Phylogeny of six naviculoid diatoms based on 18S rDNA sequences

Bánk Beszteri,1 Éva Ács,1 Judit Makk,1 Gábor Kovács,1 Károly Márialigeti1 and Keve T. Kiss2

Author for correspondence: Bánk Beszteri. Tel: +36 266 1148. Fax: +36 266 1148. e-mail: bbank@ims.u-szeged.hu

18S rDNA sequences of six Naviculaceae species [Amphora montana, Gomphonema parvulum, Eolimna minima (syn. Navicula minima), Eolimna subminuscule (syn. Navicula subminuscule), Navicula veneta and Phaeodactyli um tricornutum] were determined in order to assess the monophyly of this important group of diatoms, to date not included in 18S rDNA databases, and also that of the recently described genus Eolimna. Phylogenetic trees were constructed using other known diatom 18S rDNA sequences, and best tree topologies obtained were tested against alternative trees for their reliability. The analyses do not reject the monophyly of Naviculaceae and strongly support the separation of the genus Eolimna from Navicula sensu lato. The two species of Eolimna, however, do not appear to be each other's closest relatives among the species investigated: rather, E. subminuscule shows affinities to A. montana, and E. minima to P. tricornutum. A. montana, a species which has been proposed should be transferred into a separate taxon from the other five species, was found to have grouped well within them in all analyses.

Keywords: Bacillariophyceae, Naviculaceae, 18S rDNA, phylogeny

INTRODUCTION

Investigations into diatom phylogeny and taxonomy were based almost exclusively on morphological features of the silica frustules (and to a much smaller extent on structural features of the living cells) until the end of the 1980s, when molecular phylogenetic investigations also started. Research in this area has mainly been carried out by Medlin and co-workers (e.g. Medlin et al., 1995, 1996, 1997a, b, Kooistra & Medlin, 1996) using 18S rDNA sequences from a broad variety of diatom species, clarifying basic outlines of diatom molecular phylogeny. As a result of these studies, 18S rDNA sequences are available from representatives of most major diatom lineages. It was found through them that the traditional division of diatoms into centric and pennate (and within the latter into rapid and araphid) forms is not a true phylogenetic grouping since centrics form a paraphyletic lineage, and so do araphid pennates. In some other aspects, however, this molecular phylogeny supported the traditional classification of diatoms: both pennates and the family Bacillariaceae (and thus rapid and pennate, represented only by the Bacillariaceae species in these studies) appeared as monophyletic groups. Sorhannus et al. (1995) determined and analysed partial 28S rDNA sequences from a smaller selection of species and their results were congruent with those of the above authors. Thus the molecular data seem to contradict the proposition of Kociolek & Stoermer (1989) based on chromosome counts, suggesting an early divergence of diatoms into centrics and pennates, and also that of the pennates into rapid and araphid groups (e.g. Medlin, 1997a). However, a large, ecologically important and taxonomically interesting group of diatoms, the family Naviculaceae, has so far not been included in these molecular phylogenetic investigations.

We determined 18S rDNA sequences of six Naviculaceae species: Amphora montana, Gomphonema parvulum sensu lato, Eolimna minima (syn. Navicula minima), Eolimna subminuscule (syn. Navicula subminuscule), Navicula veneta and Phaeodactyli um...
tricormatum and present the molecular relationships of these with other pennate diatoms.

The family Naviculaceae in the sense used here includes biraphid pennate diatoms with median raphe (see Krammer & Lange-Bertalot, 1986). The taxonomy of this group has been changing constantly in the last decade; the whole family, as well as many of its genera, have undergone revisions based on morphology (e.g., Round et al., 1990). The genus Navicula, for instance, has been fundamentally revised. One of the new genera described for members of Navicula sensu lato is Eolimna (Schiller & Lange-Bertalot, 1991). Another taxonomic innovation of Round et al. (1990) concerning Naviculaceae sensu Krammer & Lange-Bertalot (1986) is that the family has been raised to the level of order and split into two groups: the orders Naviculales and Thalassiosirales. One of the species investigated in this study, A. montana, has been placed within the Thalassiosirales, and our other species within the Naviculales.

In this study we try to answer the following questions. (1) Is the family Naviculaceae a monophyletic group? (2) Is the order Naviculales sensu Round et al. (1990) a monophyletic group; do our molecular data support the separation of the order Thalassiosirales from the Naviculales? (3) Do these sequences support the separation of Eolimna from Navicula and is Eolimna monophyletic based on the (certainly limited) information available from the 18S rDNA sequences of our two Eolimna species? In other words, are the morphological features supposed to be specific to Eolimna phylogenetically informative, or do they appear to be homoplasies instead? (4) Does the family Bacillariaceae remain a monophyletic group when their presumably close relatives, Naviculaceae species, are also included in the analyses? (5) Do raphid pennates, including the Naviculaceae besides the Bacillariaceae, still form a monophyletic group? To try to answer these questions, we searched for best trees under different optimality criteria, and also tested alternative phylogenies against the best trees obtained.

METHODS

Isolation, purification and cultivation of strains. Our strains were isolated from epilithon samples from the stream Laskó (A. veneta) and the river Danube (except P. tricormatum) by plating onto agar-solidified diatom medium (Hughes, Gottham & Zehnder's medium as modified by Allen; see Hold & Wyman, 1978). The cultures were separated from eukaryotic contaminants by streaking and resolution, and the strains were maintained in liquid medium with the same composition as the medium used for their isolation. In addition to the strains isolated in our laboratory, the strain UTEX 640 of P. tricormatum was used. The strains are maintained in the collection of the Department of Microbiology, Eötvös Loránd University.

Identification of strains. Strains were treated with 30% H2O2 to clean the frustules, and identified by transmission and scanning electron microscopy based on the identification book of Krammer & Lange-Bertalot (1986).

DNA extraction and amplification of 18S rDNA. Cells were harvested by centrifugation, frozen in liquid nitrogen and ground in a mortar. The ground cells were suspended in 500 μl lysis buffer (150 mM NaCl, 10 mM Na3EDTA, pH 8), 10 μl proteinase K (10 mg ml-1, Merck) and 20 μl 25% SDS was added and samples were incubated at 65°C for 45 min. DNA was extracted with equal volumes of phenol and chloroform, and purified with a Prep-A-Gene DNA Purification Kit (Bio-Rad). Amplification of the 18S rDNA was carried out by standard PCR protocols using the eukaryotic-specific primers described by Medlin (1990) without the polylinker sites.

Sequencing of PCR products. PCR products were purified with a Prep-A-Gene DNA Purification Kit (Bio-Rad) and directly sequenced with an ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit with AmpliTaq FS DNA Polymerase FS (Applied Biosystems), and purified according to the instructions of the manufacturer. The product was then electrophoresed with an ABI Prism 377 Genetic Analyzer (Applied Biosystems). For the sequencing reactions (Sandra & Drostel, 1992), the primers 18S/445r (5'-RCGRSCCTGTGCTGCTTCTCCG-3') and 18S/997r (5'-AAAAACATCTTGGWARTGC-3') designed in our laboratory, were used.

Phylogenetic analysis. The sequences were aligned with previously published diatom 18S rDNA sequences (e.g., Manhart et al., 1995; Medlin et al., 1996) using the ARB program package (Strunk et al., 1995). The alignments were corrected manually, considering primary and secondary structural similarity.

Maximum-likelihood (ML) analyses were performed using the fastDNAnal program (version 1.1.1; Olsen et al., 1994), with empirical base frequencies, a transition-transversion ratio of 2:1, random addition of taxa and global rearrangements.

For unweighted maximum-parsimony analyses, the PAUP program (version 4.0b4a; Swoford, 1998) was used. Informative characters were treated as multistate unordered, and gaps were treated as missing data. Phylogenetic trees were obtained using the tree-bisection-reconnection branch swapping option and random addition of the taxa in a heuristic search in 10 sequence addition replicates. For bootstrap analysis, multiple data sets were obtained by generating 500 replicates, analyzed with DNA pars with random taxa addition, and a consensus bootstrap tree was produced by CONSENSE (Felsenstein, 1985).

Parsimony analysis with Gtccess's implied weights method was performed using PAUP with the default option K = 2 (with random addition of taxa, heuristic search). For bootstrap analysis with the above settings in 100 replicates, PAUP was used.

User-defined tree topologies (prepared with REFREE, PHYLIP) were compared to the ML tree with the Kishino-Hasegawa test (Kishino & Hasegawa, 1989; Swoford et al., 1996) using fastDNAnal to assess how much stronger the sequence data supported the best tree than alternative trees.

RESULTS AND DISCUSSION

Phylogenetic relationships of the Naviculaceae

The electron microscopic identification of the strains was straightforward except for the closer identification of the G. parvulum isolate: because of its small size it
could not be placed unambiguously into either of the subgroups of this morphological group with recently constantly changing taxonomy. In this paper we refer to this strain simply as G. parvulum.

18S rDNA sequences ranged from 1744 to 1754 nucleotides in length excluding amplification primers, except for A. montana which was only 1729 nucleotides long: loops 10 and 11 (following the nomenclature of De Rijk et al., 1992) in this molecule are 6 and 9 nucleotides shorter than in most diatom species, respectively. Alignment of the sequences using primary as well as secondary structures was relatively clearcut, except for the hypervariable regions of the gene (e.g. loop E21-1), which were excluded from subsequent analyses. Thus 1702 positions, containing 236 informative sites for parsimony analyses, were used for construction of phylogenetic trees.

Dissimilarity of the two Eolimna sequences was surprisingly high considering the presumed close relationship between them: 75%, while, for example, dissimilarity of the E. minima sequence to the P. tricornutum was only 54%.

The best tree obtained under the ML criterion is presented in Fig. 1. When fastDNAml was run with the K (keep multiple trees) option, three different topologies were encountered with respect to the Naviculaceae species among the trees found with almost as high probability as that of the ML tree. In one of these (tree A, see Fig. 2 and below) Naviculaceae and Bacillariaceae were monophyletic sister groups, together forming the (equally monophyletic) group of raphid pennates. In the other trees G. parvulum was found at the base of the remaining raphid pennates, just as in the ML tree. The branching order of the four Naviculaceae species besides G. parvulum (described above) and N. veneta (the position of which did not vary among trees) was either identical to that in the ML tree or E. minima was connected to the branch leading to A. montana and E. subminucula instead of P. tricornutum. The likelihoods of these trees differed minimally and the differences were not significant (see also below).

In the unweighted maximum-parsimony analysis one most parsimonious tree (MPT) was found which differed from the ML tree only in the branching order within the Bacillariaceae and that of the four above-mentioned Naviculaceae species, which were again grouped as described at the end of the last paragraph. Four trees required only one more step than the MPT. In three of these, the branching order of the Naviculaceae species was identical to that in the MPT. In
the fourth one, the branching order of the Naviculaceae was identical to that in the ML tree.

In the tree obtained with the Goloboff implied weights method, *G. parvulum* was again a sister group to all other raphid pennates (as in the ML tree and the MP1 of the unweighted analysis); the branching order of the five other members of the Naviculaceae was also identical to that in the MP1.

**Alternative phylogenies**

As seen from these results, the trees obtained under different optimality criteria all agreed in showing monophyly of the sequences from raphid pennates and from the class Bacillariaceae, and paraphyly of those from the order Naviculales sensu Round et al. (1990), and in not showing the two *Eolinna* species to be the closest relatives either of each other or of *N. veneta*. The answer they give to the question regarding the monophyly of Naviculaceae, however, is hazier. The objective of the following comparisons was to further test the reliability of these results. The alternative topologies were obtained starting from tree A (Fig. 2). This tree shows monophyly of the Naviculaceae, with a likelihood that is hardly smaller than that of the ML tree. This likelihood difference does not raise serious doubts about the monophyly of Naviculaceae as hypothesized based on morphology. We also noticed that minor changes in the selection of species included in the analyses could result in the topology of tree A having a higher likelihood than the topology presented in Fig. 1. In conclusion, the sequence data available do not make it possible to decide which topology is closer to the real branching order of species, and until further data accumulate which question the monophyly of Naviculaceae we consider this group as monophyletic.

The following alternative topologies were compared to the ML topology (the part of tree A differing from the ML tree is presented in Fig. 2, in the rest of the cases, differences from this tree are described). (A) *G. parvulum* was transferred to form the first diverging branch within the clade of Naviculaceae (Fig. 2); this tree was also found in the ML analysis, as mentioned above, and its likelihood differed from that of the ML tree to such a small extent that this difference is not significant. This tree is included in this description and in Table 1 to show this difference numerically. (B) The branch leading to *Fragilaria envirens* and *Thalassionema nitzschioides* was moved "up" in the tree, to the base of Naviculaceae (thus testing for the monophyly of the fragment of the tree).
of raphids). (C) The branch leading to Bacillaria
parviflora and the two Peridinium endosymbionts was
transferred to the root of Naviculaceae (testing for the
monophyly of Bacillariaceae). (D) E. minutum was
transferred to be the closest relative of E. subminutula
(testing for the monophyly of the genus Eolimna). (E)
As in the case of tree D, but also N. veneta was
transferred to the base of the Eolimnas, testing for the
monophyly of the old genus Navicula, also containing
our two Eolimna species. (F) A. montana and G.
parvulatum were exchanged to test for the monophyly of
Naviculales sensu Round et al. (1990).

The results of the Kishino–Hasegawa tests for these
trees against the ML tree are shown in Table 1. The
trees B, F were also compared to tree A, and the results
were identical to those seen in Table 1 with respect to
the significance of the differences. As seen there, trees
B, F and A are significantly worse than the ML tree,
while the others are not. The log likelihood of tree A is
almost as high as that of the ML tree, as noted above.
These results further reinforce the assumptions of the
monophyly of raphids, of the paraphyly of Naviculales
sensu Round et al. (1990) and of the paraphyly of the
old genus Navicula as well. They also show that the
sequence data contradict the paraphyly of the
Bacillariaceae, and the monophyly of the genus
Eolimna to a lesser extent.

Conclusions

The taxonomy of naviiculoid diatoms has been sub-
ject to constant change in the last decade. Many of
these changes were proposed by Round et al. (1990),
including the restriction of the genus Navicula to the
members of the section Lineolatae of Navicula sensu
lato. This genus formerly contained a heterogeneous
selection of bilaterally symmetric pennates that were
not classified according to a special distinctive mor-
phological feature. Now, it only contains the members
of the former Lineolatae, and new genera have been
described for members of the other sections (e.g.
Selliaphora, Fallacia, Lyrella, Petronetes, etc.) based on
differences in the fine structure of the frustule, and to
a much smaller extent on features of the living cell. One
of these new genera is Eolimna, described on the basis
of the fossil species Eolimna martini (Schiller & Lange-
Bertalot, 1997). Distinctive features of this new genus
include cell size, structure of the areolae and the
perforated valvocopula. The two Eolimna species in
our present study were formerly placed in the genus
Navicula sect. Minimaculata (Krammer & Lange-
Bertalot, 1986; Moser et al., 1998).

Another new feature in the classification of Round
et al. (1990) was the division of the group of naviiculoid
diatoms into the orders Thalassiophysales and Navi-
culales. The former contains the genera Catenula,
Amphora, Undabiella and Thalassiophysa, while the
remaining naviiculoid genera were placed in the latter.

In conclusion, the taxonomy of naviiculoid diatoms is
far from being resolved. Some of the questions raised
in the Introduction clearly require more molecular
data in order to be reliably answered. However, our
analyses gave reasonably unambiguous answers to
some of the questions: the obtained ML and
maximum-parsimony trees, as well as the comparison
of tree B to the ML tree, strongly support the notion
that the group of raphid pennates (now including the
Naviculaceae as well as the Bacillariaceae) form a
monophyletic group. Our results also strongly support
the paraphyly of the old genus Navicula (and thus the
separation of the two Eolimna species from this genus)
and of the order Naviculales sensu Round et al. (1990).
Considering the former, we refer to best trees under all
of the optimality criteria used in this study, the
bootstrap values and the result of the Kishino–
Hasegawa test of tree E against the ML tree. Con-
cerning the latter, A. montana, the species in our
analyses that was moved into the order Thalassio-
physales by Round et al. (1990), was found well within
the rest of our species (all of which would be placed in
the order Naviculales in this classification) in all
analyses and with high bootstrap support, and the
transfer of the branch of A. montana to the base of the
clade of Naviculales (tree F; this tree topology would
be expected if the order Naviculales was a mono-
phyletic group) resulted in a significantly worse tree
than the ML tree.

The family Bacillariaceae appeared as a monophyletic
group in all analyses, and the Goloboff bootstrap
value of the clade of the Bacillariaceae also supports
their monophyly. Though in the Kishino Hasegawa
test tree C was not significantly worse than the ML
tree, this fact in itself should not raise serious doubts
about the monophyly of this group.

Although in the ML tree Naviculaceae sensu Krammer
& Lange-Bertalot (1986) appeared as a paraphyletic
group, our results only slightly contradict the
monophyly of the Naviculaceae. Reasons for thinking
so are the following: the minimal (far from significant)
difference in the likelihoods of the ML tree (Fig. 1) and
tree A (Fig. 2), the fact that branch lengths of
Naviculaceae are longer than those of Bacillariaceae
(indicating a higher rate of substitution within this
lineage that can result in the larger evolutionary
distance of G. parvulatum from the other naviiculoid
species), and, especially, the fact that the branch of G.
parvulatum was found within the naviiculoid clade, for
instance when N. veneta was not included in the
analysis. In conclusion we consider that further data
could possibly reinforce or reject the hypothesis about
Naviculaceae being a monophyletic group; however, it
is present the reasons for doubting it are rather weak.
From this point of view, further 18S rDNA sequences
from close relatives of G. parvulatum could be especially
interesting.

The genus Eolimna quite unequivocally seems to be
paraphyletic from these analyses. As statistical support
for this we point to the bootstrap value of the node
connecting A. montana and E. subminutula. Though
the Kishino–Hasegawa test did not indicate that tree E
is significantly worse than the ML tree, apart from bootstrap values, the fact that the two *Eolimna* species were not grouped together in any of the analyses performed raises serious doubts about the monophyly of *Eolimna*. However, this question requires further examination; to be able to really assess the molecular relationships of the species of this genus, 18S rDNA sequences from more of them would be necessary.

**ACKNOWLEDGEMENTS**

We would like to thank Dr Linda Medlin for her corrections, advice and practical help, István Mikk for his valuable help in sequence analysis and Frigó Klement for her linguistic help. This study was supported by grant no. AKP 98-793-3,3 of the Hungarian Academy of Sciences and grant FKFP-0154/2000 of the Ministry of Education.

**REFERENCES**


