Evolution of the diatoms: insights from fossil, biological and molecular data

P. A. SIMS, D. G. MANN AND L. K. MEDLIN

1 Botany Department, Natural History Museum, Cromwell Road, London, SW7 5BD, UK
2 Royal Botanic Garden, Edinburgh EH3 5LR, Scotland, UK
3 Alfred Wegener Institute for Polar and Marine Biology, Am Handelshafen 12, D. 27570 Bremerhaven, Germany


Molecular sequence analyses have yielded many important insights into diatom evolution, but there have been few attempts to relate these to the extensive fossil record of diatoms, probably because of unfamiliarity with the data available, which are scattered widely through the geological literature. We review the main features of molecular phylogenies and concentrate on the correspondence between these and the fossil record; we also review the evolution of major morphological, cytological and life cycle characteristics, and possible diatom origins. The first physical remains of diatoms are from the Jurassic, and well-preserved, diverse floras are available from the Lower Cretaceous. Though these are unequivocally identifiable as centric diatoms, none except a possible *Stephanopyxis* can be unequivocally linked to lineages of extant diatoms, although it is almost certain that members of the Coscinodiscophyceae (radial centrics) and Mediopyceae (polar centrics) were present; some display curious morphological features that hint at an unorthodox cell division mechanism and life cycle. It seems most likely that the earliest diatoms were marine, but recently discovered fossil deposits hint that episodes of terrestrial colonization may have occurred in the Mesozoic, though the main invasion of freshwaters appears to have been delayed until the Cenozoic. By the Upper Cretaceous, many lineages are present that can be convincingly related to extant diatom taxa. Pennate diatoms appear in the late Cretaceous and raphid diatoms in the Palaeocene, though molecular phylogenies imply that raphid diatoms did in fact evolve considerably earlier. Recent evidence shows that diatoms are substantially underclassified at the species level, with many semicryptic or cryptic species to be recognized; however, there is little prospect of being able to discriminate between such taxa in fossil material.

KEY WORDS: Bacillariophyta, Cretaceous, diatom, evolution, fossil record, frustule structure, Jurassic, silica wall, Tertiary

CONTENTS

1. Introduction .............................................. 362
2. The origins of diatoms .................................... 362
   2.1. The systematic and ecological context of diatom evolution .......... 362
   2.2. Evolution of the silica wall .................................. 363
   2.3. Evolution of cell division, the life cycle and auxosporulation ........ 364
3. Evolution of the major groups of diatoms ......................... 366
   3.1. The impact of molecular data on diatom systematics ................ 366
   3.2. Correlations between molecular phylogenies and nonmolecular characteristics ........................................ 367
   3.3. The palaeoecological context of diatom origins ..................... 367
4. Diatom evolution in the Mesozoic ................................ 371
   4.1. Jurassic and Cretaceous: overview and fossil localities ............ 371
   4.2. Lower Cretaceous diatoms ................................... 372
   4.3. Upper Cretaceous centric diatoms ................................ 377
   4.4. The evolution of penrate and penrate-like diatoms .................. 381
5. Diatom evolution in the Cenozoic ................................ 386

* Corresponding author (lkmedlin@awi-bremerhaven.de).
1. INTRODUCTION

The diatoms are one of the most easily recognizable groups of major eukaryotic algae, because of their unique silicified cell walls (frustules), which consist of two overlapping thecae, each in turn consisting of a valve and a number of hoop-like or segmental girdle bands. This structure is present in all living diatoms, except following secondary loss in endosymbiotic diatoms living in some dinoflagellates (e.g. Tamura et al. 2005) and some foraminifera (Lee & Correia 2005). Normal frustules are also found in the earliest known deposits of well-preserved fossil diatoms, from the early Albanian (Lower Cretaceous) of what is now the Weddell Sea, Antarctica (Gersonde & Harwood 1990).

Molecular sequence data show that diatoms belong to the heterokont algae (Medlin et al. 1997b), a relationship long established from morphology studied at the electron microscopy level (see mini-review in Medlin et al. 1997b). Heterokonts are chlorophyll a+c containing algae that typically have zooids with two heterodynamic flagella, one covered with tripartite mastigonemes and the other smooth (van den Hoek et al. 1995). In diatoms, the flagellar apparatus is reduced or absent; indeed, only the spermatozoids of the oogamous ‘centric’ diatoms are flagellated and these are uniflagellate (Manton & von Stosch 1965), lacking all trace of a smooth posterior flagellum or basal body. Nevertheless, the characteristic heterokont mastigonemes are present and diatoms also possess similar plastid ultrastructure (with four bounding membranes, lamellae of three thylakoids, and usually a peripheral ring nucleoid) and pigment composition to the brown algae (Christensen 1989).

Many workers have speculated upon diatom origins. Diatoms may have been derived from a spherical, uniformly scaled monad (Round 1981; Round & Crawford 1981, 1984) with an anterior flagellum (Cavalier-Smith 1986), or from a cyst-like form similar to the extant Parmales in the chrysophyte algae (Mann & Marchant 1989). A scaly ancestor likely existed at some point in their phylogeny because of the presence of scales on the reproductive cells of diatoms and on the reproductive stages of labyrinthulids, an earlier divergence in the heterokont lineage (Medlin et al. 1997b). Recent phylogenies constructed from nuclear-encoded small-subunit ribosomal RNAs (SSU rRNA) place the diatoms within the pigmented heterokont algal lineages (Fig. 1; Bhattacharya et al. 1992; Leipe et al. 1994; Medlin et al. 1997b). The closest relatives discovered so far are the Bolidophyceae, which are picoplanktonic algae with a simplified cellular organization (Guillou et al. 1999). Both diatoms and bolidomonads possess similar pigments and two transverse plates at the base of each flagellum. A molecular clock constructed from four genes has placed the average age of the diatoms c. 135 Ma ago, with their earliest possible age being 240 Ma ago (Kooistra & Medlin 1996; Medlin et al. 1997a).

Here we present an integrated review of diatom evolution. We focus on the fossil record and its correspondence to molecular phylogenies, but we also make limited reference to cytological and reproductive characteristics.

2. THE ORIGINS OF DIATOMS

2.1. The systematic and ecological context of diatom evolution

Today, the diatoms are found in almost all aquatic and most wet terrestrial habitats. Given the unusual morphologies found among early Cretaceous diatoms, which bear little resemblance to modern taxa, it is interesting to speculate upon what kind of habitats could support these strange forms. Four main ideas have been proposed recently on the likely origins of the diatoms and the environments in which they developed.

Round & Crawford (1981) came to the conclusion that diatoms are monophyletic in origin and proposed that they arose from a naked photosynthetic cell that acquired a coating of siliceous scales. In the next stage, the ‘Ur-diatom’, the siliceous scales differentiated to form valves and girdle bands – a simple diatom. This, they suggested, happened sometime between the Proterozoic (Precambrian) and the late Jurassic (c. 650–140 Ma), perhaps in a ‘swampy habitat of unknown salinity, which would have been devoid of substratum for attaching algae and too unstable for planktonic forms’. They considered that an open ocean would not have provided the environment or adequate silica necessary to promote the interactions necessary to produce complex eukaryotic forms requiring large amounts of silica. Katz et al. (2005) suggest that modern diatoms thrive best in regions where nutrients are supplied with high pulse frequencies and that the evolution of grasses provided an extra pulse of silica into the ocean when
the diatoms were first rapidly expanding in the Cretaceous. They also suggest that diatom diversity responds to environmental factors, whereas calcareous nanoplankton and dinoflagellates respond to sea level changes. Round & Crawford (1981) concluded that the ancestral diatom developed in a shallow-water benthic environment, possibly associated with mucilaginous flocs: ‘probably a region of fluctuating water levels over sediment/rock in which forms associated with mucilage masses predominate’. The habitat could have been slightly saline or even freshwater and adaptation to marine habitats could have occurred at a later date, although an origin in a freshwater habitat appears to conflict with molecular data (but see below). It is highly unlikely that the diatoms arose as early as in the Precambrian (see section 4.1).

Mann & Marchant (1989), addressing the problem of the diploid nature of the diatom, proposed that a silicified cyst derived position of the Xanthophyceae in all phylogenetic analyses (see Medlin et al. 1997b), it is unlikely that the fossils resembling modern Xanthophyceae are closely related to them; they are more likely an extinct lineage. However, the Cretaceous fossils do suggest that the ability to make a cell wall of two overlapping parts appeared very early in the evolution of the pigmented heterokonts.

Most recently, Harwood et al. (2004) proposed that diatoms arose in terrestrial habitats, because a new early diatom deposit (175 Ma) seems to be nonmarine. This appears to conflict with molecular data, because the current sister group, the Bolidophyceae, are an exclusively marine group of picoplankton. However, a scenario could be proposed in which ‘Ur-diatoms’, abundant as nonsilicifying unicells in coastal waters, became stranded in isolated pools as eustatic seas flooded the continents. When these saline pools began to dry up, the diatom ancestors, if they survived, must have had to adapt to a semiterrestrial habitat. The ability to metabolize silica and the production of thick silica walls could have evolved at this time as protection against desiccation and also to put the cells into a temporary resting state (see Medlin 2002 and section 2.2) until the areas were reflooded. With further sea-level changes, the early diatoms might decolonize coastal waters but their newly developed silica walls would make them too heavy for a planktonic existence, which relegated them to a benthic

nearshore community. In this respect, Paralia sol, as the first diatom lineage in the radial centric clade in the molecular trees, is a reminder of these early heavily silicified diatoms that are benthic, only rarely being suspended up into the plankton with increased turbulence. Paralia sol is now transferred to Ellerbeckia (Crawford & Sims, 2006), and now the first divergence in the radial centrics is of a marine representative of a terrestrial genus.

Thus, existing hypotheses of diatom origins tend to agree that the prediatom or ‘Ur-diatom’ developed, not in pelagic habitats, but in shallow marine environments, perhaps with the intercalation of a freshwater or terrestrial phase. Tappan (1980), when discussing the origin of the diatoms, noted that Palaeozoic cherts (659–c. 250 Ma) have been found to contain the remains of radiolarians and, locally, sponges, but there are no diatoms present. She concluded that, because radiolarians and sponge spicules are well-preserved throughout the Palaeozoic in limestones and concretions, a similar preservation of diatoms might be expected had they too been present. She therefore suggested that diatoms probably made their first appearance in the Mesozoic oceans (c. 230–70 Ma), where both diatoms and chrysomonads appeared and thereafter showed rapid evolutionary divergence. Both have modern freshwater representatives, although their Mesozoic ancestors were entirely marine.

2.2. Evolution of the silica wall

At first sight, the evolution of silicon metabolism and the capacity to deposit solid, patterned silica appear to have occurred since the divergence of the diatoms from the Bolidophyceae, because no traces of silica scales have been found in the Bolidophyceae. However, several other heterokont groups, including the Dictyochophyceae, Synurophyceae, Chrysophyceae, Parmophyceae, and Xanthophyceae (van den Hoek et al. 1995; Graham & Wilcox 2000) produce silica structures. These are spread across the phylogenetic tree of heterokonts (see Guillou et al. 1999; Medlin et al. 2000; Kawachi et al. 2002; unfortunately, no molecular data are yet available for the Parmophyceae) and so it is possible that the ability to metabolize silica is primitive among autotrophic heterokonts. However, even if this were so, it still would be necessary to examine why diatoms have retained the capacity to metabolize silica and acquired the ability to produce a silicified wall, when their closest relatives, the Bolidophyceae, have not. Production of a wall might be an evolutionary response to desiccation, as discussed above (see section 2.1), but most algal walls are organic, not siliceous.

Exactly why diatoms and other algae use silica has frequently been debated among phycologists (Volcani 1981). Defence has sometimes been invoked as the most likely reason for silica metabolism and wall evolution in diatoms, because of the compressive strength of wall silica and its possible immunity to enzymatic attack. Hamm et al. (2003) have given support for such a function. On the other hand, Medlin (2002) has reviewed the evidence from mammalian cell lines that suggest that silica is necessary for the prevention of cell ageing and that it allows normal formation of connective tissues, reduces the effects of toxic metals and inhibits fungal attack, as well as placing cells in a prolonged resting state. In her hypothesized scenario of why diatoms produce silica walls,
initially a simple naked biflagellate cell evolved or retained silicon metabolism, which conferred advantages to the cell of preventing ageing and aiding survival in a prolonged resting state. Silicon became involved in the metabolic processes of the cell, presumably by supplying bioactive surfaces for reactions to take place. Silicon was an ‘inorganic enzyme’. As silica accumulated in the cells, it was sequestered in the endoplasmic reticulum (ER) and eventually found its way into a vacuole whose internal pH was acidic. Once inside the vacuole, the silica began to polymerize automatically and the polymerized silica was extruded from the cell because, in a polymerized state, the silica was inaccessible to the chemical reactions needed to prevent cell ageing. Thus, the cell evolved a continuous need to replenish its internal silica pool and this could explain the absolute requirement diatoms have for silica before the cell can divide (Darley & Volcani 1969). In a proposed intervening terrestrial habitat (Harwood et al. 2004), a silica wall could help prevent desiccation and put the cells into a resting state.

Raven & Waite (2004) have recently suggested that the silica wall increased the sinking rate of ancestral diatoms, moving parasitized cells out of the population and thus ensuring survival of the species: ‘regulation of sinking rate could have been the first function of externalization of internal silica if the earliest silicified diatoms were planktonic’. However, as already noted, the earliest diatoms may not have been planktonic. These ideas suggest how silicon metabolism might have evolved or been retained by the ancestral diatom cell, leading to the production of a scaly covering. However, it does not account for the transformation of a scaly covering into a true cell wall. Part of the explanation may be that walls represent a once-only investment in shape maintenance and osmoregulation and can be particularly advantageous when energy capture is slow and also in freshwater, where the cytoplasm will always be hypertonic to the medium and thus tends to take up water (Raven 1982). Furthermore, Raven (1983) has calculated that, for a given wall volume, it costs less energy to make the wall of silica than of carbohydrate. Following evolution of silica metabolism and the capacity to deposit silica selectively in special vesicles, diatom cells have evolved membrane-mediated systems and macromorphogenetic mechanisms (involving organelles and the cytoskeleton) to mould the wall features (see Kröger & Sumper 1998), thus creating a huge variety of cell wall morphologies and structures, which are often species-specific.

2.3. Evolution of cell division, the life cycle and auxosporulation

The special construction of the diatom cell wall in two overlapping pieces is correlated with special features of diatom cell division and their life cycle. In diatoms, cell division is semiconservative, in that each daughter cell inherits one of the parental thecae and forms only one new theca. This is true also of desmids and some dinoflagellates (van den Hoek et al. 1995), but in diatoms there is the extra complication that cytokinesis and the formation of the major wall elements – the valves – take place within the confines of the parental cell wall; the new daughter cells are not released until the shape and size of their new valves essentially are fixed. The invariable consequence of the constrained, internalized development of the diatom cell wall is that the new wall elements are slightly smaller than the parts of the parental cell wall that surround them as they form. In turn, this usually means that the new thecae are smaller overall than the parental thecae, so that the mean size of cells in a diatom population decreases inexorably (and the variance increases) with time, as the cells divide. Size is restored through one of two special processes, in which cells jettison the old, much-diminished cell walls and expand. The best understood process is auxosporulation (Round et al. 1990; van den Hoek et al. 1995; Chepurnov & Mann 2004; Chepurnov et al. 2002, 2004), which is usually linked to sexual reproduction (oogamy in centrics, anisogamy or isogamy in pennates) and involves the formation of a special zygote, the ‘auxospore’, capable of expansion in a highly controlled way to regenerate the large size and normal morphology of the vegetative cells. The other, much less well-known process is ‘vegetative enlargement’, in which there appears to be far less control over cell form and which results in a much smaller enlargement (Chepurnov et al. 2004). Both processes occur only in cells of a restricted range of size relative to the whole range exhibited during the life cycle. Thus, the largest vegetative cells, formed by division of the expanded auxospore, are often incapable of a further round of sexual reproduction and auxosporulation until they have reduced somewhat in size.

As far as is known, all diatoms are diplonts and the only haploid cells produced during the life cycle are the gametes. In culture, development of unfertilized, haploid auxospores (haploid parthenogenesis) has been observed in a few cases (e.g. Mann 1994; Chepurnov & Rosschin 1995; Chepurnov et al. 2002), but there is currently no evidence to suggest that this or other changes in ploidy level occur as a normal part of any diatom’s life history. In contrast, few other autotrophic heterokont algae are diplonts, except for the siphonaceous Vaucheria de Candolle and its close relatives within the Xanthophyceae (which are not closely related to diatoms: Medlin et al. 2000, Kawachi et al. 2002) and at least some Raphidophyceae (Yamaguchi & Imai 1994; Cronberg 2005). Instead, autotrophic heterokonts and other ‘chromophytes’ are haplonts (e.g. in Chrysophyceae-Synurophyceae) or have an ‘alternation of generations’, as in most brown algae (van den Hoek et al. 1995; Graham & Wilcox 2000), coccolithophorids (Billard 1994) and likely cryptomonads (Hill & Wetherbee 1986; Hoef-Emden & Melkonian 2003). Mann & Marchant (1989) therefore suggested that the evolution of the cell wall and of the diplontic life cycle were coupled, the diploid, walled vegetative diatom cell having arisen originally as an intermittently dormant cyst stage in the life cycle of haplontic or haplo-diplontic scaly flagellates. If the modern diatom cell did arise from a dormant cyst stage, then the involvement of Si to prevent the cell from ageing may play an even greater role in the evolution of the modern diatom cell.

However, if the diatoms diverged as a monophyletic lineage sister to all other autotrophic heterokonts (Medlin et al. 2000), it is more important to look at the life histories of the basal heterotrophic heterokonts, such as the oomycetes and labyrinthsulids, than at those of the autotrophic heterokonts, if we are to judge whether the diplontic diatom life cycle is plesiomorphic or apomorphic. Medlin et al. (1993, 1997b) suggested that the diplonty of diatoms is plesiomorphic, on the basis...
that basal heterotrophic heterokonts are also diplonts. The main basis for this idea was an outgroup comparison, in which Medlin et al. (2000) claimed that all oomycetes are diplotypid and implied that the labyrinthulids might be also. Certainly, many well-studied oomycete pathogens and saprophytes belonging to the Saprolegniaceae and Peronosporales are diplotypic (Alexopoulos et al. 1996). However, the life cycles of the simplest organisms currently assigned to the Oomycota (e.g. Olpidiopsisidae, Ectrogellaceae) are still unknown (Dick 1990), and in the labyrinthulids there is a round of haplo-haploid mitotic division after meiosis (Porter 1990), so that this life cycle is not truly diplontic but haplo-diplontic (this is true also of Fucus Linnaeus, contra Guillou et al. 1999). At the moment, therefore, it is unclear whether the ancestors of the diatoms were diplontic, haplo-diplontic or haplontic. A particular frustration is the lack of any information about the ploidy of what appears to be the diatoms’ sister group, the Bolidophyceae, though one might expect these organisms to be haploid, because of their extremely small size.

The universality of the internalized cell division mechanism among living diatoms suggests strongly that the ancestor of all extant diatoms also divided in this way. However, there are certain characteristics of some of the Cretaceous diatoms that are difficult to reconcile with how extant diatoms grow and divide, and it is by no means clear that other features of the cell cycle and life cycle are as universal, and therefore ancestral, in diatoms as they generally are assumed to be. The most problematic fossils are Gladius Forti & Schulz (Figs 11, 12) and Gladiopsis Gersonde & Harwood (Forti & Schulz 1932; Gersonde & Harwood 1990; Sims 1994; Harwood & Nikolaev 1995; Figs 29–32), in which the valves are long cylinders (length up to 10× greater than their diameter in Gladius antiquus Forti & Schulz; Fig. 11), instead of the cake- or hat-box–like shapes or shallow dish morphologies of extant centric diatoms. Long cylindrical (‘solenioid’) diatoms exist today in the genera Rhizosolenia Brightwell and Proboscia Sundström and, as in these two modern genera, Gladius species apparently had girdles composed of many diamond-shaped segmental bands (e.g. Gersonde & Harwood 1990, pl. 9, fig. 3, pl. 11, fig. 6; Sims 1994, figs 47, 48; compare Round et al. 1990, pp. 319, 321). However, in modern Rhizosolenia and Proboscia, the elongate cylinder of the cell is composed very largely of girdle, not valve. Two further peculiarities of G. antiquus are that the mantles (i.e. the sides of the deep cylinder of the valve) taper from the area of the valve face (i.e. the closed apex of the valve) toward the valve margin (Fig. 11) – indeed, the top of the valve is bulbous – and that the mantle pores are chambered (pseudoloculate) throughout. Modern diatoms grow by sliding apart of the two thecae, with simultaneous addition of new girdle elements to the inner, younger theca. Throughout this process, the thecae remain intimately linked, forming a single integrated cell wall. It is difficult to imagine how the same processes could have operated in G. antiquus because, during cell cycle elongation, a major inflation of the girdle diameter (almost a doubling) would have had to occur to provide space for the thecae of the following cell generation. Overall, the valve shape of Gladius species, especially G. antiquus, is so different from that of extant (or other fossil) diatoms that one has to question whether the mechanism of cell division and the life cycle of these diatoms can have resembled the ‘typical’ pattern seen in most diatoms today. The Lower Cretaceous Kreaagro Gersonde & Harwood, with its widely flaring, flanged spines (Fig. 33; Gersonde & Harwood 1990, plate 16), is also difficult to explain on the basis of current division patterns. Perhaps here, and in Gladius, valve formation was incomplete at cell separation and therefore not wholly internalized. That this is not impossible is shown by some modern genera, such as Chae-toceros Ehrenberg and Bacteriastrum Shadbolt, which have highly elongate projections, termed setae, that extend radially in two or more directions from the valve plane, well beyond the limits of the girdle. These setae are formed as projections from the cell, which are initially naked and extend out through openings between the girdle bands (Pickett-Heaps 1998a, b). The flexibility of the morphogenetic machinery even in extant diatoms is also demonstrated in Corethron Castracane, in which elaborate spines are produced totally within the confines of the girdle bands and then rotate into place once the new valve is free of the girdle bands (Crawford & Hinz 1995; Attheya T. West setae form coiled up inside the parent theca and then unroll after being freed of the parental girdle bands (Crawford et al. 1994). Similar flexibility of morphogenesis could have operated in early Cretaceous diatoms.

Another pointer to a possibly nonstandard division mechanism in some early diatoms is the chambered mantle of Gla- dius (Figs 11, 12) and several other Lower Cretaceous diatoms, e.g. Gladiopsis (Figs 29, 30, 32), Rhynchopsis Gersonde & Harwood (Figs 26, 27), Basilicostephanus Gersonde & Harwood (Fig. 13), Anclyopyrgus Gersonde & Harwood (Fig. 15) and Amblypyrgus Gersonde & Harwood (Gersonde & Harwood 1990). In these taxa, the mantles (and valve face) consist of an inner layer containing small pores, covered by an outer layer of large hexagonal chambers (pseudoloculate). The small pores of the inner layer are arranged in rows that, if traced across the mantle-face junction, can be seen to originate at an annular pattern centre (e.g. Gersonde & Harwood 1990, pl. 3, fig. 3; pl. 6, fig. 1), indicating, by analogy with extant diatoms (Pickett-Heaps et al. 1990), that this layer was formed first, spreading out just beneath the plasmalemma from the centre of the cleavage furrow. The hexagonal chambers of the outer layer bear no fixed spatial relationship to the inner rows of pores (see Sims 1994, figs 42, 47) and must have been superimposed substantially later during valve formation, just as in present-day Stephanopyxis (Ehrenberg) Ehrenberg (Pickett-Heaps et al. 1990, fig. 83). But how could this have been achieved if the whole process occurred within the girdle of these diatoms, unless the inner layer was formed at a distance from the parental girdle, allowing space for intercalation of the outer layer later in valve development? In modern diatoms with a chambered mantle, either the chambering is internal to the first-formed layer (e.g. Coscinodiscus Ehrenberg), in which case its formation is unproblematic, or it is external, but this is allowed for by a contraction of the protoplast during cleavage, so that the new valves develop beneath an area of naked plasmalemma that is unconstrained by the girdle except at its very margins, as in Stephanopyxis. However, the cylin- drical shape of Gladius antiquus, G. trispinosus Gersonde & Harwood, Anclyopyrgus and other similar Lower Cretaceous diatoms makes the existence of a contraction unlikely. Altogether, there seems enough oddity in some of the Lower Cre- taceous species to make it unsafe to assume that they divided in the same way that modern diatoms divide. Therefore it is
also unsafe to assume that they necessarily must have had a size reduction–restitution cycle like that occurring today.

In addition, the link between size reduction, sexual reproduction and auxospore formation may not be as universal in diatoms as was thought, and even where present, it may not be as strong as has usually been assumed. Lewis (1984) pointed out that the cell division mechanism per se is an inadequate explanation for the evolution of the size reduction–restitution cycle, because it has been known for many years (e.g. Geitler 1932; Wiedling 1948) that some pennate diatoms (now known to be an advanced group: see section 3) are able to avoid size reduction and can even expand (see Crawford 1981 for discussion). Size reduction can be avoided if the girdle widens as it elongates during the cell cycle. Some diatoms appear to be entirely asexual. After 20 years of observations of *Caloneis amphitsibaena* (Bory) Cleve populations in Edinburgh lakes, one of us (D.G.M.) has seen no sexual stages and no evidence of size reduction. In cases like this and the *Nitzschia* Hassall species studied by Wiedling (1948), it is plain that the ‘anomalous’ life cycle has evolved secondarily (Chepurnov et al. 2004), because other members of the same genus and related genera have a ‘normal’ life cycle, with sexually produced auxospores. However, because there are exceptions that do not undergo size reduction–restitution, Lewis (1984) argued that size reduction cannot be a mere byproduct of the wall structure and cell division mechanism. Instead, it must have adaptive significance, which he suggested was as a chronometer for sex, allowing diatoms to spread the high costs of sexual reproduction over several or many years (see also Mann 1988).

If this is so, however, there can be no *a priori* expectation that internalized cell division coevolved with the size reduction–restitution cycle and it is possible that size reduction, together with the (unknown) size measurement mechanisms that would have to underlie its operation as a sex chronometer, evolved relatively late and perhaps even several times. Critical to answering such speculations is to identify exactly which phenomena are among centric diatoms.

Later in the Cretaceous and in the Tertiary (80 Ma→), the morphologies of the fossil diatoms are wholly ‘modern’ and give no indication of any unusual cell division or life cycle feature, and at least one of the Lower Cretaceous taxa (*Bilingua rossi* Gersonde & Harwood: Figs 3, 5) exhibits size and shape variation consistent with that shown during size reduction by extant bipolar diatoms. It is reasonable to assume that these diatoms mostly exhibited size reduction and regenerated their size and shape via auxospores, even though the latter and their immediate offspring are very rarely preserved.

3. EVOLUTION OF THE MAJOR GROUPS OF DIATOMS

3.1. The impact of molecular data on diatom systematics

Most diatomists have long assumed that the diatoms contain two groups, the centrics and the pennates, which can be distinguished by their pattern centres or symmetry, mode of sexual reproduction, and plastid number and structure (Round et al. 1990). The oogamous centric diatoms, with radially symmetrical ornamentation on their valves and with numerous discoid plastids, are distinct from the isogamous pennate diatoms with bilaterally symmetrical pattern centres and generally fewer, plate-like plastids. These groups are known to most aquatic and cell biologists under these terms and each term conveys a distinct image of a particular type of diatom valve.

Historically, centric and pennate diatoms have been separated as two classes or orders. Round et al. (1990), however, recognized three classes – Coscinodiscophyceae (centric diatoms), Fragilariophyceae (araphid pennate diatoms) and Bacillariophyceae (raphid pennate diatoms) – giving equal ranking to the raphid pennate diatoms [those with a slit opening (raphe) in the cell wall for movement] and the araphid pennate diatoms (those without this slit). However, this classification was not conceived to be phylogenetic (Round et al. 1990, p. 123). Simonsen (1972) clearly recognized that the traditional centric–pennate separation might not reflect an early evolutionary dichotomy (‘The Pennales . . . must have developed from the Centrales. But from which genus or genera of the Centrales the way leads to which genus or genera of the Pennales, we cannot tell’) and later illustrated a grade of clades from centrics into pennates (Simonsen 1979). However, his hypothesis was never translated into a formal classification. An alternative classification to that of Round et al. has been formally presented by Medlin & Kaczmarska (2004), based on molecular data. This separates the diatoms into two sub-divisions, comprising ‘radial centrics’ (Coscinodiscophyta), and those diatoms – ‘multipolar centrics’ and pennates – that exhibit polarity in the shape of their valves (Bacillariophyta), except where this has been lost secondarily. We discuss this below and will use the terms ‘centric’, ‘araphid’ and ‘raphid’ to refer to structural types within the diatoms, even though two of the groups of diatoms referred to are not monophyletic. Even stronger support for the three classes has also been presented (Medlin et al. 2004).

The diatoms have an extensive fossil record and a wealth of morphological features upon which to base their systematics and so they are an ideal group in which to study the congruence of molecular, morphological and fossil data sets. Molecular data have confirmed some previous ideas. For example, the idea (e.g. Simonsen 1979) that the pennate diatoms evolved from the centric forms because they appear later in the geological record (which had been noted also by Fritsch 1935, p. 642), is supported unambiguously by molecular phylogenies (see below, this section). Likewise, Simonsen’s view that centric diatoms are paraphyletic (though he did not use this terminology) is also clearly demonstrated by molecular data, which were presented first by Medlin et al. (1993). However, molecular analyses have also brought several surprises. The first molecular phylogenies (Medlin et al. 1993) showed the araphid pennate diatoms to be paraphyletic, undermining
the class-level distinction of Fragilariophyceae and Bacillariophyceae by Round et al. (1990). Furthermore, among the taxa studied by Medlin et al. (1993), the centric diatom, Skeletonema Greville was most closely related to the pennate diatoms, with high bootstrap support in phylogenetic analyses. Sørhaugg et al. (1995) confirmed that centric and araphid taxa were paraphyletic, using an analysis of partial sequences from the 28S large-subunit (LSU) rRNA coding region from eight diatoms. These initial data suggested, therefore, that existing higher-level diatom classifications were not phylogenetic and all subsequent analyses from several other genes have supported this finding (Medlin et al. 1996a, b, 2000; Ehara et al. 2000; Fox & Søraugg 2004).

Medlin and colleagues (literature summarized by Medlin & Kaczmarska 2004) have divided the diatoms into two groups on the basis of molecular sequence data: what initially was called ‘Clade 1’ contains those centric diatoms with essentially radial symmetry of valve shape and structure. ‘Clade 2’ consists of two groups, the first of which contains the bi- or multipolar centrics and the radial Thalassiosirales (‘Clade 2a’), and the second, the pennates (‘Clade 2b’) (Figs 1, 2). Morphological and cytological support for these clades was reviewed in Medlin et al. (2000) and Medlin & Kaczmarska (2004). However, whether Clade 2a is truly monophyletic is unclear (e.g. Kooistra et al. 2003a).

Clades 1 and 2 are now recognized at the subdivision level (Figs 1, 2), as the Coscinodiscophyta and Bacillariophyta, respectively, and Clades 1, 2a and 2b are now recognized at the class level, as the Coscinodiscophyceae, Mediophyceae and Bacillariophyceae (Figs 1, 2; Medlin & Kaczmarska 2004).

3.2. Correlations between molecular phylogenies and nonmolecular characteristics

There are correlations between the principal molecular clades and certain cytological features. For example, on the whole, the Bacillariophyta have a perinuclear arrangement of the Golgi apparatus, whereas in the Coscinodiscophyta the Golgi stacks are usually in Golgi-endoplasmic reticulum–mitochondrion (G–ER–M) units (Schmid 1988; available data summarized in Medlin et al. 2000, fig. 1). However, the correlation is incomplete, since Odontella C. Agardh (molecular Clade 2a) has G–ER–M units, whereas Melosira and Aulacoseira Thwaites (Clade 1) have perinuclear arrangements. There may also be correlations with pyrenoid ultrastructure and spermatoid ultrastructure (Medlin & Kaczmarska 2004), but support is weak and interpretation difficult because of the lack of data for most taxa. Schmid (2001) reviewed the structure of the pyrenoid in diatoms. In few cases were there unambiguous correlations with higher-level taxa. However, one important observation concerns the Thalassiosirales, a group that in molecular analyses falls within Clade 2, despite their radial symmetry and circular valves. Here, there are often no membranes traversing the pyrenoids, in contrast to true radial centrics (Clade 1), which have one or more sets of membranes crossing the pyrenoid, which are not continuous with the stroma thylakoids (Schmid 2001). Thus, one of the most surprising of results from rDNA and other sequence analyses — that the Thalassiosirales are not closely related to other diatoms with circular valves — receives some indepen- dent support. Again, however, the correlations are not complete. Fossil evidence summarized by Medlin & Kaczmariska (2004) suggests that the purported ancestors of Clade 1 had elaborate linking structures and no processes or tubes in the valve centre, whereas the purported ancestors of Clade 2 had reduced linking structures and a tube or process in the valve centre.

By far the best independent, nonmolecular support for three classes comes from auxospore structure (Kaczmariska et al. 2001; Medlin & Kaczmarska 2004). Isodiametric auxospores with scales are characteristic of Clade 1; anisodiametric auxospores with scales and hoops or bands (a properizonium) are found in Clade 2a; and anisodiametric auxospores that form a complex tubular perizonium, usually consisting of transverse hoops and longitudinal bands, are found in Clade 2b. Of course, these correlations would also be consistent with paraphyly of Clade 1 and Clade 2a.

3.3. The palaeoecological context of diatom origins

A molecular clock calculated from four different genes has placed the origin of the diatoms no earlier than 240 Ma (Kooistra & Medlin 1996). At this time, the Mesozoic (230–70 Ma), the world’s climate was generally warm with no polar ice-caps, a low pole-to-equator temperature gradient and high levels of CO2, giving a ‘greenhouse climate’ considerably warmer than today (Chambers 1996; Gale 2003). From the Permian to the Turonian stage of the Cretaceous (280–90 Ma), average global temperatures increased and worldwide (eustatic) sea levels rose (Douglas & Woodruff 1981; Vial & Hardenbol 1979).

The low pole-to-equator temperature gradient and high levels of CO2 (8–10 times that at present) meant that high latitude sea water was neither sufficiently cold nor saline to sink to form bottom water (Gale 2003). In the Cretaceous, ocean bottom waters were probably saline and derived from low latitude areas of high evaporation. During the late Cretaceous (c. 95–65 Ma), eustatic sea levels were very high and the interiors of the major continents were covered with shallow seas (Gale 2003). The rise in sea levels probably reached a peak during the Turonian (c. 90 Ma), when vast shelf areas were flooded, causing difficulties for subsequent palaeo-shoreline reconstruction (Haq et al. 1987). The early Albian (Lower Cretaceous) also saw at least three flooding events. From the Turonian to the Santonian (90–85 Ma), there was a period of stability when average global temperatures and sea levels stabilized before starting to fall again at the beginning of the Campanian (84–74 Ma) (Frakes 1986; Hancock & Kaufman 1979). The late Maastrichtian (66 Ma) was marked by a significant cooling of the oceans and the Cretaceous–Tertiary (K–T) event was accompanied by a sudden drop in the eustatic sea level and global temperatures. Frakes (1986) estimated that the drop in temperature might have been as much as 5°C worldwide. The fall in the eustatic sea level led to the disappearance of many epicontinental seas and subsequently some ocean basins (e.g. the Arctic Basin) became isolated from the rest of the world ocean. During this period there also was considerable volcanic activity producing volcanic ash, which was a source of soluble silica and enabled the preservation of diatom frustules in many areas.
Fig. 1. Phylogenetic reconstruction of the Heterokonta showing the major clades of the diatoms and the newly described subdivisions and classes, using maximum parsimony and the ARB database (www.arb-home.de) of 9000 small subunit (SSU) ribosomal DNA (rDNA) sequences (redrawn from Medlin & Kaczmarska 2004). The numbers for the classes and subdivisions of diatoms represent the posterior probabilities and the bootstrap values from neighbour-joining and maximum parsimony analyses as reported in Medlin and Kaczmarska (2005).

Fig. 2. Details of the small subunit (SSU) ribosomal DNA (rDNA) diatom tree reconstructed using a Bayesian analysis with Mr BAYES. Major clades are collapsed into triangles for clarity and to emphasize the relationships among the diatoms at the genus or higher level. This tree has been composed from all diatom sequences in Genbank as of January 2005 and a number of unpublished sequences from collaborators. Numbers on the nodes are the posterior probabilities from the Bayesian analysis. For the Bayesian analysis, we ran 1,000,000 generations, saving every 1000th tree, discarding the first 500 trees and making a consensus tree from the last 500 trees. We present a branch length tree having the lowest log likelihood and place the posterior probabilities on each node of this tree. An uncollapsed version of this tree was presented in Medlin et al. (2004).
Fig. 2. Continued.
4. DIATOM EVOLUTION IN THE MESOZOIC

4.1. Jurassic and Cretaceous: overview and fossil localities

Modern diatoms are rather different morphologically from the earliest known diatoms. The earliest generally accepted record of diatoms is of two species of *Pyxidicula* Ehrenberg, *P. lasius*ca and *P. bollensis* described by Rothpletz (1896, 1900) from the Liassic Boll shales of Wurttemburg, the Toarcian stage of the Jurassic (c. 190 Ma). We are unable to accept reports of Proterozoic (Precambrian) diatom fossils (Siemińska 1981; Kwiecin Âska 2000; Siemińska & Kwiecin Âska 2002), because the forms reported, such as *Nitzschia* and a naviculoid diatom (named *Horstia* by Siemińska & Kwiecin Âska 2002) that is similar to extant *Sellaphora* Mereschkowsky and *Mayamaea* Lange-Bertalot, are more modern in appearance than (and have no counterparts in) early Tertiary raphid diatoms, and that acceptance of the Precambrian fossils as autochthonous would require that the evolutionary changes and ‘progression’ so evident in the Cretaceous–Tertiary sequence are essentially meaningless. Likewise, we are unconvinced by Palaeozoic records of diatoms listed by Siemińska (2000), either because the flora is modern (e.g. Zanon’s data), or because there is no evidence that the fossils are diatoms (as with the Vologdin & Gapeev material). It also should be noted that the mean and maximum age of the diatoms calculated from molecular data, viz., 135–240 Ma (Kooistra & Medlin 1996; Medlin 1975), indicate a Mesozoic origin for the group.

The Jurassic *Pyxidicula* species recorded by Rothpletz (1896, 1900) occurred as isolated valves or less frequently as frustules among sponge spicules, the valves being 6 to 14 µm in diameter, with the frustules having two dissimilar valves and no girdle bands. Simonsen (1979) considered the primitive characters of diatoms to be spore formation, lack of girdle bands and heterovalvity. These he associated with a neritic habitat, i.e. an inshore marine environment.

Of the habitats of the late Jurassic and Lower Cretaceous diatoms, we know little. Recently, a new diatom deposit in Korea (175 Ma) has been reported from nonmarine sediments (Harwood et al. 2004). A hypothesis has been presented above as to how the presence of nonmarine, likely terrestrial diatoms might be reconciled with molecular data that indicate that the closest relatives of diatoms and most of the extant basal lineages of diatoms (in Clade 1) are marine (e.g. Medlin & Kaczmarska 2004). The diatoms in the Korean deposit are somewhat similar to the *Pyxidicula* species in Rothpletz’s Liassic deposit. However, few diatoms have been recovered so far from the Korean material and it is likely that they give a very incomplete picture of the flora.

The first well-preserved deposit of diatoms is from Aptian–Albian core material recovered from the Weddell Sea, Antarctica, which was studied extensively by Gersonde & Harwood (1990) and Harwood & Gersonde (1990); we illustrate some of these in Figs 3–34. Gersonde & Harwood (1990) identified 12 new genera and 22 new species, most of them cylindrical valves with similar chambered valve structure. However, among them was one, *Bilingua* Gersonde & Harwood, that has bipolar, nonchambered valves, showing that considerable diversification already had occurred in or by the Lower Cretaceous, possibly into Clades 1 and 2. Gersonde & Harwood (1990) proposed a neritic shallow water environment for the Weddell Sea flora and suggested that Lower Cretaceous diatoms may have been restricted to continental margins and inland seas.

Haig & Barnbaum (1978) retrieved lightly pyritized diatoms from the Aptian of the Wallumbilla Formation in Queensland and suggested they originated in a slightly hypsaline shallow sea with a high level of dissolved silica. The pyritized internal clasts are of benthic species; because a sieve was used, only those over 60 µm were recovered. Tappan (1962) has proposed that pyritized diatoms reflect rapid burial in turbid, inner, sublittoral environments, e.g. near a deltaic environment. Kemper et al. (1975) described an inshore environment affected by transgressions and regressions for the Albian diatoms described by Forti & Schulz (1932) from northwest Germany, and Foucault et al. (1986), in their paper on the lightly pyritized Albian-Cenomanian diatoms from the Italian Alps, suggested that ‘the progressive rise in abundance of these algae in the Cretaceous is tentatively linked to the flooding of the platforms by the epicontinental seas in this period’.

From the Santonian through the Campanian and Maastrichtian (c. 88–65 Ma), there is more information on the diatom flora and its environment, although only 12 sites have been studied extensively, all of them interconnected through seaways and oceanic currents (Chambers 1996). The Santonian and Campanian marked the beginning of the formation of vast diatomaceous deposits: as much as 500 m thick in the Ural Mountains and in Western Siberia (Glieser 1966; Strel’nikova 1974, 1975), c. 1400 m on Seymour Island, eastern Antarctic Peninsula (Harwood 1988), 1094 m in the Canadian Arctic (Tapia 1996), and 124 m in Dosados Canyon, Moreno, California (Nikolaev et al. 2001). Cores have been taken and studied from the South Pacific, on the southeast margin of the Campbell Plateau (DSDP Site 275) (Hajós & Stradner 1975; Chambers 1996); from the Alpha Ridge, Arctic Ocean (Kitchell & Clark 1982; Kitchell et al. 1986; d’Agnes & Clark 1994); from CESAR core 6, also in the Arctic Ocean (Barron 1985; Chambers 1996); ODP Site 758, in the equatorial Indian Ocean (Fourtandier 1991; Chambers 1996); and a dredge from the Tonga Trench (Quinterno et al. 1994; Chambers 1996). Sites also have been found in Japan (Takahashi et al. 1999) and in Israel (Moskhover and Harwood, 1988), but these specimens are either pyritized or poorly preserved.

The flora of both of the Russian deposits and that of the Marca Shale, Moreno, California, included a considerable number of benthic species plus some neritic species, suggesting a shallow nearshore habitat. Harwood (1988) proposed a shallow water habitat in relatively quiet water conditions on the continental shelf for the Seymour Island diatom flora of benthic and neritic species together with a considerable number of resting spores. The Arctic Ocean sites on the Alpha Ridge, F1/437 and CESAR 6 and those of Tapia (1996) showed strong seasonality and those samples from the Alpha Ridge display alternating layers of vegetative cells and resting spores, which have been interpreted as the diatoms’ response to Arctic summers and winters (no daylight!) respectively, at a time when no ice sheets were present in the Cretaceous oceans. These Cretaceous sites suggest that the diatoms had not yet moved into the open ocean but were restricted to shallow waters, with benthic, epiphytic and neritic species present.
No Cretaceous diatoms have been recovered from deep-water oceanic sites. All those studied and listed above would have been nearshore habitats in the Cretaceous seas.

4.2. Lower Cretaceous diatoms

Judging by valve shape, Medlin et al. (2000) and Medlin & Kaczmarska (2004) estimated that representatives of both Clade 1 and Clade 2 (Figs 1, 2) were present in the late Aptian to early Albian (115–110 Ma) floras described by Gersonde & Harwood (1990) from the Weddell Sea. However, four morphologies can be distinguished among these early diatoms (Gersonde & Harwood 1990). Of these, one does not contain any process in the valve centre (probably equivalent to molecular Clade 1), whereas another group does (probably equivalent to molecular Clade 2: Medlin & Kaczmarska 2004). We presume that the remaining two groups are extinct.

Gersonde & Harwood (1990) described 25 taxa placed in 13 genera. All are centrics but two of them are bipolar, Bilingua rossii (Figs 3–5) and Kerkis bipinosa (Fig. 9), both described as new by Gersonde & Harwood (1990) and both very rare. Some of the Weddell taxa are also present in the Albian material described by Forti & Schulz (1932) from northwest Germany and in Aptian material recovered from Queensland, Australia (Dun et al. 1901; Harper 1977; Haig & Barnbaum 1978; Nikolaev & Harwood, in press). Material described as Jurassic by Rüst (1885) also contained diatoms, described under the title coproliths, with four species described as tintinnids and two as diatoms and these have affinities with the taxa described by Gersonde & Harwood (1990). It has been suggested by Gersonde & Harwood (1990) that this material is probably Albian and might even be from the same locality as that studied by Forti & Schulz (1932).

Gersonde & Harwood (1990) found several new valve structures in the Weddell Sea diatoms. The most primitive feature, they suggested, is a large, heavily silicified type of linking spine, which is often ridged, triangular in cross-section, and buttressed at its base; it slots into a deep socket (a scrobisculus) on the opposing valve. The scrobisci (Figs 11–13, 15) are positioned at the margin of the valve face and on the marginal ridge when one is present. A similar structure can be found in modern genera belonging to Clade 1 (Stephanopyxis in Haga 1997). The taxa with this type of linking mechanism are dominant in their material and most of these heavily silicified species have cylindrical valves, elongated along their pervalvar axis (Figs 11–16). Lund (1966) suggested that a strong linking process in planktonic diatoms is an indication of turbulent waters. Gersonde & Harwood (1990) consider that this group of species reached their maximum at this period and largely became extinct somewhere between the Albian and Campanian.

Other new structures, not seen in diatoms from the late Cretaceous and Tertiary, are the ‘punctate-haplo’ process on the genus Microorbis Gersonde & Harwood (Figs 17, 18); the ‘scissurate-haplo process’, a process considered by Gersonde & Harwood (1990) to be the ancestral stage of a labiate process, on Trochus elegantulus Gersonde & Harwood (Figs 10, 19) and Bilingua rossii (Figs 3–5); the ‘perforate process’ on Praetallassiosiropsis haslea Gersonde & Harwood (Figs 20–22), which is possibly related to the annular or multistrutted process on Thalassiosiropsis Hasle in Hasle & Syvertsen; the ‘sipho-shaped process’ of Gladiopsis, subsequently interpreted as an annular or strutted process by Nikolaev & Harwood (1997) (Figs 11, 12, 29, 31); and the ‘rhyncho-shaped process’ present on Rhynchopyxis siphonoides Gersonde & Harwood (Figs 26–28), its internal structure differing little from that present in the sipho-shaped process of Gladiopsis. It may be that when those species of Gladiopsis described by Strel’nikova (1974) as ‘Gladius speciosus f. poratus’ and ‘f. aculeolatus’ (with exterior tubular openings from the process) are examined, the internal openings of their processes may be found to be identical with that of Rhynchopyxis.

Medlin et al. (2000) interpreted the process of Rhynchopyxis (Figs 26, 27) to be a forerunner of the central microblade process of Clade 2 diatoms. Gersonde & Harwood (1990) speculated that the group of apical processes in Rhynchopyxis may also have linked frustules into chains, possibly by producing chitin threads, as in present members of the Thalassiosirales, and they considered them more advanced because an apical process or a process with chitin threads uses less silica and therefore less energy in their formation. Gersonde & Harwood (1990) suggested that Microorbis, Trochus Gersonde & Harwood and Bilingua (Figs 17–19, 3–5, respectively) may have lived in deeper waters, because they are less heavily silicified. If Lund’s (1966) ideas about linking spines are correct, then those lacking well-developed spines, such as Praethallassiosiropsis Gersonde & Harwood, Gladiopsis and Rhynchopyxis, may have lived in less turbulent waters than the group with heavily silicified linking spines. All, however, appear to be shallow-water species – neritic rather than oceanic – because of the large number of resting spores present in the material (Harwood & Gersonde 1990). In Aptian material from Australia, Nikolaev & Harwood (1997) found a new process, the ‘uvular process’, in the genus ‘Arcaegladiaopsis’. To date, this material remains unpublished (Nikolaev & Harwood, in press).

In most Aptian–Albian species, the valves have pseudolocate areolae (Figs 11, 12, 15, 23–25) – a reticulate pattern of raised ridges, T-shaped in cross-section, with a basal layer perforated by radially aligned rows of pores that probably would have been occluded originally by external vela. Bilingua, Trochus and Microorbis have a simpler construction, consisting of a single layer perforated by radial rows of puncate striae. Many taxa have a marginal ridge and costae. Remarkably, girdle bands (rarely preserved in fossil material) are present in a few cases; they are perforated by vertical rows of fine pores (Fig. 28). Vela have been observed on the valve face of Basilicostephanus Gersonde & Harwood (Figs 13, 14). Other morphological features of note are the sulci of Bilingua rossii Gersonde & Harwood (1990, pl. 18, figs 7, 9); the girdle view of this diatom is reminiscent of Bidulphia Gray or Briggara R. Ross & P.A. Sims.

Of particular note also is a species of Stephanopyxis with rimportulae (labiate processes) found by one of us (P.A.S.) on BRM slide Zu3/98 (E.F. B36/C36) (Figs 6–8) and also in original material kindly supplied by Gersonde & Harwood (Figs 23–25). If autochthonous, this would be the earliest record of an extant diatom genus. D.M. Harwood (personal communication) has cautioned us that this may be a contamination from upper layers. However, no other contaminants were found and the valve structure of Stephanopyxis species is pseudoloculate, just as in most other taxa recovered from the
Weddell Sea but uncommon in more recent taxa. None of the other Weddell Sea diatoms seem to have rimoportulae but a rimoportula is illustrated on a pyritized specimen of *Aulacodiscus siemenskae* Geroch (Geroch 1978, pl. 3 figs 1–6) from the Lower Cretaceous of the Polish Carpathians. Foucault et al. (1986) identified a *Lithodesmium* Ehrenberg species, *Thalas-

siosiropsis wittiana* (Pantocsek) Hasle and *Sheshukovia uralense* (Strel’nikova) Gleser from partly pyritized specimens from the Albian–Cenomanian of the Italian Alps, but the dating of this material is suspect. Strel’nikova & Martirosjan (1981) also featured a triangular pyritized specimen from the Albian of the Stavropol region of the former U.S.S.R.

Figs 11, 12. *Gladius antiquus* (Forti & Schulz 1932).

- **Fig. 11.** Girdle view of fractured valve. Bar = 10 \(\mu\)m.
- **Fig. 12.** Valve face with tri-butressed linking spines and a scrobisculus into which a spine from corresponding valve fits. Scale bar = 5 \(\mu\)m.


- **Fig. 13.** Note marginal ring of tri-butressed spines with a scrobisculus beside each. Scale bar = 5 \(\mu\)m.
- **Fig. 14.** Detail of valve face with vela (? rotae). Scale bar = 2 \(\mu\)m.

**Fig. 15.** *Ancyclopyrgus reticulatus* (Gersonde & Harwood 1990) with single tri-butressed spine and scrobisculus. Note pseudoloculate valve structure. Bar = 5 \(\mu\)m.

**Fig. 16.** *Archepyrgus melosiroides* (Gersonde & Harwood 1990). Two linked valves with interdigitating spines, a poroid valve structure and a `ringleist' on inner valve margin. Bar = 5 \(\mu\)m.

Thus, by the end of the Lower Cretaceous, cylindrical, circular, bipolar and possibly triangular forms were present, possessing linking spines and six types of processes including the rimoportula. Thus, the ancestors of both Clades 1 and 2 almost certainly were present. Sulci, costae, a marginal ridge, a pseudoloculate and perforate valve structure also have been identified, together with signs of vela. The most noticeable feature of the Lower Cretaceous diatoms is the heavy silicification – the genus *Kreagra* is especially massively silicified (Figs 33, 34). Another is the small size of the cells, when compared to late Cretaceous–Eocene diatoms.

The rate of evolution in Clade 1 diatoms has been calculated using two different means of calibrating molecular trees: (1) by fossil dates for the entire clade (Kooistra & Medlin 1996); and (2) by biomarker compounds for the clade containing *Rhizosolenia* (Sinninghe-Damsté *et al.* 2004). Both methods suggest that these diatoms are evolving very quickly (1% per 24 and 14 Ma for the rRNA gene, respectively) and
Figs 17–25. Albian diatoms, Lower Cretaceous, Ocean Drilling Program Leg 113, Site 693, Weddell Sea.


Fig. 17. Valve exterior with marginal ridge, central hyaline area with punctate-haplo processes, and radial rows of areolae that extend onto the vertical mantle. Scale bar = 10 μm.

Fig. 18. Valve interior. Scale bar = 10 μm.

Fig. 19. *Trochus elegantulus*. Valve interior. Scale bar = 5 μm (see Fig. 10 for exterior).


Fig. 20. Valve exterior with central perforated process, marginal ridge with sipo-marginalis. Scale bar = 20 μm.

Fig. 21. External view of central perforated process. Scale bar = 5 μm.

Fig. 22. Internal view of central perforated process. Scale bar = 5 μm.

Figs 23–25. *Stephanopyxis* sp.

Figs 23, 24. Valve exterior with openings from rimoportulae. Scale bar = 10 μm.

Fig. 25. Detail of opening from rimoportula. Scale bar = 2 μm.
Fig. 26. External view with rhyncho-shaped process in central position on valve face. Scale bar = 10 μm.
Fig. 27. Fractured specimen showing internal opening of valve and rhyncho-shaped process. Scale bar = 10 μm.
Fig. 28. Detail of interior opening of rhyncho-shaped process. Scale bar = 1 μm.

Fig. 29. Valve exterior with opening from central ‘sipho-shaped’ (= multistrutted or annular) process within circular area of radiating rows of fine pores. Note small spines at margin of valve face. Scale bar = 5 μm.
Fig. 30. Fractured specimen with attached girdle band. Scale bar = 5 μm.
Fig. 31. Internal view of multistrutted process. Scale bar = 2 μm.

Fig. 32. Gladiopsis lagenoides (Gersonde & Harwood 1990). Valve exterior with radial rows of poroids on valve face and central opening from multistrutted process. Scale bar = 10 μm.

Figs 33, 34. Kreagra forfex (Gersonde & Harwood 1990).
Fig. 33. Large hyaline linking process with two broken spines and central linking notch. Scale bar = 10 μm.
Fig. 34. Fractured body of valve with rows of poroids. Scale bar = 2 μm.
this could help explain why the morphology of the diatoms changes so rapidly across the Cretaceous, between Lower and Upper Cretaceous floras.

4.3. Upper Cretaceous centric diatoms

By the end of the Upper Cretaceous, the diatoms closely resembled modern forms. Very few samples have been collected in the period between the end of the Albien and the late Santonian, some 20 Ma, and all of these, apart from one set of samples from Arctic Canada examined by Tapia (1996) and Tapia & Harwood (2002), are of pyritized diatoms. Figure 35 shows a modified diagram by P. M. Chambers & A. Holburn (personal communication) illustrating the types of palaeomorphs present. All are large specimens, most having been sieved, and they are probably benthic in habitat. Tapia (1996) showed that Bilingua, Basilicostephanus Gersonde & Harwood (Figs 13, 14), Gladiopsis and Stephanopyxis were present after the Albien and before the Campanian (late Cenomanian to Santonian, c. 95–80 Ma), but the other genera described by Gersonde & Harwood (1990) did not survive the Albien. Basilicostephanus occurred until the early Campanian; Bilingua did not extend into the Campanian; Gladius survived until the end of the Upper Cretaceous; and Stephanopyxis is found living today. Tapia (1996) and Tapia & Harwood (2002) identified several new genera in the period between the Albien and the Campanian. These included Corinna Heiberg, Costopyxis Gleser (Fig. 37), Hemiaulus Heiberg, Hyalodiscus Ehrenberg, Micrampulla Hanna, Paralia Heiberg, Porpeia Bailey ex Ralfs in Pritchard, Pseudopyxilla Forti, Stellarima Hasle & P.A. Sims, Trochosoria Kitton and Trochostrophus Gleser. Ross & Sims (1987) recorded Briggera includens (Grun.) Ross & Sims from Upper Turonian (c. 90 Ma) material from Oceanograph Canyon (Ross & Sims 1987; material provided by Fenner). Very elaborate linking spines are seen in Strelnikovia R. Ross & P.A. Sims (Figs 52, 53), in Campanian material from Russia (Strel’nikova 1974) and described by Ross & Sims (1987).

Valve structure became more diverse in the Campanian, with loculate and pseudoloculoculate constructions both present and also simple, single-layered valves of various thicknesses, perforated by poroids. The areola occlusions (vela) became refined into cribrar (Fig. 51), volae (Fig. 49) and rotae (e.g. in Basilicostephanus), depending on the size of the perforations, which are mostly larger in Upper Cretaceous species. The annular or multistrutted process (Figs 31, 32 and see Sims 1994) survived in a few taxa until the early Eocene (Thalassiosiropsis) and the rimoportula became the dominant process present on most centric and araphid diatoms from the Upper Cretaceous onwards.

The rimoportulae, which may have been restricted to submarginal locations in Albien Clade 1 diatoms, occupy many different positions on the valve face by the end of the Cretaceous – central or submarginal or marginal. They also can be single or multiple, and with the external openings either flush to the valve surface (as in Stellarima and Coscinodiscus) or raised into stout spines (as in Aulacodiscus Ehrenberg, Paralia and Actinoptychus Ehrenberg). Marginal rimoportulae dominate in Clade 1 diatoms, whereas a centrally located one is present in Clade 2 diatoms. Rimoportulae characters define most modern centric diatoms at the family or generic level and by the end of the Cretaceous, most modern centric genera or their immediate ancestors can be found. Notable exceptions are the Thalassiosirales, the Hemidiscaceae, and the Chaetocerataceae (Simonsen 1979).

In valve view, the outlines of Upper Cretaceous diatoms are linear, lanceolate, triangular, quadrangular, circular and subcircular to solenoid (Fig. 39), with triangular, circular and bipolar shapes dominant. Clade 2a of the diatoms contains the bipolar centrics, the pseudo-araphid pennates (see section 4.4) and the radial Thalassiosirales, all of which now have been grouped into the new class Mediophyceae. Other features of the bipolar Mediophyceae present in the Upper Cretaceous are the elevations and undulations of the valve surface in hemialoids and biddulphioids (Figs 42–44). Internal costae are present, both in Arachnoidiscus Deane ex Shadbolt (Clade 1) (Figs 48, 49) and Eunotogramma Weisse (Clade 2a); in both, the costae are presumably strengthening devices. Heterovalvly also occurs in Arachnoidiscus (Figs 48, 49). The Upper Cretaceous diatoms mostly were larger in size than their Lower Cretaceous counterparts and the features discussed above suggest adaptations to new habitats and that a greater proportion

}\n
\"Sims et al.: Evolution of the diatoms 377\n
\"
Fig. 35. Time ranges of pyritized diatoms with basic illustrations of selected morphotypes from their assemblages. Redrawn from P.M. Chambers and A. Holburn, unpublished data.
Figs 36–44. Upper Cretaceous diatoms.

Fig. 36. *Meretosulus* sp. Ust’Manya River, Severnaya Sos’va, eastern slope of Urals, Russia. Campanian. Valve exterior with girdle bands and an ocellus at each pole. Scale bar = 20 μm.

Fig. 37. *Costopyxis* cf. *schulzii* Steinecke. River Synya, Til’tim, eastern slopes of Polar Urals. Campanian. Note long, hollow tubular processes from central rimoportulae, broad valve margin and raised ridges between groups of areolae. Scale bar = 20 μm.

Fig. 38. *Syndetocystis uralensis* Strel’nikova. Core 82, Ust’Manya, Severnaya Sos’va, Campanian. Note resemblance to Lower Cretaceous diatoms in mantle structure (see Harwood & Gersonde 1990). Single hollow process does not appear to be a rimoportula from interior view (not shown). Scale bar = 20 μm.


Fig. 39. Fractured body of valve. Scale bar = 20 μm.

Figs 40, 41. Tip of valves with linking apparatus. Scale bar = 2 μm.

Figs 42–44. *Biddulphia tuomeyi* var. *cretacea* Joussé. Core 82, Ust’Manya, Severnaya Sos’va, Campanian.

Fig. 42. Central area with rimoportulae. Scale bar = 10 μm.

Fig. 43. Elevation at pole with pseudocellus. Scale bar = 10 μm.

Fig. 44. Valve exterior with four sulci, marginal ridge, deep mantle and two elevations. Scale bar = 20 μm.
Figs 45–53. Upper Cretaceous diatoms.

Figs 45–47. *Aulacodiscus pugnalis* Hanna. Moreno Gulch, Fresno County, California. Maastrichtian.

Fig. 45. Exterior of frustule, valve face with four submarginal rimoportulae extending as large hollow spines and complete set of girdle bands. Scale bar = 20 μm.

Fig. 46. Valve interior with horse-shoe-shaped opening from rimoportula. Scale bar = 2 μm.

Fig. 47. Break through valve face showing loculate areolae with bullulate walls. Scale bar = 2 μm.

Figs 48, 49. *Arachnoidiscus* sp.

Fig. 48. Core 82, Ust’Manya, Severnaya Sos’va, Campanian. Valve interior with central ring of rimoportulae, primary and secondary costae extending toward valve centre. Scale bar = 20 μm.

Fig. 49. Akita, Japan. Miocene. Showing exterior openings of rimoportulae and areolae occluded by vela. Scale bar = 2 μm.


Fig. 50. Exterior of fractured specimen. Scale bar = 20 μm.

Fig. 51. Loculate areolae with cribra to exterior and foramina to interior. Scale bar = 2 μm.


Fig. 52. Two linked valves. Scale bar = 20 μm.

Fig. 53. Two valves linked by interlocking linking spines. Note pseudocelli at poles. Scale bar = 5 μm.
of the flora had adapted to live as individuals rather than in colonies.

4.4. The evolution of pennate and pennate-like diatoms

One of the major changes in the Upper Cretaceous is the appearance of pennate diatoms, which are found first in the Campanian (75 Ma) (Figs 54–73). Molecular clock calculations support this data (Kooistra & Medlin 1996). Mann (1984) identified three main types of pattern centres in diatoms: (1) the annulus of centrics; (2) the plain, rib-like sternum of araphid pennates; and (3) the complex, laterally asymmetrical raphe-sternum of raphid pennates. Molecular data show unambiguously that diatoms with raphes evolved from pennate diatoms with simple sterna, and that the first pennate diatoms evolved from a lineage of bipolar centric diatoms (Fig. 2: Kooistra et al. 2003a; Medlin & Kaczmarska 2004), so that the sternum of araphid pennates either was transformed from, or replaced, an annulus (Mayama and Idei 2004). However, there is no substantial fossil record of the key evolutionary transitions from centric to pennate, and from araphid pennate to raphid pennate. Morphogenesis may give a clue and Mayama & Kuriyama (2002) have documented what appears to be the fusion of many annulus-like elements during valve formation in an araphid pennate. However, none of the basal lineages of araphid pennates have been investigated with respect to valve ontogeny.

The advantage of the sternum appears to be physical, distributing mechanical stress along elongate valves more efficiently than an annulus and its radiating ribs. It permits the formation of highly elongate but relatively lightly silicified frustules. It may be no accident that centric diatoms with very elongate valves – the thick-walled, cylindrical Gladus and Gladiopsis – became extinct not long after pennate diatoms (Sceptroneis Ehrenberg, Incisoria Hajos) first appear in the fossil record in the Upper Cretaceous (e.g. Hajos & Stradner 1975). However, two other neritic centric diatoms appear to fill the same ecological niche in the plankton, namely Pyxilla Greville, an unrelated and successful genus with very elongate and heavily silicified valves, which made its first appearance in the early Eocene (Homan 1991; Fenner 1994), and the much less heavily silicified Proboscia, also with elongate valves and already present in the Upper Cretaceous (and still extant today).

Within the bi- and multipolar centric diatoms of the Class Mediophyceae, several have modified pattern centres that prompt speculation about how the sternum may have evolved. In Papillocellulosa Hasle, von Stosch & Syvertsen (Cymatosirales), for example, there is a circular annulus at the centre of the elongate valve; however, instead of radiating ribs as in most other centric diatoms, the valve has a set of parallel ribs, formed by the repeated branching of just two of the ribs borne by the annulus (Round et al. 1990, pp. 310, 311); it is tempting to see these two dominant ribs as possible precursors to the sternum. In Odontella sinensis (Greville) Grunow (Pickett-Heaps et al. 1990, fig. 40c) and Attheya species (Round et al. 1990), on the other hand, a sternum-like structure is formed by modification of the annulus itself, which is so highly elongate that its sides almost fuse. However, none of these living diatoms are in any sense ancestral to the pennates and their morphology must be regarded as providing interesting analogies, rather than having any direct phylogenetic significance.

Within the pennate diatoms, the sternum varies in width (it is particularly narrow in some forms, such as Fragilariforma D.M. Williams & Round or Asterionella Hassall, but wide in Delphineis G.W. Andrews or Staurosirella D.M. Williams & Round) and thickness (unthickened in e.g. Tabellaria Ehrenberg ex Kützing, internally prominent in Podocystis Bailey). At one or both poles, the sternum is often closely associated with a rimoportula (= labiate process). However, no significant evolutionary trends can yet be distinguished among araphid pennate diatoms in the organization of the sternum, rimoportula and ribs.

The first pennate diatoms to appear in the fossil record, for example Sceptroneis (Figs 54–56, 61–65) and Incisoria (Hajós & Stradner 1975), have a simple central sternum, uniserrate striae containing large pores, and rimoportulae positioned at one or both poles. Some have an apical pore field, which was presumably involved in mucilage secretion. These early pennate diatoms mostly belong to the Rhaphoneidaceae (sensa Round et al. 1990). Rhaphoneis itself is said to be present in Upper Cretaceous material (e.g. Strel'nikova 1974) and this genus is also among the first clade of pennate diatoms to diverge among the araphid diatoms in the rRNA tree, making the araphids paraphyletic (Fig. 2).

Some araphid pennate taxa are planktonic and colonial, whereas others are benthic, and the ancestral state is unclear. Extant members of the Rhaphoneidales are benthic, but the related Asterionelloides Round in Round, R.M. Crawford & D.G. Mann is planktonic. The evolution of araphid pennate diatoms, some with highly elongate valves, must have had profound consequences for the ecology of benthic habitats, because of the ability of these organisms to colonize surfaces and cover them with a dense ‘palisade’ of cells, each attached at one end by a pad.

Mann (1984) noted the existence of a few other types of pattern centre in diatoms besides the three main ones (annulus, sternum and raphe-sternum). These include the pattern centres of Ardissondea De Notaris and its allies, Toxarium Bailey, Psammodiscus Round & D.G. Mann, the eunotioid diatoms (a raphid group: see section 5.5), and a few others. Psammodiscus, though it has circular valves like many centrics, appears not to have an annulus and Round & Mann (1980) and Mann (1984) suggested that it is in fact a pennate diatom related to Rhaphoneis, but with a highly reduced, boss-like sternum. The main evidence for this, besides the absence of an identifiable annulus, is the areola structure and occurrence of other Raphoneis relatives with reduced sternum. However, Psammodiscus has not yet been subjected to a molecular phylogenetic analysis.

Determining exactly which genera should be included in the araphid pennate group is problematic. For example, Round et al. (1990) placed the Plagiogrammaceae in the centric diatoms because, despite the presence of a sternum-like structure along the midline of the valve, there also are well-defined areas of pores resembling the ocelli of some biddulphioid centrics; furthermore, the circular central boss of Plagiogramma Greville and Glyphodesmis Greville could be plausibly interpreted (in the absence of ontogenetic data) as reflecting the presence of an annulus. However, Kooistra et al. (2004) have
used 18S rDNA data to show that at least one genus of Plagiogrammaceae is in fact an araphid pennate diatom.

In contrast, Round et al. (1990) classified Ardissonea, Climacosphenia Ehrenberg and Synedrosphenia (H. Peragallo) Azpeitia Moros as araphid pinnates. All of these are large, highly elongate diatoms, but there is no central sternum, despite their 'pennate'-like appearance. Instead, there appear to be two lateral sterna, subtending transapical ribs and striae on either side. In Climacosphenia and Synedrosphenia, the lateral sterna appear to be continuous with each other at least at one
pole, so that together they may form an elongate annulus. This, however, differs from the normal centric annulus in that it subtends ribs and striae inwardly as well as outwardly. Mann (1984) therefore referred to it as a ‘bifacial annulus’. So far, no sequences are available from *Ardissonea, Climacosphenia* and *Synedrosphenia*, but 18S rDNA data have been obtained for *Toxarium*, which can be tentatively allocated to the same group as *Ardissonea*, on the basis of a similar mechanism of motility and girdle structure, the absence of rimoportulae, and absence of an axial sternum (Kooistra et al. 2003b). The 18S rDNA data indicate that *Toxarium* is not a pennate diatom but belongs near *Lampriscus* in the bi/multipolar centrics (Fig. 2; Kooistra et al. 2003a, b). If this is correct, and if *Toxarium* is indeed closely related to *Ardissonea*,
then the two lateral sterna (bifacial annuli) of *Ardissonea*, *Climacosphenia* and *Synedrosphenia* must be nonhomologous with the sterna of true araphid diatoms, though they might fulfill the same mechanical function, allowing the formation of highly elongate valves (some species in all of these genera produce valves over 500 μm long, and *Climacosphenia* cells are some of the largest of all bipolar diatoms, being up to 1.3 mm long: Hustedt 1931, 1932). Kooistra et al. (2003b) proposed that the araphid diatoms should be characterized by the ‘possession of a single sternum, subtending parallel ribs and intervening rows of pores; apical (rarely subcentral or axial) labiate processes, or both’. Most araphids also have distinct pore fields (pseudocellus) at one pole or both, but these are lacking in the *Ardissonea–Toxarium* group. Diatoms with bifacial annuli are present in the late Eocene at Oamaru (e.g. Schrader 1969, pl. 10, fig. 5), but their earlier fossil history is unclear. There is a chain-forming, elongate pennate-like diatom with scattered poroid areolae in the Maastrichtian (Upper Cretaceous) of the Moreno Shale (Figs 66, 67), which could possibly be a species of *Toxarium*.

Chambers (1966) lists 21 araphid pennate species as present in the Campanian/Maastrichtian (Upper Cretaceous), 11 of which were illustrated and placed in a genus but were given no description or specific name. Since then, an additional four species have been identified by Tapia (1996). These species have been placed in the genera *Sceptroneis* (12 sp.), *Incisoria* (5 sp.), *Rhaphoneis* (7 sp.) and *Grunowiella* Schmitz in Engler & Prantl (1 sp.). All have a central sternum, and a rimoportula and a pore field at one or both poles. All these genera are simple araphid genera whose low morphological diversity and shared features might suggest they had evolved recently, probably in the poorly documented time between the Albian and Santonian, specimens from which are unlikely to be recovered in sieved pyritized gatherings. Araphid pennes can be abundant in Upper Cretaceous deposits. Harwood (1988) gave percent abundances of *Incisoria*, possibly a benthic and epiphytic genus, of as much as 4% for *I. inordinata* Hajós, 3% for *I. lanceolata* Hajós & Stradner and 2% for *Incisoria* ‘sp. A’.

Also present in the Upper Cretaceous gatherings are some bipolar diatoms whose affinities to centric and pennate line-
Fig. 74. Phylogenetic reconstruction of the Thalassiosirales, showing the major clades of the order: Bayesian analysis of the ARB database of 9000 small subunit (SSU) ribosomal RNA (rRNA) sequences. Redrawn from Kaczmarska et al. (2006).
ages are unclear. They have a bipolar, highly elongate shape but lack a single, discrete sterna, and we list them here to indicate the diversity of pennate-like diatoms already present by the end of the Mesozoic. They include the following species: *Tabularia* (= *Tabulariella*) *antarctica* Hajós, *Longinata acuta* Hajós, *Helminthopsis (= Helminthopedella*) *wornardtii* Hajós, *Euotomogramma producta* var. recta Long, Fuge & Smith (Long et al. 1946; Harwood 1988) and *Drepanotheca bivittata* (Grunow & Pantocsek in Pantocsek) Schrader (= *Euotomogramma bivittata* Grunow & Pantocsek in Pantocsek and *E. margino-punctatum* Long, Fuge & Smith). Included in our illustrations (Figs 68–71) is an undescribed *Tabellaria*-like diatom (lacking a sterna) from the Maastrichtian of the Moreno Shale, which has also been figured by Chambers (1996) from the Maastrichtian of DSDP site 216-30-04-27/29. Chambers placed it in the centric genus *Anaulus* Ehrenberg and both this diatom and *Drepanotheca* (Figs 68–73) are indeed probably centric diatoms related to *Anaulus* or *Euotomogramma*. However, the generic limits and affinities of these genera require examination and probable revision.

5. DIATOM EVOLUTION IN THE CENOZOIC

5.1. New habitats after the Cretaceous

The Cenozoic (c. 66–16 Ma) has been a time of major change in the global climate, the distribution of the continents, land surface area and oceanic basins. Throughout the Cenozoic, plate tectonic processes have both destroyed pathways of oceanic circulation and created new ones, and changes in temperature have caused major flooding events and recessions (Pickering 2003). Commencing at 37 Ma (late Eocene), there was a transition from the generally warm climate of the Cretaceous, Palaeocene and early Eocene, with ice-free oceans, to an Oligocene climate with continental ice-sheets and/or glaciers reaching sea level (Barron & Baldauf 1989; Miller et al. 1991). By the late Miocene to Pliocene (c. 23–5 Ma), permanent Antarctic ice-sheets had been established, accompanied by strong cooling (Barron & Baldauf 1995). These global changes are correlated with a decline in many old genera and families of diatoms and the appearance of many new ones, adapting to the new environments, both on land and in the seas.

No major extinction events occurred during the Cenozoic, except at the beginning, at the Cretaceous–Tertiary (K–T) boundary (c. 65 Ma), when a distinct drop in global temperature (by as much as 5°C according to Frakes 1986) was associated with a sudden drop in eustatic sea levels and the loss of many epicontinental seas created during the Turonian–Campanian. Extinction of diatom taxa occurred at the K–T boundary, but this was not as savage as in many other groups of marine biota. Perhaps the ability to form dormant resting stages is more widespread in diatoms than is currently thought (available data were reviewed by McQuoid & Hobson 1996), enabling taxa to withstand long periods of unfavourable conditions. Harwood (1988) estimated that 84% of the species survived in a section from the Sobral Formation (Campbell Plateau, N. Seymour Island, Antarctica), but Chambers (1996), using a wider selection of material, estimated it to be far less, at 37%. The majority of K–T survivors were indeed neritic species, capable of forming resting spores. Benthic species suffered a severe decline in numbers that lasted through the early Palaeocene (c. 65–60 Ma), during which very few new diatom species evolved (Strel’nikova 1990). Hollis et al. (1995, 2003), however, found that siliceous phytoplankton thrived in the earliest Palaeocene of Marlborough, New Zealand, where there was no radiolarian mass extinction but instead an abrupt increase in the number of diatoms relative to radiolarians, causing persistent plankton blooms over the first million years of the Palaeocene. They attributed this to enhanced upwelling or a more efficient recycling of nutrients. MacLeod & Huber (1996) suggested that the cooler climate, beginning in the mid- to late Maastrichtian, led to a reorganization of deep water circulation, whereas Haq et al. (1987) believed that the rapid sea-level fall at the end of the Upper Cretaceous was followed by similarly rapid transgressions, creating an expanded zone of relatively shallow waters and hence a more efficient recycling of nutrients. Sea temperatures rose from the mid-Palaeocene (MacLeod et al. 2003) and there is evidence that, at the end of the Palaeocene into the early Eocene (c. 55–50 Ma), temperatures rose further still, causing a steady rise in eustatic sea levels, which, in turn, led to the creation of many inland seas. It was in and just after this period that a large number of new taxa evolved, especially benthic species, in the shallow waters of the inland seas.

Several major events and innovations seem to have occurred during Cenozoic diatom evolution: the invasion of freshwaters, the evolution of the Thalassiosirales, and the appearance of rapidly motile diatoms, following evolution of the raphe system. We will deal with each of these in turn, after reviewing changes in the marine flora of centric diatoms and araphid pennates.

5.2. Development of the marine centric and araphid pennate flora

In the Cenozoic, the fossil record suggests that diatoms moved from shallow shelf waters into deeper oceanic waters. However, there can be a problem in interpretation throughout the marine record, because of the loss of oceanic crust through subduction. Possibly, the increased representation of pelagic species in Cenozoic material is only an artefact of tectonic processes. Fenner (1991) described Palaeocene (c. 65–55 Ma) assemblages from four sites, three in the Antarctic and one in the Indian Ocean, in which resting spores are rare and chain-forming species of the genera *Hemiaulus*, *Sheshukovia* Gesler, *Medlinia* Sims (then listed as *Triceratium* Ehrenberg), *Stephanopyxis*, *Sceptroneis*, *Grunowiella* and *Trochosira* are dominant. Fenner also recorded *Chaetoceros* spp., possibly the first recorded appearance of this genus, which she regarded as an indicator of oceanic environments. Her sites were probably in upwelling areas. In the early Eocene (c. 55–50 Ma), Homann (1991) and Fenner (1994) recorded the genera *Sticodiscus* Greville, *Craspodiscus* Ehrenberg, *Rutillaria* Greville, *Cymatosira* Grunow, *Pseudotriceratium* Grunow, *Pseudostictodiscus* Grunow *et A. Schmidt* and *Pyxilla* from the Fur Formation, Denmark; some forms were free-living and others chain-forming, and most probably lived in the mero-plankton. Of interest is Fenner’s record of *Aulacoseira granulata* (Ehrenberg) Simonsen and a possible *Meridion* C.
Agardh, both earlier records than their first appearance in the fossil freshwater flora.

Late in the early Eocene, the centric family Asterolampracea appeared (Gombos 1980), apparently evolving from a member of the Coscinodiscaceae; like the Coscinodiscaceae, Asterolampracea are marine phytoplankton. The function of their conspicuous tubular rays is unknown; possibly they have a role in flotation, or they may be chambers for symbiotic organisms. By the middle Eocene (c. 50–45 Ma), the number of diatom species had reached a peak; it appeared to be a time of experimental development and the middle to late Eocene saw an explosive diversification, during which many bizarre genera evolved (Gombos 1980), many of them monospecific, e.g. Tubaeformis Gombos.

The Palaeocene and Eocene (c. 65–40 Ma) marine floras were dominated by relatively robust centric genera, e.g. Hemiaulus, Trinacria Heiberg, Sheshukovia, Triceratium, Stephanopyxis and Pyxilla. By the end of the Eocene and the early Oligocene, however, many features of the Upper Cretaceous and early Palaeocene (c. 60–35 Ma) floras had disappeared (e.g. Trinacria at the end of the Eocene and Pyxilla by the end of the Oligocene, Jousé 1978), and in the colder waters of the early Oligocene, these genera were gradually replaced by smaller and more delicate genera of centrics and araphid pennates, e.g. Coscinodiscus, Cestodiscus Greville, Azpeitia Peragallo in Tempère & H. Peragallo, Thalassiosira, Cavitatus D.M. Williams (formerly in Synedra Ehrenberg), and Thalassium (H. Peragallo & Peragallo) Peragallo. The more rounded outlines of early Cenozoic diatoms were replaced by more elongate forms (Barron & Baldauf 1995), both in the ocean and in neritic environments.

In the middle–late Eocene, new genera of bipolar centrics evolved. Most species were short-lived and some are used as stratigraphic marker species. The new genera included Bogorovia Jousé, Cladogramma Ehrenberg, Clavularia Greville, Clavicula Pantocsek, Cussia Schrader, Helminthopsisidella Silva, Lamella Brun, Leptoscaphos Schrader, Leudageria Tempère ex Van Heurck and Rossiella Desikachary & Maheshwari, and new species of other genera also appeared, e.g. of Tubariella Silva (first recorded in the Maastrichtian, see Schrader 1969). The pennate-like genera Arideassocia De Notaris and Toxarium Bailey, both placed under Synedra, also made their first appearance (see section 4.4).

After the Cretaceous–Tertiary extinction, the number of true araphid taxa and their significance in the neritic flora (meroplankton) increased significantly. Chambers (1996) recorded a further 17 species in the Danish (Palaeocene–early Eocene) and Strel’nikova (1990) estimated that araphids formed as much as 4% of the total species numbers in the Palaeocene. Sceptroneis and Grunowiuella were the most common araphid diatoms in the meroplankton and possibly in oceanic plankton. Fenner (1991) noted the presence of Grunowiuella palaeoceanica Jousé and related species in considerable numbers in faecal pellets in the Palaeocene, indicating very fertile, possibly upwelling conditions in the South Atlantic and Indian Oceans. The plankton at this time was dominated by chain-forming species. Grunowiuella, in particular, flourished in the Palaeocene, increasing from one species in the Cretaceous to five species (8 taxa) in the Danian. A single species could form 2% or more of the total diatom population.

Gombos (1976) recorded Cavitatus (Synedra) joueana (Sheshukova) D.M. Williams ‘var. B’ as common and confined to the late Palaeocene, and Homann (1991) identified the benthic genera Opephora Petit and Rhaphoneis, (probably misidentified because it is not included in Fenner’s 1994 survey of the same formation) as present in the early Eocene of the Fur Formation. Thus there were only four or five araphid genera recorded from the Upper Cretaceous to early Eocene (c. 70–55 Ma) and it was not until the late Eocene that additional araphid genera appeared. However, by the Oligocene (c. 38–22 Ma), Strel’nikova (1990) estimates that the araphid genera formed 33% of the total population. It has to be remembered, however, that very few depositional sites have been found in the early and middle Palaeocene, most of these being deep-sea cores, which could distort figures, particularly of benthic species. It was a period of relative stability and low evolutionary rates (Strel’nikova 1990), which changed when temperatures rose in the middle Palaeocene, leading to a rise in the eustatic sea levels and major flooding events in the late Palaeocene–early Eocene, when vast inland areas were inundated. This event lasted some six million years and in the middle Eocene (c. 50–45 Ma) a distinct change in the species composition took place. At least five new araphid genera appeared in the late Eocene (c. 45–40 Ma) (see Desikachary & Sreelatha 1989; Edwards 1991): Entopyle Ehrenberg, Grammatophora Ehrenberg, Plagiochroma, Rhabdonema Kützing and Fragilaria Ljungby. All, except possibly some Fragilaria, are and probably were benthic, epiphytic genera. All of the genera (except Fragilaria) have septa or other bar- or partition-like thickenings of valve or girdle (apparently for strengthening the frustules) and fall into a monophyletic clade (Fig. 2). Some of the earlier araphid genera were meroplanktonic (neritic), and possibly oceanic, but some – Rhaphoneis, Opephora (1 species) and Sceptroneis – were benthic. All the new araphid genera have an elaborate girdle band structure, so that the girdle axis became considerably broader than that of the valvar axis, and often form flat ribbon-like or zig-zag chains that lift the cells well above the substratum. We speculate that this enabled taxa with narrow bipolar valve faces to present a greater area for photosynthesis in benthic niches. These diatoms filled a niche previously unoccupied by araphids and all appear to evolve very slowly, with many of the species looking remarkably like some living today. Similar diatoms (e.g. Tetracyclus Ralfs, Tabellaria), again with septa, evolved later, independently, in freshwater environments.

In the plankton, changes took place as cooler conditions arrived in the later Cenozoic. Incisoria and Grunowiuella appear to have become extinct in the early Eocene, but lose generic definitions, difficulties in identifying genera with so few characters, erosion of specimens and changes in nomenclature may mean that these genera persisted longer. For example, Schrader & Fenner (1976) reported a planktonic araphid genus Neuromonas († Rhaphoneis) and all appear to evolve very slowly, with many of the species looking remarkably like some living today. Similar diatoms (e.g. Tetracyclus Ralfs, Tabellaria), again with septa, evolved later, independently, in freshwater environments.
Eocene–Oligocene before decreasing from the Miocene onwards, until possibly only one species survives today.

Cavitatus jouseana became frequent to dominant in the plankton of the Oligocene and Miocene (Akiba et al. 1993), whereas the important needle-like planktonic genera of modern seas, Thalassiothrix Cleve & Grunow and Thalassionema (which constitute a monophyletic clade in molecular analyses: Fig. 2), first appeared in the early Oligocene and early Miocene, respectively, and then formed a major constituent of the marine phytoplankton. There was a concomitant increase in species numbers of the centric genera Rhizosolenia, Chaetoceros and members of the Hemidiscaceae (Fenner 1985).

There is a good Cenozoic record of the Rhaphoneidaceae and this was analysed in a series of papers by Andrews (1970, 1975, 1977, 1986, 1990; Andrews & Stoelzel 1982; Andrews & Rivera 1987). The Rhaphoneidaceae is one of the earliest araphid pennate families, recorded first in the Upper Cretaceous, and it is also one of the most basal araphid pennate lineages according to rDNA data (Fig. 2). In a revision of the genus Rhaphoneis, Andrews (1975) considered that the genus was present from the late Eocene (c. 45–40 Ma) onwards and flourished in the late early Miocene until the late Miocene, with 20 species identified within this period. However, as noted earlier, Strel’nikova (1994) lists Rhaphoneis as being present already in the Cretaceous, but the specimens illustrated (pl. 53, figs 5–7) are oval to ovoid in outline without any central sternum but with a large central area, radial rows of areolae and no restricted areas of pore fields. Andrews (1975) removed two species from Rhaphoneis and a single species from Dimeregramma Ralfs in Pritchard, and, together with several new species, erected the genus Delphineis (Andrews 1977, 1986), the earliest record of which is in the early Miocene (c. 20–15 Ma). In 1990 Andrews removed further fossil species from Rhaphoneis, creating the genus Lancinieis, each species of which has a restricted stratigraphic range, from the early Miocene to the Pliocene. Dickensoniaformia (Scherer 1997), another genus placed in the Rhaphoneidaceae (sensu Round et al. 1990) is restricted to the Oligocene, possibly only in the Norwegian Sea.

Perissonoë (Andrews & Stoelzel 1982) is found from the middle Miocene onwards (c. 15 Ma–) and has radial symmetry, a triangular or quadrangular valve face, and three or four sterna; otherwise it has morphological features similar to Rhaphoneis and Delphineis. Living genera described as belonging to the Rhaphoneidaceae are Adoneis Andrews & Verata (1987), Diplomenora Blazé (1984), Neodelphineis Takanon (1982) and Psammodiscus (Round & Mann 1980). All these genera have narrow mantles, areolae with rotae and ri-moporulae. All occur in shallow marine waters, either free or attached to the substratum. Delphineis species, however, have adapted to brackish waters, shallow shelf waters and possibly ranged into deeper waters.

The evolution of marine Fragilariophyceae (sensu Round et al. 1990) remained almost static throughout the late Eocene to the end of the Miocene and, in molecular analyses, are distinct from their freshwater counterparts (Fig. 2). Long-term stasis in ecological assemblages is increasingly being inferred from fossil records (DiMichele et al. 2004). In the Oligocene, Pseudodimerogramma Schrader (in Schrader & Fenner 1976) and Ikebea S. Komura (in Barron and Mahood 1993) made their first appearance and in the Miocene, Dimeregramma Ralfs (in Schrader & Fenner 1976; Rehákova 1980) and Podocystis J.W. Bailey (Reinhold 1937). In more recent times, a fossil species Licmophora C. Agardh has been identified by D.M. Williams (personal communication). Both Licmophora and Podocystis are epiphytic or epilithic and are attached by means of a substantial mucilage stalk (which sometimes is branched), an adaptation that enables taxa to use space and light away from the substratum (host) to which they are attached.

5.3. Invasion of freshwaters

Following the Cretaceous flooding events, sea levels rose again in the Paleogene (in the early Paleogene, early Eocene, and early Oligocene) flooding inland areas, freshwaters, and numerous inland seas. It appears that species of several recently evolved lineages of araphid and raphid pennate diatoms, and later also members of the Hemidiscaceae (Actino-cyclos Ehrenberg), adapted to lower salinity in inland waters. It is noticeable that the older centric diatom lineages failed to adapt (e.g. Strel’nikova & Lastivka 1999), apart from Anaulus (P.A. Sims, unpublished observations) and ‘Archeopyrgus’, if the freshwater Aulacoseira evolved from the Lower Cretaceous marine genus Archeopyrgus, as suggested by Harwood & Nikolaev (1995). The distribution of genera of extant diatoms between marine, brackish and freshwaters has been analysed by Mann (1999a) and shows clear evidence of systematic bias in the ability of lineages to transgress the marine–freshwater boundary.

Most diatomists believe the freshwater diatom flora evolved from marine taxa invading freshwaters, largely because the records of fossil marine diatoms extend some 65 Ma before those of the freshwater (VanLandingham 1967; Skabichevskij 1981; Krebs 1994; Bradbury & Krebs 1995; Mann 1999a; Strel’nikova & Lastivka 1999). VanLandingham (1967) suggested that there are only a few genera that are known to have evolved wholly in nonmarine waters, e.g. Neidium Pfitzer. Bradbury & Krebs (1995) considered that ‘the evolution of marine diatoms may have played a crucial role in lacustrine diatom evolution by supplying new taxa with the potential of colonising continental environments’. Examples given were those of Actinocyclus and Thalassiosira, both originating in the marine environment and experiencing a burst of species radiation before moving into freshwaters.

There were six major flooding events in the period between the Upper Cretaceous and the Pliocene. Sea levels were particularly high from the Turonian until the Campanian (91–75 Ma), very high in the Danian (c. 66–22 Ma) for c. 4 my, and in the early Eocene (c. 55–50 Ma) for c. 6 my. In the early Oligocene (c. 40–35 Ma) the period of flooding lasted for c. 5 my, the middle Miocene (c. 18–13 Ma) for c. 3 my and in the Pliocene (c. 5.3–1.8 Ma) for c. 1.5 my (Haq et al. 1987). Smith et al. (1994) supply maps that show the areas inundated. All of these events must have had a considerable effect on diatom evolution, creating and then destroying large numbers of freshwater bodies. Strel’nikova (1990) estimated that in every million years, 1.6 genera were created in the period from the Upper Cretaceous to the end of the Oligocene. Theriot (1992) showed that new diatom species can evolve in as little as 4000 years and so it is reasonable to assume that there would be plenty of time for adaptation and speciation during
flooded episodes, when presumably salinity levels varied from marine to brackish to freshwater, with creation also of hypersaline environments in arid zones.

Freshwaters provide a chemically diverse habitat for diatoms, with greater variation in nutrient concentrations, major and minor elements, salinity, temperature and pH than in the sea; a less turbulent environment for planktonic diatoms (e.g. Thalassiosirales in lakes, pools, etc) but great variation in flow rate in lotic habitats; and many new niches for raphid and araphid diatoms to colonize (Round & Sims 1981; Bradbury & Krebs 1995; Mann 1999a). Diatom species numbers expanded rapidly in the new habitats and the first lineages to invade freshwater might have been augmented by others during subsequent flooding events. However, many freshwater habitats are unlikely to leave a fossil record, because conditions are unsuitable for preservation. Such habitats include many favoured by raphid taxa, such as the muddy shorelines of lakes and rivers, springs, temporary pools, exposed wet rocks, amongst bryophytes, and in terrestrial soils; in fact, freshwater diatoms are found in virtually all environments that combine moisture and sunlight (Patrick 1948), but almost all nonmarine deposits of fossil diatoms are lacustrine.

In a review of continental localities, Bradbury & Krebs (1995) state that the earliest nonmarine deposits are Eocene in age and, until recently, the earliest known freshwater diatoms were those recorded from the mid-Eocene (1995) state that the earliest nonmarine deposits are Eocene and that the earliest nonmarine diatoms are Eocene and that the earliest nonmarine deposits are Eocene and that the earliest nonmarine deposits are Eocene. Diatom species numbers expanded rapidly in the new habitats and the first lineages to invade freshwater might have been augmented by others during subsequent flooding events. However, many freshwater habitats are unlikely to leave a fossil record, because conditions are unsuitable for preservation. Such habitats include many favoured by raphid taxa, such as the muddy shorelines of lakes and rivers, springs, temporary pools, exposed wet rocks, amongst bryophytes, and in terrestrial soils; in fact, freshwater diatoms are found in virtually all environments that combine moisture and sunlight (Patrick 1948), but almost all nonmarine deposits of fossil diatoms are lacustrine.

In a review of continental localities, Bradbury & Krebs (1995) state that the earliest nonmarine deposits are Eocene in age and, until recently, the earliest known freshwater diatoms were those recorded from the mid-Eocene (c. 50–45 Ma) Dewey Beds of Idaho (Leonard & Marvin 1982). However, Chacón-Baca et al. (2002) have reported finding fossils in carbonaceous cherts in Mexico, dated to c. 70 Ma, which suggests that freshwater, or at least brackish penate diatoms, were already present in the Upper Cretaceous. The micrographs of Chacón-Baca et al. do not allow confident identification, because no detail is visible in the frustules. The forms present include chain-forming organisms that resemble centric and araphid pennate diatoms. Two micrographs (Chacón-Baca et al. 2002, fig. 3A, B) show solitary lanceolate cells that, out of context and hence without a priori bias, many phycologists would happily identify as raphid diatoms allied to Navicula Bory or Sellaphora, although no other raphid diatom fossils are known before the Palaeocene (see section 5.5). The araphid-like organisms in the sample resemble Tetracyclus. The modern diatoms that these fossil diatoms resemble evolved relatively late in the molecular tree (Fig. 2). Thus, if this deposit is dated correctly, there would be a major conflict between molecular phylogeny and the fossil record. However, the lack of frustule detail makes it impossible to identify the organisms precisely and they could represent extinct lineages, any resemblance to modern taxa reflecting parallel evolution. Furthermore, one of us (P.A.S.) doubts that the age of the fossil freshwater material is correct or that contamination from modern material has occurred, because there is so little change in the morphological features of the genera illustrated – if one accepts the identifications given by Chacón-Baca et al. – whereas, during the last 70 million years, the features of fossil marine centric diatoms have changed considerably (cf. our previous comments on Proterozoic diatom records).

In a nonmarine, late Eocene (c. 45–40 Ma) deposit in Wyoming, Lohman & Andrews (1968) identified seven species of centric diatoms, including six Aulacoseira species and seven araphid pennates (5 Fragilaria species and 2 of a new genus, Ambistria Lohman & G.W. Andrews). In addition, there are 20 species of raphid diatoms, including 13 Pinnularia Ehrenberg, 3 Stauroeis Ehrenberg, 2 Nitzschia, 2 Navicula and 1 Anomoeoneis Pfitzer species. The two ‘Navicula’ species will need to be reallocated to other genera, because they do not correspond to Navicula sensu stricto (Round et al. 1990), and one of the Nitzschia species belongs to the ‘Trybionella’ group, recognized as a separate genus by Round et al. (1990). On the whole, the Wyoming assemblage has a very ‘modern’ appearance, suggesting a period of prior radiation of raphid diatoms in freshwaters. Among the centric diatoms in the Wyoming deposit are a species of Anaulus (this genus is present in the Cretaceous, but then it is marine) and six species of Melosira (= Aulacoseira), consistent with the suggestion of Gersonde & Harwood (1990) that Aulacoseira could have evolved from the Lower Cretaceous Archeephyrus Gersonde & Harwood. Further evidence for this line is provided by Ambwani et al. (2003) and Wolfe & Edlund (2005) who have recorded members of the Aulacoseiraceae from the nonmarine sediments of the uppermost Upper Cretaceous of India and the Middle Eocene of Canada, respectively. One might speculate that Anaulus and Archeephyrus could have migrated into freshwater some time between the Turonian and the Campanian when the eustatic sea levels were very high and inland seas had formed. This was also a period of explosive radiation and evolutionary change (Tapia 1996).

Krebs (1994) and Bradbury & Krebs (1995) defined four distinct communities in the fossil freshwater plankton of temperate lakes in North America, beginning with an Aulacoseira–pennate (e.g. Fragilaria) dominance in the Eocene–Oligocene, followed by Actinocyclus–Thalassiosira–Aulacoseira dominance in the early to middle Miocene. After that was a flora dominated by Mesodictyon–Cyclotella–Aulacoseira in the late Miocene and lastly Stephanodiscus–Cyclotella–Cyclostephanus–Aulacoseira dominance in the Pliocene–Holocene (c. 5–0.1 Ma). VanLandingham (1967) found a species of Stephanodiscus Ehrenberg earlier, in a Miocene flora. Several species of Actinocyclus persisted in nonmarine floras until the late Miocene (c. 10 Ma), possibly later, and Actinocyclus normannii f. subsalsa (Juhlin-Dannfeldt) Hustedd is present in today’s fresh- to brackish waters (Hasle 1977). Records show, however, that species of Actinocyclus first appeared in the marine flora of the late Eocene of Oamaru (c. 45–40 Ma) (Dekshachary & Sreeleatha 1989; Edwards 1991), where there are also two marine Cyclotella (Kützing) Brébisson species and one species of Cyclodiscus Round in Theriot, Håkansson, Kociolek, Round & Stoermer recorded in their own material by Desikachary & Sreeleatha (1989). Thus, marine Actinocyclus, Thalassiosira, Cyclotella and Cyclodiscus probably appeared in the fossil record before their freshwater counterparts, assuming that the identifications and sources of material are correct.

Very little attention has been paid to the araphid pennate diatoms of freshwater (nonmarine) fossil localities, perhaps because difficulties in identification have made them of little use to stratigraphers or ecologists. Purely taxonomic studies are few. As mentioned previously, Lohman & Andrews (1968) identified five species of Fragilaria and two species of a new araphid genus Ambistria from the late Eocene nonmarine diatoms from the Beaver Divide, Wyoming, and a new species of Anaulus that they compared to and distinguished from Tetracyclus. One new species of Tetracyclus, two of Tabellaria,
one of *Fragilaria* (and another new species of *Anaulus*) were described by Lupkina & Dologmatova (1975) from Kamchatka in the late Eocene to Oligocene of the Cape Astronomicheskij Formation. They also list the araphid genera *Meridion, Diatomaria*, *Hannaea* R.M. Patrick in R.M. Patrick & Reimer (as *Cerateonis*), and *Synedra* (?). Of the 82 species listed, members of the Fragilariophyceae dominate.

Skabichevskij (1981) believed that representatives of the Fragilariophyceae crossed the border between the sea and continental waters repeatedly, but Mann’s (1999a) analysis suggests, on the contrary, that this is a most difficult ‘rubicon’ for the araphids to cross, relative to the raphid diatoms (see also Krebs 1994). However, Patrick & Reimer (1966) noted that estimates of the number of endemic araphid species ranging from 10 to 30% and, in some large lakes up to 60%, Mann speculated that the reason why it was easier for the raphids to move into freshwater might be because proper functioning of the raphe may require well-developed homeostatic mechanisms for maintaining turgor at a particular level. However, the fragilarioids were apparently the most successful group in nonmarine middle Eocene–Oligocene environments. Molecular analysis of the araphids shows that the clades within the groups primarily are defined by habitat (freshwater vs marine) coupled with morphological features (Jung & Medlin, unpublished observations; Kooistra et al. 2004; Medlin and Kaczmarska 2004; see also Fig. 2). Recent taxonomic studies of freshwater araphid genera include those on *Diatoma* and *Meridion* (Williams 1985), on *Synedra* (Williams 1986; Williams & Round 1986), on *Fragilaria* (Williams & Round 1987; Morales 2001; Morales et al. 2001) and on *Asterionella* (Körner 1970; Pappas & Stoermer 2001).

Among other published records, VanLandingham (1985) found *Opephora* in the Early Late Miocene, whereas Reháková (1980) illustrated *Fragilariaforma costatum* (Reháková) Williams (recorded as *Diatoma costatum* Z. Reháková) and *Staurosirella* Williams and Round (recorded as *Fragilaria*) from the Miocene of Czechoslovakia. Hajós (1968) illustrated species of *Staurosira* Petít & Pelletan and *Ctenophora* (Grun.) Williams and Round (recorded as *Fragilaria* and *Synedra*, respectively) from brackish sites from the Miocene of Hungary and Schrader & Geronde (1978) identified *Asterionella japonica* Cleve (= *Asterionellopsis glacialis* Castracane (Round)) from late Miocene (Messinian, 6.5–5.3 Ma) brackish water sites in the Mediterranean at Deep Sea Drilling Project sites 124 and 374, where it formed 61.5% of the total population at site 374.

Many of the taxa identified from nonmarine sites are found living today and the majority of the araphid genera are already present by the late Miocene, with the genus *Tetracyclus* having a particularly good fossil record. The taxonomy and biogeography of *Tetracyclus* have been studied extensively (for earlier publications, see Williams 1996; also Williams et al. 2001; Williams 2003), with many fossil species showing interesting, restricted patterns of distribution, e.g. around the Pacific rim (Williams 1996). According to D. W. Williams (personal communication), fossil species of *Tetracyclus* have been found mostly in freshwater sites but occasionally in brackish.

5.4. The Thalassiosirales

Clade 2a, the class Mediophyceae, contains the important planktonic order Thalassiosirales, whose key morphological synapomorphy is the fultoportula (strutted process). This order has been placed in molecular Clade 2a consistently by all molecular analyses (Ehara et al. 2000; Medlin et al. 2004; Sörhannus et al. 1995; Fox & Sörhannus 2004). Despite this, the Thalassiosirales possesses only a few of the morphological features that unify the new class to which it belongs. It does contain perinuclear Golgi bodies and species often possess structures in the valve centre, like many other Clade 2a centrics. However, it possesses the relatively simple, scaly auxospores and isodiametric auxospore development that are more characteristic of Clade 1 (Kaczmarska et al. 2001) and it possesses a pyrenoid structure that sets it apart from other Clade 2a diatoms, but also from Clade 1 taxa (Schmid 2001).

In other members of the Mediophyceae (Clade 2a), any structure present centrally, within the annulus, is a rimoportula (= labiate process) but any central structures in the Thalassiosirales are fultoportulae (strutted processes); if a labiate process is located centrally, it usually is subcentral. Fultoportulae are also present in a ring around the valve mantle and sometimes also elsewhere. In molecular analysis, the sister group of the Thalassiosirales has consistently been the Lithodesmiales, when members of this order have been included (Sörhannus et al. 1995; Ehara et al. 2000; Fox & Sörhannus 2004; Medlin et al. 2004). Removal of the Thalassiosirales to the radial centrics leads to a less robust tree (Medlin et al. 1996a). The Lithodesmiales possess a central rimoportula, although this has a characteristic bilabiate internal morphology, which one easily can envisage having been formed by the closure of a single open tube. This fact led Medlin et al. (1993) and Medlin & Kaczmarska (2004) to suggest that the central strutted process could be homologous with the bilabiate process of the Lithodesmiales. Kaczmarska et al. (2006) have suggested further that the external tubes of the marginal strutted processes could be formed by a fragmentation, then encircling and fusing of the marginal fringe in the Lithodesmiales to form hollow tubes above the pore structure of the fultoportula. Two different types of strutted processes are known in the Thalassiosirales (Kaczmarska et al. 2006), but they could all be derived in this fashion. Thus, from this scenario, perhaps the central strutted process is evolutionarily distinct from the marginal ones. One of us (D.G.M.), however, doubts that fultoportulae and rimoportulae show such high levels of homoplasy.

Another possibility is that the central fultoportula and the central rimoportula of bipolar centrics are both derived from the central structure consisting of a simple tube as seen in *Archeopyrgus* of Geronde & Harwood’s group 4 of the fossil diatoms. The fact that cytoplasmic threads anchor the nucleus to either the central rimoportula (= labiate processes) in the bipolar centrics (Schmid 1994) or to the central strutted process of the thalassiosoroid lineages (Schmid 1984) may indicate that these structures have functional features in common, despite their different morphology, and may indicate that they have evolved from the same ancestral structure, or perhaps only the position in the cell is necessary, or the orientation, and the type of structure is irrelevant (see discussion in Medlin and Kaczmarska 2004). A new genus of living diatoms, *Mediopyxis*, has been discovered that has a variable central process: either a tube with a simple tube process, a microlaibate or a bilabiate process (Kuhn et al. 2006). Not all types of processes are found on a single valve and this may provide
support for the theory that the microlabiate evolved from a central tube.

It was in the late Eocene or Oligocene (c. 40 Ma) that the fultoportula (strutted process) evolved (Gombos & Ciesielski 1983; Monjanel 1987; Desikachary & Sreeath 1989), and hence this is the minimum age for the Thalassiosirales. Strel’nikova (1990) gave the age of the Thalassiosirales as late Oligocene, when this order achieved 1.5% of the total population in marine floras. However, a molecular clock calculation placed the origin of the entire Thalassiosirales lineage much earlier, at approximately 125 Ma (Medlin et al. 1996b). This estimated date is the divergence of the Thalassiosirales from the Lithodesmiales. This date greatly predates the first appearance of the fultoportula in the fossil record at 40 Ma, suggesting that this feature did not evolve in the earliest members of the Thalassiosirales lineage; rather, it may be a feature that defines the major diversification of the order.

In a molecular analysis of all extant genera of Thalassiosirales (Kaczmarska et al. 2006; Fig. 74), Porosira E. Jorgensen and Lauderia Cleve root the lineage. Significantly, of these genera, Lauderia has occluded processes, which structurally would be the most likely candidate as the immediate predecessor of the fultoportula itself (Syvertsen & Hasle 1982). The genus Thalassiosira occurs in three separate molecular clades. Some species of Thalassiosira may be closer relatives of Stephanodiscus, whereas others are more closely related to some species presently assigned to the genus Cyclotella than they are to other species of Thalassiosira. Cyclotella and Stephanodiscus are in separate clades, which would imply that the Thalassiosirales invaded freshwater at least twice. The branching order of the molecular tree agrees quite well with first appearances in the fossil record. Stephanodiscus first appears in the Pliocene (5 Ma), whereas Cyclotella first appears (in freshwater) earlier in the Miocene (24 Ma) (Bradbury & Krebs 1995). Both of these radiations into freshwater postdate the radiation of the pennates into freshwater (see section 4).

Fultoportulae enable cells to link by means of chitin threads, allowing the formation of chains of cells that can occupy open-water habitats, though some Thalassiosirales occupy loose communities associated with surfaces (‘periphyton’). The first Thalassiosirales were probably marine, as indicated by the fossil record; later, in the early to middle Miocene (c. 22–15 Ma), freshwater habitats were colonized, in lakes and rivers, when Thalassiosira itself was largely replaced by other genera, e.g. Cyclotella, Cyclostephanos, Mesodietyon Theriot & Bradbury and Stephanodiscus.

5.5. Evolution of the raphe

Another key event in Cenozoic evolution was the appearance of diatoms with a raphe system, the first certainly known being found at Palaeocene sites in Russia. Witt (1886) described and illustrated two raphe diatoms he referred to Navicula praetexta Ehrenberg and Navicula henrydii W. Smith, from Archangelsk in Simbirsk, and Pantocsek (1889) described Navicula simbirskiana from Ananino, Ulyanovsk Province (the dating of these sites is discussed by Ross & Sims 1985). The two species recorded by Witt are now classified in Lyrella Karayeva, which is not a basal lineage within the raphid diatoms, either in 18S rDNA or rbcL analyses (Behnke et al. 2004; Jones et al. 2005). Therefore, if the molecular data are correct, the fossil record must significantly underestimate the age of the raphe group, making a Cretaceous origin not impossible and pushing back the origin of the whole pennate group.

The evolution of the raphe system allowed diatoms to establish themselves in a wide range of new benthic habitats, particularly mobile sediments. It would not be too extravagant to equate this in importance with the evolution of flight in birds. Some nonraphid diatoms are motile, such as the centrics Actinocyclus and Odontella, the pseudopennates Ardissonnea and Toxarium (for references, see Kooistra et al. 2003b), and the true araphid pennate Leciophora (Sato & Medlin 2006), but none of these can match the raphe diatoms for speed, manoeuvrability, or economy of movement (for data, see Edgar & Pickett-Heaps 1984).

There are two basic types of raphe system, distinguished by the relationship between the raphe slits and the sternum (Mann 1984). In the typical raphe diatom, the raphe slits are integrated within the pattern-centre, so that the raphe system appears to bear transapical ribs and striae on either side, just like the sternum of araphid pennate diatoms. As might be expected, therefore, during valve formation, the raphe system is formed first, followed by the transapical elements. However, there is an ontogenetic peculiarity that seems to have phylogenetic significance: one side of the raphe slits (the primary side) is formed first, as a single longitudinal rib. Then the other (secondary) side is formed by the outgrowth of secondary ribs from the centre and poles simultaneously, the last part of the pattern centre to be completed being two points, each approximately halfway from centre to pole (Chiappino & Volcani 1977; Mann 1984); a particularly good ontogenetic sequence is illustrated by Mayama & Kobayasi (1989). The only traces of the asymmetrical development of the valve visible in the completed valve are slight irregularities in the stria pattern (the Voigt discontinuities: Mann 1981) marking the points of fusion between the secondary ribs; however, these are often correlated with other asymmetries, such as the direction of curvature of the central and polar raphe endings. The composite pattern centre in this first type of raphe system is referred to as a ‘raphe-sternum’.

The other type of raphe system is restricted to a few genera, of which the largest and best known is Eunotia. Here, there is a recognizable sternum like that in araphid pennate diatoms, clearly separate from the raphe slits and bearing ribs on either side. The raphe slits are short and placed to one side of the sternum, being in contact with it only near the poles; they interrupt the striae and ribs subtended by the sternum and often subtend striae and ribs of their own, which sometimes have a different spacing to those subtended by the sternum (e.g. Round et al. 1990). Micrographs published by Mayama & Kuriyama (2002) show clearly that the lateral position of the raphe, relative to the sternum, reflects its ontogeny: the sternum is complete in Eunotia, with striae on both sides, before the raphe is formed. This type of raphe system, unlike the other, more fully integrated ‘raphe-sternum’ type of raphe system, is usually accompanied by rimoportulae.

The lesser integration of the sternum and raphe in Eunotia has been taken by some (e.g. Mann 1984) to signal that the eunotioid raphe system is primitive. Others (e.g. Kociolek 2000) have questioned this and regard the eunotioid raphe as...
possibly reduced and modified. Two key questions are: (1) whether *Eunotia* and its allies are basal within the raphid group; and if they are not, (2) how is it possible that the Eunotiidae, alone among known extant or fossil diatoms, possess rimoportulae, which otherwise are restricted to centric and araphid pennate diatoms. Most analyses (of 18S rDNA) do show *Eunotia* as basal within the raphid diatoms (Fig. 2, Medlin & Kaczmarska 2004) or show no strong support for any other group as basal. Species of *Eunotia* are first recorded in the late Eocene, occurring in marine sediments at Oamaru, New Zealand (Grove & Sturt 1886; Schrader 1969; Desikachary & Sreechetha 1989) and on the Falkland Plateau (Gombos 1976). As yet there are no confirmed reports of extant marine species either of *Eunotia* or of any marine Eunotiidae; but no fossil *Eunotia* occur in freshwaters until after the Eocene. Furthermore, the araphid *Tabulariella tabellarioides* (Schrader) Silva, also from Oamaru, bears a striking resemblance to members of the entirely freshwater genus *Tabellaria*. It has sometimes been suggested, therefore, that some of the Oamaru diatoms have been washed in from freshwater habitats, or that the Oamaru *Eunotia* do not belong to the genus, any resemblance reflecting convergent evolution. However, Novitski & Kociolek (2005) make a convincing case that two species in the Eocene Oamaru deposits, *Eunotia grovei* Desikachary & Sreechetha and *E. marina* Schrader, are indeed *Eunotia* species, very similar to many extant species, with one rimoportula per valve. They also conclude that these diatoms were most likely autochthonous and hence truly marine.

However, this does not necessarily mean that *Eunotia* and its allies evolved in the sea, nor that the late Eocene *Eunotia* were early members of the Eunotiidae, nor that freshwater *Eunotia* also may not have existed in the Eocene. As already noted, the presence of *Lyrella* in Palaeocene deposits implies either that much of the early history of the raphid group has been lost, or that we have not yet searched in the right places to find well-preserved material, and the remarkably ‘modern’ appearance of the late Eocene freshwater diatom flora described by Lohman & Andrews (1968) (see section 5.3) also lends credence to the idea that the fossil record may often be deceptive for raphid diatoms.

The placement of the eunotioid diatoms in the molecular trees seems to be rather unstable. In Medlin et al. (2000), with only one species in the tree, *Eunotia* was the basal divergence of the raphid diatoms. In Medlin and Kaczmarska (2004), with two species, they fell at the base of the naviculoid lineage, together being sister to the nitzschioid lineage. In our present tree with 28 eunotioid species, they are again the basal divergence. Following the evolution of the eunotioid diatoms, there appears to be a divergence of the pennate diatoms into two clades. One contains primarily nitzschioid diatoms and the other, the naviculoid *sensu stricto* diatoms as sister to other naviculoid diatoms (Fig. 2). The separation of the nitzschioid and naviculoid taxa is a basal divergence in nearly all the trees regardless of the placement of the eunotioid diatoms.

The origin of the raphé system has been discussed by Hasle (1974) and in more detail by Mann (1984) and Round et al. (1990, p. 41). All equate the raphé slits with modified rimoportulae, and this link is made easy to believe by images of the eunotioid diatom *Peronia* (Round et al. 1990, p. 459, fig. n), in which a reduced raphé slit occurs alongside a rimoportula, both looking very much alike. Raphé slits and rimoportulae are both unoccluded, narrow slits through the valve, with a lipped aperture internally (in the rimoportula, the lips surround the whole structure; in the raphé slit they are present at one or both ends of the slit). There also are similarities between the structures during development: the ‘labiate process apparatus’, underlying the rimoportula during ontogeny, resembles the raphé fibre present beneath the raphé at equivalent stages of valve formation (Pickett-Heaps et al. 1990). If the raphé slits are modified rimoportulae, then the lateral asymmetry of the completed raphé system in eunotioid diatoms and the developmental asymmetry of the raphé-sternum in other raphid diatoms suggests that the ancestral raphid diatoms were dorsiventral, with both proto-raphés on the same side of the sternum, as in modern eunotioids. This is not implausible. As an analogy, molecular data (Fig. 2) show that the dorsiventral genus *Amphora* Ehrenberg ex Kützing is basal to bilaterally symmetrical diatoms such as *Entomoneis* Ehrenberg, *Surirella* and *Stenopterobia* Brébisson ex Van Heurck (Fig. 2).

Another requirement of the hypothesis that the rimoportule is the precursor of the raphé is that there must have been a differentiation of the rimoportules within the cell into two ‘populations’, one of which remained unaltered in morphology while the other was transformed into raphé slits. This is required because the Eunotiidae possess both raphé slits and rimoportulae. That such differentiation is possible is illustrated by the (quite independent) differentiation of the rimoportules that has occurred in some *Coscinodiscus* species, between macrormioportulae and micrormioportulae (e.g. Round et al. 1990, p. 176).

In our current molecular phylogeny, the closest relative to the raphid diatoms found is *Rhabdonema*. In Kooistra et al. (2003a), *Striatella* lies at the base of the raphid diatoms. Unfortunately, this takes us no closer to understanding how raphid diatoms evolved. Both genera, as well as several other araphid pennates (e.g. *Grammatophora*), have a pore field (through which a mucilage stalk is secreted) and either a single, large central rimoportula at each pole (*Striatella*) or multiple rimoportulae along the sides of the sternum (*Rhabdonaema*). No features of the sternum, rimoportulae, striae or areolae suggest any link to either eunotioid or ‘normal’ raphid diatoms.

5.6. Diversification of raphid diatoms

By the late Eocene, e.g. in the Oamaru deposit, the number of raphid genera present is considerable (Schrader 1969; Desikachary & Sreechetha 1989; Edwards 1991) and most of the major lineages of raphid diatoms present today are represented. There has been considerable diversification in raphé morphology and structure (e.g. Round et al. 1990). One of the principal innovations has been to link the two sides of the raphé sternum by a series of silica bridges (‘fibulae’), thus removing a structural weakness introduced by the raphé, which usually is compensated for in nonfibulate diatoms by extra thickening of the valve at the centre (the so-called ‘central nodule’) between the two slits. Almost by definition, the possession of fibulae must be an ‘advanced’ characteristic, postdating the evolution of the raphé itself. However, in molecular trees (Fig. 2), one of the principal groups of fibulate diatoms, the Bacillariaceae, often appears as one of the first lineages to diverge and *Nitzschia* species are already present.
in the late Eocene Oamaru deposits. In fact, fibulae have evolved at least twice, judging by molecular phylogenies, once in the ancestors of the Bacillariaceae/Bacillariales (Bacillaria J.F. Gmelin, Pseudo-nitzschia H. Peragallo in H. Peragallo & Peragallo, Nitzschia, Cylindrotheca Rabenhorst and related genera) near the base of the raphid diatom tree, and once in the ancestors of Entomoneis and the Surirellaceae/Surirellales, one of the last divergences in the rRNA tree (Fig. 2).

The presence of fibulae in turn has allowed two further developments: (1) in many of the Bacillariaceae, including Bacillaria itself, the two raphe slits have merged into one (see Mann 1984), presumably allowing maximal displacement of the cell per unit of trail material used (for the current model of diatom locomotion, see Edgar & Pickett-Heaps 1984); and (2) in the Surirellaceae, the raphe has been extended around the whole circumference of the valve (Pickett-Heaps et al. 1990; Round et al. 1990; Mann 2000), so that whatever the orientation of the cell, at least one raphe slit will be in contact with the substratum. The Surirellaceae are noteworthy by their absence from the Eocene Oamaru sediments, despite the fact that many genera and life forms are present that today would be found growing alongside Surirella and Campylodiscus in shallow marine sediments. Surirellaceae do not appear until the middle Miocene (Reinhold 1937; Hajós 1968, 1986) and this lineage is one of the last major groups to diverge in the rRNA tree (Fig. 2). Today they generally occupy an epipelagic niche (Round et al. 1990) and occur in waters of all salinities.

The evolutionary significance of other aspects of raphe morphology generally is unclear. Among the raphid diatoms as a whole, there is considerable variation in the course of the raphe slits and especially in the course of the raphe endings at the centre and poles, both internally and externally. On the other hand, among members of a single species, apparently minor details of raphe morphology are usually highly constant, suggesting that they are under strong selection (e.g. Mann et al. 2004). Hence, variation in raphe morphology must ‘mean something’. Presumably, because motility is mediated by trail material secreted through and moved along the raphe slits, there must be a relationship between, on the one hand, the shape of the raphe and its endings and, on the other, the orientation and re-orientation of movement that can occur. However, this aspect of diatom motility has thus far been given little attention, and studies on planar substrata, such as glass slides (e.g. Edgar & Pickett-Heaps 1984, figs 6–9; Round et al. 1990, fig. 70), may give a poor picture of how diatoms manoeuvre in nature in sedimentary environments (illustrated well by Paterson 1986). New insights into diatom motility and the structure of the mucilage layer are being gained by atomic force microscopy (Chiiovitti et al. 2003; Higgins et al. 2003) and may throw light on how details of frustule morphology affect diatom movement.

Raphid diatoms have colonized some habitats in which their motility is of reduced importance. The most striking examples are the planktonic members of the Bacillariaceae, many of which have adopted a colonial growth-form and are either nonmotile (Fragilariopsis Hustedt) or only transiently motile (Pseudo-nitzschia) in their natural habitat. These planktonic Bacillariaceae have simpler raphes than their benthic counterparts, notably in the absence of terminal fissures, which suggests that terminal fissures may be essential for directional movement. Adaptation of Bacillariaceae to a pelagic existence occurred by the Miocene (c. 22–5 Ma), with the evolution of Denticulopsis Simonsen and Crucidendula Akiba & Yanagisawa, which link in chains by means of their valve faces and raphe: Akiba (1985) and Akiba & Yanagisawa (1986) give further information on these genera, which prefigure (and may be closely related to) Fragilariopsis and Pseudo-nitzschia.

There also are many raphid diatoms that, although they can move actively if detached, spend most of their lives attached to solid substrata, in the epiphyton, epilithon or epipsammon, or in chain-like colonies. Reduction of the terminal fissures occurs in some of these also, e.g. in Cocconeis Ehrenberg, Achnanthidium Kützing and Diadesmis Kützing, and these three genera also exhibit facultative or obligate loss of the raphe in one or both valves (Round et al. 1990). A further correlate of an attached existence is modification of cell symmetry, either to a heteropolar ‘gomphonemoid’ form, or to a laterally asymmetrical ‘cymbellloid’ or ‘amphoroid’ form. As Medlin (1991) noted, heteropolarity brings the two valves and their mucilage-secreting pore fields close together at one end of the cell and so it allows clones of cells to build up closely packed rosette-like colonies, attached to the substratum by a common pad or branching stalk. This is also true of the cymbellloid shape. It already was clear from studies of valve ultrastructure that the gomphonemoid and cymbellloid morphologies have evolved many times in raphid diatoms, and also in araphid pennate diatoms (e.g. Round et al. 1990; Medlin 1991), and this has been confirmed by molecular data (e.g. Jones et al. 2005). Similarly, the extreme dorsiventrality of Amphora is paralleled in the unrelated Seminavis D.G. Mann in Round, R.M. Crawford & D.G. Mann, and heterovalvy like that in Cocconeis and Achnanthidium also has evolved a second time in the unrelated marine Achnanthes (Medlin & Kaczmarska 2004).

6. CHANGES IN SEXUAL REPRODUCTION AND CYTOLOGY DURING DIATOM DIVERSIFICATION

Accompanying the evolution of the sternum and raphe, there have been changes in sexual reproduction and chloroplast morphology. The oogamous reproduction of centric diatoms, involving fertilization of large immobile egg cells by small anteriorly unflagellate sperm, has been replaced in pennates by (1) gametangiogamy (copulation between gametangia) and (2) isogamy or anisogamy involving (3) large nonflagellate gametes (Chepurnov et al. 2004). The evolution of an effective motility mechanism in raphid diatoms might be expected to have consequences for sexual reproduction, because mating could now be facilitated through movement of the vegetative cells (potential gametangia), as well as by the gametes themselves. A change from oogamy in centric diatoms to gametangiogamy in raphid pennates is therefore understandable. What is odd is that this change predated the evolution of the raphe, taking place instead in the araphid pennate diatoms, which (now that Ardissonaea and Toxarium have been linked to the centric diatoms), as far as we know, are largely immobile (an exception, recently discovered, is Licmophora [Chepurnov & Mann 2004; Sato & Medlin, 2006]). It seems likely, therefore, that evolution of gametangiogamy must have
taken place in crowded benthic habitats, where the chances of encounter between compatible, sexualized vegetative cells were high and dispersal of gametes was not necessary to ensure outbreeding. It is noteworthy that, in most cases investigated so far, araphid pennate diatoms have well-developed mechanisms to avoid inbreeding, through bipolar heterothallism (Chepurnov & Mann 2004).

There are differences among centric diatoms in the number of egg cells produced per oogonium (e.g. one in Bellerochea; two in Biddulphia and Chaetoceros; two in Lithodesmium, and Odontella: see von Stosch 1982) and in the method of formation of the sperm (merogenous or hologenous: see Round et al. 1990). There are also differences in whether the egg cells are released before fertilization, as in Bellerochea and Helicotheca Ricard in Sournia (von Stosch 1977, 1982); retained in the oogonium during fertilization but then released, as in Biddulphia, Lithodesmium and Odontella aurita (von Stosch 1982); retained for fertilization and developing (as an auxospore) in situ, as in Stephanopyxis and Amphitetras Ehrenberg (Liebsch 1928; von Stosch & Drebes 1964); or retained during fertilization but then developing laterally, as in Chaetoceros (von Stosch et al. 1973; von Stosch 1982). Undoubtedly, some phylogenetic patterns will become detectable in the distribution of such characteristics and others, such as details of spermatogonial development or the timing of plasmogamy in relation to meiosis (in Chaetoceros; for example, the sperm may enter while the oogonium is in meiotic prophase: von Stosch et al. 1973, whereas elsewhere the egg cell is fully mature at fertilization: see Drebes 1977), However, until many more species and genera are studied, the full significance of variation will remain unclear.

The position is a little better with respect to raphid pennate diatoms, thanks to intensive studies over many years by Geitler (lists of references in Geitler 1973, 1984), who also summarized earlier work (Geitler 1932), and more recently by Mann and others (e.g. see Edlund & Stoermer 1997). The frequency of auxospore formation in pennates may be lower than that in centrics and could help explain the calculated lower rate of evolution in pennates (Kooistra & Medlin 1996). Patterns of gametogenesis and auxospore development have prompted taxonomic revision (e.g. the resurrection of Sellaphora by Mann 1989), and particular types of auxospore formation distinguished by Geitler (1973) can now be seen to characterize major clades, such as ‘type IAla’ auxosporulation (physiologically anisogamous, each gametangium producing one active and one passive gamete; the gametes becoming rearranged and rounded off within the gametangia; auxospores expanding parallel to the gametangia; no copulation tubes between the gametangia), which is found in allogamous members of the Cymbellaceae and Gomphonemataceae. These families group together in molecular phylogenies (Fig. 2).

The ways in which auxospores reconstruct shape after the loss of the old gametangial frustules during sexual reproduction are proving to have major phylogenetic significance, as indicated in an earlier section. The coverings of auxospores vary from apparently homogeneous organic walls, through organic walls with embedded scales (two- or three-dimensional), to elaborate systems of silica bands that constrain auxospore expansion, allowing the formation of many kinds of bipolar or multipolar shapes (Mann 1982; von Stosch 1982; Round et al. 1990). The broad picture of evolution is clear. Radial centrics (Clade I) have organic or scale-containing auxospore walls and the auxospores are spherical or ellipsoidal. Members of Clade 2, in contrast, generally produce auxospores containing strips or bands of silica, together called properizonia or perizonia, which allow the formation of elongate or multipolar cells. The exception is the Clade 2 order Thalassiosirales (although these do contain some interesting bi- or tripolar forms in Cyclotella, which need ontogenetic study), where the auxospore walls are scaly.

However, beyond these important generalities, the pattern of variation in auxospore coverings is poorly documented and hence can rarely be interpreted in a phylogenetic context. One small exception is the discovery of silicified caps over the ends of the linear auxospores in Neidium and Biremis D.G. Mann & E.J. Cox in Round, R.M. Crawford & D.G. Mann (Mann 1993). This seems to be a synapomorphy and aspects of cell structure (central internal raphe endings, chambered walls, chloroplast morphology) and rbcL data (G.E. Simpson & D.G. Mann, unpublished observations) also suggest that these genera closely are related. The function of the caps seems to be to exert tight control on auxospore width during expansion, creating strictly linear auxospores (and hence linear initial cells).

Aspects of protoplasm structure also correlate with phylogeny. For example, the Cymbellaceae and Gomphonemataceae, besides sharing a similar pattern of auxosporulation, are also alike in that the nucleus always divides on the same side of the cell (Mann & Stickle 1988). In the Naviculaceae sensu stricto, on the other hand, the nucleus oscillates from one side of the cell to the other at successive divisions, except in the bilaterally asymmetrical, amphoroid genus Seminavis, where oscillation has probably been lost secondarily. The Naviculaceae is a well-supported monophyletic group (Fig. 2; Jones et al. 2005).

Particular plastid and pyrenoid morphologies often characterize genera or larger groupings and have been used to support taxonomic revision (e.g. Cox 1987; Mann 1989, 1996; Schmid 2001). Overall, however, we have disappointingly little idea of the adaptive significance of protoplasm characters. Centric and araphid pennate diatoms generally have many chloroplasts per cell (unless the cells are small), whereas raphid pennate diatoms are oligoplastidic, usually with one, two or four chloroplasts (Mann 1996). This probably reflects the fact that many centric and araphid pennate species cannot orientate themselves in relation to changes in incident light and often occupy habitats where the cells are reoriented passively by water movements; hence, the optimal distribution of the chloroplasts probably is to have them spaced evenly around the cell periphery (or around the nucleus). Raphid diatoms, on the other hand, are motile and can move and swivel in response to a perceived light gradient. Here, cells can compensate for any heterogeneity in the distribution of the chloroplasts (such as may be necessary to allow optimal functioning of the raphe system, for example) by adopting a preferred orientation in relation to the prevailing light climate. However, there are many exceptions to the general ‘rules’ concerning plastid number and size (e.g. many and small in centrics), and many instances within particular well-defined taxa where one or a few species have an aberrant cytology. For example, in Nitzschia longissima (Brébisson ex Kützing) Grunow, there
has apparently been secondary evolution of polyplastidy (almost all other Nitzschia species have just two chloroplasts per cell), perhaps in response to its planktonic habitat. However, the amazing diversity of chloroplast morphologies among rapid diatoms (Round et al. 1990; Cox 1996; Mann 1996) cannot as yet be interpreted. The silica valves and girdle bands must refract and diffract light incident upon the cell, focusing it in certain areas and shading others, but as far as we know, no analysis has been made of the optical properties of particular wall structures and so there is no basis for assessing whether chloroplast morphology is correlated in any way with internal light distribution.

7. SPECIATION AND MICROEVOLUTION

The clear message from recent molecular, morphometric and reproductive compatibility studies is that diatoms are greatly underclassed at the species level (reviewed by Mann 1999b). There can be no *a priori* expectation that speciation in diatoms (in the sense of essentially permanent cessation of gene exchange between two lineages) will always have been accompanied by significant morphological change, because diatoms do not use visual or geometric clues for mating. And recent data obligingly confirm that, indeed, speciation is not always accompanied by significant morphological change: good examples of semicryptic and cryptic speciation are now available (e.g. Mann 1999b; Lundholm et al. 2003; Montresor et al. 2003; Behnke et al. 2004; Sarno et al. 2005). Data from ‘model systems’, such as *Sellaphora pupula* (Kützing) Mereschkowsky (Mann 1999b; Mann et al. 1999; Behnke et al. 2004), the *Diploneis smithii-fusca* complex (Droop 1994; Droop et al. 2000), *Pseudo-nitzschia* (e.g. Lundholm et al. 2003; Montresor et al. 2003), and *Skeletonema* (Gallagher 1982; Medlin et al. 1991; Sarno et al. 2005) show that apparently minor morphological differences may be significant, marking the boundaries between species. Surprising levels of genetic diversity have been demonstrated in a number of planktonic diatoms, both marine and freshwater (Lewis et al. 1997; Mann 1999b; Rynearson & Armbrust 2000, 2004; Evans et al. 2005). In addition, the ease with which hitherto unknown species can be discovered (e.g. Lange-Bertalot et al. 2003) shows how inadequate the conventional taxonomy is, even for Europe, which has been studied intensively for more than 150 years.

Species-level classification in diatoms is therefore a poor and uneven foundation for studies in ecology or biogeography, or even phylogeny, and all such research should as far as possible be validated by voucher samples until such time as diatom taxonomy finally stabilizes. The semicryptic and cryptic species that are beginning to emerge cannot be assumed to have similar ecologies, even when they occur sympatrically – indeed conventional niche theory, backed by classic studies such as MacArthur’s warblers (MacArthur 1958), demands that coexisting species must be incomplete competitors and show niche separation. Hence, although it may be difficult to identify them, the semicryptic and cryptic species cannot be ignored if we are to have a meaningful science of diatom ecology. Round (1972) took the trouble to separate two very similar demes of *Sellaphora pupula* in a study of two small freshwater ponds and showed that they had quite different seasonal patterns of appearance in the epipelon. Furthermore, it is unwise to make sweeping statements about the ubiquity of diatom species (Finlay et al. 2002) until we are more confident about true species boundaries. Nevertheless molecular and breeding data are providing vast insights into our knowledge of diatom biology and evolution, and this knowledge is expanding at a rapid rate.

8. CONCLUSIONS

Although we do not have living representatives of many of the fossil genera, the general pattern of molecular evolution of the extant genera agrees reasonably with the fossil record, with very few exceptions. Some sister group relationships are surprising, e.g. the Lithodesmiales and the Thalassiosirales.

However, such discoveries prompt new and more detailed structural and cytotaxonomic investigations to reveal whether the differences are more apparent than real and to understand their basis (Kaczmarska et al. 2006). We still are lacking much detail about the life cycle and protoplast structure in many diatom genera, so that some of the sister relationships recovered in the molecular analyses remain ‘bald’ cladistic hypotheses, with little biological meaning. Auxospore formation has been the single feature most supportive of the deep branches in the molecular tree and therefore it is not unlikely that more ‘minor’ details of the reproductive cycle may be phylogenetically informative, but at lower taxonomic levels. Sequence data from common, cosmopolitan ‘species’ suggest that cryptic and semicryptic species are widespread in the diatoms, supporting earlier indications from mating studies (Mann 1999b). However, it will be extremely difficult or impossible to detect such species among fossil diatoms, because crossing experiments and molecular analyses cannot be performed.

There are still huge gaps in our knowledge of fossil diatoms. Some important fossil localities have been lost (e.g. the Hanover material studied by Forti & Schulz 1932) and new localities containing Jurassic or Lower Cretaceous diatoms are reported sporadically: there are few resources to support an intensive, directed search. Even when information on diatom evolution is available, from fossils or molecular sequence analyses, much of it cannot be fully interpreted because we still lack much basic information on the morphology, cytotaxonomy and reproduction of living species. An integrated approach is therefore desirable but it is rarely achieved. One of the few examples where a rigorous attempt has been made to combine fossil and morphological data is the study of evolution in some *Thalassiosira* species by Julius & Tanimura (2001), but in this case no molecular sequence data were available. On the whole, however, it is fossil data that are most often omitted from discussions of diatom evolution, partly because the relevant literature is widely scattered. We hope that one outcome of our review will be to make this literature more accessible and better used.

ACKNOWLEDGEMENTS

P.A.S. would like to thank David Williams and Eileen Cox for helpful discussion and David Williams for the use of his scanning electron micrographs of *Sceptroneis*. L.K.M. would like
REFERENCES


Cox E.J. 1987. Placoneis Mereschkovsky: the re-evaluation of a di-


Koosstra W.H.C. & Medlin L.K. 1996. Evolution of the diatoms (Bacillariophyta). IV. A reconstruction of their age from small sub-
ture of the male gamete of the marine centric diatom _Lithodesmium undulatum_. *Journal of the Royal Microscopical Society* 8: 119–134.


**Mayama S. & Kobayasi H. 1989.** Sequential valve development in the monoraphid diatom _Achnanthus minuttissima var. saprophila_. *Diatom Research* 4: 111–117.


**Medlin L.K. 2002.** Why silica or better yet why not silica? Speculations as to why the diatoms utilize silica. *Diatom Research* 17: 453–459.


**Nikolaev V.A. & Harwood D.M. In press.** Lower Cretaceous diatoms of Eromanga Basin, Queensland, Australia. *Alcheringa*.


**Novitski L. & Kociolek J.P. 2005.** Preliminary light and scanning electron microscope observations of marine fossil _Eunotia_ species with comments on the evolution of the genus _Eunotia_. *Diatom Research* 20: 137–143.


**Quintero P.J., Barron J.A., Burry D.L. & Blome C.D. 1994.** Late Cretaceous microfossils in a single dredge haul from the Tonga Trench: evidence of Late Cretaceous environment at the Louisville Hotspot. In: *Geology and submarine resources of the Tonga-Lau-

---

400 *Phycologia*, Vol. 45 (4), 2006


Received 10 April 2005; accepted 16 January 2006

Associate editor: T. Horiguchi