Evolution of the Diatoms (Bacillariophyta). II. Nuclear-Encoded Small-Subunit rRNA Sequence Comparisons Confirm a Paraphyletic Origin for the Centric Diatoms

L. K. Medlin, W. H. C. F. Kooistra, R. Gersonde, and U. Wellbrock

Alfred Wegener Institute for Polar and Marine Research

A phylogeny of the diatoms was inferred from comparisons of nuclear-encoded small subunit ribosomal RNA coding regions using maximum likelihood, weighted maximum parsimony, and neighbor-joining distance methods with Jukes and Cantor, Kimura, Gamma, van de Peer, and LogDet evolutionary models. Analyses of 30 taxa in 11 orders recovered two clades (Clades I and II). Neither of these clades correspond to the three classes of diatoms presently recognized or to the traditionally recognized radially symmetrical centric diatoms or bilaterally symmetrical pennate diatoms. All analyses show that the centric diatoms are a paraphyletic lineage. Tests of alternative phylogenies that address existing hypotheses regarding diatom systematics with the maximum likelihood and maximum parsimony methods support the two clades. Clade I is defined by centric diatom orders with specialized tubes, termed labiate processes, located peripherally in the cell wall. Clade II contains (1) bi(multi)polar centric diatoms with centrally located labiate processes, (2) centric diatoms with other central tubes termed strutted processes, and (3) pennate diatoms. Morphological evidence from fossil assemblages and cytological architecture support the results of the molecular analyses, whereas morphological features of extant diatoms are too derived to resolve the deeper branches in the tree.

Introduction

The diatoms (Bacillariophyta) are one of the most successful groups of microalgae, with over 10,000 known species occurring in both aquatic and terrestrial habitats. They are renown for the intricate and precise geometric designs of their silica cell walls (valves and girdle bands) and the unusual pattern of reduction in cell size of one of the daughter cells following mitosis (Mann and Marchant 1989). Small-subunit ribosomal RNA (ssu rRNA) sequence comparisons (Bhattacharya et al. 1992; Medlin, Williams, and Sims 1993) support a monophyletic origin of the diatoms within the pigmented protistan heterokont lineage (those organisms bearing tubular mastigonemes or hairs on their leading flagellum). Recent rRNA analyses have shown the diatoms to be related specifically to the Pelagophyceae, i.e., heterokont algae with a reduced flagellar apparatus that lacks the central two microtubules (Saunders et al. 1995).

Central to diatom taxonomy and phylogeny has been an assumption that the group contains two forms based on mode of sexual reproduction, pattern centers or symmetry, and plastid number and structure (Round, Crawford, and Mann, 1990, pp. 58–100). Familiar to most aquatic and cell biologists are the oogamous, radially symmetrical centric diatoms with numerous discoid plastids, and the isogamous, bilaterally symmetrical pennate diatoms with fewer plate-like plastids. These characters have led most systematists to treat the centric and pennate diatoms as two separate classes. It is generally agreed that the pennate diatoms evolved from the centric forms due to their later appearance in the fossil record. In a recent systematic revision (Round, Crawford, and Mann 1990, p. 125), the raphid pennate diatoms (those with a slit [raphe] in the cell wall for movement) and the araphid pennate diatoms (those without this slit) were given equal taxonomic ranking. Thus, the present classification system of the diatoms recognizes three classes: Coscinodiscophyceae (centric diatoms), Fragilariophyceae (araphid pennate diatoms), and Bacillariophyceae (raphid pennate diatoms).

Classification of the diatoms at the order level and below is almost exclusively based on morphological features of the siliceous cell wall (Round, Crawford and Mann 1990, pp. 37–49). Key characters limiting centric diatom orders are valve structure and symmetry, and the type, number, and arrangement of tubes through the valve. These features, plus the presence or absence of the raphe and its structure, figure prominently in the subdivision of raphid and araphid taxa. The raphe and the tubes (labiate or strutted processes) through the cell wall are probably the most important features used to infer phylogenetic relationships among the diatoms. The labiate process normally has elaborate lip-like internal extensions, but some labiate processes are reduced, with only a simple slit-like internal opening. The function of the labiate process is not well established, although in a restricted number of centric and pennate taxa mucilage

Key words: diatoms, Bacillariophyta, molecular evolution, phylogeny, small-subunit ribosomal RNA.

Address for correspondence and reprints: L. K. Medlin, Alfred Wegener Institute for Polar and Marine Research, Postbox 120161, D-27515 Bremerhaven, Germany. E-mail: lmedlin@awi-bremerhaven.de.

Mol. Biol. Evol. 13(1):67–75. 1996

is secreted through the labiate process for movement (Medlin, Crawford, and Andersen 1986; Pickett-Heaps, Hill, and Weatherbee 1986; Pickett-Heaps, Hill, and Blazé 1991), and the internal part of the tube is used as a cytological anchor for the nucleus during interphase and new valve formation (Schmid 1994). It has been hypothesized that the labiate process evolved into the raphe of the pennate diatoms (Hasle 1974). The strutted process, through which chitan threads are secreted for chain formation and flotation, is restricted to the centric order Thalassiosirales and is a simple tube with adjacent pores (Round, Crawford, and Mann 1990, pp. 35–36).

Relationships at higher taxonomic levels in the diatoms are virtually unexplored with modern phylogenetic and morphometric tools. This likely reflects the unique, highly derived morphologies of the diatom orders. In this regard, RNA sequence comparison has proven to be a powerful tool for resolving phylogenetic relationships at all taxonomic levels because it is a functionally stable evolutionary marker that is independent of morphometric characters (Woese, 1987; Bhattacharya et al. 1992). It is functionally conserved in all cells with no documented evidence of lateral gene transfer (Woese 1987).

Using rRNA sequence comparisons, Medlin, Williams, and Sims (1993) presented the first evidence that suggested centric and araphid pennate diatoms had a paraphyletic origin. One centric diatom, *Skeletonema costatum*, was most closely related to the pennate diatoms, with high bootstrap support in phylogenetic analyses; the araphid taxa were a paraphyletic lineage within the pennates. An analysis of partial sequences from the 28S large-subunit (lsu) rRNA coding region from eight diatoms also showed that centric and araphid taxa were paraphyletic (Sörhannus et al. 1995). These preliminary data suggest that higher level diatom systematics do not reflect their evolutionary history.

To investigate the phylogenetic relationships among the diatoms at higher taxonomic levels we compared the nucleotide sequence of the nuclear-encoded ssu rRNA coding regions from 30 taxa representing 11 orders and compared these sequences to homologous coding regions from other heterokonts. Mapping of morphological characters onto the rRNA phylogenies allowed deeper insights into the origins of taxonomically important features of the diatoms.

Materials and Methods

Cultures

Representatives of each diatom class and at least one species from most major orders of centric diatoms were selected (Appendix). New isolates were grown as described (Medlin, Crawford, and Andersen 1986) and deposited in the Culture Centre for Algae and Protozoa (Oban, Scotland) or in the Antarctic algal collection at the Alfred Wegener Institute. Permanent slides of each taxon are available from L.K.M.

DNA Methods

Nucleic acids were extracted as described (Medlin et al. 1988) or with a 3% CTAB (hexadecyltrimethylammonium bromide) procedure (Doyle and Doyle 1990). Small-subunit rRNA coding regions were amplified using the polymerase chain reaction (PCR; Saiki et al. 1988) with eukaryote-specific primers containing multiple restriction sites for cloning (Medlin et al. 1988; Medlin, Lange, and Baumann 1994). Alternatively, ssu rRNA coding regions were amplified with biotin-labeled primers, and single-stranded templates were produced for sequencing (Dynabeads M-280 Streptavidin, DYNAL A.S. Oslo, Norway). Both coding and noncoding strands were completely sequenced (Sanger, Nicklen, and Coulsen 1977; Elwood, Olsen, and Sogin 1985). No fewer than six PCR reactions in each orientation were pooled for the direct sequencing reactions.

Sequence Analysis.

The ssu rRNA sequences of 16 centric diatoms were aligned with previously published diatom (Medlin et al. 1991; Bhattacharya et al. 1992; Medlin, Williams and Sims 1993) and other chromophyte/oomycete, dinoflagellate, and prymnesiophyte rRNA sequences (Neefs et al. 1991; Andersen et al. 1993; Medlin, Lange, and Baumann 1994), using maximum primary and secondary structural similarity. Positional homology was assumed for 1,739 positions, of which 528 were informative for parsimony analyses. The final data set contained 34 taxa (Appendix). Three oomycetes and one pelagophyte were used to root the trees. The use of multiple outgroups stabilizes the phylogenetic analyses (Swofford and Olsen 1990). The oomycetes, sister group to the pigmented heterokont lineage (Saunders et al. 1995), are not included in the figures for clarity. Complete phylogenies are available from L.K.M.

Maximum parsimony analyses were implemented with the PAUP computer program (Swofford 1993). Introduced gaps were treated as missing data; informative characters were treated as multistate unordered. Unweighted maximum parsimony trees were obtained using the tree-bisection-reconnection (TBR) branch swapping option in a heuristic search with random taxon addition. The most parsimonious trees (MPT) and the data matrix were loaded into the MacClade computer program (Maddison and Maddison 1992) to infer a weighted data set in which the frequency of nucleotide substitutions at each position was inversely related to the number of changes at that position (scale 1–100). The

type of substitution, i.e., purine → purine, purine → pyrimidine, (i, j), at each position was also weighted (K_{ij}) as the cost of going from one state to another: $K_{ij} = -\ln(X_{ij}/X_i)$ where X_{ij} is the number of i → j changes, X_i is the number of changes from i to any state, and X is the number of changes on the tree (scale 1–100). These weighting schemes greatly enhance the ability of the maximum parsimony analyses to recover the correct tree when multiple substitutions have occurred over sites (Hillis, Huelsenbeck, and Cunningham 1994). Stability of monophyletic groups in weighted maximum parsimony trees was estimated with a bootstrap analysis (100 replicates; Felsenstein 1985). The information content of this data set was measured with the G_1 statistic, which tests for the distribution skewness of 10,000 random trees (Hillis and Huelsenbeck 1992).

Distance analyses were performed using the PHYLIP computer program (Felsenstein 1993). Dissimilarity values (Fitch and Margoliash 1967), based on pairwise comparisons of sequences, were transformed into distances using both the Jukes and Cantor (1969) and the Kimura two-parameter models (Kimura 1980). Distance matrices were converted into trees using the neighbor-joining method (Felsenstein 1993); a bootstrap analysis (100 replicates; Felsenstein 1985) was also performed. Weighted distance analyses were performed (a) by using the weighting scheme for the frequency of base substitution from MacClade to generate a data set with each position weighted, (b) by using a weighting scheme in which the positions were grouped into variability classes and only four possible weights assigned to any position (van de Peer et al. 1993), and (c) by using the gamma distribution option (Jin and Nei 1990) in the MEGA computer program (Kumar, Tamura, and Nei 1993). Neighbor-joining methods were used to construct trees from each of the weighted distance matrices.

Maximum likelihood analyses were done using the fastDNAml program (version 1.0, Larsen et al. 1993). The maximum likelihood tree was used to construct user-defined trees (RETREE, PHYLIP) to constrain maximum parsimony (PAUP) and maximum likelihood (DNAML, PHYLIP) analyses. Three alternative hypotheses of relationships (trees) were compared with the likelihood ratio test (Kishino and Hasegawa 1989). This test uses the mean and variance of log-likelihood differences between trees, taken across sites. If the mean is >1.96 standard deviations different between any two trees, then these trees are declared significantly worse (Felsenstein 1993).

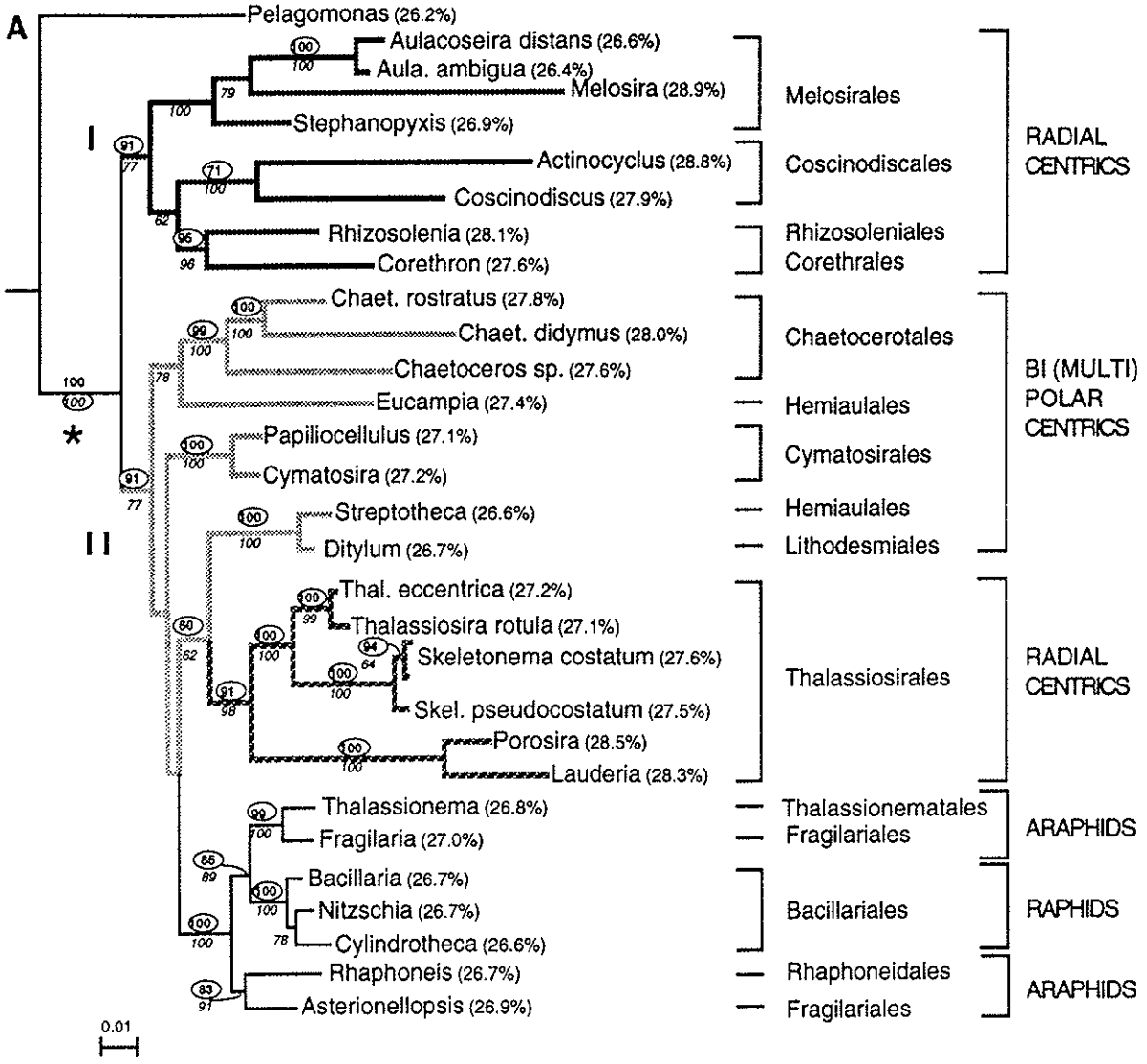
A concern regarding phylogenetic analyses is that relationships may reflect shared base compositional bias rather than a monophyletic origin. Therefore, a LogDet transformation analysis, which attempts to correct for base compositional bias (Lockhart et al. 1994), was also done to determine if this had any affect on the phylogenetic analyses. The percentage composition of the four nucleotides was calculated for each taxon.

Results

Of the 16 complete ssu rRNA sequences, 14 varied little from approximately 1,800 nucleotides (nt) in length, whereas two centric taxa, *Aulacoseira ambigua* and *Cymatosira belgica*, had insertions that increased the length of their ssu rRNA coding regions to 1,847 and 2,317 nt, respectively, including amplification primers (Medlin et al. 1988). *Cymatosira belgica* has three separate insertions of size 152, 199, and 122 nt. Comparison of these insertions with ssu rRNA Group I intron sequences from green algae showed no significant similarity. These insertions were excluded from the data analysis.

Results of maximum likelihood, neighbor-joining, and weighted maximum parsimony analyses of nuclear-encoded small subunit rRNA coding regions are summarized in figure 1. Bootstrap values (>50%) from weighted maximum parsimony and neighbor-joining analyses are placed at the internal nodes that were shared by these and the maximum likelihood analysis. Two MPTs, differing only in the position of the two *S. costatum* strains, were generated in the unweighted and weighted maximum parsimony analyses. The consistency index improved from 0.492 to 0.718 in the weighted analysis.

The analyses show that the diatoms diverge initially into two clades (fig. 1, Clades I and II). These clades do not correspond to either centric or pennate diatoms or to the three classes defined in Round, Crawford, and Mann (1990, p. 125). These clades are strongly supported in both bootstrap analyses. Other lineages within these two major clades with high bootstrap values correspond, in general, to order level systematics in the diatoms. The phylogenetic relationships inferred from the maximum likelihood methods were identical to those found in the weighted maximum parsimony bootstrap analyses; whereas the neighbor-joining bootstrap analysis showed minor rearrangements. In Clade I, the branching order of *Melosira* and *Stephanopyxis* was reversed, and the Coscinodiscales was sister taxon to the Melosirales. In Clade II, *Eucampia* was sister taxon to the Cymatosirales. The three weighting schemes used in the distance analysis also recovered Clades I and II with minor rearrangements of the branching orders within each clade and minor differences in branch lengths (data not shown). The G_1 test provided further support for the resolving power of this data set (Hillis and Huelsenbeck 1992). A G_1 value of −0.938660 is highly significant ($P < 0.01$) and indicates a strong phylogenetic signal.



B

Tree	Log-Likelihood	Difference in Log-Likelihood	Standard Deviation	Significantly Worse	Number of steps in MPT
1. Best tree	14443.799	-	-	-	2258
2. Monophyly of the Centrics	14479.662	-35.862	17.877	Yes	2270
3. " plus 3 centric clades	14482.788	-38.989	18.980	Yes	2269
4. Thalassiosirales and Pennates as sister clades	14469.000	-25.201	10.723	Yes	2268

FIG. 1.—(*A*) Phylogeny of the diatoms from ssu rRNA sequence comparisons of 30 diatom taxa with *Pelagomonas calceolata* and three Oomycetes as outgroup with the ML methods (fastDNAml). Oomycetes were pruned from the far left of the tree. Clades I and II are labeled. Percentages of adenine are shown in ( ). Encircled figures above the internal nodes are bootstrap values based on neighbor-joining analysis. Italic figures below the internal nodes are bootstrap values based on a weighted maximum parsimony analysis. Centric taxa with peripheral labiate processes are in hatched lines; those with a central labiate process are in stippled lines; those with central strutted processes are in striped lines. The period of time before the diversification into the two major clades is represented by a "*." (*B*) Results of user-defined tree analyses using the DNAML program (PHYLIP) and maximum parsimony analyses (PAUP) in which the tree in figure 1*A* is compared to alternative hypotheses concerning the monophyly of the centrics and lineages therein. The log-likelihoods of the "best" and the user-defined trees were statistically compared (Kishino and Hasagawa 1989).

Clade I contains radial centric diatoms of the orders Coscinodiscales, Rhizosoleniales, Corethrales, and Melosirales (fig. 1). These taxa have tubes (i.e., labiate processes) located, in most cases, in a peripheral position in the cell wall. These tubes can have large and elaborate internal openings, such as those in the Coscinodiscales and Rhizosoleniales, or openings that are small and slit-like, as in the Melosirales.

Clade II is composed of three groups (fig. 1). The bi(multi)polar centric diatoms are polyphyletic and can be found in all three groups. The Chaetocerotales and true Hemiaulales diverge together in the first lineage. The second lineage contains the Cymatosirales. The third lineage gives rise to the pennates, which appear to diverge at the same time as another clade containing the Lithodesmiales, *Streptotheca* (presently misclassified as a hemiauloid diatom), and the Thalassiosirales. Taxa in the bi(multi)polar orders have simple, usually centrally located tubes, i.e., labiate processes. The radial centric diatoms of the order Thalassiosirales also have centrally located tube(s), but these are strutted processes. They do have a labiate process, but it is elaborate and lies in a peripheral position. Significantly in all analyses, *Streptotheca* and *Ditylum*, two taxa with a central tube-like labiate process and similar cytological detail (von Stosch 1977), diverge with the Thalassiosirales. Among the pennate diatoms, only the araphid taxa possess tubes through the cell wall (labiate processes) usually at the cell apices. Weak bootstrap support for the branching order within Clade II in both the neighbor-joining and weighted parsimony analyses implies that a preferred branching order cannot be determined. This can reflect either a rapid radiation in these diatoms or a slow-down in sequence divergence at the time of the radiation (see Leipe et al. 1994).

An analysis of the base composition of the rRNA coding regions showed that some taxa had biased nucleotide compositions (fig. 1). The difference between the taxa with the two most varying adenine compositions is 3%. When taxa with adenine values greater than 28% were deleted from distance and maximum parsimony analyses, tree topologies did not change (data not shown). In the LogDet transformation analysis, the same two primary clades were recovered (data not shown). The topology of the first clade is identical to that in the maximum likelihood tree. However, in the second clade, there are minor rearrangements of the branching order of the major lineages as found in the distance and the weighted maximum parsimony analyses. The differences in the branching order in each of the analyses is reflected by the short branch lengths separating the major lineages and the lack of bootstrap support for their order in the distance and parsimony analyses. Thus, elevated adenine compositions in some taxa may not be significant because the inclusion of these taxa in the analysis does not appear to alter significantly the phylogenetic relationships recovered in the other analyses. In all analyses, the same two major clades are recovered.

Given that the recovered clades do not correspond to morphological-based systematics of the diatoms, three alternative topologies were tested. To address the monophyly of the centrics, we exchanged Clade I with the pennate diatoms. To address the monophyly of the centrics with three independent lineages, we maintained the first rearrangement and then grouped the centric diatoms into three lineages: the radial centrics, the bipolar centrics and the centrics with strutted processes. The third user-defined tree maintained the original maximum likelihood tree topology but removed *Ditylum* and *Streptotheca* to a separate branch so that the pennates and the Thalassiosirales were sister clades. These trees were used to constrain the maximum likelihood analysis and the maximum parsimony analysis. Statistics associated with each analysis were compared to those obtained in the original analyses, which were assumed to represent the best tree. The user-defined tree analyses (fig. 1*B*) support the divergence of the diatoms into these two major clades because the disruption of these two clades result in trees whose log likelihoods were significantly worse than that found for the best tree in the original maximum likelihood analysis (fig. 1*B*) and which required more steps in the maximum-parsimony analyses.

Discussion

The diatoms represent an important monophyletic radiation within the heterokont algae (Mann and Marchant 1989; Round, Crawford, and Mann 1990; Bhattacharya et al. 1992; Leipe et al. 1994) and, taken together with recent evidence for their monophyletic grouping with other heterokonts with a reduced flagellar apparatus (Saunders et al. 1995) may constitute a new algal taxon. The presumed dichotomy of the diatoms into centric and pennate taxa based on pattern centers, reproduction, and plastid morphology is not supported by either 18S or 28S rRNA sequence comparisons (Sörhannus et al. 1995). User-defined trees also do not support a dichotomy into centric and pennate taxa.

The divergence in the diatoms revealed by the rRNA tree is difficult to reconcile using morphological features that currently define diatom classes and orders. Valve polarity has been useful in defining several of the major lineages within each clade. However, our analyses also suggest that certain morphological features are not useful markers for delineating the deeper branches in the rRNA tree. A mapping of the presence of labiate processes onto the tree (fig. 1) shows that they are found in both clades, although its position in the cell wall can

separate the two clades. The strutted process has a unique derived origin within Clade II taxa. However, details of cellular architecture, where available, can distinguish the two clades.

In Clade I, the labiate processes are located peripherally in the cell wall. In many taxa assigned to this clade, cytoplasmic strands extend from the nucleus to the peripheral labiate processes to anchor the nucleus during interphase and new valve formation (Schmid 1994). In *Coscinodiscus, Stephanopyxis,* and *Ellerbeckia* (Melosirales), Golgi bodies are associated with a mitochondrion in an endoplasmic reticulum cisternae (Schmid 1988 and personal communication), a character also found in the Oomycetes and, therefore, likely a primitive feature of all diatoms.

The bi(multi)polar centrics (except for *Ditylum/ Streptotheca*) are the first two divergences in Clade II, although their branching order is not resolved. Their labiate processes are in a central position in the cell wall and are a simple tube structure with a slit-like internal opening. The Thalassiosirales, which have centrally located strutted processes, diverge at the same time as *Ditylum* and *Streptotheca,* which have a simple tube-like central labiate process. Among the pennate diatoms, only the araphid taxa have labiate processes. These are generally located at the cell apices and are separated by a siliceous rib extending the length of the cell wall. Cytoplasmic threads anchor the nucleus to the simple central labiate processes of the bipolar taxa (Schmid 1994) or to the central strutted processes in the Thalassiosirales (Schmid 1984). The marginally located labiate process of the Thalassiosirales is not involved in these anchoring mechanisms and is likely of an independent origin. There is no available information for the possible cytoplasmic anchoring function of the labiate process in araphid pennate diatoms; however, in the raphid pennate diatoms, cytoplasm remains attached to the central siliceous raphe ribs (Höfler 1940) during plasmolysis, indicating that this structure functions as an anchoring site for the cytoplasm. In the Thalassiosirales and the pennates, the Golgi bodies are arranged in a perinuclear shell or 'Plattenband' (Schmid 1989), but we are unaware of any documentation of the position of the Golgi bodies in the bipolar centrics.

These data suggest that the strutted process is derived from a labiate process or a tube-like structure that has been termed a labiate process. When *Ditylum* and *Streptotheca,* taxa with central labiate processes, were removed as sister taxa to the Thalassiosirales, the tree was significantly worse. There is a structural similarity between the perforations in the cell walls of *Ditylum* and the adjacent pores associated with the strutted processes of the Thalassiosirales (see illustrations in Li and Volcani 1985; Round, Crawford, and Mann 1990, pp. 36, 293).

A monophyletic origin for the pennate diatoms is supported by our analyses. Within the pennates, the araphid taxa remain paraphyletic. A monophyletic origin for the raphid taxa cannot be determined because only one raphid order has been sampled; however, the phylogenetic relationships inferred from the rRNA sequence comparisons suggest that the raphe (slit) in the raphid pennate diatoms is likely derived from the simple slit-like labiate process of the bipolar centrics of Clade II. This supports Hasle's (1974) original hypothesis that the raphe evolved from a labiate process and also provides insight into which type of labiate process was the likely progenitor of the raphe.

In our phylogenetic analyses, it is clear that the diatoms had a long evolutionary history prior to the diversification into the two major clades (see * in fig. 1). In an hypothesis of the early evolution of the diatoms, Round and Crawford (1981) and Mann and Marchant (1989) suggested that the earliest progenitor for the diatoms was a scaly unicell. In their evolutionary schemes, two of the cell scales become polarized and form the major components of the diatom cell wall; whereas the remaining scales form the girdle bands that encircle the mid-region of the diatom cell. There is no fossil record of such a potential ancestor; however, the presence of siliceous scales on the zygote of the centric diatoms indicate that a scaly ancestor was present at some point in the group's ontogeny (Round, Crawford, and Mann 1990, p. 96). Instead, the oldest unambiguous diatom fossils are of several domed-shaped centric diatoms from early Jurassic deposits (Rothpletz 1896, ca. 185 Mya). Although few taxa are present in this deposit, the wall structure is already complex but lacks any structure in the wall center. Following a gap in the fossil record, diatoms are next reported in high abundance and diversity from many shallow water environment deposits from the Lower Cretaceous (115–110 Mya) (Gersonde and Harwood 1990). A few, very poorly preserved deposits occur after this time, but by the Upper Cretaceous (87–65 Mya) diatomaceous deposits are widespread, and a more or less continuous record occurs until the present day. The assemblages occurring after the Upper Cretaceous (87–65 Mya) possess labiate/strutted processes and the raphe currently used to define diatom orders and below. However before this time, these morphological features are absent, and it appears that the evolution of many of these modern-day features must have occurred during this gap in the fossil record. Despite this, many taxa in the Lower Cretaceous closely resemble modern taxa in terms of overall valve structure. Certain taxa during this time period have a central structure in the cell wall (Gersonde and Harwood 1990). This central

structure begins as an invagination of the cell wall and gradually transforms into a distinct tube (Gersonde and Harwood 1990; Nikolaev and Harwood 1995). This tube may be a predecessor for either the modern-day strutted or labiate processes; however, the poor preservation of the diatom assemblages between the Lower (115–110 Mya) and Upper Cretaceous (87–65 Mya) preclude the establishment of a direct link between the tube structures found in the Lower Cretaceous taxa and the tube structures found in modern-day taxa. Nevertheless, overall valve structure would suggest that the fossil taxa with a central tube from the Upper Cretaceous are related to Clade II taxa with a central labiate or strutted process in the cell wall. Fossil taxa lacking any central tube from the Upper Cretaceous have an overall valve structure like Clade I taxa and more closely resemble the taxa from the Jurassic. Thus, this early, best preserved diatom deposit provides speculative ancestors for both clades.

The taxonomic implications of the phylogeny inferred from the ssu and the limited data set from the lsu rRNA coding regions (Sörhannus et al. 1995) support a major reclassification of the diatoms at higher taxonomic levels. Clearly, the features considered essential in present-day systematics of the diatoms represent derived characters that cannot be used as they are presently defined to characterize the separation of the two major clades of diatoms. Cladistic analyses of extant morphological features would not, for example, recover the molecular tree because structures, such as the labiate processes, believed to be homologous in current systematics are shown to have a paraphyletic origin. However, if labiate process are redefined, then the morphological and molecular data correlate well. Our analyses have highlighted the incongruencies in the two data sets that should be further studied. It is clear that a modern ultrastructural study of labiate process development/morphogenesis for the two clades is needed. The position of the labiate process appears to be highly significant and may be the result of a basic difference in valve morphogenesis between the two clades. The strutted process of the Thalassiosirales, the central labiate process of the bipolar centrics, and the raphe are likely derived from the same ancestral structure. Furthermore, the molecular analysis provide insights into the major period of diatom radiation that cannot be reconstructed on the basis of the existing fossil record.

Acknowledgments

This work was supported by the Deutsche Forschungs Gemeinschaft (Sm 22/7-1, Ge 516/4). Technical assistance of S. Makedanz is gratefully acknowledged. Dr. D. Bhattacharya provided computer assistance and many helpful discussions. Drs. D. Harwood and A.-M. M. Schmid and Ms. P. A. Sims kindly shared unpublished observations on the diatoms. This is contribution No. 943 from the Alfred Wegener Institute.

APPENDIX

OOMYCETES: *Lagenidium giganteum* Couch, X54266; *Phytopthora megasperma* Drech., X54265; *Achlya bisexualis* Coker M32705. **PELAGOPHYCEAE:** *Pelagomonas calceolata* And. and Saund., U14389. **BACILLARIOPHYTA:** *COSCINODISCOPHYCEAE: Thalassiosira eccentrica* (Ehrenb.) Cl., X85396; *T. rotula* Meunier, X85397; *Porosira glacialis* (Grun.) Jørg, X85398; *Skeletonema costatum* (Grev.) Cl., X52006; *S. costatum*, X85395; *S. pseudocostatum* Medl., X85393; *S. pseudocostatum* Medl., X85394; *Lauderia borealis* Gran, X85399; *Melosira varians* C. Ag., X85402; *Stephanopyxis* cf. *broschii* Grun., M87330; *Aulacoseira distans* (Ehrenb.) Sim., X85403; *A. ambigua* (Grun.) Sim., X85404; *Coscinodiscus radiatus* Ehrenb., X77705; *Actinocyclus curvatulus* Jan., X85401; *Eucampia antarctica* (Castr.) Mang., X85389; *Streptotheca thamesis* Shrubs., X85385; *Ditylum brightwelli* (West) Grun., X85386; *Corethron criophilum* Castr., X85400; *Cymatosira belgica* Grun. in V. H., X85387; *Papiliocellulus elegans* Hasle, von Stosch and Syv., X85388; *Rhizosolenia setigera* Brightw., M87329; *Chaetoceros didymus* Ehrenb., X85392; *C. rostratus* Laud., X85391; *Chaetoceros sp.*, X85390. *FRAGILARIOPHYCEAE: Fragilaria striatula* Lyngb., X77704; *Asterionellopsis glacialis* (Castr.) Round, X77701; *Rhaphoneis* cf. *belgica* (Grun.) Grun., X77703; *Thalassionema nitzschioides* (Grun.) V.H., X77702. *BACILLARIOPHYCEAE: Nitzschia apiculata* (Greg.) Grun., M97334; *Bacillaria paxillifer* (Müll.) Hend., M87325; *Cylindrotheca closterium* (Ehrenb.) Reim. and Lewin, M87326.

LITERATURE CITED

ANDERSEN, R. A., G. W. SAUNDERS, M. P. PASKIND, and J. P. SEXTON. 1993. Ultrastructure and 18S rRNA gene sequence for *Pelagomonas calceolata* gen. et sp. nov. and the description of a new algal class, the Pelagophyceae classis nov. J. Phycol. **29**:701–715.

BHATTACHARYA, D., L. MEDLIN, P. O. WAINWRIGHT, E. V. ARIZTIA, C. BIBEAU, S. K. STICKEL, and M. L. SOGIN. 1992. Algae containing chlorophylls *a* + *c* are paraphyletic: molecular evolutionary analysis of the Chromophyta. Evolution **46**:1808–1817.

DOYLE, J. J., and J. L. DOYLE. 1990. Isolation of plant DNA from fresh tissue. Focus **12**:13–15.

ELWOOD, H. J., G. J. OLSEN, and M. L. SOGIN. 1985. The small-subunit ribosomal RNA gene sequences from the hy-

potrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. Mol. Biol. Evol. **2**:399–410.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

———. 1993. PHYLIP manual, version 3.5. Department of Genetics, University of Washington, Seattle, Wash.

FITCH, W. M., and E. MARGOLIASH. 1967. Construction of phylogenetic trees: a method based on mutation distances as estimated from cytochrome *c* sequences is of general applicability. Science **155**:279–284.

GERSONDE, R., and D. M. HARWOOD. 1990. Lower Cretaceous diatoms from ODP Leg 113 site 693 Weddell Sea. Part 1: Vegetative cells. Pp. 403–425 *in* P. F. Barker, J. P. Kennett et al., eds. Proceedings of the Ocean Drilling Program, Scientific Results. Vol. 113. Ocean Drilling Program, College Station, Tex.

HASLE, G. R. 1974. The mucilage pore of pennate diatoms. Nova Hedwigla Beih. **45**:167–186.

HILLIS, D. M., and J. P. HUELSENBECK. 1992. Signal, noise, and reliability in molecular phylogenetic analyses. J. Hered. **83**:189–195.

HILLIS, D. M., J. P. HUELSENBECK, and C. W. CUNNINGHAM. 1994. Application and accuracy of molecular phylogenies. Science **264**:671–677.

HÖFLER, K. 1940. Aus der Protoplasmatik der Diatomeen. Ber. Dsch. Bot. Ges. **58**:97–120.

JIN, L., and M. NEI. 1990. Limitations of the evolutionary parsimony method of phylogenetic analysis. Mol. Biol. Evol. **7**:82–120.

JUKES, T. H., and C. R. CANTOR. 1969. Evolution of protein molecules. Pp. 21–132 *in* H. N. Munro, ed. Mammalian protein metabolism. Academic Press, New York.

KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.

KISHINO, H., and M. HASEGAWA. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order of the Hominoidea. J. Mol. Evol. **29**:170–179.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetics analysis, version 1.0. Pennsylvania State University, University Park, Penn.

LARSEN, L., G. J. OLSEN, B. L. MAIDAK, M. J. MCCAUGHEY, R. OVERBEEK, R. MACKE, T. L. MARSCH, and C. R. WOESE. 1993. The ribosomal database project. Nucleic Acids Res. **21**(suppl.):3021–3023.

LEIPE, D. D., P. O. WAINWRIGHT, J. H. GUNDERSON, D. PORTER, D. J. PATTERSON, F. VALOIS, S. HIMMERICH, and M. L. SOGIN. 1994. The stramenopiles from a molecular perspective: 16S-like rRNA sequences from *Labyrinthuloides minuta* and *Cafeteria roenbergensis*. Phycologia **33**:369–377.

LI, C.-W., and B. E. VOLCANI. 1985. Studies on the biochemistry and fine structure of silica shell formation in diatoms. VIII. Morphogenesis of the cell wall in a centric diatom *Ditylum brightwellii*. Protoplasma **124**:10–29.

LOCKHART, P. J., M. A. STEEL, M. D. HENDY, and D. PENNY. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. **11**:605–612.

MADDISON, W. P., and D. R. MADDISON. 1992. MacClade. Analysis of phylogeny and character evolution, version 3. Sinauer Assoc., Sunderland, Mass.

MANN, D. G., and H. J. MARCHANT. 1989. The origin of the diatom and its life cycle. Pp. 307–303 *in* J. C. Green, B. S. C. Leadbeater, and W. L. Diver, eds. The chromophyte algae: problems and perspectives. Clarendon Press, Oxford, UK.

MEDLIN, L. K., R. M. CRAWFORD, and R. A. ANDERSEN. 1986. Histochemical and ultrastructural evidence for the function of the labiate process in the movement of centric diatoms. Br. Phycol. J. **21**:297–301.

MEDLIN, L. K., H. J. ELWOOD, S. STICKEL, and M. L. SOGIN. 1988. The characterization of enzymatically amplified eukaryotic 18S rRNA-coding regions. Gene **71**:491–499.

———. 1991. Morphological and genetic variation within the diatom *Skeletonema costatum* Bacillariophyta: evidence for a new species, *Skeletonema pseudocostatum*. J. Phycol. **27**: 514–524.

MEDLIN, L. K., M. LANGE, and M. E. M. BAUMANN. 1994. Genetic differentiation among three colony-forming species of *Phaeocystis:* further evidence for the phylogeny of the Prymnesiophyta. Phycologia **33**:199–212.

MEDLIN, L. K., D. M. WILLIAMS, and P. A. SIMS. 1993. The evolution of the diatoms (Bacillariophyta). I. Origin of the group and assessment of the monophyly of its major divisions. Eur. J. Phycol. **28**:261–275.

NEEFS, J.-M., Y. VAN DE PEER, P. DE RIJK, A. GORIS, and R. DE WACHTER. 1991. Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res. **19**(supp.):1987–2015.

NIKOLAEV, V. A., and D. M. HARWOOD. 1995. New process, genus and family of Lower Cretaceous diatoms from Australia. Diatom Res. (in press).

PICKETT-HEAPS, J., D. R. A. HILL, and R. WEATHERBEE. 1986. Cellular movement in the centric diatom *Odontella sinensis*. J. Phycol. **22**:234–239.

PICKETT-HEAPS, J., D. R. A. HILL, and K. L. BLAZÉ. 1991. Active gliding motility in an araphid marine diatom, *Ardissonea* (formerly *Synedra*) *crystallina*. J. Phycol. **27**:718–725.

ROTHPLETZ, A. 1896. Über die Flysch-Fucoiden und einige andere fossile Algen, sowie über liasische, Diatomeen führende Horenschwämme. Z. Dsch. Geol. Ges. **48**:854–914.

ROUND, F. E., and R. M. CRAWFORD. 1981. The lines of evolution of the Bacillariophyta. I. Origin. Proc. R. Soc. Lond. B **211**:237–260.

ROUND, F. E., R. M. CRAWFORD, and D. G. MANN. 1990. The diatoms: morphology, and biology of the genera. Cambridge University Press, Cambridge, UK.

SAIKI, R., D. H. GELFAND, S. STOFFEL, S. J. SCHARF, R. HIGUCHI, G. T. HORN, K. B. MULLIS, and H. A. ERLICH. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA-polymerase. Science **239**:487–491.

SANGER, F., S. NICKLEN, and A. R. COULSEN. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.

SAUNDERS, G. W., D. POTTER, M. P. PASKIND, and R. A. ANDERSEN. 1995. Cladistic analyses of combined traditional and molecular data sets reveal an algal lineage. Proc. Natl. Acad. Sci. USA **92**:244–248.

SCHMID, A.-M. M. 1984. Wall morphogenesis in *Thalassiosira eccentrica:* comparison of auxospore formation and the effects of MT-inhibitors. Pp. 42–70 *in* D. G. Mann, ed. Proceedings of the 7th International Diatom Symposium, Philadelphia 1982. O. Koeltz, Koenigstein, Germany.

———. 1988. The special Golgi–ER–mitochondrium unit in the diatom genus *Coscinodiscus.* Pl. Syst. Evol. **158**:211–223.

———. 1989. Geitler's "Plattenband" in the diatom *Synedra* cf. *ulna* in the light of TEM investigations. Plant Syst. Evol. **164**:239–252.

———. 1994. Aspects of morphogenesis and function of diatom cell walls with implications for taxonomy. Protoplasma **181**:43–60.

SÖRHANNUS, U., F. GASSE, R. PERASSO, and A. BAROIN-TOURANCHEAU. 1995. A preliminary phylogeny of diatoms based on 28S ribosomal RNA sequence data. Phycologia **34**:65–73.

SWOFFORD, D. L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign, Ill.

SWOFFORD, D. L., and G. J. OLSEN. 1990. Phylogeny reconstruction. Pp. 411–501 *in* D. M. Hillis and C. Moritz, eds. Molecular systematics. Sinauer Associates, Sunderland, Mass.

VAN DE PEER, Y., J.-M. NEEFS, P. DE RIJK, and R. DE WACHTER. 1993. Reconstructing evolution from eukaryotic small-ribosomal-subunit RNA sequences: calibration of the molecular clock. J. Mol. Evol. **37**:221–232.

VON STOSCH, H. A. 1977. Observations on *Bellerochea* and *Streptotheca,* including descriptions of three new planktonic diatom species. Nova Hedwigia Beih. **54**:113–166.

WOESE, C. R. 1987. Bacterial evolution. Microbiol. Rev. **51**: 221–271.

MITCHELL L. SOGIN, reviewing editor.

Accepted August 7, 1995