Molecular Phylogenetics and Evolution 61 (2011) 866–879

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

# A time-calibrated multi-gene phylogeny of the diatom genus *Pinnularia*

Caroline Souffreau [a,1], Heroen Verbruggen [b], Alexander P. Wolfe [c], Pieter Vanormelingen [a], Peter A. Siver [d], Eileen J. Cox [e], David G. Mann [f], Bart Van de Vijver [g], Koen Sabbe [a], Wim Vyverman [a,*]

[a] *Laboratory of Protistology and Aquatic Ecology, Ghent University, Krijgslaan 281 – S8, 9000 Gent, Belgium*
[b] *Phycology Research Group, Ghent University, Krijgslaan 281 – S8, 9000 Gent, Belgium*
[c] *Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton AB T6G 2E3, Canada*
[d] *Department of Botany, Connecticut College, New London, CT 06320, USA*
[e] *Department of Botany, The Natural History Museum, Cromwell Road, London SW7 5BD, United Kingdom*
[f] *Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, Scotland, United Kingdom*
[g] *Department of Cryptogamy, National Botanic Garden of Belgium, Domein van Bouchout, 1860 Meise, Belgium*

ARTICLE INFO

*Article history:*
Received 16 March 2011
Revised 17 August 2011
Accepted 31 August 2011
Available online 10 September 2011

*Keywords:*
Molecular phylogenetics
Relaxed molecular clock
Fossil record
Raphid diatoms
*Pinnularia*
Bacillariophyceae
Eocene

ABSTRACT

*Pinnularia* is an ecologically important and species-rich genus of freshwater diatoms (Bacillariophyceae) showing considerable variation in frustule morphology. Interspecific evolutionary relationships were inferred for 36 *Pinnularia* taxa using a five-locus dataset. A range of fossil taxa, including newly discovered Middle Eocene forms of *Pinnularia*, was used to calibrate a relaxed molecular clock analysis and investigate temporal aspects of the genus' diversification. The multi-gene approach resulted in a well-resolved phylogeny of three major clades and several subclades that were frequently, but not universally, delimited by valve morphology. The genus *Caloneis* was not recovered as monophyletic, confirming that, as currently delimited, this genus is not evolutionarily meaningful and should be merged with *Pinnularia*. The *Pinnularia–Caloneis* complex is estimated to have diverged between the Upper Cretaceous and the early Eocene, implying a ghost range of at least 10 million year (Ma) in the fossil record.



## 1. Introduction

Diatoms are an extremely diverse group of unicellular algae that are uniquely characterized by a siliceous cell wall (the frustule) consisting of two valves and a number of girdle bands (Round et al., 1990) and a diplontic life cycle involving gradual size reduction during vegetative divisions and rapid size restitution, usually through sexual reproduction (Chepurnov et al., 2004). In the so-called pennate diatoms the valve pattern is organized bilaterally around the longitudinal axis, and in most cases the valve is elongate. Raphid pennate diatoms possess a pair of longitudinal slits along the apical axis (the raphe system) (Fig. 1), from which extracellular polymeric substances are exuded and used in locomotion and for adhesion to the substratum (Round et al., 1990). The raphe is considered a derived character state that distinguishes the raphid diatoms from the more ancestral araphid pennate forms that lack this structure and from the oldest known forms, the radially organized "centric" taxa (Sims et al., 2006). Based on fossil remains, the araphid pennates first appeared in the Upper Cretaceous (ca. 75 Ma; Chambers, 1966; Hajós and Stradner, 1975), and the raphe-bearing forms soon thereafter, around 70.6–55.8 Ma (Chacon-Baca et al., 2002; Pantocsek, 1889; Singh et al., 2006; Witt, 1886). Monophyly of pennate diatoms as a whole, as well as the raphid pennates, has been documented using SSU rDNA and *rbc*L sequences (e.g. Kooistra et al., 2003; Sorhannus, 2004, 2007). Since their origin, raphid pennate diatoms have diversified enormously and account for the majority of the over 200,000 extant species estimated to exist (Mann and Droop, 1996), indicating the evolutionary advantages conferred by the raphe (Sims et al., 2006).

Despite the diversity and ecological success of raphid pennate diatoms, relatively few detailed molecular phylogenetic reconstructions exist. Phylogenies applied at genus to ordinal levels have yielded partly unsupported taxon relationships (e.g. Bruder and Medlin, 2008; Trobajo et al., 2009), in part due to very limited taxon sampling and/or the use of a limited number of genetic markers (Mann and Evans, 2007). In addition, molecular phylogenies of individual genera have focused largely on the identification of

* Corresponding author. Fax: +32 9 264 85 99.
*E-mail addresses:* caroline.souffreau@ugent.be (C. Souffreau), heroen.verbruggen@ugent.be (H. Verbruggen), awolfe@ualberta.ca (A.P. Wolfe), pieter.vanormelingen@ugent.be (P. Vanormelingen), pasiv@conncoll.edu (P.A. Siver), e.cox@nhm.ac.uk (E.J. Cox), d.mann@rbge.org.uk (D.G. Mann), vandevijver@br.f-gov.be (Bart Van de Vijver), koen.sabbe@ugent.be (K. Sabbe), wim.vyverman@ugent.be (W. Vyverman).
[1] Present address: Laboratory of Aquatic Ecology and Evolutionary Biology, Katholieke Universiteit Leuven, Charles Debériotstraat 32, 3000 Leuven, Belgium.

1055-7903/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2011.08.031

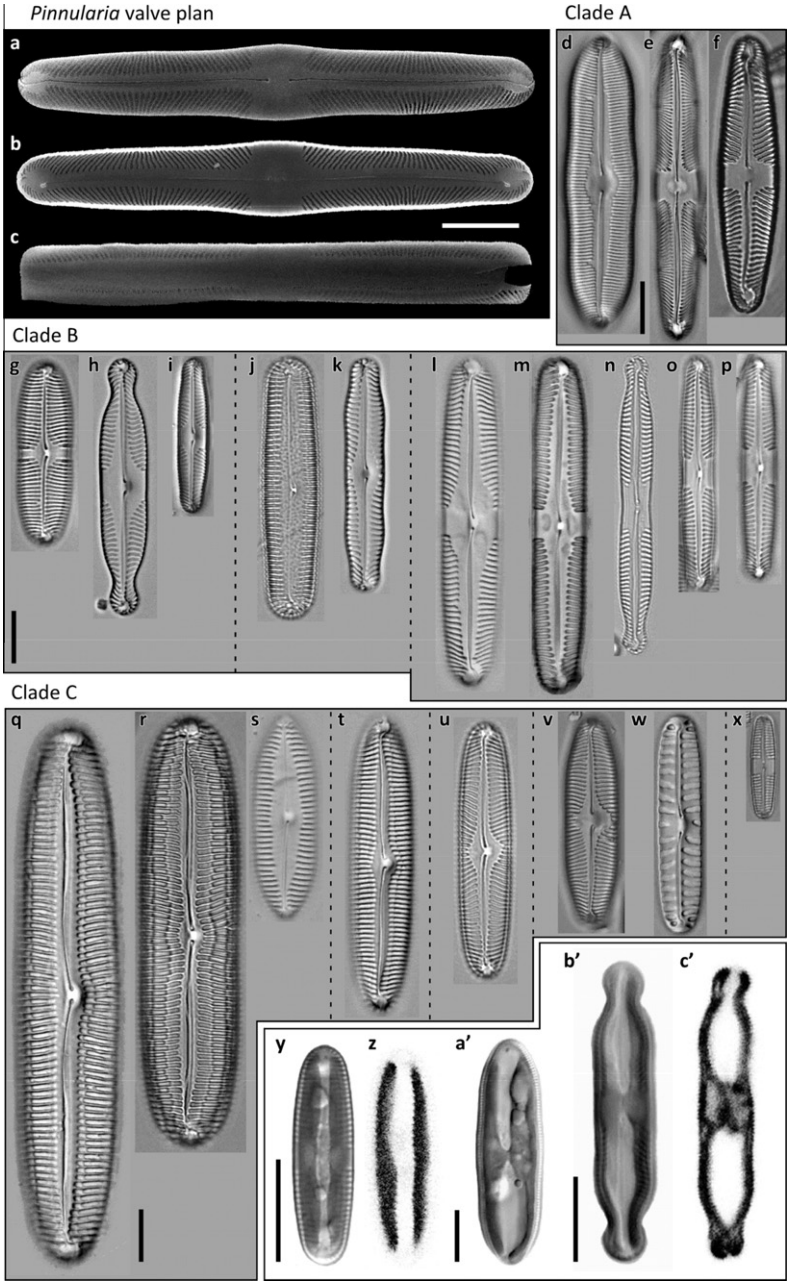


**Fig. 1.** Morphological variation in *Pinnularia* and representative illustrations of strains included in the multi-gene phylogeny. The valve construction of typical frustules belonging to the *P. divergens* group [strain (Tor7)c] is shown by scanning electron micrographs [outside (a), inside (b) and side (=girdle) view (c)]. Cultured and sequenced strains are illustrated by a series of light micrographs (d–x), divided, where appropriate, into subclades recovered in the phylogeny with vertical dashed lines. Clade A (d–f) includes *Caloneis lauta* (d) as well as *P. divergens* grade representatives (Tor7)c in (e) and (Tor1)b in (f). Clade B (g–p) includes the "*grunowii*" subclade (g–i) with *Pinnularia sp.* (Tor4)i in (g), *P. subanglica* Pin650 in (h), and *P.* cf. *marchica* (Ecrins4)a in (i); the "*nodosa*" subclade (j–k) with *P. acrosphaeria* (Val1)b in (j) and *P. nodosa* Pin885 in (k); and the "*subgibba*" subclade (l–p) represented by *P. parvulissima* Pin887 in (l), *Pinnularia* sp. "*gibba*-group" (Tor7)f in (m), *P. subcapitata* var. *elongata* (Wie)c in (n), *P.* sp. (Tor4)r in (o), and *Pinnularia* sp. "*gibba*-group" (Tor8)b in (p). Clade C (q–x) comprises the "*viridis*" subclade (q–r) represented by *P. neglectiformis* Pin706 in (q) and *P. viridiformis* (Enc2)a in (r); with its sister *P. acuminata* Pin876 in (s); the "*subcommutata*" subclade (t) represented by *P. subcommutata* var. *nonfasciata* Corsea10 in (t); forms that do not readily fit into well-defined subclades represented by *P.* sp (Wie)a in (u); the "*borealis-microstauron*" subclade (v–w) including *P.* cf. *microstauron* (B2)c in (v) and *P. borealis* Alka1 in (w); and subclade C1 (x) represented by *P.* cf. *altiplanensis* (Tor11)b in (x). Live cells of *Pinnularia sensu lato* (*i.e.* including *Caloneis*) also show two distinct plastid arrangements, which are illustrated by light (y, a′, and b′) and laser-scanning confocal microscopy of the autofluorescent organelles (z and c′). For example, representatives of the *P. subcommutata* and *gibba* taxa have parallel plastids on either side of the apical axis (y, z), while *Caloneis silicula* (a′) and *P. grunowii* (b′ and c′) have plastids that are joined by a central bridge. See text for details. All scale bars are 10 μm, and images (d–x) are reproduced at the same magnification to facilitate size comparisons between taxa.

cryptic diversity rather than the elucidation of evolutionary relationships between lineages (e.g. Beszteri et al., 2007; Evans et al., 2008; Lundholm et al., 2006). There are a few explicit time-calibrated phylogenies at the ordinal level (e.g. Medlin et al., 1996), the diatoms as a whole (e.g. Kooistra and Medlin, 1996; Medlin et al., 1997a; Sorhannus, 2007), the wider heterokont group (e.g. Brown and Sorhannus, 2010; Medlin et al., 1997b), or even eukaryotes (e.g. Berney and Pawlowski, 2006), but there are none at the generic level; furthermore, few analyses formalize the evolutionary associations between the timing of lineage splitting and ecological, morphological, physiological and/or reproductive strategies, life cycles and geographical distributions (but see Casteleyn et al., 2010). Furthermore, despite important recent micropaleontological discoveries, some of which confirm that particular genera

of raphid diatoms are older than previously suspected (Siver and Wolfe, 2007; Siver et al., 2010), numerous gaps remain in the fossil record. As a result, the overall course of evolution in raphid pennate diatoms is not known and the relationships between many groups remain uncertain (Mann and Evans, 2007).

*Pinnularia* Ehrenberg (1843) is one of the most species-rich genera of raphid pennate diatoms, with 2465 taxon names recorded in Algaebase of which 412 are currently accepted (Guiry and Guiry, 2011). The genus occurs globally in freshwater habitats of varying pH and trophic status, and to a lesser extent in moist soils, peatlands, spring seeps and marine coastal environments (Krammer, 2000; Round et al., 1990). Members of *Pinnularia* and the closely related genus *Caloneis* Cleve (1894) have linear-lanceolate, blunt-ended or occasionally capitate valves with a central raphe system (Fig. 1) that terminates internally in helictoglossae at the poles (Figs. 1b and 2b, d, f and h). They possess either two plate-like plastids (Fig. 1y and z) or a single H-shaped plastid (Fig. 1a′–c′). *Pinnularia* is characterized by the presence of a chambered,

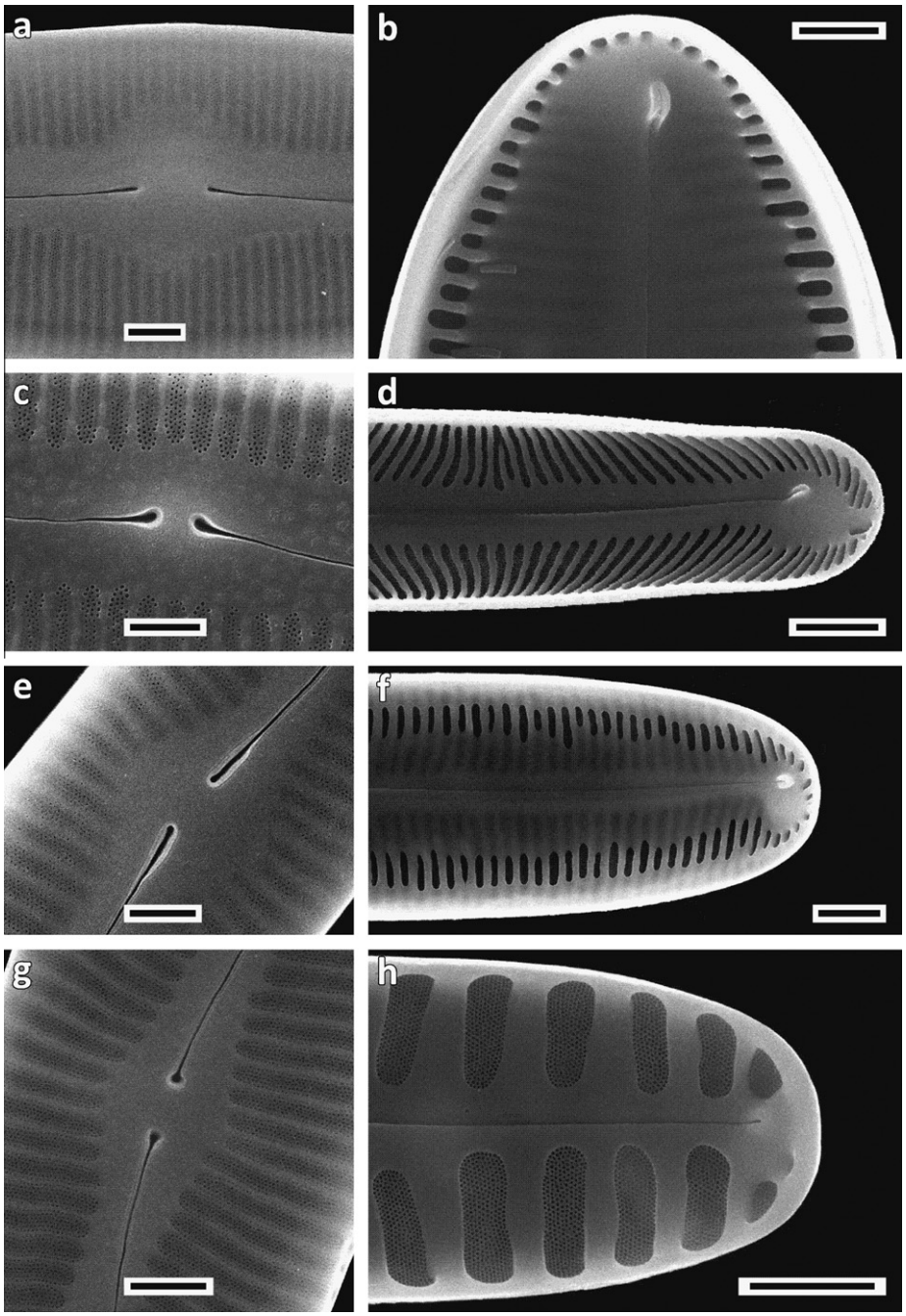


**Fig. 2.** Supporting SEM micrographs showing the shape of the central raphe endings (left column: a, c, e and g) and the extension of the alveolar opening (right column: b, d, f, h). Central raphe endings can be linear (a and e), drop-like (c) or round (g). Linear raphe endings occur in *Caloneis silicula* (a), the *subcommutata* clade (e), *P. acuminata*, (Wie)a and *P.* cf. *altiplanensis*. Drop-like endings are present in clade B (c), the *borealis-microstauron* clade and the *divergens* group; while round endings occur in the *viridiformis* clade (g). The alveolar openings can be small (b), large (d and h) or intermediate (f). Small openings are typical for *Caloneis silicula* (b) but also occur in *P. acrosphaeria*; intermediate openings (f) occur in the *subcommutata* and *viridiformis* clades and *P. acuminata*; while large openings (d and h) are present in all other species, including the *divergens* group, clade B (without *P. acrosphaeria*), the *borealis-microstauron* clade with *P. borealis* (h) and the isolates (Wie)a and *P.* cf. *altiplanensis*. Scale bars represent 2 μm (a, b, c, e and g) and 5 μm (d, f and h).

double-walled valve structure in which the outer surface is ornamented by multiple rows of small pores forming multiseriate striae (detail in Fig. 2c), whereas the inner wall of each chamber or alveolus is perforated by a large transapically elongate aperture (Figs. 1b and 2b, d, f and h) (Round et al., 1990). *Caloneis* is similar, except that the inner aperture is smaller and often circular, and sometimes there are two apertures per stria. Despite the species richness, morphological diversity and ecological significance of *Pinnularia* and *Caloneis* in freshwater and terrestrial ecosystems, the evolutionary relationships among species are poorly known. Bruder et al. (2008) constructed a molecular phylogeny of *Pinnularia* and *Caloneis* using 18S, 28S and *rbc*L genes for 15 species. Whereas some well-supported *Pinnularia* lineages were recovered and *Caloneis* was scattered among *Pinnularia*, overall support for the clades was low, suggesting that more exhaustive sampling, with respect to both taxa and genetic markers, is needed to produce a well-resolved phylogeny.

The temporal aspects of the diversification of *Pinnularia* are also poorly documented, despite the fact that numerous fossils have been reported for the genus. To date, the oldest known diatoms reliably assigned to *Pinnularia* originate from the Wagon Gap Formation of Wyoming, USA (Lohman and Andrews, 1968). The age of these sediments is not known precisely because diatom-containing sediment clasts have been redeposited in a carbonate conglomerate, but their age lies between the Late Eocene and Early Oligocene (35–32 Ma). Dating from the same period, the Oamaru diatomite deposits also contain two species of *Pinnularia* (Desikachary and Sreelatha, 1989). From the Early Miocene onwards, diverse morphological types of *Pinnularia* are reported from freshwater deposits or as freshwater inwash in marine deposits (for Miocene deposits including *Pinnularia* see for e.g. Hajós, 1986; Héribaud, 1902; Lewis et al., 2008; Li et al., 2010; Ognjanova-Rumenova and Vass, 1998; Pantocsek, 1886; Saint Martin and Saint Martin, 2005; Servant-Vildary et al., 1988; VanLandingham, 1991; Yang et al., 2007). The origin of *Pinnularia* thus predates the Late Eocene, and the occurrence of several taxa within the Wagon Gap material suggests an even earlier origin for the genus. Thus, despite the well-established fossil record dating from after the diversification of *Pinnularia* into different morphological types, the earlier fossil record is very scarce and too fragmentary to provide detailed information about the timing of early diversification events. With this in mind, and given that ghost lineages are anticipated to be common in algae in general and diatoms in particular (Brown and Sorhannus, 2010), it is particularly relevant to produce a well-resolved, time-calibrated phylogeny for this genus.

The goal of the present study is to reconstruct a molecular phylogeny for a representative selection of *Pinnularia* taxa spanning the morphological variability of the genus, and to infer a time-calibrated phylogeny constrained by accurately dated fossil representatives. To achieve this, we sequenced two nuclear markers (18S rDNA, 28S rDNA), two plastid markers (*rbc*L, *psb*A) and a mitochondrial marker (*cox*1) from 36 species of the genus and inferred phylogenies using partitioned models in a likelihood framework. We also present new *Pinnularia* fossils from the Middle Eocene of Canada, which are included as constraints in the relaxed molecular clock analyses.

## 2. Material and methods

### 2.1. Taxon sampling

In a first step, the *rbc*L and LSU of 139 cultured *Pinnularia* and *Caloneis* strains were sequenced as described below and all available additional *rbc*L and LSU sequences downloaded from GenBank. From a maximum parsimony guide tree produced from these sequences, one strain per species was selected (in a few cases, sequence divergence was so high that we selected two or even more strains) for sequencing of three additional loci (18S rDNA, *psb*A and *cox*1). Sequences from additional species that were not available in our strain collection were included from GenBank. *Caloneis amphisbaena* (AM710596 and AM710507) was removed at a later stage because its deviant 28S sequence resulted in a long branch. However, this did not affect the phylogenetic relationships (data not shown).

This selection procedure resulted in a total of 36 included *Pinnularia* taxa (Tables 1 and 2), covering the range of morphological variation within the genus (Fig. 1). Because the relationships between *Pinnularia* and *Caloneis* remain ambiguous (Bruder and Medlin, 2008; Bruder et al., 2008; Cox, 1988; Mann, 2001; Round et al., 1990), we added three sequences of *Caloneis*. Five taxa from the genera *Sellaphora*, *Eolimna* and *Mayamaea* were selected as outgroups based on their apparent phylogenetic relatedness (Bruder and Medlin, 2008; Bruder et al., 2008; Evans et al., 2008; Kooistra et al., 2003). A list of the cultured and sequenced taxa is provided in Table 2, together with their geographical origin and morphometric data. Summary photomicrographs of the strains considered are shown in Fig. 1. Identifications are based on Krammer (2000) for *Pinnularia* and Krammer and Lange-Bertalot (1986) for *Caloneis*. Voucher slides of oxidized material of all natural samples and cultures and samples of extracted and purified DNA are held in the Laboratory of Protistology & Aquatic Ecology (Ghent University) and are available upon request.

### 2.2. DNA extraction, amplification and sequencing

DNA was extracted from centrifuged diatom cultures following Zwart et al. (1998) using a bead-beating method with phenol extraction and ethanol precipitation. After extraction, DNA was purified with a Wizard® DNA Clean-up system (Promega). Sequences of the nuclear 18S and the D1–D2 region of 28S, the two plastid genes *rbcL* and *psbA*, and the mitochondrial gene *cox1* were amplified using standard PCR primers and protocols (Daugbjerg and Andersen, 1997; Elwood et al., 1985; Evans et al., 2007; Guillou et al., 1999; Gunderson et al., 1986; Jones et al., 2005; Saunders, 2005; Scholin et al., 1994; van Hannen et al., 1999; Yoon et al., 2002). PCR products were cleaned using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The sequencing reaction was performed by cycle sequencing (initial step of 1 min at 96 °C, 30 cycles of 10 s at 96 °C, 10 s at 50 °C and 1 min 15 s at 60 °C) using the ABI Prism BigDye V 3.1 Terminator Cycle Sequencing kit (Applied Biosystems). The resulting sequencing reaction products were analyzed on a Perkin–Elmer ABI Prism 3100 automated DNA sequencer (Applied Biosystems). Primer sequences of both PCR and sequencing reactions and PCR temperatures are listed in Appendix A. All newly generated sequences have been deposited in GenBank (Table 1).

### 2.3. Sequence alignment

The sequences of 18S, 28S, *rbc*L, *psb*A and *cox*1 were edited separately and automatically aligned using ClustalW (Thompson et al., 1994), as implemented in BioEdit 7.0.3 (Hall, 1999). Plastid and mitochondrial markers aligned unambiguously without any gaps. The 18S and 28S alignments were corrected manually using the secondary structure of *Toxarium undulatum* (Alverson et al., 2006) and *Apedinella radians* (Ben Ali et al., 2001), respectively, after which ambiguously aligned regions were removed. All alignments are available from TreeBase (URL: http://purl.org/phylo/treebase/phylows/study/TB2:S11769).

**Table 1**
Taxon list with Genbank accession numbers. Missing sequences due to failed PCR are indicated with a dash, sequences which were not available in GenBank are indicated by N.A. Data taken from GenBank are indicated in bold.

| Strain | Taxon | 18S | 28S | *psb*A | *rbc*L | *cox*1 |
|---|---|---|---|---|---|---|
| | ***Caloneis budensis* (Grunow) Krammer** | AM502003 | AM710559 | **N.A.** | AM710470 | **N.A.** |
| | ***Caloneis lauta* Carter and Bailey-Watts** | AM502039 | AM710595 | **N.A.** | AM710506 | **N.A.** |
| Cal 890 TM | *Caloneis silicula* (Ehrenberg) Cleve | JN418593 | JN418626 | JN418728 | JN418663 | – |
| | ***Eolimna minima* (Grunow) Lange-Bertalot** | AM501962 | AM710516 | **N.A.** | AM710427 | **N.A.** |
| (Wes2)f | *Mayamaea atomus* var. *permitis* (Hustedt) Lange-Bertalot | JN418600 | JN418633 | JN418735 | JN418670 | JN418700 |
| | ***Mayamaea atomus* var. *permitis* (Hustedt) Lange-Bertalot** | AM501969 | AM710524 | **N.A.** | AM710435 | **N.A.** |
| (Enc2)b | *P. acrosphaeria* W. Smith | – | JN418635 | JN418737 | JN418672 | JN418701 |
| Pin 876 TM | *P. acuminata* W. Smith | JN418597 | JN418630 | JN418732 | JN418667 | JN418697 |
| (W045)b | *P. australogibba* var. *subcapitata* Van de Vijver, Cahttová and Metzeltin | – | JN418636 | JN418738 | JN418673 | JN418702 |
| Alka 1 | *P. borealis* Ehrenberg | JN418592 | JN418625 | JN418727 | JN418662 | – |
| (Tor12)d | *P. borealis* Ehrenberg *cf.* var. *subislandica* Krammer | JN418570 | JN418603 | JN418705 | JN418640 | – |
| (Ecrins7)a | *P. borealis* Ehrenberg var. *borealis* Krammer | – | JN418634 | JN418736 | JN418671 | – |
| (Tor3)a | *P. borealis* Ehrenberg var. *subislandica* Krammer | JN418575 | JN418608 | JN418710 | JN418645 | – |
| (Tor11)b | *P. cf. altiplanensis* Lange-Bertalot | JN418573 | JN418606 | JN418708 | JN418643 | JN418678 |
| Cal 878 TM | *P. cf. isselana* Krammer | JN418594 | JN418627 | JN418729 | JN418664 | JN418694 |
| (Ecrins4)a | *P. cf. marchica* Ilka Schönfelder | JN418569 | JN418602 | JN418704 | JN418639 | JN418676 |
| (B2)c | *P. cf. microstauron* (Ehrenberg) Cleve ("southern *microstauron*") | JN418568 | JN418601 | JN418703 | JN418638 | JN418675 |
| Pin 889 MG | *P. grunowii* Krammer | JN418588 | JN418621 | JN418723 | JN418658 | JN418690 |
| | ***P. mesolepta* (Ehrenberg) W. Smith** | AM501994 | AM710550 | **N.A.** | AM710461 | **N.A.** |
| Pin 706 F | *P. neglectiformis* Krammer | JN418596 | JN418629 | JN418731 | JN418666 | JN418696 |
| (Tor1)a | *P. neomajor* Krammer | JN418571 | JN418604 | JN418706 | JN418641 | JN418677 |
| Corsea2 | *P. neomajor* Krammer | JN418585 | JN418618 | JN418720 | JN418655 | JN418687 |
| Pin 885 TM | *P. nodosa* (Ehrenberg) W. Smith | JN418587 | JN418620 | JN418722 | JN418657 | JN418689 |
| Pin 877 TM | *P. parvulissima* Krammer | JN418591 | JN418624 | JN418726 | JN418661 | JN418693 |
| (Tor4)i | *P. sp.* | JN418580 | JN418613 | JN418715 | JN418650 | JN418683 |
| (Tor4)r | *P. sp.* | JN418581 | JN418614 | JN418716 | JN418651 | JN418684 |
| (Wie)a | *P. sp.* | JN418578 | JN418611 | JN418713 | JN418648 | – |
| Pin 873 TM | *P. sp.* | JN418590 | JN418623 | JN418725 | JN418660 | JN418692 |
| PinnC7 | *P. sp.* | JN418583 | JN418616 | JN418718 | JN418653 | JN418685 |
| (Tor1)b | *P. sp.* (divergens-group) | JN418572 | JN418605 | JN418707 | JN418642 | – |
| (Tor7)c | *P. sp.* (divergens-group) | JN418582 | JN418615 | JN418717 | JN418652 | – |
| (Tor7)f | *P. sp.* (gibba-group) | JN418576 | JN418609 | JN418711 | JN418646 | JN418680 |
| (Tor8)b | *P. sp.* (gibba-group) | JN418577 | JN418610 | JN418712 | JN418647 | JN418681 |
| Pin 649 K | *P. sp.* (subcommutata-group) | JN418595 | JN418628 | JN418730 | JN418665 | JN418695 |
| Pin 883 TM | *P. sp.* (subcommutata-group) | JN418586 | JN418619 | JN418721 | JN418656 | JN418688 |
| Pin 650 K | *P. subanglica* Krammer | JN418598 | JN418631 | JN418733 | JN418668 | JN418698 |
| (Wie)c | *P. subcapitata* var. *elongata* Krammer | JN418579 | JN418612 | JN418714 | JN418649 | JN418682 |
| Corsea10 | *P. subcommutata* var. *nonfasciata* Krammer | JN418584 | JN418617 | JN418719 | JN418654 | JN418686 |
| | ***P. substreptoraphe* Krammer** | AM502036 | AM710592 | **N.A.** | AM710503 | **N.A.** |
| (Enc2)a | *P. viridiformis* Krammer | JN418574 | JN418607 | JN418709 | JN418644 | JN418679 |
| Pin 870 MG | *P. viridiformis* Krammer | JN418589 | JN418622 | JN418724 | JN418659 | JN418691 |
| | ***P. viridis* (Nitzsch) Ehrenberg** | AM502023 | AM710579 | **N.A.** | AM710490 | **N.A.** |
| (Bfp5x8)F1–3 | *Sellaphora blackfordensis* D.G. Mann and S. Droop | JN418599 | JN418632 | JN418734 | JN418669 | JN418699 |
| (Bfp04)02 | *Sellaphora blackfordensis* D.G. Mann and S. Droop | – | JN418637 | – | JN418674 | – |

### 2.4. Model testing and phylogenetic analyses

Nine alternative partitioning strategies were tested: partitioning into genes, codon positions, stem and loop regions of rDNA regions, functionality and some combinations of these. Selection of a suitable partitioning strategy and models for the different partitions was based on the Bayesian Information Criterion (BIC, Schwarz, 1978). For each partitioning strategy six substitution models were optimized (JC, F81, K80, HKY, SYM and GTR) with or without a proportion of invariable sites and/or a gamma distribution to accommodate rate variation across sites. All parameters were unlinked between partitions. The preferred model and partitioning strategy was a GTR + $\Gamma_4$ in which 18S and 28S each formed a separate rDNA partition, whereas plastid and mitochondrial genes were separated by genome and were both partitioned into three codon positions.

Single genes and the complete concatenated dataset were analyzed by maximum likelihood phylogenetic inference using RAxML 7.2.6 (Stamatakis, 2006) under the preferred model and partition strategy with 10,000 independent tree searches from randomized MP starting trees. Maximum likelihood bootstrap analyses (Felsenstein, 1985) consisted of 1000 replicates. Bayesian phylogenetic inference was carried out with MrBayes v.3.1.2 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) under the same model and partitioning scheme. Using default settings, two independent runs with four incrementally heated Metropolis-coupled Monte-Carlo Markov Chains were run for five million generations for individual genes and 70 million generations for the concatenated dataset. Runs were sampled every 1000th generation, and convergence and stationarity of the log-likelihood and parameter values was assessed using Tracer v.1.5 (Rambaut and Drummond, 2007). The first million generations and 10 million generations were discarded as burn-in for the single-gene and multi-gene analyses, respectively. Runs for *psb*A converged only when using a heating factor of 0.1. We therefore also analyzed the concatenated dataset with a heating factor of both 0.2 and 0.1, but no differences in topology, likelihood or posterior probabilities were detected. The post-burn-in trees from the different runs were summarized and posterior probabilities (PPs) were calculated in MrBayes using the sumt command. All analyses were performed using the computer cluster Bioportal (Kumar et al., 2009).

### 2.5. Fossil Pinnularia

In order to assist with the temporal calibration of the multigene phylogeny, we document a new fossil occurrence of the genus

**Table 2**
List of the 38 taxa cultured and sequenced for this study, with source locality and date of collection, identity of collector and morphometric measurements (length, width and stria density from 10 valves per taxon, ± standard deviation, n.m. = not measured).

| Strain | Taxon | Locality and date of collection | Collector | Length (μm) | Width (μm) | # Stria per 10 μm |
|---|---|---|---|---|---|---|
| Cal 890 TM | *Caloneis silicula* (Ehrenberg) Cleve | Threipmuir Resr. (Scotland) – 02.III.2008 | D.G. Mann | 48.8 ± 3.7 | 13.0 ± 0.9 | 17.4 ± 0.5 |
| (Wes2)f | *Mayamaea atomus* var. *permitis* (Hustedt) Lange-Bertalot | De Panne (Belgium) – 03.III.2008 | C. Souffreau | 8.7 ± 0.1 | 3.9 ± 0.2 | 22.4 ± 1.9 |
| (Enc2)b | *Pinnularia acrosphaeria* W. Smith | Ovalle (Chile) – 22.I.2007 | C. Souffreau | 22.8 ± 1.3 | 10.6 ± 0.4 | 12.1 ± 0.3 |
| Pin 876 TM | *P. acuminata* W. Smith | Threipmuir Resr. (Scotland) – 02.III.2008 | D.G. Mann | 38.9 ± 2.6 | 11.5 ± 0.8 | 10.0 ± 0.0 |
| Alka 1 | *P. borealis* Ehrenberg | Podkova (Czech Republic) – 19.IV.2006 | A. Poulíčková | 37.4 ± 1.3 | 8.2 ± 0.6 | 4.6 ± 0.3 |
| (Tor12)d | *P. borealis* Ehrenberg *cf.* var. *subislandica* Krammer | Seno Otway (Chile) – 02.II.2007 | C. Souffreau | 36.8 ± 1.6 | 9.5 ± 0.5 | 5.0 ± 0.0 |
| (Ecrins7)a | *P. borealis* Ehrenberg var. *borealis* Krammer | Ailfroide (France) – 23.IX.2006 | C. Souffreau | 33.0 ± 1.7 | 8.1 ± 0.3 | 5.3 ± 0.4 |
| (Tor3)a | *P. borealis* Ehrenberg var. *subislandica* Krammer | Torres del Paine (Chile) – 30.I.2007 | C. Souffreau | 27.3 ± 0.6 | 9.6 ± 0.3 | 6.0 ± 0.0 |
| (Tor11)b | *P. cf. altiplanensis* Lange-Bertalot | Torres del Paine (Chile) – 01.II.2007 | C. Souffreau | 17.9 ± 0.5 | 4.5 ± 0.3 | 20.0 ± 0.0 |
| Cal 878 TM | *P. cf. isselana* Krammer | Threipmuir Resr. (Scotland) – 02.III.2008 | D.G. Mann | 33.4 ± 1.1 | 7.4 ± 0.3 | 12.1 ± 0.3 |
| (Ecrins4)a | *P. cf. marchica* Ilka Schönfelder | Ailfroide (France) – 22.IX.2006 | C. Souffreau | 27.3 ± 0.4 | 5.4 ± 0.4 | 15.6 ± 0.7 |
| (B2)c | *P. cf. microstauron* (Ehrenberg) Cleve ("southern *microstauron*") | Beak Island (Antarctic Peninsula) – I.2006 | E. Verleyen | 38.9 ± 1.4 | 9.9 ± 0.3 | 14.0 ± 0.4 |
| Pin 889 MG | *P. grunowii* Krammer | Balerno (Scotland) – 02.III.2008 | D.G. Mann | 41.9 ± 0.8 | 7.9 ± 0.2 | 13.0 ± 0.8 |
| Pin 706 F | *P. neglectiformis* Krammer | Figgate Loch (Scotland) – 23.I.2008 | D.G. Mann | 105.3 ± 3.9 | 18.5 ± 0.5 | 8.1 ± 0.2 |
| (Tor1)a | *P. neomajor* Krammer | Torres del Paine (Chile) – 30.I.2007 | C. Souffreau | 151.4 ± 3.6 | 24.9 ± 0.4 | 8.2 ± 0.4 |
| Corsea 2 | *P. neomajor* Krammer | Glen Corse Resr. (Scotland) – 06.IV.2008 | P. Vanormelingen | 169.2 ± 1.9 | 28.4 ± 0.5 | 7.8 ± 0.3 |
| Pin 885 TM | *P. nodosa* (Ehrenberg) W. Smith | Threipmuir Resr. (Scotland) – 02.III.2008 | D.G. Mann | 42.1 ± 1.2 | 6.7 ± 0.2 | 10.0 ± 0.0 |
| Pin 877 TM | *P. parvulissima* Krammer | Threipmuir Resr. (Scotland) – 02.III.2008 | D.G. Mann | 58.3 ± 1.5 | 9.8 ± 0.3 | 10.0 ± 0.0 |
| (Tor4)i | *P. sp.* | Torres del Paine (Chile) – 31.I.2007 | C. Souffreau | 33.3 ± 1.3 | 9.4 ± 0.5 | 14.0 ± 0.0 |
| (Tor4)r | *P. sp.* | Torres del Paine (Chile) – 31.I.2007 | C. Souffreau | 42.6 ± 0.4 | 5.8 ± 0.3 | 12.4 ± 0.5 |
| (Wie)a | *P. sp.* | De Wieden (The Netherlands) – 26.II.2007 | P. Vanormelingen | 46.7 ± 0.9 | 9.7 ± 0.2 | 12.1 ± 0.2 |
| Pin 873 TM | *P. sp.* | Threipmuir Resr. (Scotland) – 02.III.2008 | D.G. Mann | 28.3 ± 2.8 | 10.2 ± 0.4 | 12.3 ± 0.5 |
| PinnC7 | *P. sp.* | Iles Crozet (Sub-Antarctica) – XII.2004 | B. Van de Vijver | 11.8 ± 0.8 | 6.3 ± 0.3 | 13.8 ± 0.6 |
| (Tor1)b | *P. sp.* (divergens-group) | Torres del Paine (Chile) – 30.I.2007 | C. Souffreau | 51.5 ± 2.1 | 10.4 ± 0.3 | 10.8 ± 0.8 |
| (Tor7)c | *P. sp.* (divergens-group) | Torres del Paine (Chile) – 31.I.2007 | C. Souffreau | 88.4 ± 1.7 | 12.5 ± 0.4 | 10.4 ± 0.5 |
| (Tor7)f | *P. sp.* (gibba-group) | Torres del Paine (Chile) – 31.I.2007 | C. Souffreau | 61.2 ± 1.6 | 8.9 ± 0.2 | 10.1 ± 0.5 |
| (Tor8)b | *P. sp.* (gibba-group) | Torres del Paine (Chile) – 31.I.2007 | C. Souffreau | 40.0 ± 0.5 | 5.9 ± 0.3 | 11.8 ± 0.4 |
| (W045)b | *P. australogibba* var. *subcapitata* Van de Vijver, Cahttová and Metzeltin | Amsterdam Island – 06.XII.2007 | B. Van de Vijver | 37.1 ± 0.8 | 8.7 ± 0.2 | 11.2 ± 0.6 |
| Pin 649 K | *P. sp.* (subcommutata-group) | Kew Billabong (Australia) – 02.XII.2007 | K.M. Evans | 53.0 ± 3.2 | 11.7 ± 0.3 | 10.1 ± 0.2 |
| Pin 883 TM | *P. sp.* (subcommutata-group) | Threipmuir Resr. (Scotland) – 02.III.2008 | D.G. Mann | 63.1 ± 2.8 | 14.1 ± 0.1 | 10.3 ± 0.5 |
| Pin 650 K | *P. subanglica* Krammer | Kew Billabong (Australia) – 02.XII.2007 | K.M. Evans | 46.2 ± 0.3 | 8.2 ± 0.2 | 11.1 ± 0.7 |
| (Wie)c | *P. subcapitata* var. *elongata* Krammer | De Wieden (The Netherlands) – 26.II.2007 | P. Vanormelingen | 53.3 ± 0.6 | 6.3 ± 0.3 | 11.9 ± 0.3 |
| Corsea 10 | *P. subcommutata* var. *nonfasciata* Krammer | Glen Corse Resr. (Scotland) – 06.IV.2008 | P. Vanormelingen | 50.1 ± 1.8 | 10.7 ± 0.1 | 11.8 ± 0.4 |
| (Enc2)a | *P. viridiformis* Krammer | Ovalle (Chile) – 22.I.2007 | C. Souffreau | 74.9 ± 1.8 | 17.7 ± 0.5 | 8.8 ± 0.4 |
| Pin 870 MG | *P. viridiformis* Krammer | Balerno (Scotland) – 02.III.2008 | D.G. Mann | 80.5 ± 2.3 | 16.7 ± 0.3 | 9.2 ± 0.3 |
| (Bfp04)02 | *Sellaphora blackfordensis* D.G. Mann and S. Droop | Culture collection BCCM/DCG | – | 38.1 ± 13.6 | 9.0 ± 1.4 | 20.2 ± 1.1 |
| (Bfp5x8)F1-3 | *Sellaphora blackfordensis* D.G. Mann and S. Droop | Culture collection BCCM/DCG | – | n.m. | n.m. | n.m. |

*Pinnularia* from Middle Eocene lacustrine facies in northwestern Canada. The Giraffe kimberlite fossil locality (64°48′N, 110°04′W) contains unpermineralized diatom-rich organic sediments, which constitute the post-eruptive infilling of the site's maar crater. These sediments are dated between 40 and 48 Ma (Lutetian Stage), and have already revealed first occurrences for a number of freshwater diatom lineages (Siver and Wolfe, 2007; Siver et al., 2010; Wolfe and Siver, 2009). The excellent preservation of Giraffe *Pinnularia* specimens is illustrated by a range of light and scanning electron micrographs (Fig. 3).

### 2.6. Relaxed molecular clock analysis

A time-calibrated phylogeny was inferred using a relaxed molecular clock method as implemented in BEAST v.1.5.4 (Drummond and Rambaut, 2007). An uncorrelated lognormal clock model and Yule tree prior were specified along with the same partitioning scheme and models of sequence evolution used for the phylogenetic reconstruction. Different calibration strategies based on the fossil record were carried out to assess congruence between calibration points. The root node calibration prior was varied (uniform,

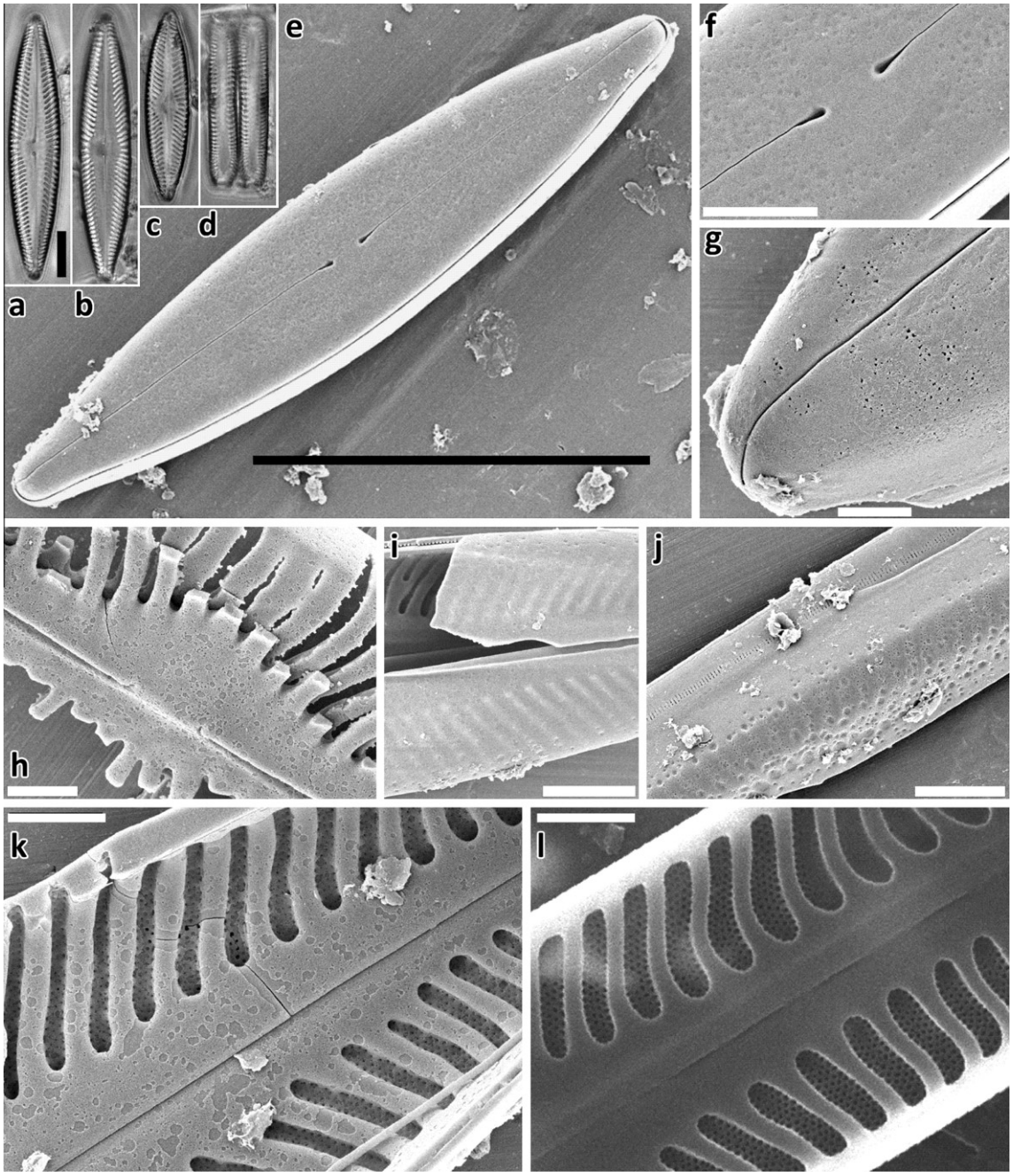


**Fig. 3.** Middle Eocene *Pinnularia* diatoms from the Giraffe kimberlite pipe fossil locality with a minimum age of 40 Ma. In light microscopy, valve (a–c) and girdle (d) views illustrate morphologies consistent with the genus as currently circumscribed. Under scanning electron microscopy, the valve overview (e), close-ups of proximal (f) and distal external raphe endings (g), as well as details of the outer (h–j) and inner (k) valve surfaces all confirm considerable affinity with extant forms. The alveolate striae of the Eocene diatom (k) are directly comparable with numerous modern congeners, here shown is a representative of the *P. subgibba* subclade [(Tor4)r] in (l). Scale bars are 10 μm (a–d), 20 μm (e), 5 μm (f, i and j), and 2 μm (g, h, k and l).

**Table 3**
List of calibration points and alternative calibration schemes used to date the phylogenetic tree, with indication of the minimum (min) and maximum (max) age constraints. Numbers refer to the calibration points on the phylogenetic tree (Fig. 4).

| Node | | Calibration scheme with age constraints (Ma) | | | | | Bauplan | Reference |
|---|---|---|---|---|---|---|---|---|
| | | Root | A | B | C | E | | |
| Root node | Min | 40 | 40 | 40 | 40 | 40 | *Pinnularia spp.* | This study |
| | max | 75 | 75 | 75 | 75 | 100 | | |
| *neomajor–neglectiformis* clades | Min | – | 11.7 | 11.7 | 11.7 | 11.7 | *P. viridis* | Saint Martin and Saint Martin (2005) |
| | Max | – | 40 | 40 | 40 | 40 | | |
| *nodosa–acrosphaeria* | Min | – | 14.5 | 14.5 | 14.5 | 14.5 | *P. nodosa* | Pers. comm. A. Menicucci |
| | Max | – | 40 | 40 | 40 | 40 | | |
| *grunowii–mesolepta* | Min | – | – | 11.7 | 11.7 | 11.7 | *P. mesolepta* | Saint Martin and Saint Martin (2005) |
| | Max | – | – | 40 | 40 | 40 | | |
| *borealis–microstauron* | Min | – | – | – | 13 | 13 | *P. borealis* | Servant-Vildary et al. (1988) |
| | Max | – | – | – | 70 | 70 | | |

gamma and truncated normal probability distribution), and all internal calibration points had a uniform prior (Ho and Phillips, 2009). All calibration points and strategies are summarized in Table 3 and fossil constraints are explained below. For each calibration strategy, three independent runs were carried out, starting from a user-defined, linearized time-constrained starting tree constructed in r8s (Sanderson, 2003) based on the ML tree of the phylogenetic analysis. Markov chains were run for 50 million generations and sampled every 1000 generations. Convergence and stationarity of log-likelihood and parameter values were assessed using Tracer v.1.5 (Rambaut and Drummond, 2007) and 10% of the generations were discarded as burn-in. The post-burn-in trees of three independent runs were combined, after which a maximum clade credibility chronogram with mean node heights was calculated with TreeAnnotator v.1.5.

No strict molecular clock model was used. MrBayes analyses were run under the unconstrained and strict clock models (single run, 1 million generations) and the Bayes factor $2*\log(B_{10})$, which weighs the evidence in favor of model 1 (unconstrained) compared to model 0 (strict clock), was 317.12, indicating that the rate of molecular evolution was not constant and the strict molecular clock should not be applied to the data. In order to evaluate potential impacts of choosing the uncorrelated lognormal model in BEAST, the molecular clock analysis was repeated in PhyloBayes 3.3f (Lartillot et al., 2009) using an autocorrelated model of rate evolution (see Supplementary Fig. S1 for further information).

Fossil diatoms were used to calibrate the tree in geological time using the following constraints. The root node of *Pinnularia* was given a minimum age of 40 Ma following the recovery of *Pinnularia* in sediments from the Giraffe locality (Fig. 3). Two different maximum age scenarios were considered. A first maximum age of the root node was set at 100 Ma (Albian, Lower Cretaceous), based on the occurrence of rich diatom floras in southern hemisphere ocean sediments containing a range of centric morphologies but neither raphid nor araphid pennate taxa (Gersonde and Harwood, 1990; Harwood and Gersonde, 1990). This date can be assumed to be highly conservative, and therefore we also considered an alternate upper boundary for *Pinnularia* at 75 Ma (Campanian, Upper Cretaceous), the period for which the earliest araphid pennate diatoms have been confirmed (Chambers, 1966; Hajós and Stradner, 1975). Because these fossils are abundant but not diverse, it is believed that araphid pennate forms had evolved only recently by this time (Sims et al., 2006). Following this reasoning, we assume that raphid pennates had not yet evolved.

Four internal calibration points were also used (Fig. 4), based on first-appearance dates for characteristic morphological types or Baupläne recovered from late-Oligocene to mid-Miocene freshwater deposits. The following morphological types were used: diatoms morphologically similar to extant *P. viridis* (11.7 Ma; Saint Martin and Saint Martin, 2005), *P. mesolepta* (11.7 Ma; Saint Martin and Saint Martin, 2005), *P. nodosa* (14.5 Ma; pers. comm. A. Menicucci) and *Pinnularia borealis* (13.0 Ma; Servant-Vildary et al., 1988). These fossil occurrences were used as minimum estimates for the corresponding Bauplan, thus constraining the MRCA (most recent common ancestor) of clades containing species within each of these (Fig. 4). The maximum age for these different Baupläne was set at 75 Ma when using the conservative calibration scheme, and 40 Ma for the less conservative calibration strategy. Although specimens originating from the Wagon Gap Formation clearly belong to *Pinnularia* (Lohman and Andrews, 1968), in the absence of electron microscopical observations they cannot be assigned with confidence to any of the modern morphological types used in our phylogeny. Linked to the uncertain dating of this locality, Wagon Gap *Pinnularia* were not used as an explicit constraint for our phylogeny.

## 3. Results

### 3.1. Dataset properties

Our 44-taxon, five-locus dataset is 87% filled (192 sequences; Table 1) and includes 4852 sites of which 1012 (21%) are parsimony-informative. The most complete markers are 28S rDNA and *rbc*L (0 of 44 missing), followed by 18S rDNA (4 sequences or 9% missing), *psb*A (8 or 18% missing) and *cox*1 (16 or 36% missing). Missing sequences were due to failed PCR or because they were not available in GenBank (Table 1). Previous studies revealed that taxa with missing sequence data (up to 95% of missing data) can still be accurately placed as long as the overall number of characters is large (Philippe, 2004; Wiens, 2003; Wiens and Moen, 2008), a requirement that we believe is met in this study. The 18S rDNA sequences provided most characters (1706) followed by *rbc*L (1386), *psb*A (762), cox1 (615) and 28S rDNA (383). The percentage of parsimony-informative characters varied between genes, with the highest percentages for 28S rDNA (42%) and *cox*1 (37%), the lowest for *psb*A (10%), and 18S rDNA (18%) and *rbc*L (17%) being intermediate.

### 3.2. Phylogenetic relationships

Phylogenies of individual genes revealed no strong conflicts; therefore we show only the results of the concatenated analysis. Similarly, ML and BI analyses of the concatenated dataset produced identical topologies so we show only the ML phylogeny with both ML bootstrap support (BS) and BI posterior probabilities (PP)

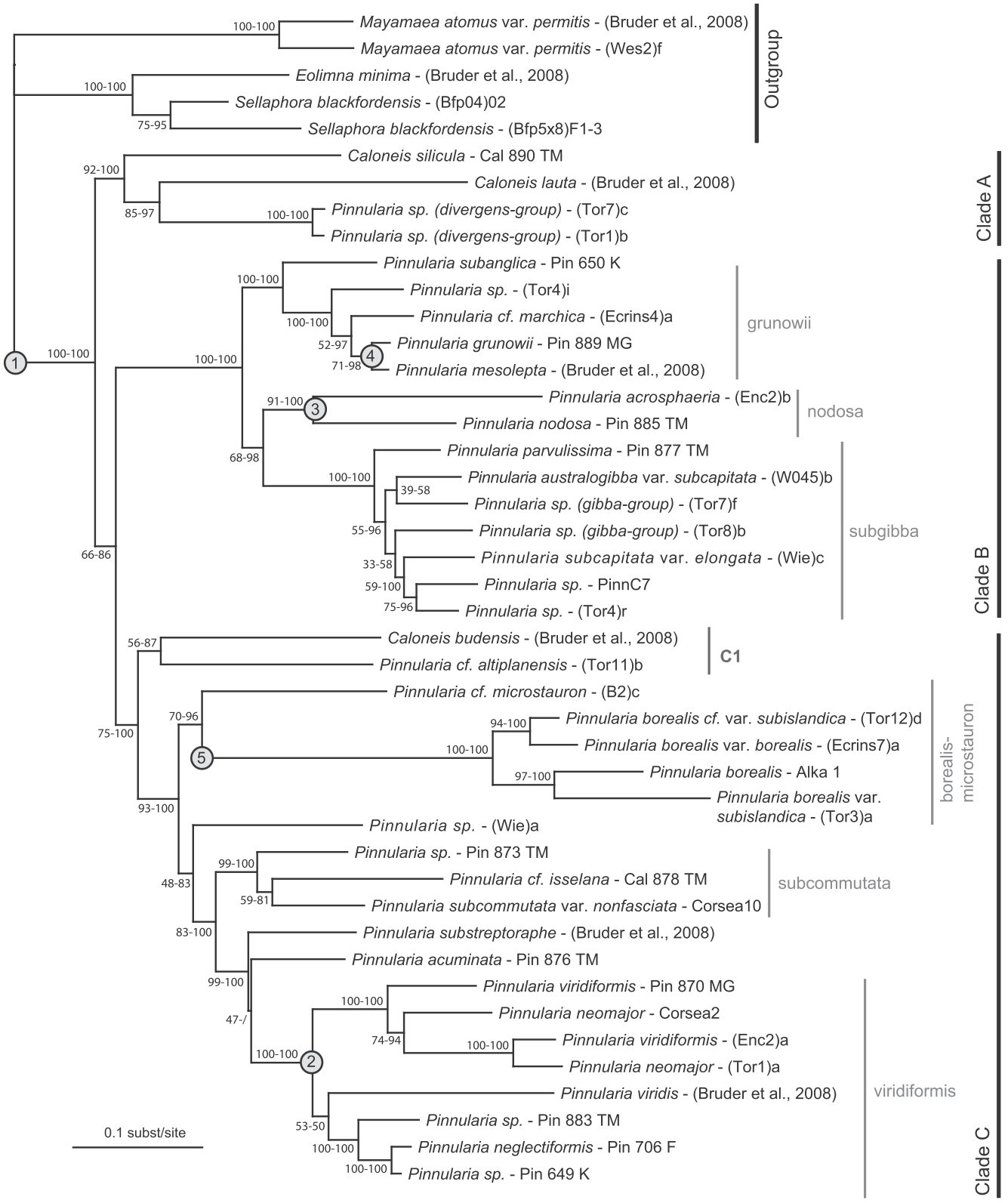


**Fig. 4.** Phylogenetic relationships within the genus *Pinnularia* inferred from a five-locus DNA alignment using maximum likelihood under a partitioned model. Numbers at nodes indicate statistical support, ML bootstrap proportions - BI posterior probabilities (both given as percentages). Encircled numbers represent nodes constrained in the relaxed molecular clock analysis (see Table 3).

(Fig. 4). The BI-based tree is available through TreeBase (URL: http://purl.org/phylo/treebase/phylows/study/TB2:S11769). Based on our concatenated dataset, taxa representing *Pinnularia* and *Caloneis* form a monophyletic group comprising three robustly supported clades that each contain several well-supported subclades. Clade A comprises *Caloneis silicula*, *Caloneis lauta*, and two species of *Pinnularia* cf. *divergens*. No apparent morphological synapomorphies unite these taxa, *C. silicula* being characterized by

linear external central raphe endings (Fig. 2a), small alveolar apertures (Fig. 2b) and parallel striae (Fig. 1d), whereas the *divergens*-group has drop-like external central raphe endings (Fig. 2c), large alveolar apertures (Fig. 2d) and oblique striae (Fig. 1e and f).

Clade B includes small, linear *Pinnularia* species (Fig. 1g-p) that have drop-like external central raphe endings (Fig. 2c). Clade B is subdivided into three well-supported subclades: the *grunowii*, *nodosa* and *subgibba* subclades. The *grunowii* subclade contains *P. grunowii*, *P. subanglica* and *P. marchica*. Species in this group contain a single H-shaped plastid with two pyrenoids (Fig. 1b′, c′), whereas, with the exception of some species (*C. silicula, P. cf. altiplanensis*, *P. sp.* (Wie)a, *P. cf. isselana*, *P. microstauron*) all other clades of the tree are characterized by two linear, strip-like plastids (Fig. 1y and z). The *nodosa* subclade includes *P. nodosa* and *P. acrosphaeria*, which are both characterized by irregular wart-like structures on the external valve face (Fig. 1j and k; detail in Fig. 2c). The *subgibba* subclade contains different species resembling *P. subgibba*, *P. parvulissima*, and *P. subcapitata*, all characterized by a combination of an elongated shape often with ghost striae (striae that are partially filled in with silica) and a broad, non-porous, central area (fascia) (Fig. 1l–p).

Clade C is composed of two subclades. Subclade C1 contains *Caloneis budensis* and *Pinnularia* cf. *altiplanensis* (Fig. 1x), but is poorly supported (BS = 56; PP = 87). Subclade C2 is further subdivided into moderately to highly supported groups, all of which have characteristic morphologies. The *borealis–microstauron* group includes the *P. borealis* species complex, which is easily recognizable by its relatively broad cells and coarse striae (Fig. 1w), and *Pinnularia microstauron* with its relatively robust shape and wedge-shaped ends (Fig. 1v). The remaining groups (the *subcommutata* and *viridiformis* clades) contain large, elongated-elliptical species, with almost parallel striae and small central areas (Fig. 1q–t), undulate external raphe fissures, intermediate-sized alveolar openings (Fig. 2f), and plastids without pyrenoids. The *subcommutata* group contains *Pinnularia* species with linear central raphe endings (Fig. 2e) whereas species of the *viridiformis* group have round central raphe endings (Fig. 2g) and a complex raphe system, visible as three parallel (Fig. 1q) or twisting lines (Fig. 1r) with light microcopy (Krammer, 1992, 2000; Round et al., 1990). Sister to the *viridiformis* and *subcommutata* groups is the isolate (Wie)a (Fig. 1u), an undescribed *Pinnularia* species morphologically similar to *P. microstauron*. In the single-gene trees, (Wie)a is always retrieved in clade C, but with a closer (unsupported) relationship to the *viridiformis*-group based on 18S and 28S rDNA, or with a sister relationship to *P.* cf. *divergens* and *P. borealis* based on *rbc*L and *psb*A.

### 3.3. Newly-observed fossil Pinnularia specimens

*Pinnularia* specimens from the Middle Eocene Giraffe kimberlite locality show considerable affinity to modern representatives of the genus. Although they are never abundant in this material, sufficient numbers have been observed to document their morphology in light and scanning electron microscopy (Fig. 3). The specimens are characterized by an alveolate valve structure (Fig. 3k), placing them confidently in the genus *Pinnularia*. Moreover, the Giraffe *Pinnularia* forms have large alveolar openings (Fig. 3k), identical to the *Pinnularia* species from clades A, B, C1 and the *microstauron–borealis* group of clade C (e.g. Figs. 3l and 2d and h). However, the particular combination of morphological characters of the fossil specimens, including wart-like structures on the valves (similar to *P. acrosphaeria*), the very wide axial area (similar to *P. acuminata*) and the distinctly lanceolate or "naviculoid" shape (not encountered in our species) meant that we were not able to place these specimens within a known *Pinnularia* Bauplan in the present phylogeny nor, to our knowledge, with any currently described species. Their distribution in samples spanning tens of meters of core, typically associated with chrysophyte and euglyphid siliceous microfossils and benthic eunotioid diatoms, and their occurrence within the mudstone matrix discounts the possibility that they are contaminants. A more complete analysis of the authenticity of Giraffe microfossils is given elsewhere (Wolfe and Siver, 2009). Being $\geqslant$40 Ma in age, these *Pinnularia* specimens represent the earliest known forms of *Pinnularia*, providing us with a fossil calibration point to constrain the timing of the basal node of our phylogeny.

### 3.4. Time-calibrated phylogeny

The different calibration schemes analyzed in BEAST all recover identical relationships to the RAxML and MrBayes analyses with comparable posterior probabilities. The average node ages and 95% HPD (Highest Posterior Density) resulting from the different calibration schemes are given in Fig. 5. Estimates using the internal calibration points based on the frustule morphologies of *viridis*, *nodosa* and *borealis* were consistent with estimates using only a root calibration. However, adding the internal calibration point based on *grunowii–mesolepta* resulted in larger 95% HPD intervals and higher average node ages (plus 5–12 Ma) at all nodes. Using only the root constraint, the *grunowii–mesolepta* node is estimated as being 3.5–12.6 Ma. Constraining this node to have a minimum age of 11.7 Ma therefore pushes its own time-estimate, and that of all other nodes, back in time (Fig. 5).

Using the conservative maximum root age of 100 Ma resulted in a 7–11 Ma higher average root node age and a 12–20 Ma higher 95% HPD maximum root age compared with the alternative 75 Ma maximum age boundary, but for the internal nodes, changes in mean age and 95% HPD were only 1–6 Ma. Furthermore, the use of different probability distributions on the root calibration (uniform, gamma or truncated normal) resulted in predictable differences in average node ages and 95% HPD up to 6 Ma for the less-conservative 75 Ma schemes and up to 10 Ma for the conservative 100 Ma scheme. Because the Eocene fossil record is so fragmentary, we chose the conservative uniform prior probability distribution as the preferred calibration strategy.

Fig. 6 presents our preferred relaxed molecular clock as applied to the *Pinnularia* phylogeny, constrained by all four internal calibration points (scheme C = *viridis*, *borealis*, *nodosa* and *mesolepta*), with the root age constrained between 40 and 75 Ma using a uniform distribution. ESS values of this analysis are given in Table S1, and exploration of the posterior node age estimates with and without the dataset show that the DNA sequences contain a clear signal about the inferred node ages (see Appendix B). This estimate places the origin of *Pinnularia* around 64 Ma, between the Campanian and Early Eocene [75–50]. The *Pinnularia* complex began to diversify between the Maastrichtian and Middle Eocene 60 Ma [73–45] and continued through to the Pliocene, with no obvious evidence for increased rates of diversification over any time period. Clades A and C diversified mostly during the Paleocene–Eocene (respectively, 49 Ma [65–33] and 52 Ma [64–38]), while clade B started to diverge between the Eocene and Oligocene (39 Ma [50–28]).

## 4. Discussion

### 4.1. Relationships within Pinnularia

The current study presents a multi-gene phylogeny of the genus *Pinnularia* and extends the sampling of Bruder et al. (2008) threefold. Our five-locus analysis yielded a well-supported phylogenetic tree, with *Pinnularia* and *Caloneis* resolved in a single monophyletic clade, in accordance with the results of Bruder and Medlin (2008)

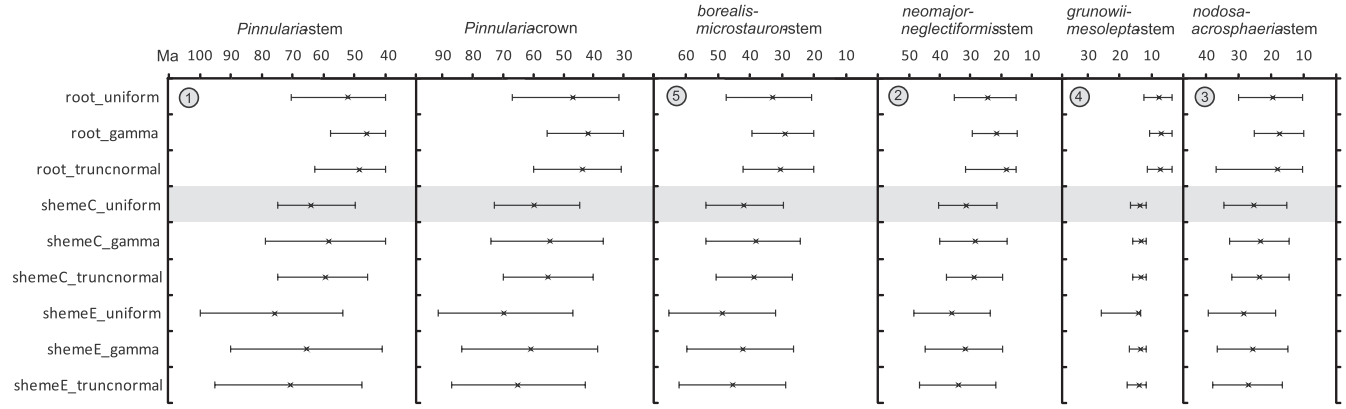


**Fig. 5.** Resulting mean node ages (indicated as "*x*"; in Ma) and 95% HPD (error bars; in Ma) for nine of the alternative calibration shemes (see Table 3), visualized for the different calibration points (numbered circles refer to the constrained nodes in Fig. 4). The preferred calibration sheme used for Fig. 6 is indicated in grey. "Root" shemes were only constrained on the root. Sheme C used a less conservative calibration, with a maximum of 75 Ma and all internal calibration points. Sheme E used a very conservative calibration with a maximum of 100 Ma and all internal calibration points. The influence of the constraint based on the *grunowii–mesolepta* Bauplan is clearly visible as the differences between age estimations based on "root" and "sheme C" or "sheme E". "Truncnormal" means a truncated normal probability distribution.

and Bruder et al. (2008). Our phylogeny shows that *Caloneis* (as currently defined) is not monophyletic, but rather that its species were recovered in various places among the *Pinnularia* species. This finding supports the earlier work of Cox (1988), Round et al. (1990) and Mann (2001) who concluded that the current separation of *Caloneis* from *Pinnularia* was not supported by plastid and frustule structure, and confirms and extends the molecular phylogenetic results of Bruder and Medlin (2008) and Bruder et al. (2008). Cleve (1894) delimited *Caloneis* from *Pinnularia* using light microscopy, but he admitted that their discrimination, based on stria orientation, the path of the terminal raphe fissures, plastids and the presence of "longitudinal lines" (the margins of the alveolar apertures on the inner side of the valve) was not consistent. While Krammer and Lange-Bertalot (Krammer, 1992; Krammer and Lange-Bertalot, 1986) argued that *Caloneis* can indeed be distinguished from *Pinnularia* based on a specific combination of characters, Round et al. (1990) suggested that, although *Pinnularia–Caloneis* is a natural group characterized by the double-walled, chambered (alveolate) valve structure, any split of the *Pinnularia–Caloneis* complex would not follow the traditional boundary between the two genera. Despite the fact that our molecular phylogeny contained only three *Caloneis* species, there is no apparent simple genetic boundary between *Pinnularia* and *Caloneis*. Nevertheless, ultrastructural characters, e.g. the size of the alveoli, and/or plastid characters can apparently be used, alongside molecular synapomorphies, to delimit groups or genera within this complex. However, much more genetic and morphological investigation of the *Pinnularia–Caloneis* complex is required before any attempt is made to subdivide it formally. Alternatively, subdivision could be considered unnecessary, given the highly characteristic general morphology of the complex, all species being included in an expanded *Pinnularia*.

*Pinnularia* is a morphologically diverse genus and the interspecific relationships based on our five-locus dataset generally follow morphological patterns. Our dataset revealed three large, well-supported clades, which were also recovered in the concatenated (18S, 28S, *rbc*L) dataset of Bruder et al. (2008) but without bootstrap support. With the exception of clade A, which contains two *Caloneis* species and the *Pinnularia divergens* group, the remaining clades are morphologically well-defined. Clade B includes *Pinnularia* taxa of intermediate size that have a fascia and is subdivided into three highly supported and morphologically well-defined subclades, of which the *grunowii* and *subgibba* clades were already recovered with good support by Bruder et al. (2008). The structural similarities between *P. acrosphaeria* and *P. nodosa* were noted by Cleve (1895) and more recently by Krammer and Lange-Bertalot (1986). The *grunowii* and *subgibba* subclades are also well-defined by H-shaped plastids and a wide fascia, respectively. Clade C contains the larger and more robust forms of *Pinnularia* and is subdivided into several well-supported subclades, each with a distinct morphology. The combined *viridiformis* and *subcommutata* group, characterized by intermediate alveolar openings (Fig. 2f), was previously described by Cleve (1895) as two groups, the Maiores and Complexae, and more recently as a single group by Krammer and Lange-Bertalot (1986). The phylogenetic signal in these different morphological characters was already suggested by the phylogeny of Bruder et al. (2008), and the present data confirm and give further support to their findings.

Besides these well-defined and well-supported clades, some clusters are less easily interpreted. Some enigmatic relationships within clade A can only be further resolved and understood through additional taxon sampling. This is also the case for the positions of the undescribed species (Wie)a and *P.* cf. *altiplanensis,* and the relationship between *P. microstauron* and *P. borealis*. Because only a small subset of the described *Pinnularia* species were included in this analysis, it is plausible that they represent more elaborate clusters. Adding taxa in those parts of the tree could therefore improve our understanding of the evolution of morphological features.

### 4.2. Evolution of the genus Pinnularia

Based on our preferred time-calibration, *Pinnularia* originated between the Campanian (Late Cretaceous, *ca.* 75 Ma) and Early Eocene (*ca.* 50 Ma). This estimate is 10–35 Ma older than the *ca.* 40 Ma origin deduced from the chronogram of Sorhannus (2007). However, the goal of the latter study was to elucidate deeper divergences within major diatom lineages, and furthermore only two *Pinnularia* taxa were included and analyzed with a single genetic marker (nuclear-encoded SSU rRNA). More comprehensive taxon sampling (Sanderson, 1990), the use of multiple genes (Rodríguez-Trelles et al., 2003), and full integration of the fossil record (Donoghue and Benton, 2007) have all been shown to improve the accuracy of molecular dating. In the case of *Pinnularia*, the $\geqslant$40 Ma Giraffe specimens convincingly support the older range of our estimated ages for the origin of the genus (*i.e.* Late Cretaceous). Given the morphological similarity of Giraffe *Pinnularia* forms to extant taxa, we are confident that such older forms exist. Although the Giraffe specimens are, so far, the oldest representatives from well-dated sediments, they do not inform the search for morphologically primitive *Pinnularia*, or its immediate ancestor.

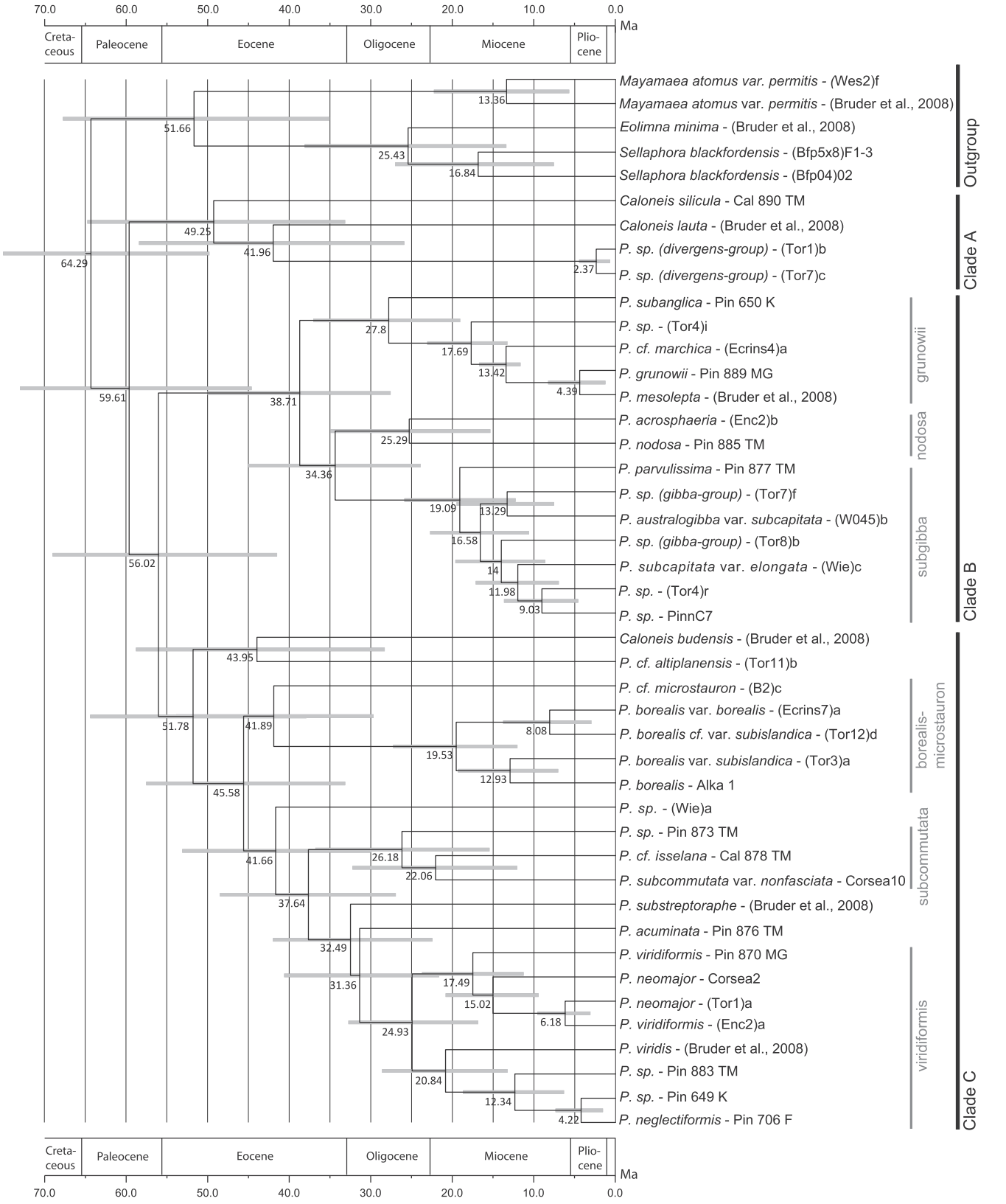


**Fig. 6.** Chronogram of the genus *Pinnularia* from a Bayesian relaxed molecular clock analysis performed with BEAST (Drummond and Rambaut, 2007), time-constrained by four internal fossil calibration points (*viridis*, *borealis*, *nodosa* and *mesolepta*) and a 40 Ma minimum and 75 Ma maximum root age with uniform probability distribution (scheme C) based on the fossil data from the Giraffe kimberlite Pipe. Values at nodes are mean node ages and grey bars represent 95% HPD (highest probability density) intervals.

The new estimate for the origin of *Pinnularia* implies a considerable gap in the fossil record of freshwater diatoms. The magnitude of this ghost lineage (*ca.* 10–35 Ma) should not be surprising given that freshwater deposits are underrepresented and understudied

in the paleontological record. Similar gaps apply to nearly all extant raphid diatom lineages, and the scarcity of fossil raphid diatoms of Late Cretaceous to Paleocene age has resulted in a lack of consensus on plausible minimum ages for most genera (but see Singh et al., 2006). It is therefore difficult to place the proposed age for *Pinnularia* in a broader context, although generally our observations are consistent with the preferred model of Berney and Pawlowski (2006), in which pennate diatoms arose some 98 (110–77) Ma ago, leaving the entire Late Cretaceous for the evolution of innovations such as the raphe. Apart from *Pinnularia*, no raphid diatom genera have yet been dated by phylogenetic methods, but the time-calibrated phylogeny of Sorhannus (2007) suggests that several raphid genera, e.g. *Lyrella* (see also Jones et al., 2005) are as old as, or even older than, our estimate for *Pinnularia*.

Given the available fossil and phylogenetic records, our estimated timing of 60 Ma for the origin of *Pinnularia* seems plausible. Since its origin, *Pinnularia* has diversified widely into a large number of species and a range of morphologies. The evolution of particular morphological innovations appears to have sparked the appearance of new lineages, such as the evolution of the "complex" raphe, which may have allowed the development of larger cells, as seen in the *viridiformis* group, or the evolution of H-shaped plastids in the *grunowii* group. Further research will be needed to integrate the ecological, morphological, genetic and evolutionary processes underlying the diversification of diatom genera such as *Pinnularia*. Our time-calibrated multi-gene phylogeny of *Pinnularia* forms a first important step towards this goal by providing a temporal context in which to interpret the diversification of the genus, and has only been possible by integrating molecular and paleontological information.

## Acknowledgments

We thank Aloisie Poulíčková for providing cultures of *Pinnularia borealis* and Elie Verleyen for providing live environmental samples. Alex Ball enthusiastically assisted with confocal microscopy. Sequencing was done by Andy Vierstrate. Funding was provided by the Fund for Scientific Research-Flanders (FWO-Flanders, comprising a PhD fellowship to CS, post-doctoral fellowships to HV and PV, and Research Grant 3G/0533/07). Research on fossil diatoms is partially supported by the US National Science Foundation (DEB-0716606 to PAS) and the Natural Sciences and Engineering Council of Canada (APW). Morphological analyses at the Natural History Museum, London, were supported by an EU framework 7 SYNTHESYS Grant (GB-TAF-77) to CS.

## Appendix A and B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2011.08.031.

## References

Alverson, A.J., Cannone, J.J., Gutell, R.R., Theriot, E.C., 2006. The evolution of elongate shape in diatoms. J. Phycol. 42, 655–668.
Ben Ali, A., De Baere, R., Van der Auwera, G., De Wachter, R., Van de Peer, Y., 2001. Phylogenetic relationships among algae based on complete large-subunit rRNA sequences. Int. J. Syst. Evol. Microbiol. 51, 737–749.
Berney, C., Pawlowski, J., 2006. A molecular time-scale for eukaryote evolution recalibrated with the continuous microfossil record. Proc. Roy. Soc. B. Biol. Sci. 273, 1867–1872.
Beszteri, B., John, U., Medlin, L.K., 2007. An assessment of cryptic genetic diversity within the *Cyclotella meneghiniana* species complex (Bacillariophyta) based on nuclear and plastid genes, and amplified fragment length polymorphisms. Eur. J. Phycol. 42, 47–60.
Brown, J.W., Sorhannus, U., 2010. A molecular genetic timescale for the diversification of autotrophic Stramenopiles (Ochrophyta): substantive underestimation of putative fossil ages. Plos One, 5.
Bruder, K., Medlin, L.K., 2008. Morphological and molecular investigations of naviculoid diatoms. II. Selected genera and families. Diatom Res. 23, 283–329.
Bruder, K., Sato, S., Medlin, L.K., 2008. Morphological and molecular investigations of naviculoid diatoms IV. *Pinnularia* vs. *Caloneis*. Diatom 24, 8–24.
Casteleyn, G., Leliaert, F., Backeljau, T., Debeer, A.E., Kotaki, Y., Rhodes, L., Lundholm, N., Sabbe, K., Vyverman, W., 2010. Limits to gene flow in a cosmopolitan marine planktonic diatom. Proc. Natl. Acad. Sci. USA 107, 12952–12957.
Chacon-Baca, E., Beraldi-Campesi, H., Cevallos-Ferriz, S.R.S., Knoll, A.H., Golubic, S., 2002. 70 Ma nonmarine diatoms from northern Mexico. Geology 30, 279–281.
Chambers, P.M., 1966. Late Cretaceous and Palaeocene Marine Diatom Floras. University College London, London, p. 498.
Chepurnov, V.A., Mann, D.G., Sabbe, K., Vyverman, W., 2004. Experimental studies an sexual reproduction in diatoms. In: International Review of Cytology – A Survey of Cell Biology, vol. 237. Elsevier Academic Press Inc., San Diego, p. 91.
Cleve, P.T., 1894. Synopsis of the naviculoid diatoms. Part I. Kongliga Svenska Vetenskapsakademiens Handlingar 26, 1–194.
Cleve, P.T., 1895. Synopsis of the naviculoid diatoms. Part 2. Kongliga Svenska Vetenskapsakademiens Handlingar 27, 1–219.
Cox, E.J., 1988. Variation within the genus *Pinnularia* Ehrenb.: further evidence for the use of live material in diatom systematics? In: Proceedings of the 9th International Diatom Symposium, Bristol, 1986, pp. 437–447.
Daugbjerg, N., Andersen, R.A., 1997. A molecular phylogeny of the heterokont algae based on analyses of chloroplast-encoded *rbc*L sequence data. J. Phycol. 33, 1031–1041.
Desikachary, T.V., Sreelatha, P.M. (Eds.), 1989. Oamaru Diatoms. J. Cramer, Stuttgart.
Donoghue, P.C.J., Benton, M.J., 2007. Rocks and clocks: calibrating the Tree of Life using fossils and molecules. Trends Ecol. Evol. 22, 424–431.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 8.
Ehrenberg, C.G., 1843. Mittheilungen übre 2 neue asistische Lager fossiler Infusorie-Erden aus dem russischen Trans-Kaukasien (Grusien) und Siberien. Be. Bekanntm. Verh. Königl. Preuss. Akad. Wiss. Berlin 1834, 43–49.
Elwood, H.J., Olsen, G.J., Sogin, M.L., 1985. The small-subunit ribosomal RNA gene sequences from the hypothrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. Mol. Biol. Evol. 2, 399–410.
Evans, K.M., Wortley, A.H., Mann, D.G., 2007. An assessment of potential diatom "barcode" genes (*cox*1, *rbc*L, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). Protist 158, 349–364.
Evans, K.M., Wortley, A.H., Simpson, G.E., Chepurnov, V.A., Mann, D.G., 2008. A molecular systematic approach to explore diversity within the *Sellaphora pupula* species complex (Bacillariophyta). J. Phycol. 44, 215–231.
Felsenstein, J., 1985. Confidence-limits on phylogenies – an approach using the bootstrap. Evolution 39, 783–791.
Gersonde, R., Harwood, D.M., 1990. Lower Cretaceous diatoms from ODP Leg 113 site 693 (Weddell Sea). Part 1: vegetative cells. In: Proceedings of the Ocean Drilling Program, Scientific Results, pp. 365–402.
Guillou, L., Chretiennot-Dinet, M.J., Medlin, L.K., Claustre, H., Loiseaux-de Goer, S., Vaulot, D., 1999. *Bolidomonas*: a new genus with two species belonging to a new algal class, the Bolidophyceae (Heterokonta). J. Phycol. 35, 368–381.
Guiry, M.D., Guiry, G.M., 2011. AlgaeBase. World-wide Electronic Publication, National University of Ireland, Galway. <http://www.algaebase.org> (accessed 17.08.11).
Gunderson, J.H., McCutchan, T.F., Sogin, M.L., 1986. Sequence of the small subunit ribosomal RNA gene expressed in the blood-stream stages of *Plasmodium berghei* – evolutionary implications. J. Protozool. 33, 525–529.
Hajós, M., 1986. Stratigraphy of Hungary's Miocene diatomaceous earth deposits. Geol. Hungarica, Ser. Palaeontol. 49, 1–339.
Hajós, M., Stradner, H., 1975. Late Cretaceous Archaeomonadaceae, Diatomaceae and Silicoflagellatae from the South Pacific Ocean. Deep Sea Drilling Project Leg 29, site 275. Init. Rep. Deep Sea Drilling Proj. 29, 913–1109.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
Harwood, D.M., Gersonde, R., 1990. Lower Cretaceous diatoms from ODP Leg 113 Site 693 (Weddell Sea). Part 2: resting spores, chrysophycean cysts, an endoskeletal dinoflagellate and notes on the origin of diatoms. Proc. Ocean Drilling Program, Sci. Results 113, 403–425.
Héribaud, J., 1902. Les diatomées fossiles d'Auvergne, Clermont-Ferrand.
Ho, S.Y.W., Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst. Biol. 58, 367–380.
Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.
Jones, H.M., Simpson, G.E., Stickle, A.J., Mann, D.G., 2005. Life history and systematics of *Petroneis* (Bacillariophyta), with special reference to British waters. Eur. J. Phycol. 40, 61–87.
Kooistra, W., Medlin, L.K., 1996. Evolution of the diatoms (Bacillariophyta). 4. Reconstruction of their age from small subunit rRNA coding regions and the fossil record. Mol. Phylogen. Evol. 6, 391–407.
Kooistra, W., De Stefano, M., Mann, D.G., Salma, N., Medlin, L.K., 2003. Phylogenetic position of *Toxarium*, a pennate-like lineage within centric diatoms (Bacillariophyceae). J. Phycol. 39, 185–197.
Krammer, K., 1992. *Pinnularia*, eine Monographie der europäischen Taxa. J. Cramer, Berlin.
Krammer, K., 2000. The genus *Pinnularia*. A.R.G. Gantner Verlag, Ruggell.
Krammer, K., Lange-Bertalot, H., 1986. Bacillariophyceae 1.Teil: Naviculaceae. Gustav Fisher Verlag, Stuttgart.

Kumar, S., Skjaeveland, A., Orr, R.J.S., Enger, P., Ruden, T., Mevik, B.H., Burki, F., Botnen, A., Shalchian-Tabrizi, K., 2009. AIR: a batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. BMC Bioinf. 10, 7.

Lartillot, N., Lepage, T., Blanquart, S., 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25, 2286–2288.

Lewis, A.R., Marchant, D.R., Ashworth, A.C., Hedenas, L., Hemming, S.R., Johnson, J.V., Leng, M.J., Machlus, M.L., Newton, A.E., Raine, J.I., Willenbring, J.K., Williams, M., Wolfe, A.P., 2008. Mid-Miocene cooling and the extinction of tundra in continental Antarctica. Proc. Natl. Acad. Sci. USA 105, 10676–10680.

Li, Y.M., Ferguson, D.K., Wang, Y.F., Li, C.S., 2010. Paleoenvironmental inferences from diatom assemblages of the middle Miocene Shanwang Formation, Shandong, China. J. Paleolimn. 43, 799–814.

Lohman, K.E., Andrews, G.W., 1968. Late Eocene Nonmarine Diatoms from the Beaver Divide Area, Fremont County, Wyoming. United States Geological Survey Professional Paper 593-E.

Lundholm, N., Moestrup, O., Kotaki, Y., Hoef-Emden, K., Scholin, C., Miller, P., 2006. Inter- and intraspecific variation of the *Pseudo-nitzschia delicatissima* complex (Bacillariophyceae) illustrated by rRNA probes, morphological data and phylogenetic analyses. J. Phycol. 42, 464–481.

Mann, D.G., 2001. A discussion of *Caloneis* and related genera. Diatom 17, 29–36.

Mann, D.G., Droop, S.J.M., 1996. Biodiversity, biogeography and conservation of diatoms. Hydrobiologia 336, 19–32.

Mann, D.G., Evans, K.M., 2007. Molecular genetics and the neglected art of diatomics. In: Brodie, J., Lewis, J.M. (Eds.), Unraveling the Algae – The Past, Present and Future of Alga Molecular Systematics. CRC Press, London, pp. 265–321.

Medlin, L.K., Kooistra, W.H.C.F., Gersonde, R., Wellbrock, U., 1996. Evolution of the diatoms (Bacillariophyta): III. Molecular evidence for the origin of the Thalassiosirales. Nova Hedwigia Beihefte 112, 221–234.

Medlin, L.K., Kooistra, W., Gersonde, R., Sims, P.A., Wellbrock, U., 1997a. Is the origin of the diatoms related to the end-Permian mass extinction? Nova Hedwigia 65, 1–11.

Medlin, L.K., Kooistra, W., Potter, D., Saunders, G.W., Andersen, R.A., 1997b. Phylogenetic relationships of the 'golden algae' (haptophytes, heterokont chromophytes) and their plastids. Plant Syst. Evol. 187, 219.

Ognjanova-Rumenova, N., Vass, D., 1998. Paleoecology of the Late Miocene maar lakes, Podrecany Basalt Formation, Southern Slovakia, on the basis of siliceous microfossils. Geol. Carpath. 49, 351–368.

Pantocsek, J., 1886. Beiträge zur Kenntniss der fossilen Bacillarien Ungarns. I. Theil: Marine Bacillarien. J. Platzko, Nagy-Tapolscány, Hungary.

Pantocsek, J., 1889. Beiträge zur Kenntniss der fossilen Bacillarien Ungarns. Teil II Brackwasser Bacillarien. Julíus Platzko, Nagy-Tapolcsány, Hungary.

Philippe, H., 2004. Phylogenomics of eukaryotes: impact of missing data on large alignments. Mol. Biol. Evol. 21, 1740–1752.

Rambaut, A., Drummond, A.J., 2007. Tracer. <http://beast.bio.ed.ac.uk/tracer/>.

Rodríguez-Trelles, F., Tarrío, R., Ayala, F.J., 2003. Molecular clocks: whence and whither? In: Donoghue, M.J., Smith, M.P. (Eds.), Telling the Evolutionary Time: Molecular Clocks and the Fossil Record. Taylor & Francis, London, pp. 5–26.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Round, F.E., Crawford, R.M., Mann, D.G., 1990. The Diatoms: Biology and Morphology of the Genera. Cambridge University Press.

Saint Martin, S., Saint Martin, J.P., 2005. The diatom assemblages as environmental evolution recording of the Paratethys area during Sarmatian times. C. R. Palevol. 4, 191–201.

Sanderson, M.J., 1990. Estimating rates of speciation and evolution – a bias due to homoplasy. Cladistics-Int. J. Willi Hennig Soc. 6, 387–391.

Sanderson, M.J., 2003. R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19, 301–302.

Saunders, G.W., 2005. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. Philos. Trans. R. Soc. B. Biol. Sci. 360, 1879–1888.

Scholin, C.A., Herzog, M., Sogin, M., Anderson, D.M., 1994. Identification of group-specific and strain-specific genetic-markers for globally distributed *Alexandrium* (Dinophyceae). 2. Sequence-analysis of a fragment of the LSU ribosomal RNA gene. J. Phycol. 30, 999–1011.

Schwarz, G., 1978. Estimating dimension of a model. Ann. Stat. 6, 461–464.

Servant-Vildary, S., Paicheler, J.C., Semelin, B., 1988. Miocene lacustrine diatoms from Turkey. In: Proceedings of the 9th International Diatom Symposium, Bristol, 1986. Biopress Ltd., Koenigstein, pp. 165–180.

Sims, P.A., Mann, D.G., Medlin, L.K., 2006. Evolution of the diatoms: insights from fossil, biological and molecular data. Phycologia 45, 361–402.

Singh, R.S., Stoermer, E.F., Kar, R., 2006. Earliest freshwater diatoms from the deccan intertrappean (Maastrichtian) sediments of India. Micropaleontology 52, 545–551.

Siver, P.A., Wolfe, A.P., 2007. *Eunotia spp.* (Bacillariophyceae) from Middle Eocene lake sediments and comments on the origin of the diatom raphe. Can. J. Bot. Rev. Can. Bot. 85, 83–90.

Siver, P.A., Wolfe, A.P., Edlund, M.B., 2010. Taxonomic descriptions and evolutionary implications of Middle Eocene pennate diatoms representing the extant genera *Oxyneis*, *Actinella* and *Nupela* (Bacillariophyceae). Plant Ecol. Evol. 143, 340–351.

Sorhannus, U., 2004. Diatom phylogenetics inferred based on direct optimization of nuclear-encoded SSU rRNA sequences. Cladistics – Int. J. Willi Hennig Soc. 20, 487–497.

Sorhannus, U., 2007. A nuclear-encoded small-subunit ribosomal RNA timescale for diatom evolution. Mar. Micropaleontol. 65, 1–12.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. Clustal-W – improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Trobajo, R., Clavero, E., Chepurnov, V.A., Sabbe, K., Mann, D.G., Ishihara, S., Cox, E.J., 2009. Morphological, genetic and mating diversity within the widespread bioindicator *Nitzschia palea* (Bacillariophyceae). Phycologia 48, 443–459.

van Hannen, E.J., Mooij, W.M., van Agterveld, M.P., Gons, H.J., Laanbroek, H.J., 1999. Detritus-dependent development of the microbial community in an experimental system: qualitative analysis by denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. 65, 2478–2484.

VanLandingham, S.L., 1991. Precision dating by means of traditional biostratigraphic methods for the middle Miocene diatomaceous interbeds within the middle Yakima (Wanapum) Basalt of south-central Washington (USA). Nova Hedwigia 53, 349–368.

Wiens, J.J., 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52, 528–538.

Wiens, J.J., Moen, D.S., 2008. Missing data and the accuracy of Bayesian phylogenetics. J. Syst. Evol. 46, 307–314.

Witt, O.N., 1886. Ueber den Polierschiefer von Archangelsk-Kurojedowo im Gouv. Simbirsk., Verhandlungen der Russisch-Kaiserlichen Mineralogischen Gesellschaft zu St Petersburg, Series 2, pp. 137–177.

Wolfe, A.P., Siver, P.A., 2009. Three extant genera of freshwater Thalassiosiroid diatoms from middle Eocene sediments in northern Canada. Am. J. Bot. 96, 487–497.

Yang, J., Wang, Y.F., Spicer, R.A., Mosbrugger, V., Li, C.S., Sun, Q.G., 2007. Climatic reconstruction at the miocene shanwang basin, China, using leaf margin analysis, clamp, coexistence approach, and overlapping distribution analysis. Am. J. Bot. 94, 599–608.

Yoon, H.S., Hackett, J.D., Pinto, G., Bhattacharya, D., 2002. The single, ancient origin of chromist plastids. Proc. Natl. Acad. Sci. USA 99, 15507–15512.

Zwart, G., Huismans, R., van Agterveld, M.P., Van de Peer, Y., De Rijk, P., Eenhoorn, H., Muyzer, G., van Hannen, E.J., Gons, H.J., Laanbroek, H.J., 1998. Divergent members of the bacterial division Verrucomicrobiales in a temperate freshwater lake. FEMS Microbiol. Ecol. 25, 159–169.