number of compatible sites N_{max} (or even higher) will be

$$P = \sum_{i=N_{max}}^{M} C_M^i \left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^{M-i}$$

This value may be considered as a first approximation for the statistical estimation of the choice of the best topology of elementary trees. If P for a given quartet of nucleotide sequences is smaller than a certain threshold level P_0 ($P_0 \ll 1$), this means that one of the topologies definitely exceeds two others in the number of compatible sites and may be considered as the best representation of the true pattern of divergence. For the whole tree reconstruction all possible C_N^4 quartets of sequences are considered and for each of them the value of P is calculated. The tree is constructed by adding in stepwise succession the elementary trees with increasing P values. The topologies obtained at lower P values are fixed and are not changed in the process of addition of elementary trees with higher P.

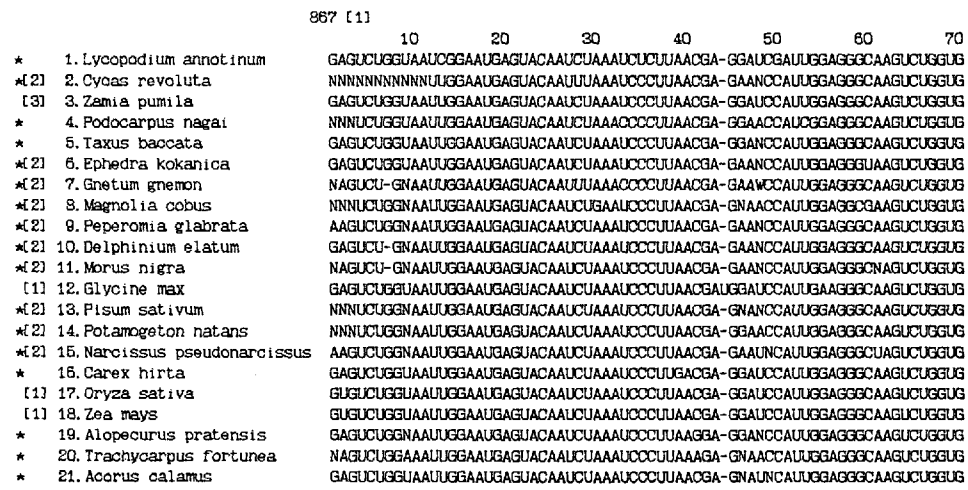
Results and Discussion

The nucleotide sequences used for phylogenetic tree building are presented in Figs. 1–5. Phylogenetic trees constructed for these five different cytosolic and chloroplast rRNA species of a number of land plants are shown in Fig. 6. Dendrograms for cytosolic 18S rRNA and chloroplast 4.5S rRNA (Fig. 6A and B) and those for 5S and 5.8S rRNAs (Fig. 6C–E) were obtained by the compatibility and the maximum parsimony methods, respectively. Besides the compatibility method, the phylogenetic dendrograms for 4.5S and 18S rRNAs were also constructed by the maximum parsimony method; their topologies proved to be very similar to those presented in Fig. 6A and B (data not shown). When the maximum parsimony method was used to analyze the complete sets of data for any rRNA studied, several dendrograms with different but equally parsimonious topologies were obtained. To overcome this difficulty the dendrograms were built stepwise. At the beginning, partially overlapping, locally optimal dendrograms were constructed, which at the next step were brought together into the global trees. Using such an approach, we supposed that the probability of similarity arising due to homoplasy is lower in sequences that have diverged more recently.

As it follows from Fig. 6, the topologies of the trees for different rRNA species are not fully congruent. Because in the case of 4.5S and 5S rRNAs it is not possible to find unambiguously the relative position for all the dicot representatives, corresponding trees include only a part of dicotyledonous species for which the sequences of these rRNAs are known. Comparing the individual dendrograms in Fig. 6 we may conclude that a unimolecular dendrogram does not allow definite conclusions to be made concerning phylogenetic interrelatedness of taxa, even if an rRNA is analyzed.

We may speculate that a global tree derived from a more representative set of the rRNA sequences would give us a better insight into the land plant genealogy, but this is as yet impossible because the data available for different rRNA species are incomplete and overlap only partially. The aim of our present research is to accumulate data and try to construct such a tree.

18S rRNA



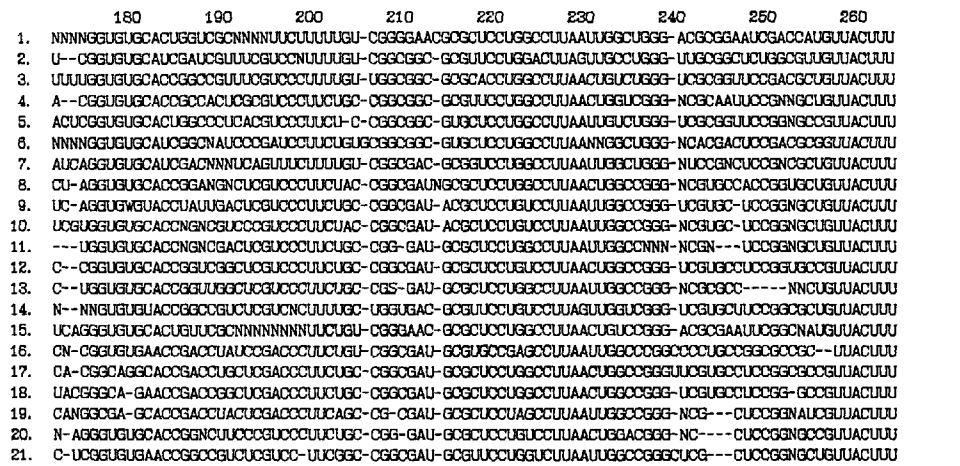


Fig. 1. Alignment of partial nucleotide sequences of plant 18S rRNAs. Numbers in brackets are the references to original papers or compilations; *, our data; if they were published earlier, the reference is given in brackets. The first nucleotide is 867 according to the alignment in compilation [1] (Dams et al. 1988). Other references: [2] Rakhimova et al. (1989), [3] Nairn and Ferl (1988).

Chloroplast 4.5S rRNA

```
                                    10        20        30        40        50        60        70
*[4]  1. Marchantia polymorpha      ---UAAGGU-GACGGCAAGACUAGCCGUUUAUUUUUA-CGAUAGGUGCCAAGUGGAAGUGCAGUAAUGUA
*     2. Ginkgo biloba              ---UAAGGU-CACGGCGAGACGAGCCGUUUAUCAUCA-CGAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
*     3. Larix sibirica             ---UAAGGU-CACGGCGAGACGAGCCGUUUAUCAUUA-CGAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
*     4. Cycas revoluta             ---UAAGGU-CACGGCGAGACGAGCCGUUUAUCAUCA-CGAUAGGUGUCAAGUGGAAGUGCAGUGUUGCA
*     5. Ephedra kokanica           ---UAAGGU-CACGGCGAGACGAGCCGUUUAUCAUCA-CGAUAGGUGUCAAGUGGAAGUGCAGUGAUGCA
*     6. Delphinium consolida       ---UAAGGU-CACGGCGAGACGAGCCGUUUAUCAUUA-CGAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
 [4]  7. Spinacia oleracea          AGAGAAGGU-CACGGCGAGACGAGCCGUUUAUCAUUA-CGAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
*[5]  8. Ligularia calthifolia      ---GAAGGU-CACGGCGAGACGAGCCGUUUAUCAUUA-CGAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
 [6]  9. Apium graveoleus           ---GAAGGU-CACGGUGAGACGAGCCGUUUAUCAUUA-CGAUAGGUGUCUAGUGGAAGUGCAGUGAUGUA
 [7] 10. Oenothera berteriana       ---GAAGGU-CACGGCGAGACGAGCCGUUUAUCAUUA-UGAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
*[4] 11. Morus nigra                ---UAAGGU-CACGGCGAGACGAGCCGUUUAUCAUUA-CGAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
 [8] 12. Glycine max                ---GAAGGU-CACGGCGAGACGAGCCGUUUCUAAUUAAUGAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
 [6] 13. Commelina communis         ---UAAGGU-AGCGGCGAGACGAGCCGUUUAUCAUUA-CGAUAGGUGUCAAGUGGAAGUGCAGUAAUGUA
 [6] 14. Allium tuberosum           ---UAAGGU-CACGGCGAGACGAGCCGUUUAUCAUUA-CGAUAGGUGUCAUGUGGAAGUGCAGUGAUGUA
*[4] 15. Narcissus pseudonarcissus  ---UAAGGU-CACGGCAAGACGAGCCGUUUAUCAUUA-CGAUAGGUGCCAGGUGGAAGUGCAGUAAUGUA
 [6] 16. Hordeum vulgare            ---UAAGGU-AGCGGCGAGACGAGCCGUUUA--------AAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
 [4] 17. Zea mays                   ---UAAGGU-AGCGGCGAGACGAGCCGUUUA--------AAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
 [4] 18. Triticum aestivum          ---UAAGGUGAGCGGCGAGACGAGCCGUUUA--------AAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
*[5] 19. Acorus calamus             ---UAAGGU-CACGGCGAGACGAGCCGUUUAUCAUUA-CGAUAGGUGCCAAGUGGAAGUGCAGUGAUGUA
 [4] 20. Spirodela oligorhisa       ---UAAGGU-CACGGCUAGACGAGCCGUUUAUCAUUA-CGAUAGGUGUCAAGUGGAAGUGCAGUGAUGUA
```

```
          80        90        100
 1.  UGUAGCUGAGGCAUCCUAACAGACCGAGAGAUUUAAAC
 2.  UGCAGCUGGAGCAUCCUAACAGACCGAGAGAUUUGAAC
 3.  UGCAGCUGAGGCAUCCUAACAGACCGAGAGAUUUGAAC
 4.  UGCAGCUGAGGCAUCCUAACAGACCGAGAGAUUUGAAC
 5.  UGCAGCUGAGACAUCCUAACAGACCGAGAGAUUUGAAC
 6.  UGCAGCUGAGGCAUCCUAAUAGACCAAUAGACUUGAAC
 7.  UGCAGCUGAGGCAUCCUAACAGACCCACAGACUUGAAC
 8.  UGCAGCUGAGGCAUCCUAACAGACCGGUAGACUUGAAC
 9.  UGCAGCUGAGGCAUCCUAACAGACCGGCAGAUUUGAAC
10.  UGCAGCUGAGACAUCCUAACAGACCGCUAGACUUGAAC
11.  UGCAGCUGAGGCAUCCUAACAGACCGGUAGACUUGAAC
12.  UGCAGCUGAGGCAUCCUAACAGACCGGUAGACUUGAAC
13.  UGCAGCUGAGGCAUCCUAACAGACCGAGAGAUUUGAAC
14.  UGCAGCUGAGGCAUCCUAAUAGACCGAGAGAUUUGAAC
15.  UGCAGCUGAGGCAUCCUAACAGACCGAGAGAUUUGAAC
16.  UGCAGCUGAGGCAUCCUAACGAACCGAACGAUUUGAAC
17.  UGCAGCUGAGGCAUCCUAACGAA-CGAACGAUUUGAAC
18.  UGCAGCUGAGGCAUCCUAACGAA-CGAACGAUUUGAAC
19.  UGCAGCUGAGGCAUCCUAACAGACCGAGAGAUUUGAAC
20.  UGCAGCUGAGGCAUCCUAAU-GACCGAGAGAUUUGAAC
```

Fig. 2. Alignment of plant chloroplast 4.5S rRNA sequences. Designations are the same as in Fig. 1. [4] Troitsky et al. (1989b), [5] Bobrova et al. (1987), [6] Cheng et al. (1986), [7] Schuster and Brennicke (1987), [8] Nazar et al. (1987).

5.8S rRNA

```
                                    10        20        30        40        50        60        70
*[9]  1. Mnium rugicum              ---AUAACCCUCAGCAACGGAUAUCUUGGCUCUUGCAACGAUGAAGAACGCAGCGAAAUGCGAUACGUAG
*     2. Taxus baccata              CUUGGCACUCUCGGCAACGGAUAUCUCGGCUCUCGC-ACGAUGAAGAACGUAGCGAAAUGCGAUACUUAG
*[10] 3. Ephedra kokanica           CUUACGACUCUCGGCAAUGGAUAUCUCGGCUCUCGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACUUAG
*     4. Gnetum gnemon              CCCACGACUCUCGACAAUGGAUAUCUCGGCUCUCGU-UCGAUGAAGAACGUAGCGAAAUGCGAUACUUGG
*     5. Picea excelsa              UAAAUGACUCUCGGCAACGGAUAUCUCGGCUCUUGU-ACGAUGAAGAACGUAGCGAAAUGCGAUACUUAG
*     6. Potamogeton natans         AUCAUGACUCUCGGCAACGGACAUCUUGGCUCUCGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACUUGG
 [11] 7. Triticum aestivum          CACACGACUCUCGGCAACGGAUAUCUCGGCUCUCGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACCUGG
 [11] 8. Oryza sativa               CACACGACUCUCGGCAACGGAUAUCUCGGCUCUCGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACCUGG
 [12] 9. Lycopersicon esculentum    CAAACGACUCUCGGCAACGGAUAUCUCGGCUCUCGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACUUGG
 [11] 10. Lupinus luteus            CUAAAGACUCUCGGCAACGGAUAUCUCGGCUCUUGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACUUGG
 [11] 11. Vicia faba                AGAAUGACUCUCGGCAACGGAUAUCUAGGCUCUUGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACUUGG
 [13] 12. Cucumis melo              -CAACGACUCUCGGCAACGGAUAUCUCGGCUCUCGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACUUGG
 [14] 13. Vigna radiata             AAAACGACUCUCGGCAACGGAUAUCUCGGCUCUUGCAUCGAUGAAGAACGUAGCGAAAUGCGAUACUUGG
```

```
          80        90        100       110       120       130       140       150       160
 1.  UGUGAAUUGCAGAAUUCCGCGAAUCAUCGAGUCUUUGAACGCAAGUUGCGCCCGAGGC---UCGUCCGAGGGCAUUUCCGUUAGAGCGUCACC-
 2.  UGUGAAUUGCAGAAUCCCGUGAAUCAUCGAGUCUUUGAACGCAAGUUG-GCCCGGAGC---UCGGCCGAGGGCACGUCUGCUUGGGCGUCGCAC
 3.  UGUGAAUUGCAGAAUCCCGUGAAUCAUCGAGUCUUUGAACGCAAGUUG-GCCCGAAGCC--UCCGCCAAGGGCACGUCUGCCUGGGCGUCGCAA
 4.  UGUGAAUUGCAGAAUCCCGUGAAUCAUCGAAUCUUUGAACGCAAGUUG-GCCCUCCGAGCCUAGGCCGAGGGCACGUCUGCCUGGGUGUUGUGG
 5.  UGUGAAUUGCAGAAUCCCGUGAAUCAUCGAGUUUUUGAACGCAAUUUG-GCCCAGAAGGCCUUGGUCGAGGGCACGUCUGUCUGGGCGUCGCAU
 6.  UGUGAAUGGCAGAAUCCCGUGAACCAUCGAGUUUUUGAACGCAAGUUGCGCCCUAAGCUUCCGGGCCGAGGGCAAGCCUGCGUGGGCGUCAAGC
 7.  UGUGAAUUGCAGAAUCCCGCGAACCAUCGAGUCUUUGAACGCAAGUUGCGCCCGAGGCCACUCGGCCGAGGGCACGCCUGCCUGGGCGUCACGC
 8.  UGUGAAUUGCAGAAUCCCGUGAACCAUCGAGUCUUUGAACGCAAGUUGCGCCCGAGGCCAUCCGGCCGAGGGCACGCCUGCCUGGGCGUCACGC
 9.  UGUGAAUUGCAGAAUCCCGUGAACCAUCGAGUCUUUGAACGCAAGUUGCGCCCGAAGCCAUUUGGCCGAGGGCACGUCUGCCUGGGCGUCACG-
10.  UGUGAAUUGCAGAAUCCCGUGAACCAUCGAGUCUUUGAACGCAAGUUGCGCCCGAAGCCAUUAGGCCGAGGGCACGCCUGCCUGGGUGUUGCAC
11.  UGUGAAUUGCAGAAUCCCGUGAACCAUCGAGUCUUUGAACGCAAGUUGCGCCCGAUGCCAUUAGGUUGAGGGCACGUCUGAAUGGGUGUCACAU
12.  UGUGAAUUGCAGGAUCCCGCGAACCACCGAGUCUUUGAACGCAAGUUGCGCCCGGAGCCUUCUGGCCGAGGGCACGUCUGCCUGGGCGUCACGC
13.  UGUGAAUUGCAGAAUCCCGUGAACCAUCGAGUCUUUGAACGCAAGUUGCGCCCAAAGCCAUUAGGCC-AGGGCACGCCUGCCUGGGUGUCACAC
```

Fig. 3. Alignment of plant 5.8S rRNA sequences. Designations are the same as in Figs. 1 and 2. [9] Troitsky et al. (1989a), [10] Melekhovets and Troitsky (1990), [11] Erdmann and Wolters (1986), [12] Kiss et al. (1988), [13] Kavanagh and Timmis (1988), [14] Schiebel and Hemleben (1989).

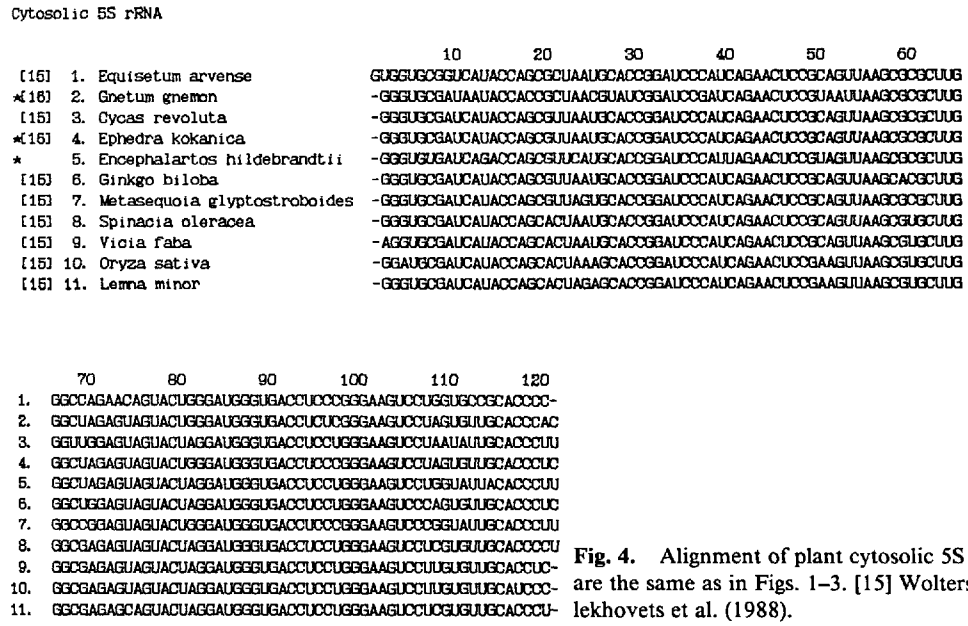


Fig. 4. Alignment of plant cytosolic 5S rRNA sequences. Designations are the same as in Figs. 1–3. [15] Wolters and Erdmann (1988), [16] Melekhovets et al. (1988).

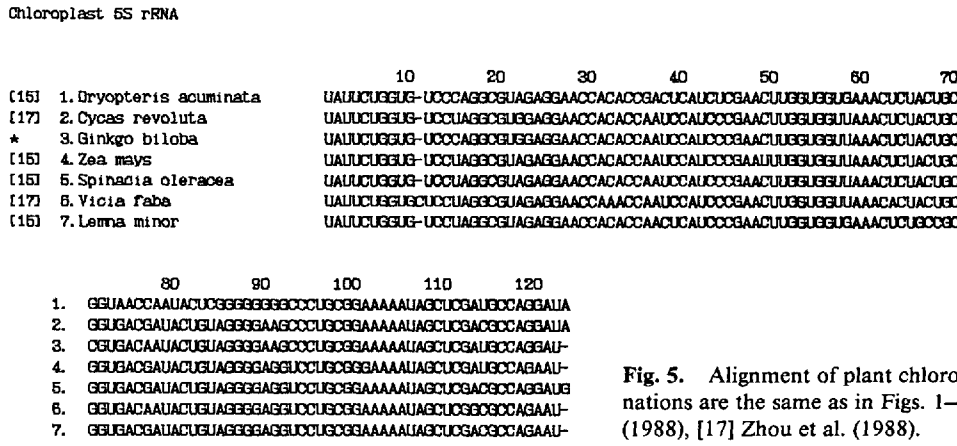
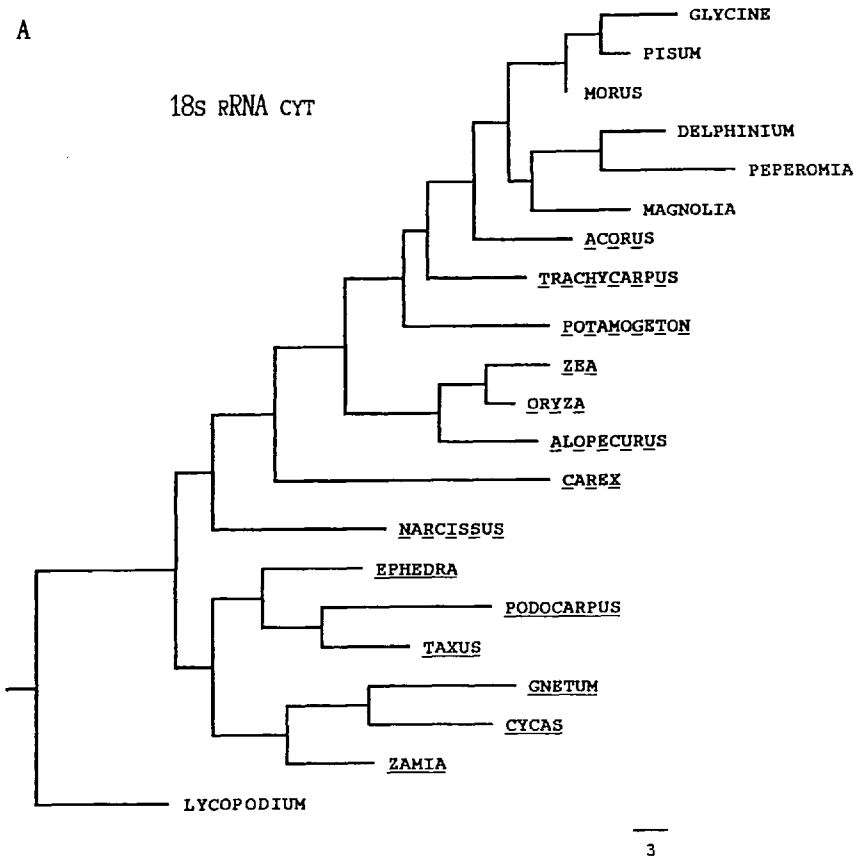


Fig. 5. Alignment of plant chloroplast 5S rRNA sequences. Designations are the same as in Figs. 1–4. [15] Wolters and Erdmann (1988), [17] Zhou et al. (1988).

Although the differences in the topology of the trees in Fig. 6 are obvious, they share some common features that will be the object of our consideration here. The topologies of the trees suggest that the divergence of all the main groups of extant gymnosperms occurred *after* the branching off of the angiosperm lineage. These groups are Cycadales (*Cycas, Zamia, Encephalartos*), Coniferales (*Metasequoia, Podocarpus, Taxus, Larix, Picea*), Ginkgoales (*Ginkgo*), Gnetales (*Gnetum*), and Ephedrales (*Ephedra*). The taxonomic status of these groups may differ in the systems suggested by botanists, but there is a general view now that these groups represent all the major gymnosperm lineages (Doyle 1978; Meyen 1984; Crane 1985; Doyle and Donoghue 1987; Beck 1988; Krassilov 1989). As the time of divergence of at least some of these taxa is traceable back to the early Carboniferous, we have concluded that the genealogical splitting of gymnosperm and angiosperm lines of descent occurred *before* this event, i.e., at the Devonian–Carboniferous boundary, approximately 360 million years (Myr) ago, shortly after the branching off of the Pteridophyta lineage (Rakhimova et al. 1989).

Naturally, proceeding from molecular data alone, one cannot imagine the morphology of these ancient extinct angiosperms. Here two possibilities may be discussed. One is that such ancestral forms had already possessed some specific angiospermous features. The absence of unequivocal angiosperm fossils in pre-Cretaceous strata may be due to the scarcity of proangiosperms or their poor preservation in some special habitats (Axelrod 1970). The second hypothesis, which seems more realistic to us, is that up to the early Cretaceous, when massive angiosperm radiation occurred, these plants have had mostly gymnospermous features, which masked their differentiation from extinct gymnosperms.



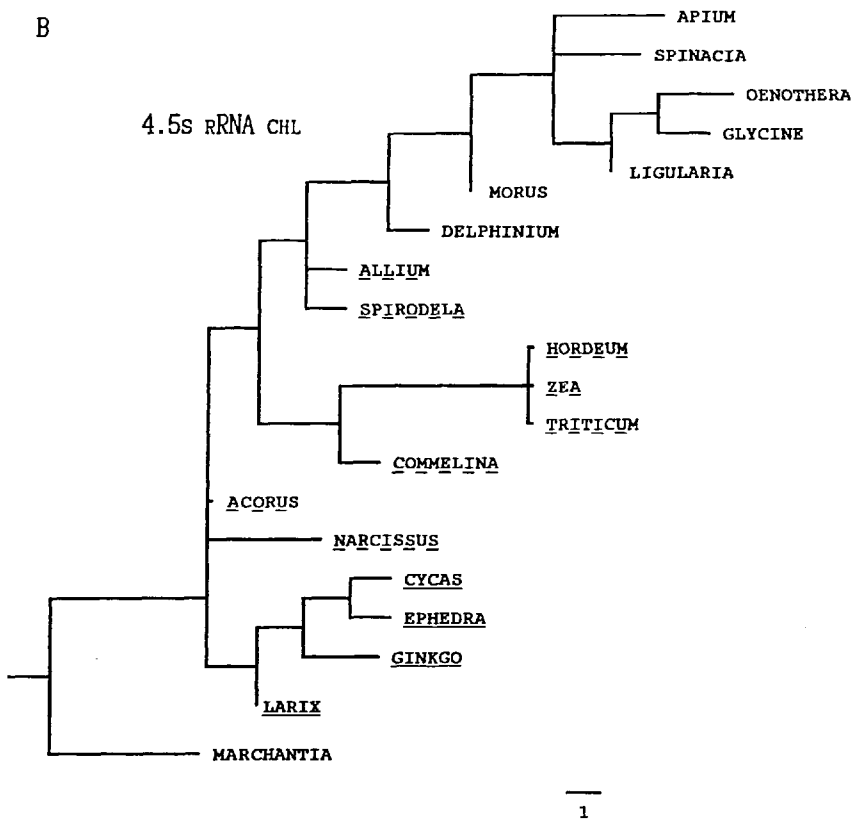


Fig. 6. Phylogenetic dendrograms constructed for different plant rRNAs: A, cytosolic 18S rRNA, a fragment of 263 nucleotides from position 867 according to numeration in the compilation of Dams et al. (1988); B, chloroplast 4.5S rRNA; C, cytosolic 5.8S rRNA; D, cytosolic 5S rRNA; E, chloroplast 5S rRNA. Alignments of sequences are presented in Figs. 1–5. Only generic names of plants are given; the full names of the species may be found in Figs. 1–5. The units of the scales are the minimal numbers of mutations (Fitch 1971). At each dendrogram the monocot species are underlined by dashed lines and gymnosperms by straight lines.

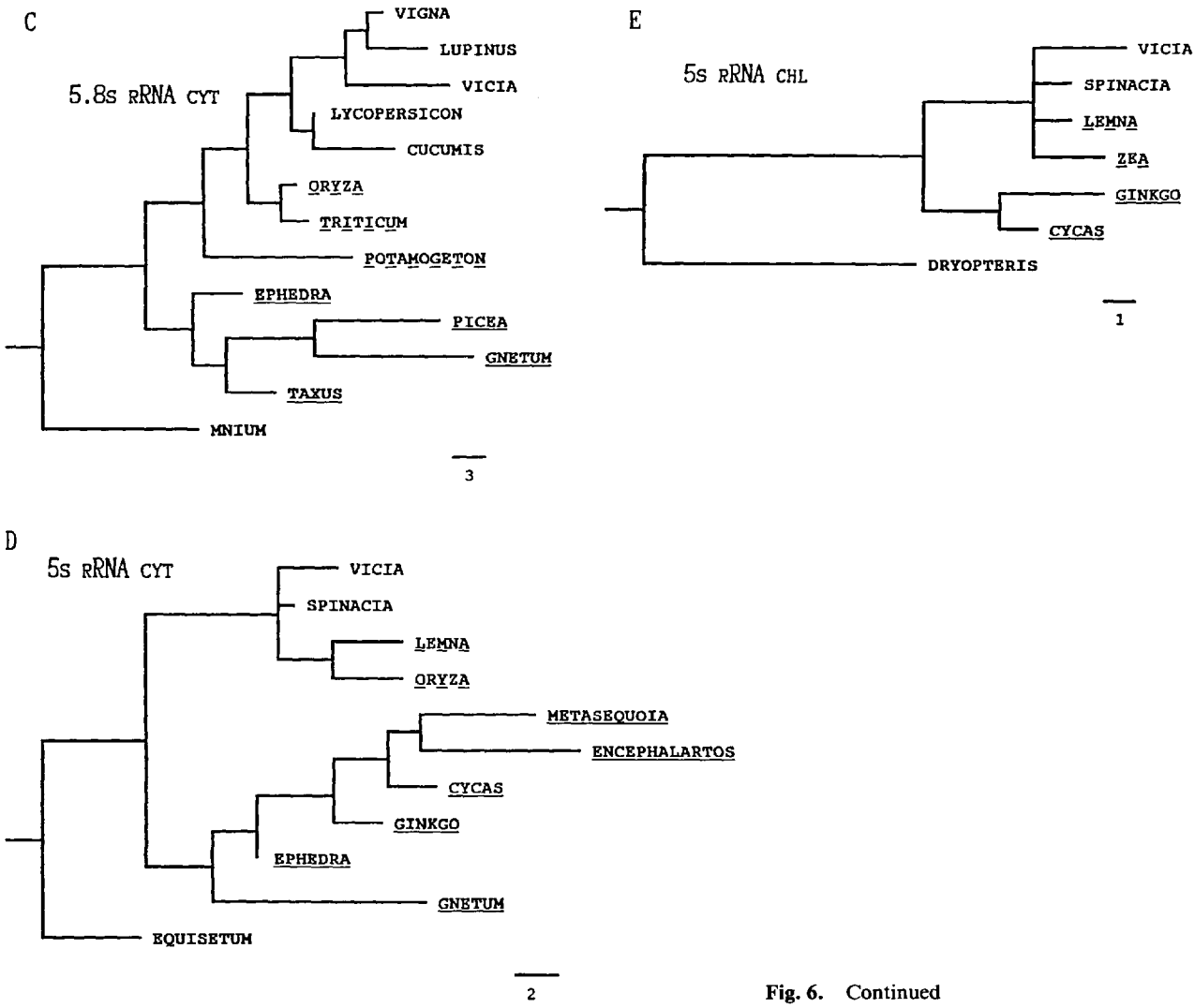


Fig. 6. Continued

In any case, the molecular data at our disposal show that none of the gymnosperm lineages could have been an ancestral one for angiosperms. We have suggested therefore that ancestral forms of angiosperms ought to be searched among Progymnospermopsida (Rakhimova et al. 1989). This conclusion may explain why attempts to deduce angiosperms from gymnosperms have failed, even though nearly all major groups of gymnosperms were considered as putative ancestors for angiosperms (Doyle 1978). So the genealogical splitting of angiosperm and gymnosperm lineages occurred long before the formation of the characteristic sets of morphological characters of these two groups. It should be noted that in the unrooted tree built from the partial ribulose bisphosphate decarboxylase sequences, gymnosperms also formed a distinct cluster separated from angiosperms (Martin and Dowd 1986).

Genealogical relationships among gymnosperm taxa cannot be deduced unambiguously on the basis of the available rRNA data (see Fig. 1). The only inference may be that the taxon Gnetopsida, including Gnetales and Ephedrales, is an artificial one, and *Gnetum* and *Ephedra* belong to quite different lineages of gymnosperms. This contradicts widely adopted schemes, specifically those inferred recently from the cladistic analyses of morphological traits of extinct and extant plants; but some authors have come to similar conclusions on the basis of traditional evidence (see Meyen 1984; Crane 1985; Doyle and Donoghue 1987; Beck 1988; Krassilov 1989).

Recently Martin et al. (1989), proceeding from the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene sequences of nine flowering plant species, postulated that the divergence of angiosperms occurred much earlier than is generally accepted, i.e., 150 Myr ago. These authors built a tree for angiosperm sequences only and placed the root between the two monocots, maize and barley, and seven dicots. According to the authors, this was the earliest branching event, which occurred about 320

Myr ago. Their estimates are based on the assumption of the constancy of molecular evolution rates in different lineages. Criticized by Goodman (1981), Antonov and Troitsky (1986), Britten (1986), and Gillespie (1986), this hypothesis cannot be taken for granted. In particular, our data in Fig. 1 indicate that the rate of rRNA evolution in plants may not be equal in different phylogenetic lineages. According to the data of Martin et al. (1989), the rate of GAPDH evolution in different eucaryotes may vary twofold. The observed nonconstancy of the rates is quite sufficient for shifting the position of nodes on the tree considerably. For more correct estimates of the tree topology, an outgroup of angiospermous plants is badly needed. Such outgroups have been used by Wolfe et al. (1989) in their analysis of chloroplast DNA sequence data, and it only confirmed the inequality of the rates of molecular evolution in different lineages and shifted the time of monocot–dicot divergence up to 200 Myr ago. Both Martin et al. (1989) and Wolfe et al. (1989) analyzed too few species. Moreover, the only monocots in their considerations were domesticated cereals. In our analysis we used the necessary outgroups, tried several kinds of rRNA from different cellular compartments, included far more species, mostly from wild flora, avoided the moot molecular clock hypothesis, and operated only with the paleobotanical datings on the appearance of first gymnosperms. Moreover, there are some grounds to believe that in general the rRNA data are more informative in plant phylogeny reconstruction than the protein data (Archie 1989).

The other point of interest in the rRNA dendrograms is the relative position of monocots and dicots. It is widely accepted now that the ancient angiosperms were close to extant Magnoliales from which other dicotyledonous groups, as well as Monocotyledones, arose. According to some of the trees in Fig. 6, extant monocotyledons are a paraphyletic group located near the root of the angiosperm branch. The other trees at least do not contradict such an evolutionary pattern. We may conclude that either monocotyledonous characters arose independently more than once in different groups of ancient Magnoliales or that monocotyledons rather than dicotyledonous Magnoliales were the earliest angiosperms. The latter suggestion seems more plausible by virtue of its greater parsimony and gives fresh impetus to the hypothesis of a monocotyledonous origin of angiosperms (Burger 1981). It is worth mentioning that no definite proof of the ranalean hypothesis (Takhtajan 1969), either neontological or paleontological, exists. On the other hand, available data indicate the early diversification of monocotyledons. The fossil leaves and pollen with characters similar to extant monocotyledons have been found in the earliest strata of the Potomac group (Beck 1976; Doyle 1978; Dahlgren and Rasmussen 1983; Krassilov 1989).

Judging by the trees in Fig. 6A and B, containing the greatest number of monocotyledon species, Magnoliales are the most ancient group among dicotyledons. The most ancient lineage among monocotyledons leads to Liliaceae.

When this paper was ready for publication, we learned about the study of Zimmer et al. (1989) concerning an attempt to reconstruct flowering plant evolution from an analysis of 18S and 26S rRNA partial sequences from 39 species. Although the conclusions from this work agree with ours in the early appearance of monocotyledons in the evolution of angiosperms (Bobrova et al. 1987; Troitsky et al. 1989b; Rakhimova et al. 1989), they differ with respect to relationships between gymnosperms and angiosperms. Zimmer et al. (1989) are inclined to think that Gnetales may be the sister group of angiosperms, even though they consider this suggestion not to be fully proven, as the branching order of gymnosperm taxa cannot be deduced unambiguously.

Finally we believe that the future progress of plant phylogenetics will depend not only on paleo- and neobotanical, but on molecular evidence as well. Yet we would stress that due to some discrepancies in the trees for different molecular species more data on various molecules from a larger set of plant species are needed in order to infer the pattern of seed plant molecular evolution.

Acknowledgments. We are indebted to I.N. Klikunova for her help in computing the sequencing results by the compatibility method.

References

Antonov AS, Troitsky AV (1986) The results of investigation of evolution of plant rRNA evoke doubt in universal fitness of the "molecular clock" hypothesis. J Evol Biochem Physiol 22:343–350 [in Russian]

Archie JW (1989) Phylogenies of plant families: a demonstration of phylogenetic randomness in DNA sequence data derived from proteins. Evolution 43:1796–1800.

Axelrod DI (1970) Mesozoic paleogeography and early angiosperm history. Bot Rev 36:277–319

Beck CB (ed) (1976) Origin and early evolution of angiosperms. Columbia University Press, New York

Beck CB (ed) (1988) Origin and evolution of gymnosperms. Columbia University Press, New York

Bobrova VK, Troitsky AV, Ponomarev AG, Antonov AS (1987) Low-molecular-weight rRNAs sequences and plant phylogeny reconstruction: nucleotide sequences of chloroplast 4.5S rRNAs from *Acorus calamus* (Araceae) and *Ligularia calthifolia* (Asteraceae). Plant Syst Evol 156:13–27

Britten RJ (1986) Rates of DNA sequence evolution differ between taxonomic groups. Science 231:39–44

Burger WC (1981) Heresy revived: the monocot theory of angiosperm origin. Evol Theory 5:189–225

Cheng Z-Q, Zhang H, Li G-Y (1986) The nucleotide sequences of chloroplast 4.5S rRNAs from four species of plants, celery (*Apium graveoleus*), barley (*Hordeum vulgare*), Chinese chive (*Allium tuberosum*), and dayflower (*Commelina communis*). FEBS Lett 200:193–196

Crane PR (1985) Phylogenetic analysis of seed plants and the origin of angiosperms. Ann Mo Bot Gard 72:716–793

Dahlgren R, Rasmussen FN (1983) Monocotyledon evolution. Characters and phylogenetic estimation. Evol Biol 16:255–395

Dams E, Hendriks L, Van de Peer Y, Neefs J-M, Smits G, Vandenbempt I, De Wachter R (1988) Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res 16S: r87–r173

Doyle JA (1978) Origin of angiosperms. Annu Rev Ecol Syst 9:369–392

Doyle JA, Donoghue MJ (1987) The importance of fossils in elucidating seed plant phylogeny and macroevolution. Rev Palaeobot Palynol 50:63–95

Erdmann VA, Wolters J (1986) Collection of published 5S, 5.8S and 4.5S ribosomal RNA sequences. Nucleic Acids Res 14S: r1–r59

Estabrook GF (1983) The causes of character incompatibility. In: Felsenstein J (ed) Numerical taxonomy. NATO ASI Series, vol G1. Springer-Verlag, Berlin, pp 279–295

Estabrook GF, McMorris FR (1980) When is one estimate of evolutionary relationships a refinement of another? J Math Biol 10:367–373

Felsenstein J (1989) PHYLIP—phylogeny inference package, Version 3.2

Fitch WM (1971) Toward defining the course of evolution: minimum change for specific tree topology. Syst Zool 20:406–416

Gillespie JH (1986) Variability of evolutionary rates of DNA. Genetics 113:1077–1091

Goodman M (1981) Decoding the pattern of protein evolution. Progr Biophys Mol Biol 37:105–164

Kavanagh TA, Timmis JN (1988) Structure of melon rDNA and nucleotide sequence of the 17–25S spacer region. Theor Appl Genet 76:673–680

Kiss T, Kis M, Abel S, Solymosy F (1988) Nucleotide sequence of the 17S–25S spacer region from tomato rDNA. Nucleic Acids Res 16:7179

Krassilov VA (1989) Origin and early evolution of flowering plants. Nauka, Moscow [in Russian]

Le Quesne WJ (1983) The uniquely derived concept as a basis for character compatibility analyses. In: Felsenstein J (ed) Numerical taxonomy. NATO ASI Series, vol G1. Springer-Verlag, Berlin, pp 296–303

Liang HQ, Bernard M, Bachellerie J-P (1983) Improved method for structure probing in large RNAs: a rapid heterologous sequencing approach is coupled to the direct mapping of nuclease accessible sites. Application of eukaryotic 28S rRNA. Nucleic Acids Res 11:5903–5920

Martin PG, Dowd JM (1986) A phylogenetic tree for some monocotyledons and gymnosperms derived from protein sequences. Taxon 35:469–475

Martin PG, Boulter D, Penny D (1985) Angiosperm phylogeny studied using sequences of five macromolecules. Taxon 34: 393–400

Martin W, Gierl A, Saedler H (1989) Molecular evidence for pre-Cretaceous angiosperm origin. Nature 339:46–48

Melekhovets YuF, Troitsky AV (1990) Comparative analysis of 5.8S rRNA from *Ephedra kokanica* Regl. (Gymnospermae) and other plant species. Biochim Biophys Acta 1048:294–296

Melekhovets YuF, Troitsky AV, Valiejo-Roman KM, Bobrova VK, Antonov AS (1988) Nucleotide sequences of cytosolic 5S ribosomal RNAs from two gymnosperms, *Gnetum gnemon* and *Ephedra kokanica*. Nucleic Acids Res 16:4155

Meyen SV (1984) Basic features of gymnosperm systematics and phylogeny as evidenced by the fossil record. Bot Rev 50: 1–111

Nairn CJ, Ferl RJ (1988) The complete nucleotide sequence of the small-subunit ribosomal RNA coding region for the cycad *Zamia pumila*: phylogenetic implications. J Mol Evol 27: 133–141

Nazar RN, McDougall J, Van Ryk DI (1987) Structure and evolution of the 4.5–5S ribosomal RNA intergenic region from *Glycine max* (soya bean). Nucleic Acids Res 15:7593–7603

Omelyanchuk LV, Kolchanov NA (1985) Algorithm for additive tree building from sets of homologous sequences. Veracity of phylogenetic reconstructions. In: Algorithmic analysis of structural information (Computer Systems), no 112. IM SO AN SSSR Novosibirsk, pp 46–55 [in Russian]

Peattie DA (1979) Direct chemical method of sequencing RNA. Proc Natl Acad Sci USA 76:1760–1764

Rakhimova GM, Troitsky AV, Klikunova IN, Antonov AS (1989) Phylogenetic analysis of 18S rRNA partial sequences from 14 higher plant species. Mol Biol 23:830–842 [in Russian]

Schiebel K, Hemleben V (1989) Nucleotide sequence of the 18S–25S spacer region from rDNA of mung bean. Nucleic Acids Res 17:2852

Schuster W, Brennicke A (1987) Plastid DNA in the mitochondrial genome of *Oenothera*: intra- and interorganelle rearrangements involving part of the plastid ribosomal cistron. Mol Gen Genet 210:44–51

Takhtajan AL (1969) Flowering plants: origin and dispersal. Smithsonian, Washington, DC

Troitsky AV, Bobrova VK, Ponomarev AG, Antonov AS (1984) The nucleotide sequence of chloroplast 4.5S rRNA from *Mnium rugicum* (Bryophyta): mosses also possess this type of RNA. FEBS Lett 176:105–109

Troitsky AV, Bobrova VK, Melekhovets YuF, Valiejo-Roman KM (1989a) The nucleotide sequence of 5.8S rRNA from the moss *Mnium rugicum* Laur. Nucleic Acids Res 17:459

Troitsky AV, Valiejo-Roman KM, Melekhovets YuF, Bobrova VK, Omelyanchuk LV, Antonov AS (1989b) Higher plant phylogeny reconstruction based on chloroplast 4.5S and 5S rRNAs. Nucleotide sequences in 4.5S rRNAs from five plant species. J Evol Biochem Physiol 25:167–175 [in Russian]

Wolfe KH, Gouy M, Yang Y-W, Sharp PM, Li W-H (1989) Date of the monocot–dicot divergence estimated from chloroplast DNA sequence data. Proc Natl Acad Sci USA 86: 6201–6205

Wolters J, Erdmann VA (1988) Compilation of 5S rRNA and 5S rRNA gene sequences. Nucleic Acids Res 16S:r1–r70

Zharkikh AA, Rzhetsky AYu, Morosov PS, Sitnikova TL, Krushkal JS, Matushkin YuG (1990) VOSTORG: package of microcomputer programs of phylogenetic analysis. In: Modelling and computer methods in molecular biology and genetics. Abstracts of the International Conference, Novosibirsk, pp 217–218

Zhou X, Liu W, Wang M (1988) Comparative study on the evolution of chloroplast ribosomal 5S RNA of a living fossil plant, *Cycas revoluta* Thunb. FEBS Lett 235:30–34

Zimmer EA, Hamby RK, Arnold ML, Leblanc DA, Theriot EC (1989) Ribosomal RNA phylogenies and flowering plant evolution. In: Fernholn B, Bremer K, Jörnvall H (eds) The hierarchy of life. Elsevier Science Publishers BV (Biomedical Division), Amsterdam, pp 205–214

Received June 1, 1990/Revised October 15, 1990