Received: 23 February 2024 | Accepted: 13 August 2024

DOI: 10.1111/jpy.13499

RESEARCH ARTICLE

The Cretaceous Diatom Database: A tool for investigating early diatom evolution

Karolina Bryłka[1] | **Matt P. Ashworth**[2] | **Andrew J. Alverson**[3] | **Daniel J. Conley**[1]

[1]Department of Geology, Lund University, Lund, Sweden

[2]Department of Molecular Biosciences, UTEX Culture Collection of Algae, The University of Texas at Austin, Austin, Texas, USA

[3]Department of Biological Sciences, University of Arkansas Fayetteville, Fayetteville, Arkansas, USA

Correspondence
Karolina Bryłka, Department of Geology, Lund University, Sölvegatan 12, 223 62 Lund, Sweden.
Email: karolina.brylka@geol.lu.se

Funding information
European Research Council, Grant/Award Number: 833454; Knut och Alice Wallenbergs Stiftelse; Division of Environmental Biology, Grant/Award Number: DEB-1651087

Editor: P.G. Kroth

Abstract

The Cretaceous period is the time of the first appearance of the diatoms in the fossil record. These fossils give us direct evidence of the age and early evolution of the diatom lineage. The fossil record, however, is incomplete and therefore often extrapolated through time-calibrated phylogenies. These two approaches offer different perspectives on the early evolution of diatoms, which is still poorly understood. We compiled the first comprehensive Cretaceous Diatom Database, a tool to investigate the taxonomy, diversity, and occurrence of the earliest known diatom lineages. To further aid the integration and use of the oldest diatom fossils in molecular clock analyses, we present a set of well-documented Cretaceous fossils that can be placed onto molecular phylogenetic trees of extant and extinct species, making them ideal candidates for the calibration of molecular clocks. The analysis of the fossil record and the Cretaceous Diatom Database revealed Cretaceous diversity is substantially greater than previously thought, yet considerable taxonomic work is still needed. The Cretaceous Diatom Database and the list of Cretaceous fossils for calibrating molecular clocks represent valuable resources for future evolutionary and taxonomic studies of modern and fossil diatoms.

KEYWORDS
calibration point, Cretaceous Diatom Database, diatoms, molecular clock

INTRODUCTION

Dated fossils offer direct evidence of the age and early evolution of a lineage. The fossil record is notoriously incomplete, however, resulting in gaps in our understanding of the timing and sequence of evolutionary changes in a lineage. Diatoms appeared in the fossil record in the Early Cretaceous, about 120 million years ago (Mya; Nikolaev, Harwood, & Samsonov, 2001). Based on the available fossil record, it appears that diatoms were present on both hemispheres by 110 Mya but only in a few isolated niches (Gersonde & Harwood, 1990). Documenting the early diatom fossil record is an uphill climb. An estimated 3% of the diatoms from a community will ever be preserved and represented in ancient sediments (Tréguer et al., 1995), and most of the earliest fossils that have been found are poorly preserved due to alteration by diagenetic processes (Geroch, 1978). Only in the Upper Cretaceous do we observe many deposits with abundant and well-preserved fossils worldwide (Hajós & Stradner, 1975; Harwood et al., 2007;

Abbreviations: cf., in Latin: confer (Eng. compared to); Ma, million years; Mya, million years ago.

Tapia & Harwood, 2002; Witkowski et al., 2011). Consequently, the fossil record leaves limited traces of the past.

Our understanding of the age and timescale of evolution can be augmented, in part, with molecular clock methods that combine fossils and DNA sequence information to estimate divergence dates across a phylogenetic tree (Lee & Ho, 2016). These methods have been used in a wide range of studies on understanding the origin and diversification of diatoms and trait evolution, (Kooistra & Medlin, 1996; Lewitus et al., 2018; Li et al., 2020; Medlin, 2015; Nakov et al., 2014, 2018a; Sorhannus, 2007; Souffreau et al., 2011), radiations in new habitats (Alverson, 2014; Pinseel et al., 2020), and genome evolution (Parks, Nakov, et al., 2018). Accurate timescales are also important for paleoecological studies, such as those on past biogeochemical cycles (Conley et al., 2017). Although powerful tools, molecular clock methods are sensitive to evolutionary rate variation (Lee & Ho, 2016), sequence saturation, where sites in the genome have experienced multiple nucleotide substitutions (Lozano-Fernandez, 2022), and the use of erroneous fossil calibration points (Angelis & Dos Reis, 2015). Methods such as relaxed clock and partitioned clock models have been developed to deal with some of these problems (Tiley et al., 2020), but several issues still remain (Budd & Mann, 2024).

The estimated age of diatoms from molecular clock studies has changed as methods have improved, datasets have gotten larger, and fossil and taxon sampling have increased. The earliest molecular clock studies of diatoms were based on a single gene, and depending on the method and calibrations used, they yielded an average age of 164–166 Ma and a maximum possible age of 266 Ma (Kooistra & Medlin, 1996), while other estimates pointed toward origin time between 400 and 330 Mya (Sorhannus, 1997). With further improvement in methods, the use of more genes and fossil calibrations, and incorporating *Bolidomonas* as an outgroup (Daugbjerg & Guillou, 2001), the date of origin was recalculated as 250–190 Mya (Medlin, 2015; Sorhannus, 2007) and later 200–190 Mya (Nakov et al., 2018a), which is still 80 Ma older than the earliest known diatom fossils (Bryłka et al., 2023).

Although molecular clocks have limitations, the fossil record is incomplete and, in some cases, misleading due to misidentification or misclassification of fossils. For example, the probability that the earliest members of a group will be discovered is vanishingly small, and diagenetic processes can erase whatever fossils do remain. Some diatom fossils, including ones used to calibrate molecular clocks, have not held up to further scrutiny. For example, the oldest reported diatom fossils from the Lower Jurassic are more likely calcareous nannofossils (Bryłka et al., 2023), whereas others are represented by pyritized specimens with equivocal taxonomic association (Geroch, 1978). As suggested previously (Bryłka et al., 2023), these cases may be connected with poor documentation of the elemental composition of the observed fossils. Moreover, the classification of pyritized diatoms is based on overall similarity rather than specific ultrastructural characters of the frustule, which can be altered or erased through diagenetic processes.

Given the consequences for understanding the evolution of diatoms and their impacts on the ancient and modern ocean, it is important to establish an accurate age and a series of divergence times for diatoms. To this end, we compiled literature on Cretaceous diatoms to create the first comprehensive Cretaceous Diatom Database. We derived a set of well-documented fossils that exhibit compelling morphological evidence, based on the presence of presumed synapomorphies, which justify the placement of the fossils in the chosen taxonomic levels and accurate dating of the sediments containing the fossils (Parham et al., 2011). In doing so, we also characterized trends of species accumulation throughout the Cretaceous. We discuss weaknesses in the documentation of fossil diatoms and suggest substantial taxonomic work is necessary on both species and genus levels. Our results highlight the importance of integrating two aspects of the same evolutionary history—fossil morphology and molecular data—to better understand early diatom evolution.

MATERIALS AND METHODS

Cretaceous Diatom Database and the analysis of the cretaceous fossil record of diatoms

For the compilation of the Cretaceous Diatom Database (Appendix S1 in the Supporting Information), we collected data from 18 scientific papers (Table 1). The reported data span in age from the Aptian (121.4 Mya) to the Maastrichtian (66 Mya). The main objectives of the database were to: (1) compile a list of all species present in the Cretaceous, (2) track the distribution and species richness of genera throughout the Cretaceous, (3) identify general trends in the emergence of morphological groups, and (4) identify the first occurrence of extant genera and genera previously used as calibration points in molecular clocks. The Cretaceous Diatom Database consists of the following: (1) species list with corresponding papers, (2) species list with corresponding age of occurrence, and (3) genera list with corresponding age of occurrence and number of species.

When compiling the database, we excluded pyritized diatoms because the identity of diagenetically transformed frustules cannot be established in many cases.

TABLE 1 Compilation table of articles used for the Cretaceous Diatom Database and fossil record analysis.

Locality	Age	Age Ma	No. of species	No. of genera	References
NE Australia	Lower Aptian	121 to 120	38	13	Nikolaev, Harwood, and Samsonov (2001)
Weddell Sea, ODP Site 693	Lower-Middle Albian	113 to 107	17	7	Harwood and Gersonde (1990)
		113 to 107	28	13	Gersonde and Harwood (1990)
Hannover, Germany	Albian	113 to 100.5	11	9	Forti and Schultz (1932)
Sinai Peninsula, Egypt	Albian-Maastrichtian	113 to 66	41	18	Zalat (2013)
Arctic Archipelago and Canadian Continental Margin	Upper Cenomanian-Campanian	~95 to 72.1	154	47	Tapia and Harwood (2002)
Omagari Formation, Japan	Uppermost Santonian -Lowermost Campanian	~84 to 82	42	22	Shimada et al. (2022)
Devon Island, Arctic Ocean	Santonian-Lower Campanian	86.3 to ~80	90	38	Witkowski et al. (2011)
Eastern Arctic Peninsula, Seymour Island	Campanian-Maastrichtian	83.6 to 66	145	39	Harwood (1988)
Worldwide (DSDP Site 216, DSDP Site 275, ODP Site 758, Alpha Ridge, Moreno Formation, Tonga Trench)	Campanian-Maastrichtian	83.6 to 66	505	75	Chambers (1997)
ODP Site 758, Indian Ocean	Upper Campanian	~75 to 72.1	18	10	Fourtanier (1991)
Alpha Ridge, Arctic Ocean	Upper Campanian	~75 to 72.1	33	14	Barron (1985)
Alpha Ridge, Arctic Ocean	Upper Campanian	~75 to 72.1	59	24	Davies and Kemp (2016)
DSDP Site 275, Pacific Ocean	Upper Campanian-Maastrichtian	~75 to 66	142	41	Hajós and Stradner (1975)
Alpha Ridge, Arctic Ocean	Middle-Upper Maastrichtian	~70 to 66	25	13	Dell'Agnese and Clark (1994)
Moreno Formation, California USA	Middle-Upper Maastrichtian	~70 to 66	37	23	Hanna (1927)
Moreno Formation, California USA	Middle-Upper Maastrichtian	~70 to 66	59	40	Nikolaev, Kociolek, et al. (2001)
Moreno Formation, California USA	Middle-Upper Maastrichtian	~70 to 66	13	9	Davies and Kemp (2016)

The Cretaceous Diatom Database described here is intended to be a comprehensive resource that expands upon previous summaries of the Cretaceous diatom record (Ballance et al., 1989; Gresham, 1985; Harwood et al., 2007; Jousé, 1948; Long et al., 1946; Moshkovitz et al., 1983; Oreshkina et al., 2013; Sims, 1998; Sims et al., 2006; Strel'nikova, 1975). The exclusion of the above does not impact the content of the database since taxa described in these references and references included in the database overlap.

From the Cretaceous Diatom Database, we derived a set of well-documented fossils for future molecular clocks. To this end, we focused on validating taxa in extant genera or ones previously used to calibrate molecular clock studies. We carefully examined all available light and/or scanning electron microscope images provided in the respective publications to determine whether the fossils matched the gross morphology or, ideally, possessed synapomorphies that supported their relationship to extant species in the genus. We classified genera into morphological categories to plot species accumulation throughout the Cretaceous. We followed the higher level classification of diatoms that divided them into the following: radial centric (Coscinodiscophyceae sensu Medlin & Kaczmarska, 2004), bi(multi)polar centric (Mediophyceae sensu Medlin & Kaczmarska, 2004), and pennate (Bacillariophyceae sensu Medlin & Kaczmarska, 2004). It should be noted, however, that abundant molecular evidence suggests that Coscinodiscophyceae is not monophyletic (Nakov et al., 2018a; Parks, Wickett, & Alverson, 2018).

RESULTS

Cretaceous Diatom Database

The Cretaceous Diatom Database includes 721 species and 128 genera (Figure 1). The number of genera and species in each reference used to compile the Cretaceous Diatom Database can be found in Table 1 and in Appendix S1. Based on the Cretaceous Diatom Database, we determined the first reliable appearance of extant genera and genera previously used as calibration points for future molecular clocks, and

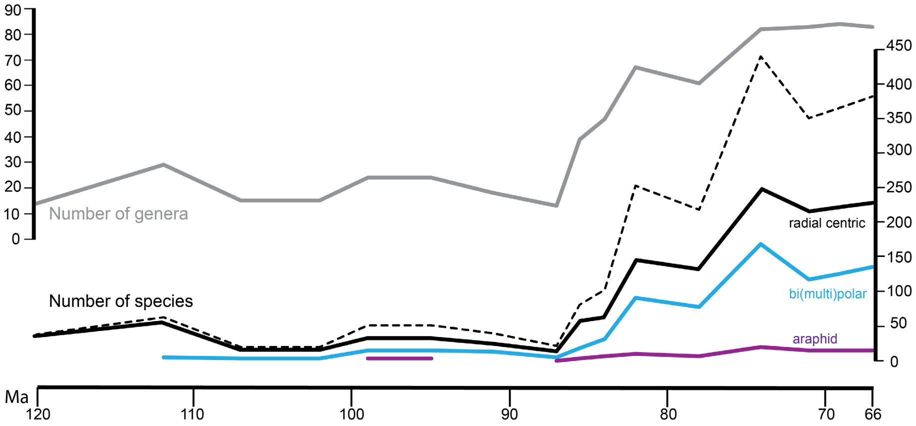

FIGURE 1 Number of genera and species during the Cretaceous period (120–66 Mya). Radial centric includes resting spore species.

characterized general trends in diversity, species richness, and stratigraphic ranges of different taxa.

Over the Cretaceous period, there was a substantial increase in a number of species and genera (Figure 1). The Early Cretaceous (Aptian–Albian period; 121.4–100.5 Mya) is represented by 86 species in 34 genera of radial centric and bi(multi)polar morphologies. Continuing to the lower part of the Late Cretaceous (Cenomanian–Santonian period; 100.5–83.6 Mya), fewer diatoms were described: 51 species and 23 genera predominantly in radial centric and bi(multi)polar morphologies, with the first record of araphid diatoms. The upper part of the Late Cretaceous (Campanian–Maastrichtian, 83.6–66 Mya) is a time of major diversification with 632 described species and 107 genera, spanning predominantly radial and polar centric taxa.

The highest species diversity was identified in the following genera: *Actinoptychus*, *Aulacodiscus*, *Biddulphia*, *Coscinodiscus*, *Hemiaulus*, *Stephanopyxis*, *Trinacria*, and *Triceratium*. It should, however, be noted that classification of taxa in these genera is not always straightforward (see Discussion). A total of 45 of the 128 Cretaceous diatom genera are monotypic (Appendix S1, Figure 2). As monotypic genera constitute fully one third of the total number of genera, we surveyed documented photographs and observed that many monotypic genera were described from LM observation only and would benefit from SEM evaluation as detailed characters of the valves are not visible. These genera include but are not necessarily limited to: *Corinna* Heiberg, *Kittonia* Grove and Strutt, *Mammidion* Long, Fuge and Smith, *Meretrosulus* Hanna, *Omphalotheca* Ehrenberg, *Peponia* Greville, *Radiodiscus* Forti and Schulz, *Rhaphidophora* Long, Fuge and Smith, *Tubularia* Brun, and *Upothema* Long, Fuge and Smith. Other genera such as *Ditylum* and *Ktenodiscus* (Forti & Schultz, 1932) were only documented with simple drawings, which is challenging for assessing their identity. The Cretaceous Diatom Database (Figure 2 and Appendix S1) can be used as a framework to understand where detailed taxonomic work is needed.

Calibration points for molecular clocks

These observations are based on analysis of the Cretaceous Diatom Database and photographs from the scientific literature, not on site-specific type slides. The purpose was to derive the maximum amount of fossil information from sources that are readily available. In this section, we assess the earliest occurrence of photographically well-documented taxa from genera that exhibit either all (crown calibrations) or some (stem calibrations) genus-level synapomorphies (shared derived characters; Figure 3). Genus-level synapomorphies were derived primarily from genus descriptions in Round et al. (1990) and supplemented by literature focused on specific genera where possible. Taxa that may serve as calibration points are also summarized in Table 2.

Genera for crown calibrations

Genus Actinoptychus Ehrenberg

Genus description and synapomorphies: radial centric valve with radially alternating elevated and depressed sectors that are almost flat and even in size; rows of loculate areolae radiate from the plain or granulated central area; prominent external tubes of rimoportulae are often located on the distal part of the elevated sector; the external opening of rimoportula is round, while internal opening is usually a linear radial slit (Ross & Sims, 1997; Round et al., 1990)

Minimum clade age: Upper Campanian (~75–72.1 Mya); plate 29, figs. 1–4 in Tapia & Harwood, 2002

Sediment dating method: biostratigraphy based on mollusks, foraminifera, and palynomorphs

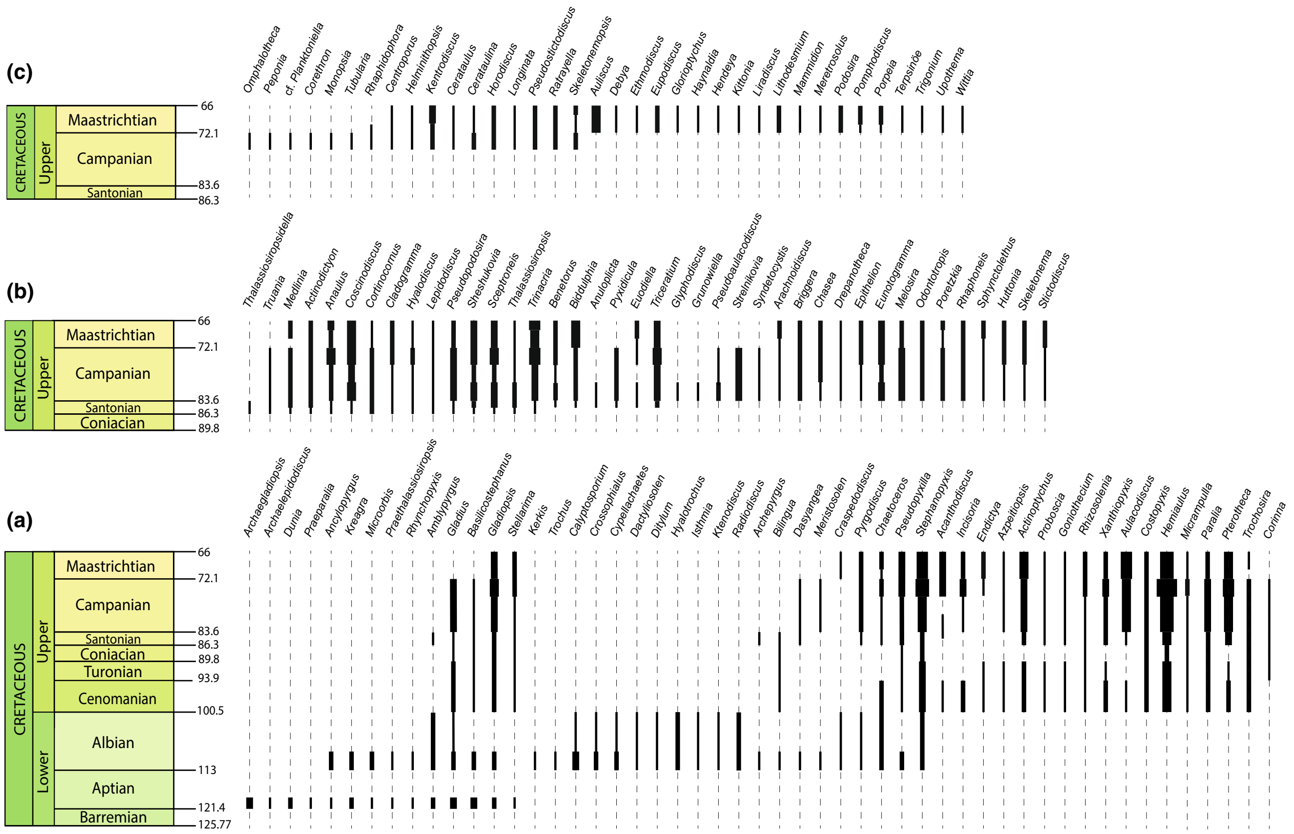

FIGURE 2 Stratigraphic ranges and species richness of Cretaceous diatom genera in stages of the Cretaceous representing taxa with first occurrence in (a) Aptian–Cenomanian, (b) Santonian–lower Campanian, and (c) upper Campanian–Maastrichtian. Thickness of the line illustrates relative variation in the number of species in the genus (for exact numbers, see Appendix S1).

Notes: There are older records of *Actinoptychus* (e.g., Cenomanian 100.5–93.9 Mya; plate 2, fig. 4–5 in Zalat, 2013); however, the rimoportulae are not visible, but this could be used as a stem calibration; some other genera such as *Trossulus* Ross and Sims and *Polymyxus* Bailey have radially alternating depressed and elevated sectors and have been placed in the same family as *Actinoptychus*, that is, Heliopeltaceae; however, these have convex elevated sectors and distinct internal rimoportulae openings (Ross & Sims, 2002); lastly, *Actinodiction* Pantocsek likewise differ in internal rimoportulae openings and its sectors are uneven in size (Ross & Sims, 1997).

Genus Arachnoidiscus Deane ex Pritchard

Genus description and synapomorphies: heterovalvar frustule where one valve has a plain central area, whereas the other valve has a plain center ringed by elongated radial slits; areolae occluded by vola are sectored by massive system of radial costae (Round et al., 1990)

Minimum clade age: Campanian (83.6–72.1 Mya); fig. 48 in Sims et al. (2006)

Sediment dating method: ammonoid biostratigraphy

Notes: Sims et al. (2006) is an overview study on the Cretaceous diatoms not included in the Cretaceous Diatom Database (see Materials and Methods).

Genus: Aulacodiscus Ehrenberg

Genus description and synapomorphies: radial centric frustule with three or more conspicuous marginal rimoportulae open externally via simple long tubes or winglike tubes, internally opened into curved slits; rows of loculate areolae radiate from the hyaline center (Round et al., 1990; Witkowski et al., 2017)

Minimum clade age: Lower Campanian 83.6–80 Mya; fig. 7A–B and fig. 11A in Witkowski et al. (2011)

Sediment dating method: diatom biostratigraphy

Genus Auliscus Ehrenberg

Genus description and synapomorphies: valves circular or slightly elliptical; two or three marginal ocelli positioned on elevations; valve face ridged with distinctive parallel external costae

Minimum clade age: Middle-Upper Maastrichtian (~70–66 Mya); plate 1, fig. 8 in Hanna (1927)

Sediment dating method: foraminifera biostratigraphy

Genus Chaetoceros Ehrenberg

Genus description and synapomorphies: elliptical valves with flat or domed valve face, bearing seta from each pole; central annulus and rimoportula may be present; radial rows of fine areolae radiate from the valve center (Round et al., 1990); *Chaetoceros*

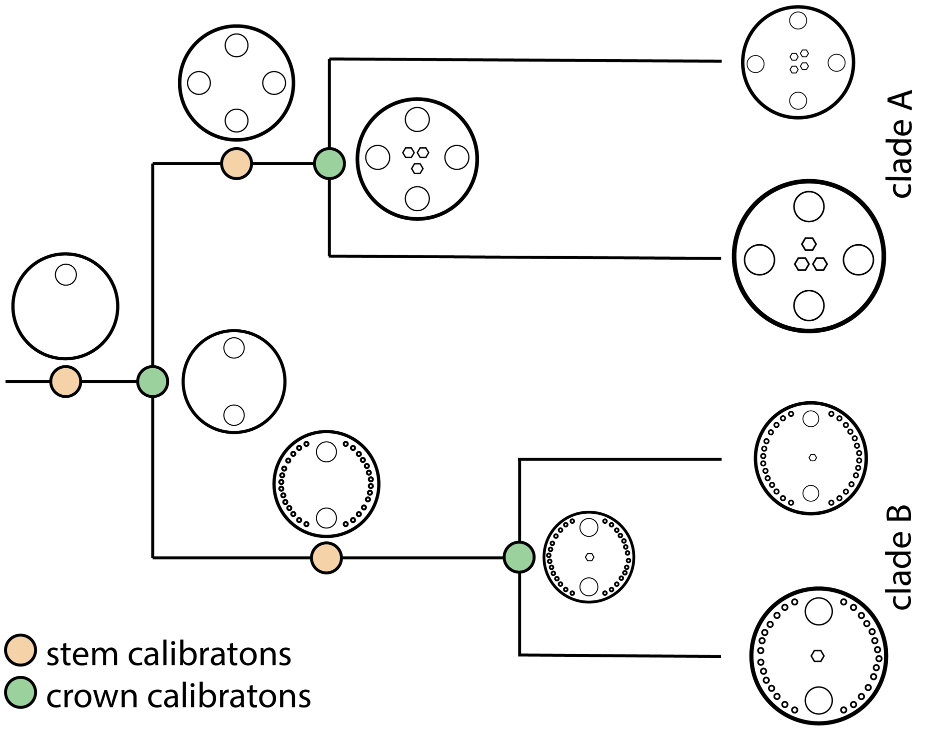

FIGURE 3 Conceptual diagram representing the placement of the crown (darker) and stem (light) calibrations. Crown calibrations exhibit all the features of the terminal taxa. In this example, all species in clade A, share four marginal ocelli and multiple central processes. For clade B, all species share two marginal ocelli, a single central process, and a ring of marginal processes. Crown calibrations in these cases would be taxa exhibiting all characters. The stem calibration for the clade A could be a taxon with four ocelli without central processes. For clade B, the stem could be a taxon with two ocelli and a marginal ring of processes without a central process. Crown calibration for both clades could be a taxon with only two ocelli and the stem could be a taxon with one ocellus. Note that this scheme is conceptual and does not represent actual diatom genera.

can from resting spores with much morphological variation (Ishii et al., 2011); valves of resting spores are typically vaulted or capitate and either smooth, or ornamented by short or elaborate spines and knobs (Ishii et al., 2011)

Minimum clade age: Upper Campanian (~75–72.1 Mya); figs. 7A–B in Davies and Kemp (2016)

Sediment dating method: biostratigraphy

Notes: This specimen is represented by a small form with vaulted valves lacking ornamentation and was classified as resting spore; this specimen has a strong resemblance to extant *Chaetoceros* spores (Sunesen et al., 2008).

Genus Coscinodiscus Ehrenberg

Genus description and synapomorphies: saucer to Petri dish-shaped valves bearing marginal ring of rimoportulae, including at least one macro rimoportula; loculate valves with small or inconspicuous central area

Minimum clade age: Middle-Upper Maastrichtian (~70–66 Mya); plate 16, figs. 1–6 in Nikolaev, Kociolek, et al. (2001)

Sediment dating method: foraminifera biostratigraphy

Notes: There are older records of *Coscinodiscus* dating back to Santonian (86.3–83.6 Mya) (Witkowski et al., 2011) and Campanian (83.6–72.1 Mya; Barron, 1985; Chambers, 1997; Hajós & Stradner, 1975; Harwood, 1988; Tapia & Harwood, 2002); *Coscinodiscus* is among most confused genera in the fossil record with many taxa transferred to, among others: *Actinocyclus* Ehrenberg, *Azpeitia* Peragallo, *Stellarima* Hasle and Sims, *Stephanopyxis* Ehrenberg, *Thalassiosira* Cleve (Guiry, 2013); therefore, a proper identification can only be achieved through SEM examination which Santonian and Campanian documentations lack, except for Hajós and Stradner (1975); however, ring or marginal rimoportulae is not documented.

Genus Hyalodiscus Ehrenberg

Genus description and synapomorphies: hemispherical loculate valves with mantle edge turned outward; central area is hyaline; ring of rimoportulae present around the valve mantle with simple external openings

Minimum clade age: Lower Campanian (~83–80 Mya); figs. 7C–D in Witkowski et al. (2011)

Sediment dating method: diatom biostratigraphy

Notes: Rimoportulae not visible on the provided photograph; however, because external openings are simple, they are often indistinguishable from areolae; some species of *Hyalodiscus* were transferred to *Podosira*—a genus of alike valves characteristics, *Podosira* however, has no distinct valve margin and

TABLE 2 Compilation of Cretaceous diatom genera that may serve as calibration points in future molecular clocks.

Genus	Occurrence Ma (lower bound)	Occurrence Ma (upper bound)	Documentation (LM/SEM)	C/S calibration	Citation
Actinoptychus	100	94	LM	S	Zalat (2013)
Actinoptychus	75	72	SEM	C	Tapia and Harwood (2002)
Amblypyrgus (*Stephanopyxis*)	113	107	SEM	S	Gersonde and Harwood (1990)
Arachnoidiscus	83	72	LM	C	Sims et al. (2006)
Aulacodiscus	83	80	SEM, LM	C	Witkowski et al. (2011)
Archepyrgus (*Aulacoseira*)	113	107	SEM, LM	S	Gersonde and Harwood (1990)
Auliscus	70	66	LM	C	Hanna (1927)
Biddulphia	83	72	SEM	S	Sims et al. (2006)
Chaetoceros	75	72	SEM	C	Davies and Kemp (2016)
Corethron	75	72	LM	S	Chambers (1997)
Coscinodiscus	70	66	SEM	C	Nikolaev, Kociolek, et al. (2001)
Eunotogramma	83	80	LM	S	Chambers (1997)
Hemiaulus	100	94	LM	S	Zalat (2013)
Hyalodiscus	83	80	LM	C	Witkowski et al. (2011)
Paralia	95	84	LM	C	Tapia and Harwood (2002)
Proboscia	83	72	SEM	S	Jordan and Ito (2002)
Praeparalia (*Ellerbeckia*)	121	120	SEM	S	Nikolaev, Harwood, and Samsonov (2001)
Rhizosolenia	75	72	SEM, LM	S	Davies and Kemp (2016)
Sceptroneis	84	83	SEM	C	Shimada et al. (2022)
Stellarima	84	82	SEM	C	Shimada et al. (2022)
Stephanopyxis	75	72	SEM	C	Davies and Kemp (2016)
First diatoms	121	120	LM, SEM	C	Nikolaev, Harwood, and Samsonov (2001)
Next diatoms	113	107	LM, SEM	C	Harwood and Gersonde (1990)

Note: We also added tie points for the first and next diatoms.

Abbreviations: C, crown; LM, light microscope; S, stem; SEM, Scanning Electron Microscope.

hyaline central area is not a genus synapomorphy (Round et al., 1990).

Genus Paralia Heiberg

Genus description and synapomorphies: circular loculate valves with radial costae and depressions on the valve face which appear to function as interlocking linking structures; linking marginal spines can also be present; valve face mostly without areolae

Minimum clade age: Upper Cenomanian to Upper Santonian (~95–84 Mya); plate 9, fig. 13 in Tapia and Harwood (2002)

Sediment dating method: biostratigraphy based on mollusks, foraminifera, and palynomorphs

Notes: *Paralia* shares characteristics with *Ellerbeckia* and *Trochosira* such as heterovalvy for both, radial markings on the valve face for the former and elaborate marginal spines that join frustules into colonies for the latter (Round et al., 1990; Sims, 1988); while other genera may share many of these characters with *Paralia*, the combination of all of these characters is unique to *Paralia*, making identification of this genus relatively straightforward.

Genus Sceptroneis Ehrenberg

Genus description and synapomorphies: elongated slightly curved heteropolar valves with indistinct central sternum; uniseriate striae with large areolae are occluded by rotae; apical pore fields present on both poles with single rimoportula

Minimum clade age: Uppermost Santonian–Lowermost Campanian (~84–82 Mya); figs. 4A–C in Shimada et al. (2022)

Sediment dating method: biostratigraphy based on ammonoids and inoceramids.

Notes: This specimen of *Sceptroneis* has only one visible rimoportulae, and rotae are not preserved, which is expected in fossil specimens.

Genus Stellarima Hasle and Sims

Genus description and synapomorphies: radial centric, weakly to strongly convex, discoid valves; areole externally with or without cribra; central rimoportulae (observed 1–15) are arranged in a ring, bearing slit-like openings on a central hyaline area (Hasle & Sims, 1986; Round et al., 1990)

Minimum clade age: Uppermost Santonian-Lowermost Campanian (~84–82 Mya); fig. 4Q in Shimada et al. (2022)

Sediment dating method: biostratigraphy based on ammonoids and inoceramids

Genus Stephanopyxis Ehrenberg

Genus description and synapomorphies: domed valves with large hexagonal areolae opened by foramina with underlying continuous sheet of silica; ring of tubular linking processes present on the valve face; ring of marginal rimoportulae (Round et al., 1990)

Minimum clade age: Upper Campanian (~75–72.1 Mya); figs. 5A and 13A in Davies and Kemp (2016)

Sediment dating method: diatom biostratigraphy

Notes: This specimen has domed valves jointed together by tubular linking spines, with hexagonal areolae but lacking underlying sheet of silica but this is expected in fossil specimens which underwent dissolution processes.

Genera for stem calibrations

Calibration for Actinoptychus Ehrenberg stem

Genus description and synapomorphies: radial centric valve with radially alternating elevated and depressed flat sectors; sectors even in size; rows of loculate areolae radiating from the plain or granulated central area; prominent external tubes of rimoportulae often located on the distal part of the elevated sector; the external opening of rimoportula is round, while internal opening is a linear radial slit (Ross & Sims, 1997; Round et al., 1990)

Minimum clade age: Cenomanian (100.5–93.9 Mya); plate 2 figs. 4–5 in Zalat (2013)

Sediment dating method: biostratigraphy based on mollusks, foraminifera, and palynomorphs

Notes: This specimen's rimoportulae are not visible.

Calibration for Aulacoseira Thwaites stem

Genus description and synapomorphies: circular valves with deep mantle that is perpendicular to the valve face; mantle with vertical or curved areolae; spines present at the valve face/mantle junction joining adjacent frustules; inside at the base of the valve mantle thickening termed "ringliest" with single rimoportulae at base (Round et al., 1990)

Minimum clade age: Lower Albian (113–107 Mya); plate 13, figs. 1–8 in Gersonde and Harwood (1990)

Sediment dating method: biostratigraphy

Notes: This calibration refers to *Archepyrgus* Gersonde and Harwood: This genus has the above characters except rimoportulae.

Calibration for Biddulphia Gray stem

Genus description and synapomorphies: bipolar valves with short apices, where each apex bears a pseudocellus; valves undulate in shape in valve and girdle view; large areole occluded by cribra; multiple central rimoportulae present, opened externally via tube and by a slit internally; valve interior costate; costae either transverse, short, or branching (Sims et al., 2023)

Minimum clade age: Campanian (83.6–72.1 Mya); figs. 42–44 in Sims et al. (2006)

Sediment dating method: ammonoid biostratigraphy

Notes: This specimen has a bipolar valve with relatively tall apices burdened by pseudocelli; valve undulate in valve view; hyaline depressions, which give an impression of costae; however, internal valve view is not shown; these can therefore be thickened hyaline folds (sulci); central rimoportulae present; the above characters suggest this specimen belongs to genus *Biddulphiella* Sims and Ashworth (Sims et al., 2023), which is a sister genus to *Biddulphia*.

Calibration for Corethron Castracane stem

Genus description and synapomorphies: heteropolar valves; one valve with marginal sockets into which are inserted toothed t-shaped spines; second valve with marginal sockets into which are inserted spines of two kinds: narrower spined and hooked spines (Crawford & Hinz, 1995; Fryxell & Hasle, 1971)

Minimum clade age: Upper Campanian (~75–72.1 Mya); plate 2, fig. 1 in Hajós and Stradner (1975)

Sediment dating method: siliceous microfossil biostratigraphy

Notes: This specimen is characterized by a dome-like valve with multiple processes projecting from sockets in the valve brim and is classified as *Skeletonema subantarctica* Hajós; Abe et al. (2021) investigated the material of Hajós and Stradner (1975) under the SEM and identified features in this specimen which differ from modern *Corethron* in the lack of hooked spines and a different arrangement of spines and sockets; consequently, Abe et al. (2021) separated Cretaceous taxa into genus *Praecorethron* Abe, Lam, Ashworth, Harwood and Jordan; here, we further consider this an extinct sister lineage to modern *Corethron*.

Calibration for Eunotogramma Weisse stem

Genus description and synapomorphies: valve elongated, often arcuate; plain cross bars across the valve face present—internally these are transverse costae; rimoportulae positioned near the central ventral margin; areolae arranged in straight lines or radiating from annulus (Round et al., 1990; Witkowski, Ashworth, et al., 2020; Witkowski, Harwood, et al., 2020)

Minimum clade age: Upper Campanian (~75–72.1 Mya); plate 5, fig. 1–2 in Fourtanier (1991)

Sediment dating method: calcareous nannofossil biostratigraphy

Notes: This specimen (*Eunotogramma polymorphus*) has an elongated, slightly arcuate valve with multiple cross bars across the valve face; rimoportulae and internal costae not documented; sectors separated by the bars are heavily inflated; large scattered areolae; Cretaceous *Eunotogramma* shares very few characteristics with extant *Eunotogramma* taxa and likely represents extinct sister lineage; some other Cretaceous *Eunotogramma* species (*Eunotogramma* sp. plate 16, fig. 4 in Long et al., 1946; *Eunotogramma* sp. plate 5, fig. 3 in Fourtanier, 1991) have multipolar morphologies (3–6 poles); the above suggest Cretaceous taxa described in *Eunotogramma* genus should likely be reclassified to separate genus.

Calibration for Hemiaulus Heiberg stem

Genus description and synapomorphies: elliptical valves with curved valve face merging imperceptibly with a deep mantle; valves with simple or complex areolae; long apical processes terminating by spines that link adjacent frustules into colonies; central rimoportulae may or may not be present; lightly silicified, delicate forms (Round et al., 1990)

Minimum clade age: Santonian (86.3–83.6 Mya); for example, figs. 8G, 8P, 8Q and fig. 9B in Witkowski et al. (2011)

Sediment dating method: diatom biostratigraphy

Notes: This specimen, like the majority of photographically documented *Hemiaulus* species in the Cretaceous, has bipolar valves with apical processes terminated by spines joining adjacent frustules; valve face undulated, crossed by thickened folds (sulci); central rimoportulae present; Cretaceous *Hemiaulus* is comprised of species that are heavily silicified and often exhibit robust ornamentation, in contrast to modern *Hemiaulus* taxa; accordingly, we suggest Cretaceous *Hemiaulus* to be extinct sister lineage to extant *Hemiaulus* or perhaps even to the family *Hemiauluaceae*.

Calibration for Ellerbeckia Crawford stem

Genus description and synapomorphies: circular heterovalvar frustules with radial markings on the valve face (ridges or grooves); valves can be convex or concave; ridges and grooves vary in length from ones that are short and about one fifth of the valve radius or long reaching the small annulus; valve face without pores; heavy and thick mantle with unique type of tubular processes; ridges and groves interlock and connect adjacent frustules into colonies (Round et al., 1990)

Minimum clade age: Lower Aptian (120 Mya); plate 15, fig. 9 in Nikolaev, Harwood, and Samsonov (2001)

Sediment dating method: strontium isotopes (Williamson et al., 2012)

Notes: This calibration refers to extinct genus *Praeparalia* Nikolaev and Harwood; documented specimen exhibits radial marking on the valve face that are one fourth and three fourth of a radius length and reach the valve mantle; center of the valve slightly convex with no ornamentation; although this specimen possesses only a few of the *Ellerbeckia* characters, we suggest it is a representative of an extinct sister lineage to *Ellerbeckia*.

Calibration for Proboscia Sundström stem

Genus description and synapomorphies: conoid valve, tapering into a slightly curved proboscis—a tube formed of partly overlapping silica flaps; groove located at the base of the proboscis where sister cell is held; rimoportula located at the tip of the proboscis with slit like opening (Jordan et al., 2019; Round et al., 1990)

Minimum clade age: Campanian (83.6–72.1 Mya); figs. 1–49 in Jordan and Ito (2002)

Sediment dating method: diatom biostratigraphy.

Notes: Long and curved valves with short curved proboscis, groove present at the base of the proboscis; in the considered specimen, proboscis is an extension of the valve face and is not made of overlapping silica flaps; although Jordan and Ito (2002) is not included in the Cretaceous database, taxa they describe are from site 275 described by Hajós and Stradner (1975).

Calibration for Rhizosolenia Brightwell stem

Genus description and synapomorphies: conical valves drawn out into a spine which is an external tube of the central rimoportula; loculate areolae externally closed by delicate plates with linear or elongate pores (Jordan et al., 2019, Round et al., 1990)

Minimum clade age: Upper Campanian (~75–72.1 Mya); figs. 14A–E in Davies and Kemp (2016)

Sediment dating method: biostratigraphy

Notes: Specimen identified as a resting spore; valves cone like with central spine, which is likely external rimoportulae opening; valves ornamented by linear ridges; the overall morphology of this specimen suggests its relationship to modern *Rhizosolenia*.

Calibration for Stephanopyxis Ehrenberg stem

Genus description and synapomorphies: domed valves; hexagonal areolae opened by large foramina with underlying continuous sheet of silica; ring of tubular linking processes present on the valve face; ring of marginal rimoportulae (Round et al., 1990)

Minimum clade age: Lower Albian (113–107 Mya); plate 12, figs. 1–7 in Gersonde and Harwood (1990)

Sediment dating method: biostratigraphy

Notes: This calibration refers to extinct genus *Amblypyrgus* Gersonde and Harwood; specimen has domed valves with hexagonal areolae with underlying continuous sheet of silica; apical processes consist of spines (not tubes like in *Stephanopyxis*) often connected together at their base; ring of small circular pores near the margin—possibly rimoportulae opening; this genus is likely an extinct sister genus to extant *Stephanopyxis*.

An additional 14 extant genera are included in the Cretaceous Diatom Database: *Anaulus* Ehrenberg, *Cerataulus* Ehrenberg, *Cerataulina* Peragallo, *Dactyliosolen* Castracane, *Endictya* Ehrenberg, *Ethmodiscus* Castracane, *Isthmia* Aghard, *Lithodesmium* Ehrenberg, *Melosira* Aghard, *Podosira* Ehrenberg, *Rhaphoneis* Ehrenberg, *Skeletonema* Greville, *Triceratium* Ehrenberg, and *Terpsinöe* Ehrenberg. However, these taxa either lack photographic documentation or lack genus-defining features (see Appendix S2 in the Supporting Information for more discussion), casting doubt on their classification.

DISCUSSION

The Cretaceous Diatom Database is a comprehensive resource for the study of the earliest diatom fossils, some of which can be used to calibrate molecular clocks. Our compilation has highlighted ambiguities in the taxonomy of Cretaceous diatoms relative to modern taxa, which highlight specific genera in need of examination and revision in future work. Importantly, we are not suggesting that the taxonomy applied in the scientific literature is incorrect but, on occasion, poorly documented, such as when the genus-defining feature is out of focus or incompletely documented by electron microscopy. Our intention is to identify directions for improvement so that diatom taxonomy and classification can better serve both morphological and molecular studies of diatoms.

Calibration points for molecular clocks

Careful scrutiny of previously used fossil calibrations for molecular clock studies rejected several fossils for various reasons and identified many more that can be incorporated into future studies. Our assessment was limited to published photographs but, in many cases, would benefit greatly from reexamination and updated photographs of original material, with an eye toward capturing key taxonomically and phylogenetically informative characters. Most of the Cretaceous diatom literature is based on light microscope images only, so some of the photographs either cannot or do not resolve key characters. A systematic reevaluation of original materials would likely lead to taxonomic revisions and provide many more reliable fossil calibrations.

The documentation of fossil species relies necessarily on morphological information alone. Although molecular character matrices exist for diatoms, no large-scale, diatom-wide morphological character matrix has been published. Short of this, we are left to classify fossil diatoms based on a phylogenetic interpretation of their characters, with reference to the distribution of those characters on molecular phylogenies. This approach can be misleading, however, by homoplasy, which has been documented for diatoms (Kociolek et al., 2019). This is also illustrated by some of the genera discussed here. For example, *Hemiaulus* and *Biddulphia* appear to share some morphological characteristics (central rimoportulae, bipolar valves, and apical processes) and in earlier studies were considered to belong to one subfamily (Schütt, 1896). A later molecular phylogenetic analysis placed *Biddulphia* and *Hemiaulus* in separate lineages, however (Ashworth et al., 2013), again highlighting ambiguities in diatom classification that should be studied further. Our analysis of the fossil record revealed that Cretaceous taxa in, among others, *Hemiaulus*, are morphologically dissimilar from extant representatives. We suggest that Cretaceous *Hemiaulus* is likely comprised entirely of extinct species and are the sister clade to modern *Hemiaulus*. Technically however, the type species for *Hemiaulus* Heiberg is a fossil specimen (*H. proteus*), so modern *Hemiaulus* would require reclassification. Likewise, Cretaceous species classified to *Biddulphia* lack some extant genus characters and were suggested to belong to a new genus, *Biddulphiella* (Sims et al., 2023). The Cretaceous diatom record likely includes many more examples like this.

Besides crown calibrations, we also listed stem calibrations that can be incorporated into molecular phylogenies. Some of these could breakup long branches, especially toward the root of the tree like those of *Corethron* and *Ellerbeckia*. These long branches were also called "ghost lineages" because their age on the phylogeny predates the first occurrence in the fossil record (Medlin, 2015, p. 220). Ghost lineages were suggested to be the result of misdiagnosed or undiscovered fossils (Sorhannus, 2007). We suggest that the diversity exists but is unacknowledged, for example, *Micrampulla* and *Praecorehtron* are extinct genera that are likely an extinct sister lineage to modern *Corethtrales* (Abe et al., 2022), and *Praeparalia*, likewise, may be ancestral or sister to modern day *Ellerbeckia*. Molecular phylogenies should be a guide as to what we are missing in the fossil record and what characters we should be seeking. Phylogenetic comparative methods can further aid these searches by identifying lineages with exceptionally high rates of extinction (Nakov et al., 2018a, 2018b).

Our focus predominantly on the taxa published under extant genus names might have biased our observations, since our search should be focused on synapomorphic characters rather than taxonomic affiliations. Since the Cretaceous Database consists of >120 genera and 700 species, it would be rather challenging to survey them all. Identifying extinct genera that may belong in the same lineage as extant genera is, therefore, challenging and, albeit labor intensive, would benefit greatly from examination of site-specific type slides. Some of those links, however, may still be tentative, especially for the extinct taxa: if a taxon fits into several lineages by morphological characters, then its potential position on the phylogeny will be ambiguous.

Moving forward, if our goal is to reconstruct the early evolution of diatoms in the least biased manner possible, then we should use both the fossil record and molecular phylogenies. For this to be efficient, when describing taxa, especially extinct fossils, our priority should be to identify synapomorphies, while molecular phylogenies should serve as a guide where such fossils could be placed on the tree. Reexamining fossil morphology of Cretaceous taxa that are classified to modern genera but excluded as reliable (e.g., *Anaulus* and *Skeletonema*) as well as those included as stem calibrations but likely representing sister lineages (e.g., *Eunotogramma* and *Hemiaulus*) are good examples of where to start to reexamine fossil morphology and analyzing it in a cladistic manner. This will ideally separate the fossil and modern taxonomy, which will further improve molecular inferences.

Cretaceous diatoms and taxonomic notes

In the Cretaceous, we observed (1) an increase in diatom diversity at all taxonomic levels (species, genus, and lineage) toward the Upper Cretaceous, (2) extant genera are the most species-rich, and (3) one third of the genera are monotypic, being represented by a single species. The disparity between the Lower and Upper Cretaceous diversity has been previously discussed. Sims et al. (2006) suggested that bi- and multipolar diatoms evolve faster than other diatoms, which might help account for the rapid morphological evolution that took place between the Lower and Upper Cretaceous. It is important to consider that the lack of intermediate pre-Aptian and Cenomanian–Coniacian deposits creates observational bias, so evolution may prove to be more gradual when that historical gap is eventually filled. Harwood et al. (2007) suggested that following the Cenomanian–Turonian extinction (likely associated with Ocean Acidification Event-OAE2), the Lower Cretaceous flora was erased as bi(multi)polar diatoms diversified, migrated into specialized niches, and dominated the marine plankton. This could explain the differences in Lower and Upper Cretaceous diatom assemblages; however, as the authors highlight, there is no continuous Albian–Coniacian section to test this hypothesis. Moreover, there is a clear preservation disparity between Lower and Upper Cretaceous floras, where many of the Lower Cretaceous deposits are pyritized (Foucault et al., 1986; Geroch, 1978; Jousé, 1949; Strel'nikova & Martirosjan, 1981; Wall, 1975), and only one study site contains well-preserved fossils (Gersonde & Harwood, 1990; Harwood & Gersonde, 1990). The Cretaceous Diatom Database excludes the pyritized records, which decreases the overall diversity. Regardless of the above biases, the Late Cretaceous is a period of major diversification with all morphological clades present (also raphid diatoms; see Siver & Velez, 2022) and dispersal to a broad range of niches (Harwood et al., 2007; Sims et al., 2006).

The highest species diversity was identified in the extant genera. These genera had a worldwide distribution by the end of the Cretaceous (see Appendix S1), which likely promoted allopatric speciation (Sobel et al., 2010). Some of these, such as *Actinoptychus* and *Stephanopyxis*, are cosmopolitan taxa with broad ecological tolerance, widespread across Cenozoic sediments (Harwood & Maruyama, 1992; Roberts et al., 2003; Witkowski, Ashworth, et al., 2020; Witkowski, Harwood, et al., 2020) and in marine plankton or coastal environments in the modern ocean (Grassle, 2000; Round et al., 1990).

Overall, the diversity patterns presented here should be treated with a great deal of caution, however, since many taxa in the Cretaceous Diatom Database, predominantly in the most diverse genera, were identified only to the genus level. For example, 40 species of *Hemiaulus*, the most diverse Cretaceous genus, are highly similar morphologically, so the genus shows little evidence of morphological diversification commensurate with the number of named species in the genus (Jewson & Harwood, 2017). We suggest thorough analysis of these large genera would reveal that many species could be combined and classified together. As suggested in the supplementary discussion (Appendix S2), species described in *Triceratium* likely do not represent *Triceratium* and potentially should be reclassified to *Trinacria* Heiberg, *Medlinia* Sims, or *Sheshukovia*, which are likewise broadly confused genera, and many species have already been moved between these genera (see Appendix S1). Similar uncertainties apply to *Biddulphia*, *Briggera* Ross and Sims, and *Hemiaulus*, all of which represent polar centric diatoms with apical processes (Round et al., 1990). Careful SEM-based systematics would help resolve the relationships between these genera and their modern relatives and, as already suggested, potentially lead to descriptions of new fossil genera. Taken together, the total number of genera and possibly species in our database may be overestimated. Nevertheless, previous estimates point toward a number of 300 described species by the end of the Cretaceous (Jewson & Harwood, 2017; Knoll & Follows, 2016; Lazarus et al., 2014); hence, even if overestimated, Cretaceous diatoms are substantially more diverse than previously suggested.

On a final note, our knowledge of the Cretaceous fossil record could be improved in several ways, including (1) a search for new records to describe and complement existing fossil information, particularly from the Cenomanian–Coniacian period, as only two studies describe diatoms of these ages (Tapia & Harwood, 2002; Zalat, 2013); (2) reinvestigation of already published studies for taxonomic evaluation, particularly those with low-quality LM photographs and drawings (Forti & Schultz, 1932); and (3) unifying the taxonomy of easily confused genera, which is a daunting task and potentially unachievable without a significant dedicated effort and resources.

AUTHOR CONTRIBUTIONS

Karolina Bryłka: Conceptualization (lead); data curation (lead); formal analysis (lead); investigation (lead); visualization (lead); writing – original draft (lead); writing – review and editing (lead). **Matt P. Ashworth:** Conceptualization (supporting); formal analysis (supporting); investigation (supporting); writing – review and editing (equal). **Andrew J. Alverson:** Conceptualization (equal); formal analysis (supporting); funding acquisition (supporting); investigation (supporting); supervision (supporting); writing – original draft (supporting); writing – review and editing (equal). **Daniel J. Conley:** Conceptualization (equal); funding acquisition (lead); project administration (lead); resources (lead); supervision (lead); writing – original draft (supporting); writing – review and editing (equal).

ACKNOWLEDGMENTS

This research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant agreement No. 833454), a grant from the Knut and Alice Wallenberg Foundation to Daniel J. Conley, and a grant from the National Science Foundation (DEB-1651087) to Andrew J. Alverson. We thank three anonymous reviewers for insightful comments and suggestions.

ORCID

Karolina Bryłka https://orcid.org/0000-0002-3302-5365
Matt P. Ashworth https://orcid.org/0000-0002-4162-2004
Andrew J. Alverson https://orcid.org/0000-0003-1241-2654
Daniel J. Conley https://orcid.org/0000-0001-9668-9284

REFERENCES

Abe, K., Harwood, D. M., & Jordan, R. W. (2022). Observations on late cretaceous *Micrampulla* (Corethrales, Bacillariophyceae) from the Campbell Plateau (Zealandia), southwest Pacific Ocean. *Diatom Research*, *37*, 241–253.

Abe, K., Lam, D. W., & Ashworth, M. P. (2021). Observations on *Praecorethron crawfordii* gen. et sp. nov. (Corethrales, Bacillariophyceae) from upper cretaceous marine sediments, southwest Pacific. *Nova Hedwiga Beihefte*, *37*(3), 273–314.

Alverson, A. J. (2014). Timing marine–freshwater transitions in the diatom order Thalassiosirales. *Paleobiology*, *40*, 91–101.

Angelis, K., & Dos Reis, M. (2015). The impact of ancestral population size and incomplete lineage sorting on Bayesian estimation of species divergence times. *Current Zoology*, *61*, 874–885.

Ashworth, M. P., Nakov, T., & Theriot, E. C. (2013). Revisiting Ross and Sims (1971): Toward a molecular phylogeny of the Biddulphiaceae and Eupodiscaceae (Bacillariophyceae). *Journal of Phycology*, *49*, 1207–1222.

Ballance, P. F., Barron, J. A., Blome, C. D., Bukry, D., Cawood, P. A., Chaproniere, G. C. H., Frisch, R., Herzer, R. H., Nelson, C. S., Quinterno, P., Ryan, H., Scholl, D. W., Stevenson, A. J., Tappin, D. G., & Vallier, T. L. (1989). Late cretaceous pelagic sediments, volcanic ash and biotas from near the Louisville hotspot, Pacific plate, paleolatitude ~42°S. *Palaeogeography Palaeoclimatology Palaeoecology*, *71*, 281–299.

Barron, J. F. (1985). Diatom biostratigraphy of the CESAR 6 core. *Geological Survey of Canada Paper*, *84*(22), 137–148.

Bryłka, K., Alverson, A. J., Pickering, R. A., Richoz, S., & Conley, D. J. (2023). Uncertainties surrounding the oldest fossil record of diatoms. *Scientific Reports*, *13*, 8047.

Budd, G. E., & Mann, R. P. (2024). Two notorious nodes: A critical examination of relaxed molecular clock age estimates of the bilaterian animals and placental mammals. *Systematic Biology*, *73*(1), 223–234.

Chambers, P. M. (1997). Late cretaceous and Paleocene marine diatom floras. [Doctoral disseration, University of London]. https://discovery.ucl.ac.uk/id/eprint/10104676/

Conley, D. J., Frings, P. J., Fontorbe, G., Clymans, W., Stadmark, J., Hendry, K. R., Marron, A. O., & de la Rocha, C. L. (2017). Biosilicification drives a decline of dissolved Si in the oceans through geologic time. *Frontiers in Marine Science*, *4*, 397.

Crawford, R. M., & Hinz, F. (1995). The spines of the centric diatom *Corethron criophilum*: Light microscopy of vegetative cell division. *European Journal of Phycology*, *30*, 95–105.

Daugbjerg, N., & Guillou, L. (2001). Phylogenetic analyses of Bolidophyceae (Heterokontophyta) using *rbc*L gene sequences support their sister group relationship to diatoms. *Phycologia*, *40*, 153–161.

Davies, A., & Kemp, A. E. S. (2016). Late cretaceous seasonal palaeoclimatology and diatom palaeoecology from laminated sediments. *Cretaceous Research*, *65*, 82–111.

Dell'Agnese, D. J., & Clark, D. L. (1994). Siliceous microfossils from the warm late cretaceous and early Cenozoic Arctic Ocean. *Journal of Paleontology*, *68*, 31–47.

Forti, A., & Schultz, P. (1932). Erste Mitteilung über Diatomeen aus dem hannoverschen Gault [First report on diatoms from the Hanoverian Gault]. *Beihefte zum Botanischen Centralblatt*, *50*(2), 241–246.

Foucault, A., Servant-Vildary, S., Nianqiao Fang, S., Fang, N., & Powichrowski, L. (1986). Un des plus anciens gisements de diatomées découvert dans l'Albien-Cénomanien inférieur des Alpes ligures (Italie). Remarques sur l'apparition de ces algues [One of the oldest diatom deposits discovered in the Lower Albian-Cenomanian of the Ligurian Alps (Italy). Remarks on the appearance of these algae]. *Comptes Rendus de l'Académie Des Sciences II*, *303*, 397–402.

Fourtanier, E. (1991). Diatom biostratigraphy of Equatorial Indian Ocean site 758. *Proceedings of the Ocean Drilling Program, Scientific Results*, *121*, 189–208.

Fryxell, G. A., & Hasle, G. R. (1971). *Corethron criophilum* Castracane: Its distribution and structure. *Biology of the Antarctic Seas IV*, *17*, 335–346.

Geroch, S. (1978). Lower cretaceous diatoms in the polish Carpathians. *Annales. Societatis Geologorum Poloniae*, *48*, 283–295.

Gersonde, R., & Harwood, D. M. (1990). 25. Lower cretaceous diatoms from ODP leg 113 site 693 (Weddell Sea). Part 1: Vegetative cells. *Proceedings of the Ocean Drilling Program, Scientific Results*, *113*, 365–402.

Grassle, J. F. (2000). The ocean biogeographic information system (OBIS): An on-line, worldwide atlas for accessing, modeling and mapping marine biological data in a multidimensional geographic context. *Oceanography*, *13*, 5–7.

Gresham, C. W. (1985). Cretaceous and Paleocene siliceous phytoplankton assemblages from DSDP sites 216, 214 and 208 in the Pacific and Indian Oceans. University of Wisconsin–Madison (466 pp).

Guiry, M. D. (2013). AlgaeBase. World-wide electronic publication. http://www.algaebase.org

Hajós, M., & Stradner, H. (1975). Late cretaceous archaeomonadaceae, diatomaceae, and silicoflagellatae from the south Pacific Ocean, deep sea drilling project, leg 29, site 275. *Initial Reports of the Deep Sea Drilling Project*, *29*, 913–1009.

Hanna, G. D. (1927). *Cretaceous diatoms from California*. California Academy of Sciences.

Harwood, D. M. (1988). Upper Cretaceous and lower Paleocene diatom and silicoflagellate biostratigraphy of Seymour Island, eastern Antarctic peninsula. *Geological Society of America Memoirs*, *169*, 55–130.

Harwood, D. M., & Gersonde, R. (1990). 26. 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. *Proceedings ODP, Scientific Results*, *113*, 403–425.

Harwood, D. M., & Maruyama, T. (1992). Middle Eocene to Pleistocene diatom biostratigraphy of Southern Ocean sediments from the Kerguelen plateau, LEG 120. *Proceeding ODP, Scientific Results*, *120*, 683–733.

Harwood, D. M., Nikolaev, V. A., & Winter, D. M. (2007). Cretaceous records of diatom evolution, radiation, and expansion. *Paleontological Society Papers*, *13*, 33–59.

Hasle, G. R., & Sims, P. A. (1986). The diatom genera *Stellarima* and *Symbolophora* with comments on the genus Actinoptychus. *British Phycological Journal*, *21*, 97–114.

Ishii, K.-I., Iwataki, M., Matsuoka, K., & Imai, I. (2011). Proposal of identification criteria for resting spores of *Chaetoceros* species (Bacillariophyceae) from a temperate coastal sea. *Phycologia*, *50*, 351–362.

Jewson, D. H., & Harwood, D. M. (2017). Diatom life cycles and ecology in the cretaceous. *Journal of Phycology*, *53*, 616–628.

Jordan, R. W., Ashworth, M. P., Uezato, Y., & Manning, S. R. (2019). Polyphyly and homoplasic structures in rhizosolenioid diatom genera: A review. *Plant Ecology and Evolution*, *152*, 142–149.

Jordan, R. W., & Ito, R. (2002). Observations on *Proboscia* species from Late Cretaceous sediments, and their possible evolution from Kreagra. In *Proceedings of the 15th International Diatom Symposium*. *Perth* (*Australia*) (Vol. 28, pp. 313–330).

Jousé, A. P. (1948). Pre-tertiary diatom algae. *Botanicheskii Zhurnal SSSR*. *33*(3), 345–356.

Jousé, A. P. (1949). Diatomovye Mezozojskih otlozhenij [Diatoms from Mesozoic deposits]. In A. Proshkina-Lavrenko (Ed.), *Diatomovyj Analysiz* (1st ed., Vol. *1*, pp. 109–114). Obshchaya i paleobotanicheskaya harakteristika diatomovyh vodoroslej. Gosdeolizdat.

Knoll, A. H., & Follows, M. J. (2016). A bottom-up perspective on ecosystem change in Mesozoic oceans. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20161755.

Kociolek, J. P., Williams, D. M., Stepanek, J., Liu, Q., Liu, Y., You, Q., Karthick, B., & Kulikovskiy, M. (2019). Rampant homoplasy and adaptive radiation in pennate diatoms. *Plant Ecology and Evolution*, *152*, 131–141.

Kooistra, W. H., & Medlin, L. K. (1996). Evolution of the diatoms (Bacillariophyta). IV. A reconstruction of their age from small subunit rRNA coding regions and the fossil record. *Molecular Phylogenetics and Evolution*, *6*, 391–407.

Lazarus, D., Barron, J., Renaudie, J., Diver, P., & Türke, A. (2014). Cenozoic planktonic marine diatom diversity and correlation to climate change. *PLoS ONE*, *9*, e84857.

Lee, M. S. Y., & Ho, S. Y. W. (2016). Molecular clocks. *Current Biology*, *26*, R399–R402.

Lewitus, E., Bittner, L., Malviya, S., Bowler, C., & Morlon, H. (2018). Clade-specific diversification dynamics of marine diatoms since the Jurassic. *Nature Ecology & Evolution*, *2*, 1715–1723.

Li, C., Ashworth, M. P., Mackiewicz, P., Dąbek, P., Witkowski, J., Górecka, E., Krzywda, M., & Witkowski, A. (2020). Morphology, phylogeny, and molecular dating in Plagiogrammaceae family focused on Plagiogramma-Dimeregramma complex (Urneidophycidae, Bacillariophyceae). *Molecular Phylogenetics and Evolution*, *148*, 106808.

Long, J. A., Fuge, D. P., & Smith, J. (1946). Diatoms of the Moreno shale. *Journal of Paleontology*, *20*, 89–118.

Lozano-Fernandez, J. (2022). A practical guide to design and assess a phylogenomic study. *Genome Biology and Evolution*, *14*, evac129.

Medlin, L. K. (2015). A timescale for diatom evolution based on four molecular markers: Reassessment of ghost lineages and major steps defining diatom evolution. *Vie et Milieu-Life and Environment*, *65*(4), 219–238.

Medlin, L. K., & Kaczmarska, I. (2004). Evolution of the diatoms: V. Morphological and cytological support for the major clades and a taxonomic revision. *Phycologia*, *43*, 245–270.

Moshkovitz, S., Ehrlich, A., & Soudry, D. (1983). Siliceous microfossils of the Upper Cretaceous Mishash Formation, Central Negev, Israel. *Cretaceous Research*, *4*, 173–194.

Nakov, T., Beaulieu, J. M., & Alverson, A. J. (2018a). Accelerated diversification is related to life history and locomotion in a hyperdiverse lineage of microbial eukaryotes (diatoms, Bacillariophyta). *New Phytologist*, *219*, 462–473.

Nakov, T., Beaulieu, J. M., & Alverson, A. J. (2018b). Insights into global planktonic diatom diversity: The importance of comparisons between phylogenetically equivalent units that account for time. *ISME Journal*, *12*, 2807–2810.

Nakov, T., Theriot, E. C., & Alverson, A. J. (2014). Using phylogeny to model cell size evolution in marine and freshwater diatoms. *Limnology and Oceanography*, *59*, 79–86.

Nikolaev, V. A., Harwood, D. M., & Samsonov, N. I. (2001). *Early cretaceous diatoms*. Nauka.

Nikolaev, V. A., Kociolek, J. P., Fourtanier, E., Barron, J. A., & Harwood, D. M. (2001). Late cretaceous diatoms (Bacillariophyceae) from the Marca shale member of the Moreno formation, California. *Occasional Papers of the California Academy of Sciences*, *152*, 1–119.

Oreshkina, T. V., Lygina, E. A., & Vozhzhova, O. A. (2013). Diatoms and silicoflagellates of the upper cretaceous from Saratov region: Biostratigraphy and sedimentation settings. *Stratigraphy and Geological Correlation*, *21*, 222–236.

Parham, J. F., Donoghue, P. C. J., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., Inoue, J. G., Irmis, R. B., Joyce, W. G., Ksepka, D. T., Patané, J. S. L., Smith, N. D., Tarver, J. E., van Tuinen, M., Yang, Z., Angielczyk, K. D., Greenwood, J. M., Hipsley, C. A., Jacobs, L., … Benton, M. J. (2011). Best practices for justifying fossil calibrations. *Systematic Biology*, *61*, 346–359.

Parks, M. B., Nakov, T., Ruck, E. C., Wickett, N. J., & Alverson, A. J. (2018). Phylogenomics reveals an extensive history of genome duplication in diatoms (Bacillariophyta). *American Journal of Botany*, *105*, 330–347.

Parks, M. B., Wickett, N. J., & Alverson, A. J. (2018). Signal, uncertainty, and conflict in phylogenomic data for a diverse lineage of microbial eukaryotes (Diatoms, Bacillariophyta). *Molecular Biology and Evolution*, *35*, 80–93.

Pinseel, E., Janssens, S. B., Verleyen, E., Vanormelingen, P., Kohler, T. J., Biersma, E. M., Sabbe, K., van de Vijver, B., & Vyverman, W. (2020). Global radiation in a rare biosphere soil diatom. *Nature Communications*, *11*, 2382.

Roberts, A. P., Bicknell, S. J., Byatt, J., Bohaty, S. M., Florindo, F., & Harwood, D. M. (2003). Magnetostratigraphic calibration of Southern Ocean diatom datums from the Eocene–Oligocene of Kerguelen plateau (ocean drilling program sites 744 and 748). *Palaeogeography Palaeoclimatology Palaeoecology*, *198*, 145–168.

Ross, R., & Sims, P. A. (1997). A revision of *Actinodyction* Pantocsek. *Diatom Research*, *12*, 321–340.

Ross, R., & Sims, P. A. (2002). Trossulus Ross & Sims, a new genus of the Heliopeltaceae. *Diatom Research*, *17*, 177–188.

Round, F. E., Crawford, R. M., & Mann, D. G. (1990). *The diatoms: Biology and morphology of the genera*. Cambridge University Press.

Schütt, F. (1896). Bacillariales. In A. Engler & K. Prantl (Eds.), *Die natürlichen Pflanzenfamilien* (1st ed., Vol. *1*, pp. 31–150). Verlag von Wilhelm Engelmann.

Shimada, C., Saito-Kato, M., Jenkins, R. G., Tanaka, Y., & Hikida, Y. (2022). Late cretaceous diatoms (Bacillariophyta) from the Teshio-Nakagawa area, Hokkaido, northern Japan: Significance for their origin and biostratigraphy. *Paleontological Research*, *26*, 301–313.

Sims, P. A. (1988). The fossil genus *Trochosira*, its morphology, taxonomy and systematics. *Diatom Research*, *3*, 245–257.

Sims, P. A. (1998). The early history of the Biddulphiales. I. The genus *Medlinia* gen. nov. *Diatom Research*, *13*, 337–374.

Sims, P. A., Ashworth, M. P., Theriot, E. C., & Manning, S. R. (2023). Molecular and morphological analysis of *Biddulphia sensu lato*: A new diagnosis of *Biddulphia*, with a description of the new genera *Biddulphiella* and *Neobrightwellia*. *Marine Micropaleontology*, *178*, 102186.

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.

Siver, P. A., & Velez, M. L. (2022). The oldest raphe-bearing diatoms: Evidence from the Upper Cretaceous of western and northern Canada. *Cretaceous Research*, *144*, 105456.

Sobel, J. M., Chen, G. F., Watt, L. R., & Schemske, D. W. (2010). The biology of speciation. *Evolution*, *64*, 295–315.

Sorhannus, U. (1997). The origination time of diatoms: An analysis based on ribosomal RNA data. *Micropaleontology*, *43*, 215–218.

Sorhannus, U. (2007). A nuclear-encoded small-subunit ribosomal RNA timescale for diatom evolution. *Marine Micropaleontology*, *65*, 1–12.

Souffreau, C., Verbruggen, H., Wolfe, A. P., Vanormelingen, P., Siver, P. A., Cox, E. J., Mann, D. G., Vijver, B. V., Sabbe, K., & Vyverman, W. (2011). A time-calibrated multi-gene phylogeny of the diatom genus *Pinnularia*. *Molecular Phylogenetics and Evolution*, *61*, 866–879.

Strel'nikova, N. I. (1975). Diatoms of the cretaceous period. *Nova Hedwigia*, *53*, 311–321.

Strel'nikova, N. I., & Martirosjan, G. N. (1981). Lower diatom algae from Stavropol. *Viestnik LGU, Seriya Biologiya*, *3*, 52–57.

Sunesen, I., Hernández-Becerril, D. U., & Sar, E. A. (2008). Marine diatoms from Buenos Aires coastal waters (Argentina). V. Species of the genus *Chaetoceros*. *Revista de Biología Marina y Oceanografía*, *43*(2), 303–326.

Tapia, P. M., & Harwood, D. M. (2002). Upper Cretaceous diatom biostratigraphy of the Arctic archipelago and northern continental margin, Canada. *Micropaleontology*, *48*, 303–342.

Tiley, G. P., Poelstra, J. W., Dos Reis, M., Yang, Z., & Yoder, A. D. (2020). Molecular clocks without rocks: New solutions for old problems. *Trends in Genetics*, *36*, 845–856.

Tréguer, P., Nelson, D. M., Van Bennekom, A. J., Demaster, D. J., Leynaert, A., & Quéguiner, B. (1995). The silica balance in the world ocean: A reestimate. *Science*, *268*, 375–379.

Wall, J. H. (1975). Diatoms and radiolarians from the Cretaceous system of Alberta: A preliminary report. In G. W. E. Caldwell (Ed.), *The Cretaceous system in the western interior of North America: The proceedings of an international symposium organized by the Geological Association of Canada and held at the University of Saskatchewan in Saskatoon* (pp. 391–410). Geological Association of Canada.

Williamson, T., Henderson, R. A., Price, G. D., & Collerson, K. D. (2012). Strontium–isotope stratigraphy of the lower Cretaceous of Australia. *Cretaceous Research*, *36*, 24–36.

Witkowski, A., Ashworth, M., Li, C., Sagna, I., Yatte, D., Górecka, E., Franco, A. O. R., Kusber, W. H., Klein, G., Lange-Bertalot, H., Dąbek, P., Theriot, E. C., & Manning, S. R. (2020). Exploring diversity, taxonomy and phylogeny of diatoms (Bacillariophyta) from marine habitats. Novel taxa with internal costae. *Protist*, *171*, 125713.

Witkowski, J., Harwood, D. M., & Chin, K. (2011). Taxonomic composition, paleoecology and biostratigraphy of late cretaceous diatoms from Devon Island, Nunavut, Canadian high Arctic. *Cretaceous Research*, *32*, 277–300.

Witkowski, J., Harwood, D. M., Wade, B. S., & Bryłka, K. (2020). Rethinking the chronology of early Paleogene sediments in the western North Atlantic using diatom biostratigraphy. *Marine Geology*, *424*, 106168.

Witkowski, J., Sims, P. A., & Williams, D. M. (2017). Typification of Eocene–Oligocene diatom taxa proposed by Grove & Sturt (1886–1887) from the Oamaru diatomite. *Diatom Research*, *32*, 363–408.

Zalat, A. A. (2013). Cretaceous diatoms biostratigraphy and taxonomy from the North-eastern Sinai, Egypt. *Micropaleontology*, *59*, 305–323.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1. Cretaceous Diatom Database.

Appendix S2. Supplementary discussion on extant genera included in the Cretaceous Diatom Database but excluded from the tie point set.

How to cite this article: Bryłka, K., Ashworth, M. P., Alverson, A. J., & Conley, D. J. (2024). The Cretaceous Diatom Database: A tool for investigating early diatom evolution. *Journal of Phycology*, *00*, 1–15. https://doi.org/10.1111/jpy.13499