INCORPORATING MOLECULAR DATA IN THE TAXONOMIC STUDY OF DIATOMS: AN EXAMPLE USING TWO WELL-KNOWN GENERA, *FRUSTULIA* AND *NAVICULA* S.S. (BACILLARIOPHYCEAE, NAVICULALES)

ANDRÉANNE BOUCHARD

Thesis submitted to the University of Ottawa
in partial Fulfillment of the requirements for the
Masters of Science degree in Biology

Department of Biology
Faculty of Science
University of Ottawa

Abstract

Diatoms are crucially important to the global ecosystem due to their role in regulating the world's carbon and silicon cycles, and their production of large amounts of organic material in aquatic environments. They are thought to comprise ca. 100,000 species, although some estimates suggest that there could be over a million. Despite their importance and high species diversity, little is known about their phylogeny due to technical issues that hinder the reconstruction of their relationships. However, owing to a new technique that allows for DNA to be amplified from a single isolated cell, it is possible to explore diatom relationships with extensive taxonomic sampling. This thesis aims to demonstrate that the integration of molecular data and morphological characters can provide a new paradigm for future phylogenetic and taxonomic studies of diatoms, especially among closely related and taxonomically complex groups. To achieve this, I examined common species from two naviculoid diatom genera, *Frustulia* and *Navicula* using sequence data from three molecular markers (*rbcL*, *atpB*, 18S), traditional and fine-scale morphological characters, and frustule shape. The molecular markers *rbcL* and *atpB* evolved at a similar rate and performed well at reconstructing species-level phylogenies, whereas 18S was more conserved and best used for resolving relationships at higher taxonomic levels. Hidden diversity was uncovered in what have traditionally been thought as well-circumscribed taxa, and three new species were described. The methods used here show promise for the future of diatom systematics.

Acknowledgments

First, I would like to thank my supervisors Paul Hamilton and Dr. Julian Starr, who have been with me since my undergraduate honour's thesis. It has been a pleasure working with you, and I truly feel that I have learned so much through your guidance – not only academically, but also in professional and personal contexts.

I also want to thank Roger Bull for his help with DNA-related matters, and for making the Canadian Museum of Nature's DNA laboratory such a friendly, dynamic place to work. A special thanks to the students and volunteers of the phylogeny lab: Hadil Sayed, Elizabeth Hulianga, Joe Holmes, Karen Cimon, and Micheline Vanier. Thank you as well to Dr. Amanda Savoie, who provided me with some guidance on the manuscript for Chapter 2.

To my parents, Jocelyne Bolduc and Francis Bouchard, my sister, Mélodie Bouchard, and my friends – thank you for your support and encouragement.

Finally, a big thank you to my partner, Sébastien Charette. Your unwavering support has been crucial for me throughout this process. I am so glad to have you by my side.

Statement of contributions of collaborators

This thesis is my original work, conducted under the supervision of Paul Hamilton and Dr. Julian Starr, and under the guidance of committee members Dr. Frances Pick and Dr. Nicolas Rodrigue.

Chapters 1 and 4 are the General Introduction and General Conclusions. They were written solely by me, and revised by Paul Hamilton and Dr. Julian Starr.

Chapter 2 was published in the journal *Diatom Research*.

I was responsible for the laboratory work and data analysis. Paul Hamilton was responsible for some of the sampling. I received guidance from Paul Hamilton regarding terminology and morphology, and from Dr. Julian Starr and Dr. Amanda Savoie regarding phylogenies and use of software.

I wrote the manuscript, which was then revised by all co-authors. Paul Hamilton was responsible for the supplemental material presented in Appendix A.

Chapter 3 will be submitted to a peer reviewed journal.

I was responsible for the laboratory work and data analysis. Paul Hamilton was responsible for some of the sampling, and processing of samples for microscopy. I received guidance from Paul Hamilton regarding culturing and primer design, and from Dr. Julian Starr regarding phylogenies.

I wrote the manuscript, which was then revised by Paul Hamilton and Dr. Julian Starr.

Table of contents

List of Tables

List of Figures

Chapter 1 – General Introduction

The term "algae" refers to a large group of photosynthetic organisms that are very diverse genetically and morphologically (Lee, 2008). These organisms are not necessarily closely related, which makes the group artificial and polyphyletic. Because of this particularity, algae as a group are difficult to define. Many definitions of the term have been proposed over the years, some of which include both photosynthetic prokaryotes (cyanobacteria) and eukaryotes, but none has been universally accepted (Chapman, 2013; Bold & Wynne, 1985). Most modern definitions exclude prokaryotes and restrict the word to a diverse group of photosynthetic eukaryotic organisms that comprises multiple large lineages, some of which, like the Chlorophyta, are closely related to land plants (Nabors, 2004). Depending on the criteria used, algae have been estimated to include anywhere from 30 000 to over 1 million species (Guiry, 2012).

As primary producers, algae form the base of the food chain in aquatic systems, providing animal life with food and energy (Bold & Wynne, 1985). Estimates suggest they contribute to over 20% of the biomass production for the global ecosystem (Ryther, 1969; Antoine et al. 1995). The products that they secrete or excrete, such as polysaccharides, amino acids, enzymes, and proteins, are also used by many other organisms (Aaronson et al. 1971).

Even within such an important group of organisms, the diatoms stand out. Diatoms are eukaryotic, unicellular microalgae enclosed within a siliceous cell wall called a frustule (Round et al. 2007). They are part of the heterokont lineage, in which they appeared ca. 250 Ma ago (Medlin, 2011). The heterokonts are characterized by the presence of two dissimilar flagella on their motile cells, with one being covered in lateral tripartite hairs and the other being smooth (Kooistra et al. 2007). The group also contains descendants of such organisms, which have

subsequently lost this feature (Riisberg et al. 2009). In diatoms, this structure has been lost or greatly reduced, as only male gametes of centric diatoms have retained a flagellum (Daugbjerg & Guillou, 2001).

The chloroplasts of heterokonts originate from a red algal endosymbiont (Kooistra et al. 2007). They contain chlorophyll a + c and the accessory pigment fucoxanthin, which gives them a characteristic golden-brown color (Kooistra et al. 2007). There are two alternative theories for the origin of chloroplasts in diatoms, namely (1) the secondary endosymbiosis of a red alga or (2) the tertiary endosymbiosis of a cryptomonad which contained a chloroplast of red algal origin (Nonoyama et al. 2019). Whether the first or second theory is true has yet to be resolved.

Diatoms are cosmopolitan, and may be found in any environment with water, such as oceans, lakes, soils, and wetlands (Smol & Stoermer, 2010). They form the majority of the phytoplankton in marine and fresh water environments, partly because they divide more rapidly than most other taxa (John, 2012; Furnas, 2010). Comprising anywhere from 20,000 to over a million species, with most reports suggesting ca. 100,000 taxa, the group is thought to be responsible for 25% to 50% of the primary production of organic material in oceans (Yool & Tyrrell, 2003). Diatoms also play an important role in regulating the global silicon cycle (Yool & Tyrell, 2003), and the oceanic carbon cycle (Smetacek, 1985). They are an ideal tool for monitoring environmental conditions and inferring paleoclimates as their silica valves tend to be abundant and well preserved in most habitats, and because of species-specific sensitivities to various water conditions, such as pH and salinity (Smol, 1988; Smol & Stoermer, 2010).

Nevertheless, despite their worldwide distribution and ecological significance, very little is known about diatom relationships. This is largely due to the technical issues that hinder both the use of morphological and molecular characters in phylogenetic reconstruction.

Few studies have used morphology to clarify diatom relationships (Cox & Williams, 2000; Kociolek et al. 1989) because phylogenetic resolution is typically poor due to a lack of characters (Cox & Williams, 2000). This is complicated by the fact that of the few characters that can be scored in studies that describe and compare diatom species (ca. 10 to 20; Siver & Baskette, 2004; Cox, 1999; Wikowski et al. 1998), many are quantitative and continuous such as valve length, width, and stria density (Siver & Baskette, 2004). Such continuous characters are ill-suited to reconstructing relationships because dividing them into finite, qualitative states, as required by modern phylogenetic methods (Wiens, 2001), is subjective (Gift & Stevens, 1997; Stevens, 2000). Slight adjustments to the coding of such characters can have a major impact on tree topologies (Cox & Williams, 2000), thus questioning whether these characters can accurately recover a species phylogeny (Stevens, 2000; Scotland et al. 2003). The current trend in diatom species identifications tends towards using narrow concepts for species, with little reference to phenotypic expression. Species are being separated from others on the basis of one or two morphological characters exclusively, without consideration for the possibility that these small differences may be phenotypic variations (Witkowski et al. 1997; Van de Vijver & Kopalová, 2014).

Diatom phylogenies reconstructed using molecular data are more common; however, they face their own challenges. Diatoms are unicellular organisms, and single cells typically do not provide enough DNA for standard molecular techniques. Consequently, studies that use DNA sequences to reconstruct diatom phylogenies (Sato et al. 2009; Souffreau et al. 2011; Ajani et al. 2013) have done so by growing large cultures for all the taxa analysed, a laborious and time-consuming process that is not possible for all diatom species. Consequently, most molecular phylogenies to date have focused on higher-level relationships where the sampling of many

individuals and extensive taxonomic sampling are not always essential (Witkowski et al. 2020; Kulikovskiy et al. 2016). However, recent technical developments have made it possible to amplify and study DNA from a single isolated cell (Hamilton et al. 2015), which has enabled the exploration of relationships within diatoms at all taxonomic levels, even at the species and infraspecific levels where taxonomic problems are most acute.

The goal of this thesis is to demonstrate that the integration of both molecular data with traditional characters can provide a new paradigm for future phylogenetic and taxonomic studies among closely related and taxonomically complex diatom groups. Specifically, the objectives of my thesis are to: (1) use the single-cell DNA extraction and amplification technique of Hamilton et al. (2015) to reconstruct the phylogeny of two well-known, but taxonomically difficult diatom genera, by sampling a wide range of their morphological, geographic, and taxonomic diversity; (2) use these phylogenies to reinterpret the morphology of these genera and clarify their taxonomy, and (3) compare trends in the phylogenetic relationships between two benthic naviculoid diatom genera. Given how poorly the relationships of diatoms and their taxonomy is known, I expect that the integration of both molecular and traditional characteristics will not only be integral to resolving some long-stranding taxonomic issues, but it will lead to the discovery of cryptic diversity. By studying two genera, I hope to see similar trends in both, thus demonstrating the importance of integrating molecular data into traditional systematic studies within diatoms.

The genera chosen are *Frustulia* Rabenhorst and *Navicula* Bory sensu stricto (s. str.). The principle morphological characters for each genus are shown in Figures 1.1 – 1.6. *Frustulia* is a diatom genus comprising 242 species and 56 additional varieties and forms (Guiry, 2020). It is primarily found in the benthic zone of acidic, but clean freshwaters, such as seepage lakes,

rivers, and streams worldwide. In North America, studies have shown that it is a common and widespread genus that can produce a significant amount of biomass in the habitats where it is found (Gaiser & Johansen, 2000; Siver et al. 2004). The genus has been the subject of much taxonomic confusion. For example, a study suggests *F. rhomboides* consists of five separate species, some being widespread and common (*F. saxonica*, *F. crassinervia*, and *F. krammeri*; Vesela et al. 2012), but they remain difficult to distinguish (Siver et al. 2004). Phenotypic plasticity, which may result from environmental conditions, often masks the informative differences of key features among taxa (Siver et al. 2004; Mann, 1999).

Frustulia was selected in part because it comprises a relatively small number of species compared to other diatom genera – therefore, a good starting point for a newly developed protocol. Although some of its species were studied using molecular data in Europe, North American species or North American representatives of species also found in Europe are almost exclusively known by morphological data alone. This data shows a wide morphological diversity of species in North America, including morphological extremes such as *Frustulia bahlsii* Edlund & Brant (large valves, distinct porte-crayon structure) and *F. vulgaris* (Thwaites) De Toni (unique central nodule structure and striae organisation). The presence of pseudo-cryptic species complexes, such as the complex between the closely related *F. saxonica,* and *F. crassinervia*, was also observed (Siver & Baskette, 2004). These considerations make *Frustulia* a good candidate for this work. The genus is studied in Chapter 2.

Navicula is one of the most commonly encountered diatom genera in the northern hemisphere (Underwood, 2001). It is a widespread genus, and has been found worldwide (Witkowski et al. 1998). It comprises about 200 species, most of which can be found in freshwater environments (ca. 150), although some are found in brackish or marine habitats

(Bruder & Medlin, 2008). With more research in marine ecosystems, it is anticipated that more marine species will be discovered.

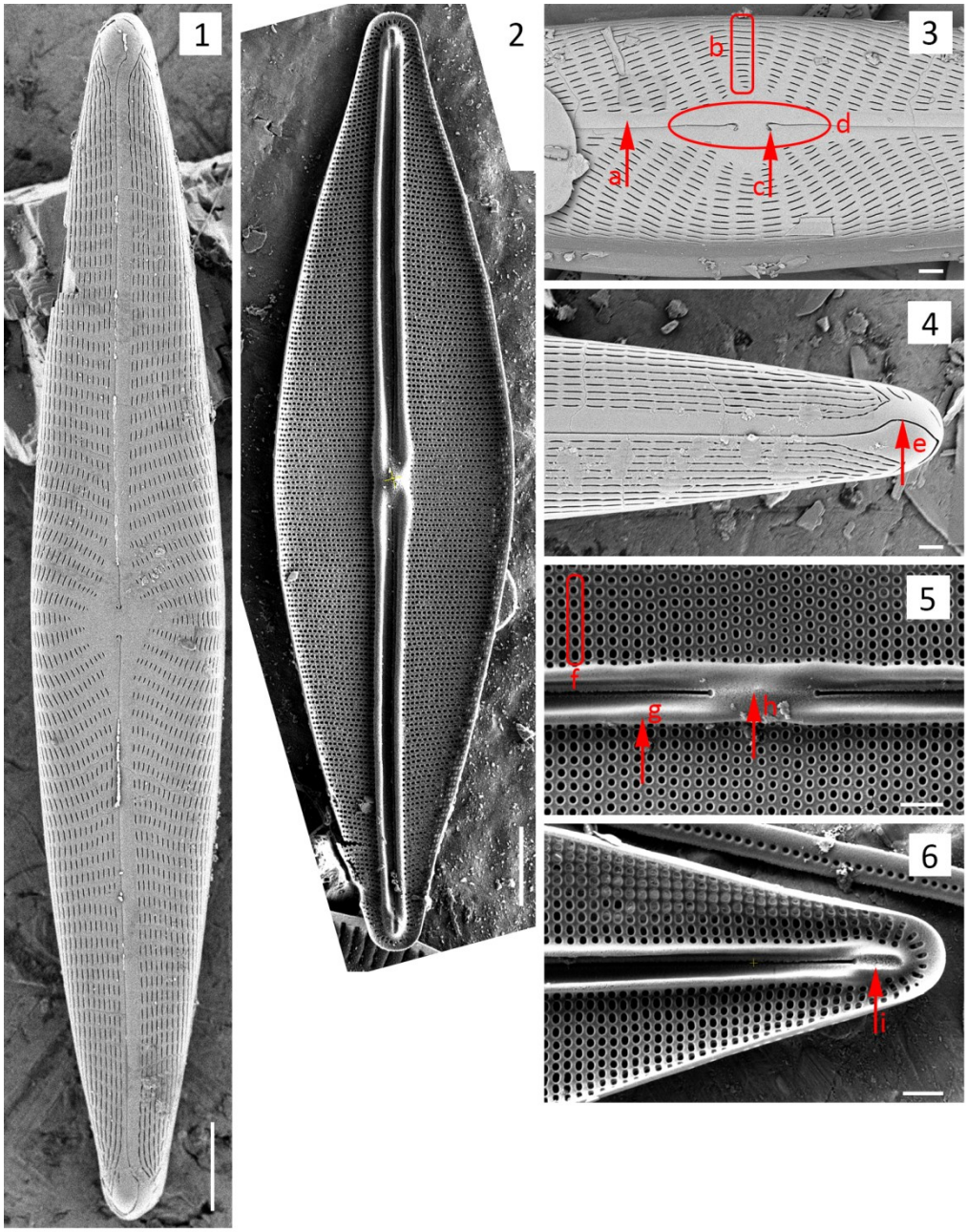
Navicula was chosen as the second genus to study as it shows a long history of taxonomic disagreements. In her review of the genus, Cox (1999) revealed a lack of consensus on the definition of *Navicula*. Many of its major species have been described as cosmopolitan (e.g., *N. cryptocephala* Kützing, *N. veneta* Kützing, *N. trivialis* Lange-Bertalot, *N. gregaria* Donkin, and *N. cryptotenella* Lange-Bertalot) (Poulíčková et al. 2010). However, evidence suggests that these could be cryptic and pseudo-cryptic species complexes, masking several species with restricted distributions (Poulíčková et al. 2010). Such hidden and poorly known diversity not only has enormous implications for the conservation of rare species, but also to our understanding of aquatic ecology worldwide. Few studies have used molecular data to explore relationships within the genus, and they have used only a handful of species that were exclusively sampled from Europe (Bruder & Medlin, 2008; Poulíčková et al. 2010). In light of these factors, *Navicula* is an ideal choice for re-examination with modern techniques. *Navicula* is studied in Chapter 3.

Few phylogenetic studies of diatoms have attempted to define a species concept for the group. Although there has been much debate about the species concept in biology, the conversation has been dominated by the biological species concept (BSC) for the past 80 years (Mann, 1999). The BSC posits that species are "groups of interbreeding natural populations that are reproductively isolated from other such groups" (Mayr, 2000).

The BSC is challenging to apply in a practical manner to diatoms, as it is difficult to observe their sexual reproduction. Although diatoms reproduce both asexually and sexually, their primary method of reproduction is asexual (Mann, 1999).

In a comprehensive review of the species concept in diatoms, Mann (1999) expanded upon the BSC and posited that species sit at the boundary between hierarchical and non-hierarchical variation. That is because reproductive barriers between species lead to discontinuous variation, while gene flow, which provides cohesion within a species, results in continuous variation. Mann further proposes that finding this cohesion, by using as many independent sources of data as possible and looking for congruence between them, is the key to recognizing species in diatoms. In this thesis, multiple sources of data were used to search for characters unique to groups of individuals. As these unique characters suggested genetic isolation from all other such groups, they were treated as species.

Despite the fact that diatoms are one of the most common and ecologically important organisms in aquatic habitats worldwide, their relationships and taxonomy are poorly resolved. Furthermore, species identification is often inaccurate due to the unreliable nature of the morphological characters currently used (Mann, 1999). Large-scale species-level phylogenetic analyses are now possible thanks to a novel molecular technique (Hamilton et al. 2015), and a subsequent reinterpretation of morphology is necessary to address old taxonomic problems and to facilitate the easy identification of its species for future studies. Additionally, by opening the possibility of easily combining DNA sequence data with traditional morphological data, this thesis could serve as a model to better understand the diversity and evolution of these key organisms.



Figs 1.1 – 1.6. Principle morphological characters for genera *Navicula* and *Frustulia*. For *Navicula*: a. sternum b. striae c. proximal pores d. central area e. axial hook. For *Frustulia*: f. striae; g. raphe ribs; h. central nodule; i. helictoglossae. In *Frustulia*, the fusion of the raphe ribs with the helictoglossae forms a porte-crayon.

Chapter 2 – Molecular and morphological Data Reveal Hidden Diversity in Common North American *Frustulia* Species (Amphipleuraceae)

This chapter is a slight modified version of the following published article:

AndréAnne J. Bouchard, Paul B. Hamilton, Amanda M. Savoie & Julian R. Starr (2019): Molecular and morphological data reveal hidden diversity in common North American *Frustulia* species (Amphipleuraceae), Diatom Research 34(4) 205 – 223.

2.1 Acknowledgements

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant to JRS), the Canadian Museum of Nature (RAC), and the Ontario Graduate Scholarship program (OGS). We wish to extend our thanks to Roger Bull for his direction and guidance in the DNA analysis and two reviewers for valuable comments and criticisms on the original submission. Collections from New York State (Department of Environment and Conservation) were under TRP permit. Collections from the Northwest Territories were under Aurora Research Institute licence numbers 14435 and 14965. Other collections from Canada (Ontario, Nova Scotia, Newfoundland) fall under a public access designation. No potential conflict of interest is here reported by the authors.

Cette recherche a été financée par le Conseil de recherches en sciences naturelles et en génie du Canada, le Musée canadien de la nature, et le programme de Bourse d'études supérieures de l'Ontario (BÉSO).

2.2 Abstract

Frustulia Rabenhorst is an established diatom genus that is common and widespread across North America. Like many diatom genera, taxa within *Frustulia* have been the subject of taxonomic confusion. Although recent studies have examined species from Europe and New Zealand, there exists no detailed genetic data for North American individuals. Using both molecular (*rbcL* and 18S rRNA sequences) and morphological (i.e. frustule characters and shape analysis) data, we investigated common taxa from *Frustulia* in North America. European and New Zealand taxa were also included to study how North American species were related. We recognized nine taxa from North America in this study including two unknowns. A new species, *F. gibsonea* sp. nov. is described. *Frustulia gibsonea* was observed in previous studies and described as *F.* cf. *krammeri* based on morphology. The use of molecular characters in this study demonstrates that this taxon is a distinct species, and not a morphotype of *F. krammeri*. Despite differences in *rbcL* and 18S, *F. gibsonea* and *F. krammeri* are quite similar morphologically, showing overlap using both traditional measurements and shape analysis. This suggests that the combination of molecular and morphological data can help in deciphering these morphologically cryptic taxa. *Frustulia saxonica* and *F. krammeri* could be weakly separated based on sequence divergences, although in this case morphological characters were more useful. As evidenced by the low sequence divergence values obtained between the two taxa, they are closely related. Future molecular research, focusing on less conserved markers, will be necessary to resolve the relationships of these taxonomic complexes.

Keywords: *Frustulia*, *rbcL*, 18S, shape analysis, morphology, North America

2.3 Introduction

Frustulia Rabenhorst is an old established diatom genus of approximately 200 species (Guiry & Guiry, 2019) that can be recognized in light microscopy (LM) by thickened axial ribs with a porte-crayon apex termination and distinct paired flattened chloroplasts with a central lenticular pyrenoid (Round et al. 2000, Cox, 1996). The genus is primarily found in benthic freshwater environments where it lives as solitary cells or in clustered mucilage tubes. Species in this genus are most abundant in acidic habitats (pH [average weight means] 4.3 – 5.2; Hustedt, 1930; Siver & Baskette, 2004), however taxa like *Frustulia vulgaris* (Thwaites) De Toni can be found in disturbed or unstable environments with a circumneutral pH (7.7; Hustedt, 1930; Siver & Hamilton, 2011; pers. obs. P. B. Hamilton). Studies have shown that the genus is widespread across North America producing a notable amount of biomass in the habitats where it is found (Gaiser & Johansen, 2000; Siver et al. 2005). It appears especially common along the eastern coast of the United States, in states such as North and South Carolina (Gaiser & Johansen, 2000; Siver & Hamilton, 2011), Massachusetts (Siver & Baskette, 2004), Maine (Edlund & Brant, 1997), and New Jersey (Siver & Hamilton, 2011). A large-scale study of waterbodies along the East coast of the United States found 11 taxa belonging to the genus with *Frustulia krammeri* Lange-Bertalot & Metzeltin, *Frustulia saxonica* Rabenhorst, and *Frustulia crassinervia* (Brébisson) Lange-Bertalot & Krammer being the most common (in over half of all study sites; e.g. Siver & Hamilton, 2011).

Like many diatom genera, taxa within *Frustulia* have been the subject of confusion. The original type species for the genus is *F. saxonica* and the type material (Rabenhort Exsicc 42 Algen Sachsens) has two morphological forms, large and small with wide and narrow sternums (Lange-Bertalot & Jahn, 2000). At present these forms are still considered to be the same

species. Another species, *Frustulia rhomboides* (Ehrenberg) De Toni, has a number of varieties which have recently been morphologically recognized as species; two of which are common in North America, *F. crassinervia* and *F. krammeri* (Siver & Baskette, 2004). The distinction between the species *F. saxonica*, *F. rhomboides, F. crassinervia* and *F. krammeri* is still primarily based on valve shape and formation of the thickened internal raphe sternum (Lange-Bertalot & Jahn, 2000; Siver & Baskette 2004). For example, *F. saxonica* is separated from *F. rhomboides* by valve size, general shape, stria density and terminal formation of the porte-crayon. However, phenotypic plasticity caused by environmental conditions can mask the informative differences of these features (Mann, 1999), and there still exists confusion around the distinctions between these taxa (e.g. Siver et al. 2004; Urbánková et al. 2016).

Although very little is known about the phylogeny of *Frustulia*, six taxonomic groups from Europe, Azores Islands, Marion Island, Greenland and New Zealand, including taxa noted above, have been studied using a combination of morphological and molecular taxonomy (Veselá et al. 2012; Urbánková & Veselá, 2013; Urbánková et al. 2016). Using the highly variable large subunit (LSU) ribosomal DNA expansion segments D1 and D2, Veselá et al. (2012) were unable to effectively separate European members of the *F. rhomboides* complex described above. This suggests that more conserved markers might not be able to resolve species flocks within *Frustulia*. Using DNA sequences (LSU and *rbcL*) and traditional and geometric morphometrics, Urbánková et al. (2016) were able to define five taxa within Europe that were potentially Holarctic in distribution. Although species such as *F. krammeri* and *Frustulia septentrionalis* Lange-Bertalot were genetically distinct in their analyses, *F. crassinervia* and *F. saxonica* could not be clearly separated. At present, no detailed genetic data are available for North American individuals in the above mentioned taxa, so the question as to how they may be

related to European conspecifics remains unresolved. Fortunately, recent technical developments in DNA extraction and PCR amplification in single isolated cells or small groups of cells have now made it possible to reconstruct phylogenies and characterise genetic diversity on wide geographic scales without the need for established cultures (Hamilton et al. 2015).

The objective of this study was to characterize, using both molecular and morphological data, the natural variation that might be expected within and between common species of *Frustulia* in North America. With this goal in mind, we reconstructed a phylogeny of morphologically similar and presumably closely related species including *F. krammeri, F. saxonica,* and *F. crassinervia*, and morphologically extreme and presumably distant species like *Frustulia bahlsii* Edlund & Brant (large valves, distinct porte-crayon structure) and *F. vulgaris* (unique central nodule structure and striae organisation) to determine whether molecular and morphological divergence were correlated. When possible, other specimens were included in order to study how North American taxa and global species were related. A complete analysis of morphological variation, including shape analysis with species descriptions, was conducted.

2.4 Materials and Methods

2.4.1 Sampling and outgroup choice

Sediment samples were collected from twenty-three freshwater sources in North America (Table 2.1). To do this, macropipetting was used to aspirate the sediment surface from the littoral zone of water bodies. This living material was brought to the lab, and a total of sixty *Frustulia* single-cells were isolated using a micropipette with contaminants removed by successive water isolations. Cells were then transferred to PCR tubes with a 10% Chelex 100 solution (Bio-Rad

Laboratories) (Hamilton et al. 2015). Three *Amphipleura pellucida* (Kützing) Kützing (= *Frustulia pellucida* Kützing) single-cells were also isolated from two of these sediment samples to act as a sister outgroup in the phylogenetic analyses. As a segregate genus of *Frustulia*, *Amphipleura* is presumed to be among the closest relatives to *Frustulia* in the family Amphipleuraceae (Cleve, 1894). Species of the genus *Climaconeis* Grunow and *Berkeleya* Greville were also chosen to be used as closely related outgroups within the Order Naviculales (Nakov et al. 2018). Voucher information for all samples used in this study can be found in Table 2.1. GenBank accession numbers for all specimens used in this study can be found in Table 2.2.

2.4.2 Genetic studies

DNA extraction was performed on each of the specimens using the Chelex-based protocol detailed in Hamilton et al. (2015). A nested PCR technique was used to amplify portions of the following molecular markers: ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) and 18S ribosomal RNA (18S rRNA). *RbcL* has been identified as the best gene for species-level taxonomy in diatoms, while 18S (SSU) represents a more conserved barcode marker (Urbánková & Veselá, 2013, and citations within). For the first PCR step, the reactions were carried out in a total volume of 15 µL, with final concentrations of: 1X DreamTaq Buffer (ThermoFisher Scientific), 2 mM dNTPs, 0.25 mM of each primer, 0.75 U of DreamTaq polymerase (ThermoFisher Scientific), 1 µL of DNA template, and DNA-free water for the remaining volume. Primers used are listed in Table 2.3.

For the first PCR amplification, thermocycling conditions were as follows, for 35 cycles: an initial step at 95ºC for 3 min, followed by a denaturing step at 95 ºC for 30 sec, an annealing step at 52 ºC for 30 sec, an extension step at 72 ºC for 1 min 30 sec, and a final step at 72 ºC for 10 min. The second PCR amplification was performed using the same procedure and cycling steps as the first one, with the slight modification that 1 µL of product from the first reaction was used instead of 1 µL of DNA template. For the 18S gene, however, a 10X dilution of the PCR product was performed between the first and second amplification in order to minimize non-specific amplification. PCR products were visualized using 1.5 % agarose gel electrophoresis with Ethidium Bromide.

The products of successful reactions were diluted in DNA-free water (from 0 µL to 60 µL, depending on the strength of the band). Sequencing was performed following the protocol in Hamilton et al. (2015). Using Geneious version 6.1.5 (Kearse et al. 2012), consensus sequences were assembled and aligned using the MAFFT (Multiple Alignment using Fast Fourier Transform) tool. Because the 18S alignment contained gaps, it was re-aligned using the method described in Starr et al. (2004). Ambiguously aligned bases were adjusted manually and assessed by parsimony analysis. If the adjustment resulted in a tree with a lesser number of steps, or produced a tree with the same number of steps but removed insertion/deletions (indels) from the matrix, it was accepted. Otherwise, the adjustment was rejected. The shortest tree obtained comprised 110 steps. The datasets were trimmed to 1151 characters for *rbcL* and 744 characters for 18S. Six sequences were missing from the *rbcL* datset, and 31 from the 18S dataset, due to lower amplification success in the latter.

Two trees were reconstructed for each molecular marker using maximum parsimony (MP) and maximum likelihood (ML) methods. Parsimony analyses were conducted in PAUP

4.0b10 (Swofford, 2002) using heuristic searches and a random addition sequence of taxa for 1000 replicates with the MULTREES option on and with tree-bisection-reconnection (TBR) branch swapping conducted on a maximum of 1000 trees per replicate. Support for internal branches was assessed via a bootstrap (BS; Felsenstein, 1985) using a simple addition sequence for 1,000 replicates with the MULTREES option "off" which gives BS values similar to those seen when the option is "on" (DeBry & Olmstead, 2000). Maximum likelihood trees were produced using the program RAxML version 8.2.11 (Stamatakis, 2014) in Geneious version 11.1.5 (https://www.geneious.com). Partitioning schemes were determined using PartitionFinder version 2.1.1, with linked branch lengths and the "all" search scheme (Lanfear, 2012). For *rbcL,* two subsets were created: one with codon positions 1 and 2, and one with codon position 3. Partitioned trees were the same as their nonpartitioned counterparts. 500 tree searches were performed, and branch support was again assessed via bootstrap, with 1000 (rapid) replicates.

2.4.3 Morphological studies

LM micrographs of the living isolated single-cells were taken at the time of isolation, at 300X. No differences could be found between the chloroplast structures of the taxa (Appendix A). Valve length, width, and shape were recorded for each specimen. Following isolations, the sediment samples were digested with boiling H_2SO_4:HNO_3 for twenty minutes then washed through five dilutions in order to remove all residual acid. Drops of digested samples were dried on aluminium foil pieces, which were mounted on pin stubs using double-sided carbon tabs. These stubs were covered in ~500 Å Au-Pd, examined, and photographed using an APREO FEI SEM. The following data was recorded from the micrographs: (1) valve length and (2) width, (3) stria density, (4) helictoglossae length and (5) shape (Fig. 2.14, red arrow), (6) raphe ribs

maximal width and (7) shape at apex, (8) degree of concavity of raphe ribs around the central nodule (Fig. 2.8, red arrow), (9) degree of separation of the central nodule (Fig. 2.7, red arrow), (10) minimal and (11) maximal width of the central nodule.

Single-cells and SEM micrographs were linked together based on information gathered from the LM pictures taken at isolation. For a match to be made between LM and SEM micrographs, the featured specimens had to share a set of morphological characteristics: valve size, valve shape, and other notable characteristics (central nodule or apex features) when they could be observed on the LM micrographs. Valve size was allowed to differ by ± 1 standard deviation (SD) of the appropriate taxa. Many samples contained only one species of *Frustulia*, and they were matched first. Then, those matched species were eliminated from other samples, and proceeding by deduction, every valve was matched.

Valve shapes were extracted from SEM images using DiaOutline (Wishkerman & Hamilton, 2018), and a PCA analysis was performed in R 3.5.0 (R Core Team, 2018) with the following packages installed: MASS (Venables & Ripley, 2002), ggplot2 (Wickham, 2016), GGally (Schloerke et al. 2018), doBy (Højsgaard & Halekoh, 2018), data.table (Dowle & Srinivasan, 2018), plyr (Wickham, 2011), gridExtra (Auguie, 2017) and Momocs (Bonhomme et al. 2014). Valve shape was also extracted from micrographs of type specimens for *F. saxonica* (Lange-Bertalot, 2001), *F. krammeri* (Lange-Bertalot, 2001), and *F. bahlsii* (Edlund & Brant, 1997), except for *F. crassinerva* (Brébisson ex Smith, 1853) in which a type line drawing was used.

A non-metric multi-dimensional scaling (NMDS) analysis was performed using Past3 (Hammer et al. 2001) in order to examine if our molecular clades would also group together based on morphological characters alone. All of the previously described morphological

characters recorded from micrographs were used in this analysis, with the exception of valve shape. The Gower similarity index was used, since this dataset contains both continuous and non-continuous characters. Three dimensions were used to minimize stress.

2.5 Results

2.5.1 Description of clades in this study

Frustulia bahlsii Edlund and Brant 1997 Figs 2.1, 2.7, 2.13

Frustules flat with shallow mantle, rhomboid, and slightly constricted near apices. Valves 148 –201.5 µm in length, 27.5 – 34 µm in width with a mean length/width ratio of 5.6 (±0.40) (Fig. 2.1). Striae parallel throughout, density 20 – 26 in 10 µm with mean 24.5 (±1.2). Areolae open externally as small circular to elongated slits-like pores aligned with valve apical axis. Internally, areolae much larger, circular and covered with an elevated hymen. Internal raphe ribs extend from apex to apex. Constriction of raphe ribs around central nodule variable – structure convex in some valves, and strongly concave in others. Most valves show separation of raphe ribs at central nodule (Fig. 2.7). At apices, raphe ribs extend completely around helictoglossae, not an open porte-crayon. Helictoglossae long and straight with a mean length of 3.4 (±0.3) µm (Fig. 2.13). At central nodule, mean separation of proximal raphe ends at largest point 5.4 (±0.5) µm, and smallest point 4.7 (±0.1) µm. At apex, mean width of porte-crayon at largest point 4.16 (±0.3) µm. Externally, raphe straight with T-shaped proximal and distal fissures. Internally, raphe simple and straight, proximal ends forming a figure-eight shape.

Frustulia krammeri Lange-Bertalot & Metzeltin 1998 Figs 2.2, 2.8, 2.14

Frustules flat with shallow mantle, rhomboid, with narrowly rounded, not constricted apices. Valves 96.0 –125.5 µm in length, 18.0 – 24.0 µm in width with a mean length/width ratio of 4.9 (±0.4) (Fig. 2.2). Striae parallel throughout, 27 – 31 in 10 µm with mean 27.2 (±1.3). Areolae open externally through small circular to slits-like pores aligned with valve apical axis. Internally, areolae larger, circular and covered by an elevated hymen. Internal raphe ribs extend from apex to apex. At apices, raphe ribs terminate halfway to three-quarters along helictoglossae and fuse with it. Terminal helictoglossae long and straight (Fig. 2.14). Constriction of the raphe ribs around central nodule pronounced in most valves. Separation of raphe ribs at heart of central nodule variable – distinct in some valves and less in others (Fig. 2.8). Mean length of the helictoglossae is 2.9 (±0.4) µm. At central nodule, mean separation of proximal raphe ends at largest point 4.0. (±0.3) µm, and smallest point 3.1 (±0.3) µm. At the apex, the mean width of the raphe at its largest point is 3.1 (±0.4) µm.

Frustulia gibsonea Bouchard, Hamilton, J.R.Starr & Savoie 2020 Figs 2.3, 2.10, 2.16, 2.19–2.23

Frustules flat with shallow mantle, rhomboid, and narrowly rounded at apices. Valves 66.8 – 136.0 µm in length, 15.7 – 22.9 µm in width with a mean length/width ratio of 5.6 (±0.4) (Fig. 2.3). Striae parallel throughout, 26 – 32 in 10 µm with mean 28 (±1.6). Areolae open externally with circular to elongated slits-like pores aligned with valve apical axis. Internally, areolae larger, circular and covered with an elevated hymen.

Internal raphe ribs extend from apex to apex. At central nodule, raphe ribs slightly to strongly concave. All valves show separation of the raphe ribs and heart of central nodule (Fig. 2.10). At apices, in most valves, raphe ribs extend halfway up and fuse with helictoglossae. Terminal helictoglossae long and straight with mean length 2.6 (±0.6) (Fig. 2.16). Mean separation of terminal raphe ends at largest point 3.0 (±0.7) μm, and smallest point, 2.4 (±0.6) μm. At apex, the mean width of the raphe at its largest point is 2.0 (±0.4) μm.

Differential diagnosis:

This species has distinctive features that are present in all valves: the raphes are angled (Fig. 2.10), and there are no areolae above the helictoglossae (Fig. 2.16). The areolae on each side of the helictoglossae have a linear shape. The raphe ribs of *F. gibsonea* are narrower around the central nodule (2.4 – 3.0 μm) and the helictoglossae (2.0 μm) than those of *F. krammeri* (3.1 – 4.0; 3.1) μm. Overall valve shape was not a significant character in the identification of this species.

Etymology: The species name ***gibsonea*** is derived from the location where most of its specimens were found: Gibson Lake, Ontario.

Type slide & material: CANA 109952, in the Canadian Museum of Nature, Ottawa, Canada, collected by K. E. Lefebvre, July 5th 2015. Holotype specimen (Fig. 2.20).

Isotype: Slide ANSP GC65332, Academy of Natural Sciences of Drexel University, Philadelphia USA.

Genotype: Specimen ID 34D8, isolated from the holotype material, is the genetic type for *F. gibsonea*. GenBank accession numbers MH789538 for *rbcL*; MH789489 for 18S.

Type habitat: Gibson Lake, ON, (lat. 46.2464, long. -78.1798). Freshwater. Slightly acidic (pH 6.4) and specific conductance of 14.0 µS cm^{-1}

Distribution: This species was also observed in eastern Canada from Nova Scotia and Newfoundland.

Comments:

Frustulia gibsonea is identical to a group of specimens that were observed in Bay Tree Lake, North Carolina, and then described as *F.* cf. *krammeri* based on morphology (Siver & Hamilton, 2011). The use of molecular characters in this study demonstrates that *F. gibsonea* is a distinct species (Tables 2.4 – 2.6, Figs. 2.33 – 2.34). This taxon may have a wider distribution than currently thought, as there is a population in Europe, as evidenced by GenBank specimen KU663126.

A full protologue is given here because this species was first described in the publication upon which this chapter is based (Bouchard et al. 2019).

Frustulia crassinervia (Brébisson) Lange-Bertalot and Krammer 1996 Figs 2.6, 2.9, 2.15

Frustules flat with shallow mantle, rhomboid with undulate margins, and slightly constricted near apices (Fig. 2.6). Valves 41.4 – 61.8 µm in length, 9.5 – 15.9 µm in width with a mean length/width ratio of 4.3 (±0.3) (Fig. 2.6). Striae parallel throughout, 30 – 34 in 10 µm with mean 32.9 (±1.4). Areolae open externally as small circular to elongated

slits-like pores aligned with valve apical axis. Internally, areolae much larger, circular and covered with an elevated hymen. Internal raphe ribs extend from apex to apex. Raphe ribs at central nodule straight or slightly constricted. Most valves show distinct separation of raphe ribs and heart of central nodule (Fig. 2.9). At apices, in most valves, raphe ribs end before reaching helictoglossae, although ribs may reach and weakly fuse with helictoglossae. Terminal helictoglossae slightly rounded to rounded (Fig. 2.15). Helictoglossae long and straight with a mean length of 1.2 (±0.2) µm. At central nodule, mean separation of proximal raphe ends at largest point 2.0 (±0.4) µm, and smallest point 1.7 (±0.3) µm. At apex, mean width of porte-crayon at largest point 1.2 (±0.2) µm.

Frustulia sp. (a) [*F.* cf. *saxonica*] Figs 2.4, 2.11, 2.17

Frustules flat with shallow mantle, rhomboid with undulate margins, and slightly constricted near apices (Fig. 2.4). Valves 54.8 – 66.7 µm in length, 9.7 – 14.0 µm in width with mean length/width ratio of 5.3 (±0.4). Striae parallel throughout, always 34 in 10 µm. Areolae open externally as small circular to elongated slits-like pores aligned with valve apical axis. Constriction of raphe ribs around central nodule pronounced. Separation of raphe ribs and heart of central nodule variable – distinct in some valves (Fig. 2.11). At apices, most raphe ribs end before reaching the helictoglossae, but some extend completely up the structure. Helictoglossae are long and straight, mean width 1.8 (±0.2) µm (Fig. 2.17). At central nodule, mean separation of proximal raphe ends at largest point 2.1 (±0.3) µm, and smallest point 1.7 (±0.4) µm. At apex, mean width of porte-crayon at largest point 1.5 (±0.2) µm.

Frustulia sp. (b) [*F.* cf. *saxonica*] Figs 2.5, 2.12, 2.18

Valves range in size from a length of 51 – 58.1 µm, a width of 13.8 – 15.2 µm, and a mean length/width ratio of 3.7 (±0.16). Valves rounded and very constricted near apices (Fig. 2.5). Stria density 28 – 30 in 10 µm with a mean 29.8 (±0.7). At the central nodule, raphe ribs straight or slightly constricted. Raphe ribs not distinctly separated from heart of central nodule (Fig. 2.12). At apices, raphe ribs cut off, often asymmetrically, before reaching helictoglossae. Helictoglossae long and straight (Fig. 2.18). Mean length of the helictoglossae 1.9 (±0.1) µm. At central nodule, mean separation of proximal raphe ends at largest point 2.3 (±0.1) µm, and smallest point 2.2 (±0.2) µm. At apex, mean width of raphe ribs at largest point 1.5 (±0.1) µm.

Frustulia sp. (c) Appendix A: A.28 – A.30

We were unable to match SEM micrographs to this clade, as specimens were only found in one location, which contained multiple *Frustulia* morphotypes corresponding to the description obtained from the live cell micrographs. For this reason, the following description is based on the live cell micrographs alone. Low quality of these micrographs, due to cell movement, did not allow us to gather much informative data.

Frustules flat, rhomboid with narrowly rounded (Appendix A: A.28–A.30).Valves range in size from 80.0 – 86.5 µm in length, 16.5 – 18.5 µm in width with mean length/width ratio of 4.8 (±0.6).

2.5.2 Morphology analysis

The results of the NMDS, performed as described above and based on 11 traits, can be found in Fig. 2.24. This analysis was performed to examine if our taxa would group together based on their morphological characters. Species were separated, but overlap was seen between *F. gibsonea* and *F. krammeri*, and the unknowns, *F.* sp. (a) [*F.* cf. *saxonica*] and *F.* sp. (b) [*F.* cf. *saxonica*]. In general variability within each taxon was consistent, with a smaller variance cluster in *F.* sp. (b) [*F.* cf. *saxonica*]. Three taxa (*F. bahlsii*, *F. gibsonea* and *F.* sp. (b) [*F.* cf. *saxonica*]) each had one shape outlier relative to the other specimens.

The results of the PCA for valve shapes can be found in Fig. 2.25. In total, 77.29% of the variation between specimens was explained by the first two PC axes – 67.50% was explained by the first axis and 9.79% by the second axis. The first PC axis represents valve width. There was overlap in shape forms for 4 of the 7 taxa. *Frustulia saxonica* (type), *F. krammerii*, *F.* sp. (a) [*F.* cf. *saxonica*] and *F. gibsonea* showed no clear discrimination between shape forms, while *F. bahlsii* and *F. crassinervia* were more distinct with some minor overlap. *Frustulia* sp. (b) [*F.* cf. *saxonica*] was clearly separated. The largest form differences were observed in *F. gibsonea*, followed by *F. bahlsii*, *F. krammeri* and *F. crassinervia*.

2.5.3 Sequence analysis

RbcL data revealed the most variation with a median divergence between taxa of 5.8%. Within the North American taxa, base pair (bp) divergence between taxa ranged from 0.0 – 5.1% for *rbcL* and 18S (Tables 2.4, 2.5). The *rbcL* results for North American taxa showed that

Frustulia gibsonea was most similar to *Frustulia* sp. (c) (15 – 18 bp divergence (1.3 – 1.6%), Table 4). *Frustulia krammeri* and *F. crassinervia* were also relatively close in *rbcL*, with 20 bp different between them (1.7%). Comparisons outside North America revealed that *F. gibsonea* was similar to *F. gondwana* (9 bp divergence, 1.4%) and an unknown specimen from Europe ("*F.* sp." 11 bp divergence, 1.8%). *Frustulia* cf. *krammeri* identified from Europe was genetically identical to *F. gibsonea* (0 bp divergence, 0%). Likewise *F.* sp. (a) [*F.* cf. *saxonica*] (a) was similar to *F. crassinervia x saxonica* from Europe (1 – 2 bp divergence, 0.2 – 0.3%) and the European taxa *F. saxonica* and *F. crassinervia* were comparable to *F. krammeri* (12 bp divergence, 1.8 – 1.9%). *Frustulia vulgaris* from Europe was distinctly separated from the other *Frustulia* species (7.7 – 9.1% divergence).

The 18S results revealed that divergence between taxa was less than for *rbcL*, ranging from 0.1 – 3.4% with a median divergence among taxa of 1.2%. *Frustulia gibsonea* was similar to *Frustulia* sp. (c) (1 – 3 bp divergence, 0.1 – 0.4%) from North America and *F. erifuga* (2 – 4 bp divergence, 0.3 – 0.5%) from Europe (Table 2.5). *Frustulia gibsonea* was also similar to *F. crassinervia* (0.7 – 0.9% divergence) and *Frustulia krammeri* (0.9 – 1.2% divergence) from North America. *Frustulia saxonica* from Europe was also similar to *F. krammeri* (0.1% divergence) from North America. *Frustulia vulgaris* was the most diverse (>2.8% divergence) relative to the other *Frustulia* taxa.

Intraspecific divergence was low within North American taxa (Table 2.6). A range of 0 – 2 bp differences (0.000 – 0.174%) for *rbcL* and 0 – 8 (0.000 – 1.075%) for 18S was recorded. *Frustulia gibsonea* (n = 16) showed zero divergence for the *rbcL* sequence from eight sites across eastern Canada. *Frustulia bahlsii* (n = 23) and *F. krammeri* (n = 10) had 0 – 2 bp differences also from twelve and five sites, respectively, across eastern Canada.

2.5.4 Phylogenetic trees

In the Maximum Likelihood (ML) analyses of *rbcL* sequence data, the new species *F. gibsonea* formed a monophyletic clade with high support with specimens identified as *Frustulia* sp. (c), *F. gondwana*, and *F. erifuga* (Figure 2.33). *Frustulia septentrionalis* and *F. gaertnerae* were associated with the North American *F. crassinervia* and the European *F. saxonica*. *Frustulia bahlsii* was associated with the New Zealand taxon *F. aotearoa*, in both *rbcL* and 18S (Figs 2.33, 2.34).

Maximum likelihood analyses of 18S sequence data (Fig. 2.34) showed the same general tree structure but with poor support. *Frustulia gibsonea* was related to *F.* sp. (c) and *F. erifuga. Frustulia vulgaris* was associated with *Berkeleya fennica* and *B. rutilans*, with no support. The sequence data for *F. vulgaris* was downloaded from GenBank and this taxon needs further study. The *rbcL* tree showed weak to no support for the separation of *F. krammeri* and *F. saxonica* (Fig. 34), including selected specimens identified as *F. crassinervia*. In contrast, the 18S tree showed a weak separation of *F. krammeri* and *F. saxonica* with some support (BS, ML/MP: 83/*).

2.6 Discussion

Eight taxa from North America are recognized in this study, along with shape images for the syntype specimens of *F. saxonica*. One species is identified as new and named *F. gibsonea.* Specimens of *F. gibsonea* were identical in *rbcL* sequence (n=16) and distinct from other species of *Frustulia* (nearest neighbour *F.* sp. (c), 1.3% difference in *rbcL*); however, this species could not be distinguished using 18S. Morphologically, specimens of *F. gibsonea* were

indistinguishable from one another. Compared to other taxa, *F. gibsonea* can be distinguished by genetics and fine morphological characters, but not by shape.

Frustulia gibsonea is morphologically identical to specimens observed in Bay Tree Lake, North Carolina, which were described as *F.* cf. *krammeri* based on morphology (Siver & Hamilton, 2011). A European specimen identified as *F. krammeri* (KU663126) also grouped with *F. gibsonea* in the *rbcL* tree. *F. gibsonea* and *F. krammeri* are morphologically similar, showing overlap using both traditional measurements and shape analysis (Figs 2.31, 2.32). Therefore taxa differences are difficult to identify using subjective identifications alone. In this study, *F. gibsonea* and *F. krammeri* are separated using molecular data and a particular set of morphological measurements that are often overlooked, namely, raphe angle, raphe ribs width, and striae organisation at the apex. This type of detailed morphology tends to be neglected; although it is required for fine-scale taxonomy. In cryptic marine *Pseudo-nitzschia* taxa, genetic differences in sympatric populations were related to fine structural morphological differences (Amato et al. 2007, Lundholm et al. 2012). In contrast, genetically distinct strains (taxa?) were identified in the freshwater *Cyclotella meneghiniana* with no clear distinctions in morphology (Beszteri et al. 2005). One publication on the freshwater benthic genus *Sellaphora* (Vanormelingen et al. 2013) found genetic morphospecies with no simple metrics that could identify the taxa; this is likely a close analogue to what we are observing in the genus *Frustulia*. This suggests that there could be many other unidentified cryptic taxa within the genus. Further in-depth exploration of morphological and molecular characters of *Frustulia* species will surely reveal this cryptic diversity.

Three *Frustulia* lineages from North America were unidentified. The lineages, *F.* sp. (a) [*F.* cf. *saxonica*] and *F.* sp. (b) [*F.* cf. *saxonica*], overlapped using traditional morphology

measurements, but were effectively separated based on valve shape (Fig. 32). *Frustulia* sp. (b) was distant from the type material for *F. saxonica* sensu lato in the shape analysis and was genetically separated from the group containing *F. saxonica*, *F. crassinervia*, and *F. krammeri* with moderate support (BS, ML/MP: 61/72) in the *rbcL* tree. This suggests *F.* sp. (b) [*F.* cf. *saxonica*] is not closely related to *F. saxonica*, despite being morphologically similar to the species using traditional characters such as valve size, stria density, and helictoglossae and raphe features. *Frustulia.* sp. (b) is not described as a new species as only three specimens were recovered, all of which were isolated from the same location. Future work with increased sampling will be required to determine where this potential taxon belongs within the genus. *Frustulia* sp. (a), however, forms a monophyletic group with European sequences identified as *F. crassinervia x saxonica* (treated as a complex) in the *rbcL* tree. In the shape analysis, it overlaps with other taxa, while the morphological characters were able to clearly separate this taxon. In contrast to *F.* sp. (b) [*F.* cf. *saxonica*], morphological characters for *Frustulia* sp. (a) [*F.* cf. *saxonica*] could separate this valve form, but shape and genetics could not. Thus the usefulness and necessity of different metrics in distinguishing between taxa is evident, with conserved genes showing some taxonomic relevance within the genus *Frustulia*.

Preliminary observations suggest that *F. cassieae* (New Zealand) (like *F. vulgaris*) is genetically unique within the genus, while *F. gondwana* (New Zealand) is associated with *F. gibsonea* (North America), *F. erifuga* (Europe), and *F*. sp. (c) (North America). Likewise *F. maoriana* (New Zealand) is linked to *F*. sp B (*F*. cf. *saxonica*) from North America. These findings are in agreement with the DNA-barcoding survey of Urbánková & Veselá (2013) who examined *rbcL*, SSU and LSU gene sequences. At present, there is no genetic biogeographic evidence for radiations and species flocks within the genus *Frustulia*.

Frustulia bahlsii, F. krammeri, and *F. crassinervia* were observed from North America with sizeable populations. These species are well separated using traditional measurements (Fig. 2.31). In addition, *F. bahlsii* was strongly supported in the *rbcL*, and 18S trees (BS, ML/MP: *rbcL*, 18S 100/100). Morphological and genetic characters were able to separate *F. bahlsii* while shape as measured by e-Fourier analysis was not a good identification character for this taxon. These results conform to the original species description for this taxon (Edlund & Brant, 1997). The North American taxon *F. krammeri* was strongly supported (BS, ML/MP: 99/99) while *F. crassinervia* was weakly supported (BS, ML/MP: 68/60) in the *rbcL* tree. Weak support was also observed for earlier nodes separating sequences for *F. saxonica*, *F.* sp. (a) [*F.* cf. *saxonica*], *F. crassinervia*, and *F. krammeri* (below 70% BS) in both trees (Figs 2.33, 2.34). As evidenced by the sequence divergence data, these four species are related (<2.6% divergence for *rbcL*; <0.9% for 18S) however, morphological characters were able to distinguish between taxa (Fig. 2.31). This suggests that less conserved molecular markers are possibly needed to resolve these taxonomic complexes or a more extensive genetic comparison of populations is required. Alternatively, this morphological variation of the studied specimens may only represent phenotypic expression for one species. It is also possible that this study is missing key taxa/specimens that could resolve the relationships.

Although 18S results might be too conserved for fine grained species-level phylogenies of diatoms, it has been found to perform better at higher taxonomic levels (Guo et al. 2015). This is observed in our tree. Although the 18S tree shows low resolution of closely related species it places *F. vulgaris* as monophyletic with the rest of the *Frustulia* species, with moderate support (BS, ML/MP: 68/77). In both the *rbcL* and 18S trees, our outgroup, *A. pellucida,* is nested within the *Frustulia* genus. Nakov et al. (2018) obtained the same result in their phylogeny. This

suggests that *A. pellucida* may be closer related to *Frustulia* than previously thought. It would be interesting to explore this relationship with increased taxon sampling for the *Amphipleura* genus and a better representation of *F. vulgaris*.

In this study, *F. crassinervia* sequences from North America were 2.1% different from European *F. crassinervia* specimens in *rbcL* (Table 2.4). Based on the available morphological data, there appears to be little difference between European and NA populations, but further sampling is required to validate these preliminary results. There are still taxonomic problems with identifications within the *F. saxonica* group, but there is now more evidence about taxonomic associations among the taxa.

2.6.1 Morphology vs. molecular data

The NMDS results show good separation between the older established taxa, suggesting that traditional measurements are powerful in distinguishing species. However, the combination of molecular and morphological data can help in deciphering potential cryptic taxa. Specimens now identified as *F. gibsonea* were identified as *F.* cf. *krammeri* when using morphology alone, while this study shows that *F. krammeri* is a separate and distinct species.

Previous studies have identified a suite of morphological characters that are most informative for distinguishing *Frustulia* species based on morphology. Urbankova et al. (2016) identified a combination of four characters – valve length, width, stria angle (towards the valve ends) and scores representing the first axis of a PCA performed on valve shape – as the best characters. Siver & Baskette (2004) used valve length, width, shape, and both striae and areola density to distinguish between North American species. Helictoglossae shape was also used,

particularly to discriminate between *F. saxonica* and *F. crassinervia.* This is consistent with the features we found to be most informative: features linked to the overall valve shape (length, width, stria density) and to the helictoglossae (length, shape) were more informative than features linked to the central nodule (concavity, separation). Some features (minimum and maximum width of the central nodule and of the ribs at the helictoglossae) were particularly useful in distinguishing between morphologically similar *F. gibsonea* and *F. krammeri*, but not useful for species identification in general.

The species *F. krammeri*, *F. saxonica*, and *F. crassinervia* could not be clearly separated using molecular analyses. In contrast *F. gibsonea* and *Frustulia* sp. (b) were separated by molecular studies, but not easily separated morphologically. Future research, with less conserved markers than those used here, will likely be needed to resolve these genetically based taxonomic relationships. Currently, *rbcL* is one of the less conserved genes commonly used in diatom phylogenies. It also has the advantages of high amplification success, which is especially critical with the single-cell technique. Researching new molecular markers that are less conserved than *rbcL*, and present a similar ease of amplification, will be necessary to address these types of species complexes.

Table 2.1. Location and collection date for sediment samples used in this study. Specimens are stored in the National Phycology Collection of Canada (Herbarium Code: CANA).

Genus	CANA Accession Number	Number of Single-Cells Isolated	Date Collected	Latitude	Longitude	Country	Province/State
Frustulia	93279	1	2014-05-10	44.44806	-74.77439	USA	New York
	93281	1	2014-05-11	44.00715	-74.49155	USA	New York
	93284	3	2014-05-10	43.81723	-74.88400	USA	New York
	93415	6	2014-06-15	46.3469	-78.1797	Canada	Ontario
	108073	4	2014-10-25	45.1363	-78.1841	Canada	Ontario
	108124	3	2014-10-16	44.4487	-74.7739	USA	New York
	108128	3	2014-10-16	43.9092	-74.4397	USA	New York
	108605	2	2015-05-21	47.2657	-52.8355	Canada	Newfoundland
	108610	4	2015-05-21	46.9585	-52.9572	Canada	Newfoundland
	108628	5	2015-05-30	44.7101	-63.8967	Canada	Nova Scotia
	108632	2	2015-05-30	44.6672	-63.9955	Canada	Nova Scotia
	108634	2	2015-05-30	44.6437	-64.0652	Canada	Nova Scotia
	108639	1	2015-05-31	44.6885	-63.6637	Canada	Nova Scotia

Table 2.1 Continued

Genus	CANA Accession Number	Number of Single-Cells Isolated	Date Collected	Latitude	Longitude	Country	Province/State
	108642	1	2015-05-31	44.7795	-63.6248	Canada	Nova Scotia
	109952	7	2015-07-05	46.1447	-78.1048	Canada	Ontario
	109953	2	2015-07-05	46.1447	-78.1048	Canada	Ontario
	115551	1	2015-08-20	46.4254	-81.0403	Canada	Ontario
	126140	10	2016-10-08	46.2464	-78.1799	Canada	Ontario
	126320	1	2017-05-06	46.2464	-78.1798	Canada	Ontario
	126321	1	2017-05-06	46.2464	-78.1798	Canada	Ontario
	126324	1	2017-05-06	46.2464	-78.1798	Canada	Ontario
Amphipleura	125993	2	2016-09-07	62.54525	-114.40205	Canada	Northwest Territories
	125958	1	2016-09-05	62.54877	-113.93135	Canada	Northwest Territories

Table 2.2. GenBank accession numbers for specimens used in this study. ON Ca: Ontario, Canada; NS Ca: Nova Scotia, Canada; NWT CA: Northwest Territories, Canada; NZ: New Zealand; NY USA: New York State, United States of America.

Taxon	Identifier	*rbcL*	18S	Source	Location
Amphipleura pellucida	ECT3568	KC309549	KC309477	Ashworth et al. 2013	USA
Craticula cuspidata	UTEX FD35	HQ912445	HQ912581	Theriot et al. 2010	USA
Craticula sp. [cf. *cuspidata*]	6C2	KM999071	KM999001	Hamilton et al. 2015	Ontario, Ca
Frustulia vulgaris	MIC10-15	HF562253		Urbánková & Veselá, 2013	Marion Is.
Frustulia vulgaris	AT-108Gel03		AM502038	Bruder & Medlin, 2007	Germany
Frustulia crassinervia x *saxonica*	WA3	KU663125		Urbánková et al. 2016	Europe
Frustulia crassinervia x *saxonica*	NZ136	HF562241		Urbánková & Veselá, 2013	South Is., NZ
Frustulia crassinervia x *saxonica*	WA2	KU663124		Urbánková et al. 2016	Europe
Frustulia crassinervia x *saxonica*	F45	HF562237	HF562291	Urbánková & Veselá, 2013	France
Frustulia crassinervia x *saxonica*	21-9B	HF562239		Urbánková & Veselá, 2013	Czech Rep.
Frustulia crassinervia x *saxonica*	20-2F	HF562238		Urbánková & Veselá, 2013	Denmark
Frustulia crassinervia x *saxonica*	19-2B	HF562240	HF562294	Urbánková & Veselá, 2013	Denmark
Frustulia saxonica	NZ39	HF562243		Urbánková & Veselá, 2013	South Is., NZ
Frustulia saxonica	NZ184	HF562254		Urbánková & Veselá, 2013	South Is., NZ
Frustulia saxonica	NZ34	HF562242		Urbánková & Veselá, 2013	South Is., NZ
Frustulia crassinervia	F237	HF562235		Urbánková & Veselá, 2013	Ireland
Frustulia krammeri	JC8	KU663126		Urbánková et al. 2016	Europe

Table 2.2 Continued

Taxon	Identifier	*rbcL*	18S	Source	Location
Frustulia crassinervia x *saxonica*	F259		HF562292	Urbánková & Veselá, 2013	France
Frustulia saxonica	NZ33		HF562293	Urbánková & Veselá, 2013	South Is., NZ
Frustulia crassinervia	F350		HF562290	Urbánková & Veselá, 2013	Ireland
Amphipleura pellucida	59B3	MH789548		This study	NWT, Ca
Amphipleura pellucida	59B6	MH789549		This study	NWT, Ca
Amphipleura pellucida	60E1	MH789550		This study	NWT, Ca
Frustulia bahlsii	13C6	MH789491	MH789468	This study	NY, USA
Frustulia bahlsii	15J5		MH789469	This study	NY, USA
Frustulia bahlsii	15K2	MH789494	MH789470	This study	NY, USA
Frustulia bahlsii	15K5	MH789499	MH789471	This study	NY, USA
Frustulia bahlsii	66E3	MH789546	MH789472	This study	ON, Ca
Frustulia bahlsii	25A6	MH789510	MH789473	This study	NS, Ca
Frustulia bahlsii	35F7	MH789505	MH789474	This study	ON, Ca
Frustulia bahlsii	35B5	MH789511	MH789475	This study	ON, Ca
Frustulia bahlsii	60F3	MH789496	MH789476	This study	ON, Ca
Frustulia bahlsii	13A5	MH789490		This study	NY, USA
Frustulia bahlsii	25A7	MH789492		This study	NS, Ca
Frustulia bahlsii	14C7	MH789493		This study	ON, Ca

Table 2.2 Continued

Taxon	Identifier	*rbcL*	18S	Source	Location
Frustulia bahlsii	25A5	MH789495		This study	NS, Ca
Frustulia bahlsii	25A1	MH789497		This study	NS, Ca
Frustulia bahlsii	25A8	MH789498		This study	NS, Ca
Frustulia bahlsii	60F7	MH789500		This study	ON, ca
Frustulia bahlsii	23C2	MH789501		This study	NF, Ca
Frustulia bahlsii	25B6	MH789502		This study	NS, Ca
Frustulia bahlsii	13C5	MH789503		This study	NY, USA
Frustulia bahlsii	23C4	MH789504		This study	NF, Ca
Frustulia bahlsii	23C8	MH789506		This study	NF, Ca
Frustulia bahlsii	23C6	MH789507		This study	NF, Ca
Frustulia bahlsii	23C5	MH789508		This study	NF, Ca
Frustulia bahlsii	35C2	MH789509		This study	ON, Ca
Frustulia crassinervia	15K1	MH789513	MH789484	This study	NY, USA
Frustulia crassinervia	13B5	MH789512		This study	NY, USA
Frustulia gibsonea (b)	25F6		MH789479	This study	NS, Ca
Frustulia gibsonea (a)	14E2	MH789529	MH789486	This study	ON, Ca
Frustulia gibsonea (a)	34D8	MH789538	MH789489	This study	ON, Ca
Frustulia gibsonea (a)	14D3	MH789528	MH789477	This study	ON, Ca

Table 2.2 Continued

Taxon	Identifier	*rbcL*	18S	Source	Location
Frustulia gibsonea (a)	14D7	MH789537	MH789478	This study	ON, Ca
Frustulia gibsonea (a)	35A4	MH789536	MH789481	This study	ON, Ca
Frustulia gibsonea (a)	14D6	MH789530		This study	ON, Ca
Frustulia gibsonea (a)	19A7	MH789531		This study	ON, Ca
Frustulia gibsonea (a)	19F2	MH789532		This study	ON, Ca
Frustulia gibsonea (a)	26C4	MH789533		This study	NS, Ca
Frustulia gibsonea (a)	23C3	MH789534		This study	NF, Ca
Frustulia gibsonea (a)	19E7	MH789535		This study	ON, Ca
Frustulia gibsonea (a)	19E2	MH789539		This study	ON, Ca
Frustulia gibsonea (a)	26E4	MH789540		This study	NS, Ca
Frustulia gibsonea (a)	37A1	MH789541		This study	ON, Ca
Frustulia gibsonea (a)	25F7	MH789542		This study	NS, Ca
Frustulia gibsonea (a)	34D5	MH789543		This study	ON, Ca
Frustulia gibsonea (b)	66E1	MH789547	MH789482	This study	ON, Ca
Frustulia gibsonea (b)	14D4	MH789544	MH789483	This study	ON, Ca
Frustulia krammeri	34C4	MH789516	MH789480	This study	ON, Ca
Frustulia krammeri	61A1	MH789515	MH789487	This study	ON, Ca

Table 2.2 Continued

Taxon	Identifier	*rbcL*	18S	Source	Location
Frustulia krammeri	66D7	MH789545	MH789488	This study	ON, Ca
Frustulia krammeri	25B5	MH789514		This study	NS, Ca
Frustulia krammeri	60F4	MH789517		This study	ON, Ca
Frustulia krammeri	60F8	MH789518		This study	ON, Ca
Frustulia krammeri	35F6	MH789519		This study	ON, Ca
Frustulia krammeri	60F2	MH789520		This study	ON, Ca
Frustulia krammeri	30B5	MH789521		This study	ON, Ca
Frustulia krammeri	60E5	MH789522		This study	ON, Ca
Frustulia sp. (a)	61A4	MH789523		This study	ON, Ca
Frustulia sp. (a)	61A5	MH789524		This study	ON, Ca
Frustulia sp. (b)	60E7	MH789526	MH789485	This study	ON, Ca
Frustulia sp. (b)	14D5	MH789525		This study	ON, Ca
Frustulia sp. (b)	15I4	MH789527		This study	ON, Ca

Table 2.3. Primers used in the nested PCR amplifications.

Primer type	Name	Sequence (5'-3')	Primer direction	Source
External	rbcL66+	TTTAAGGAGAAATAAATGTCTCAATCTG	F	Alverson et al., 2007; Daugbjerg and Anderson, 1997
	rbcLdp7	AAASHDCCTTGTGTWAGTVTC	R	
	18SP2F	CTGGTTGATTCTGCCAGT	F	Ruck and Theriot, 2011
	18SP4R	TGATCCTTCYGCAGGTTCAC	R	
Internal	rbcL40+	GGACTCGAATVAAAAGTGAACG	F	Hamilton et al., 2015
	rbcL1444	GCGAAATCAGCTGTATCTGTWG	R	
	18S 434F	ACAATAAATAACAATGCCGGGC	F	This study
	18S1449R	GVRTRCATCAGTGTAGCGCG	R	

Table 2.4. Sequence divergence between species of *Frustulia* for the rbcL gene. Absolute differences are given above the diagonal, and percent divergence is given below the diagonal.

Source		This study						
	Species	*F. bahlsii*	*F. crassinervia*	*F. krammeri*	*F. gibsonea*	*F.* sp. (c)	*F.* cf. *saxonica* (a)	*F.* cf. *saxonica* (b)
This study	*F. bahlsii*	–	37	37	53	50 – 51	38	47
	F. crassinervia	3.2	–	20	55	54 – 55	26	39
	F. krammeri	3.2	1.7	–	57	56 – 57	18	31
	F. gibsonea	4.6	4.8	5.0	–	15 – 18	55	59
	F. sp. (c)	4.3 – 4.4	4.7 – 4.8	4.9 – 5.0	1.3 – 1.6	–	52 – 53	58 – 59
	F. cf. *saxonica* (a)	3.3	2.3	2.3	4.8	4.5 – 4.6	–	36
	F. cf. *saxonica* (b)	4.1	3.4	3.4	5.1	5.0 – 5.1	3.1	–
	F. erifuga	5.8 – 6.1	6.4 – 6.8	6.1 – 6.4	2.1 – 2.4	1.4 – 1.8	6.3	6.8 – 7.1
	F. gondwana	5.9	6.2	5.9	1.4	0.9 – 1.1	5.9	6.5
	F. maoriana	4.8 – 5.0	3.2 – 3.3	2.1 – 2.3	6.8	6.1 – 6.5	2.9 – 3.1	1.6 – 1.9
	F. crassinervia x saxonica	4.3 – 5.0	3.0 – 3.4	2.1 – 2.4	6.6 – 7.2	5.9 – 6.6	0.2 – 0.3	3.2 – 3.9
GenBank	*F. crassinervia*	3.3	2.1	1.8	6.7	6.2 – 6.4	2.6	3.0
	F. saxonica	4.5	2.9	1.9	6.6	6.1 – 6.3	2.6	3.5
	F. sp.	5.9	5.9	5.9	1.8	1.1	6.1	6.1
	F. aotearoa	3.1	5.0	4.5	6.8	6.2 – 6.3	5.0	6.0
	F. septentrionalis	3.0	1.8	1.5	6.4	5.8 – 6.1	2.3	2.7
	F. gaertnerae	3.6	2.9	2.6	6.8	6.1 – 6.2	2.3	3.8
	F. krammeri	6.4	6.4	6.8	0.0	1.8 – 1.9	6.9	7.2
	F. vulgaris	8.1	9.1	8.1	8.8	8.0 – 8.1	8.1	8.5
	F. cassieae	6.2	7.6	7.7	8.0	7.1 – 7.3	7.6	8.2

Table 2.4 Continued

Source		GenBank					
	Species	*F. erifuga*	*F. gondwana*	*F. maoriana*	*F. crassinervia x saxonica*	*F. crassinervia*	*F. saxonica*
This study	*F. bahlsii*	36 – 38	39	30 – 31	27 – 31	22	28
	F. crassinervia	40 – 42	41	20 – 22	19 – 21	14	18
	F. krammeri	38 – 40	39	13 – 14	13 – 15	12	12
	F. gibsonea	13 – 15	9	42	41 – 45	44	41
	F. sp. (c)	9 – 11	6 – 7	38 – 40	37 – 41	41 – 42	38 – 39
	F. cf. *saxonica* (a)	39	39	18 – 19	1 – 2	17	16
	F. cf. *saxonica* (b)	42 – 44	43	10 – 12	20 – 24	20	22
GenBank	*F. erifuga*	–	10	35 – 39	36 – 40	37 – 39	37 – 38
	F. gondwana	1.5	–	37	36 – 40	39	36
	F. maoriana	5.6 – 6.3	6.0	–	16 – 18	15 – 17	18 – 19
	F. crassinervia x saxonica	5.8 – 6.4	5.8 – 6.4	2.6 – 2.9	–	14 – 17	15 – 16
	F. crassinervia	5.6 – 5.9	5.9	2.3 – 2.6	2.1 – 2.6	–	13
	F. saxonica	5.9 – 6.1	5.8	2.9 – 3.0	2.4 – 2.6	2.1	–
	F. sp.	1.8 – 1.9	0.6	5.1 – 5.9	5.8 – 6.3	5.8	5.8
	F. aotearoa	6.2 – 6.3	6.0	4.9 – 5.0	4.7 – 5.0	4.5	5.3
	F. septentrionalis	5.3 – 5.6	5.6	2.0 – 2.3	1.8 – 2.3	0.3	1.7
	F. gaertnerae	5.8 – 6.1	6.1	2.6 – 3.0	2.0 – 2.4	1.4	2.7
	F. krammeri	2.1 – 2.3	1.4	6.3 – 6.8	6.6 – 7.1	6.6	6.6
	F. vulgaris	7.7 – 8.1	7.8	7.3	7.7 – 8.1	8.3	8.6
	F. cassieae	7.3 – 7.6	7.4	7.0 – 7.3	7.0 – 7.4	7.1	7.4

Table 2.4 Continued

Source		GenBank						
	Species	*F.* sp.	*F. aotearoa*	*F. septentrionalis*	*F. gaertnerae*	*F. krammeri*	*F. vulgaris*	*F. cassieae*
This study	*F. bahlsii*	37	19	20	24	40	51	41
	F. crassinervia	37	31	12	19	40	57	50
	F. krammeri	37	28	10	17	42	51	51
	F. gibsonea	11	42	42	45	0	55	53
	F. sp. (c)	7	38 – 39	38 – 40	40 – 41	11 – 12	50 – 51	47 – 48
	F. cf. *saxonica* (a)	38	31	15	15	43	51	50
	F. cf. *saxonica* (b)	38	37	18	25	45	53	54
	F. erifuga	11 – 12	38 – 39	35 – 37	38 – 40	13 – 14	48 – 51	48 – 50
	F. gondwana	4	37	37	40	9	49	49
	F. maoriana	32 – 37	30 – 31	13 – 15	17 – 20	39 – 42	46	46 – 48
	F. crassinervia x saxonica	36 – 39	29 – 31	12 – 15	13 – 16	41 – 44	48 – 51	46 – 49
GenBank	*F. crassinervia*	36	28	2	9	41	52	47
	F. saxonica	36	33	11	18	41	54	49
	F. sp.	–	39	34	37	11	48	47
	F. aotearoa	6.3	–	26	26	42	47	36
	F. septentrionalis	5.2	3.9	–	7	39	52	45
	F. gaertnerae	5.6	3.9	1.1	–	42	54	45
	F. krammeri	1.8	6.8	6.3	6.8	–	52	50
	F. vulgaris	8.0	7.5	8.3	8.6	8.3	–	53
	F. cassieae	7.1	5.5	6.8	6.8	7.6	8.0	–

Table 2.5. Sequence divergence between species of *Frustulia* for the 18S gene. Absolute differences are given above the diagonal, and percent divergence is given below the diagonal.

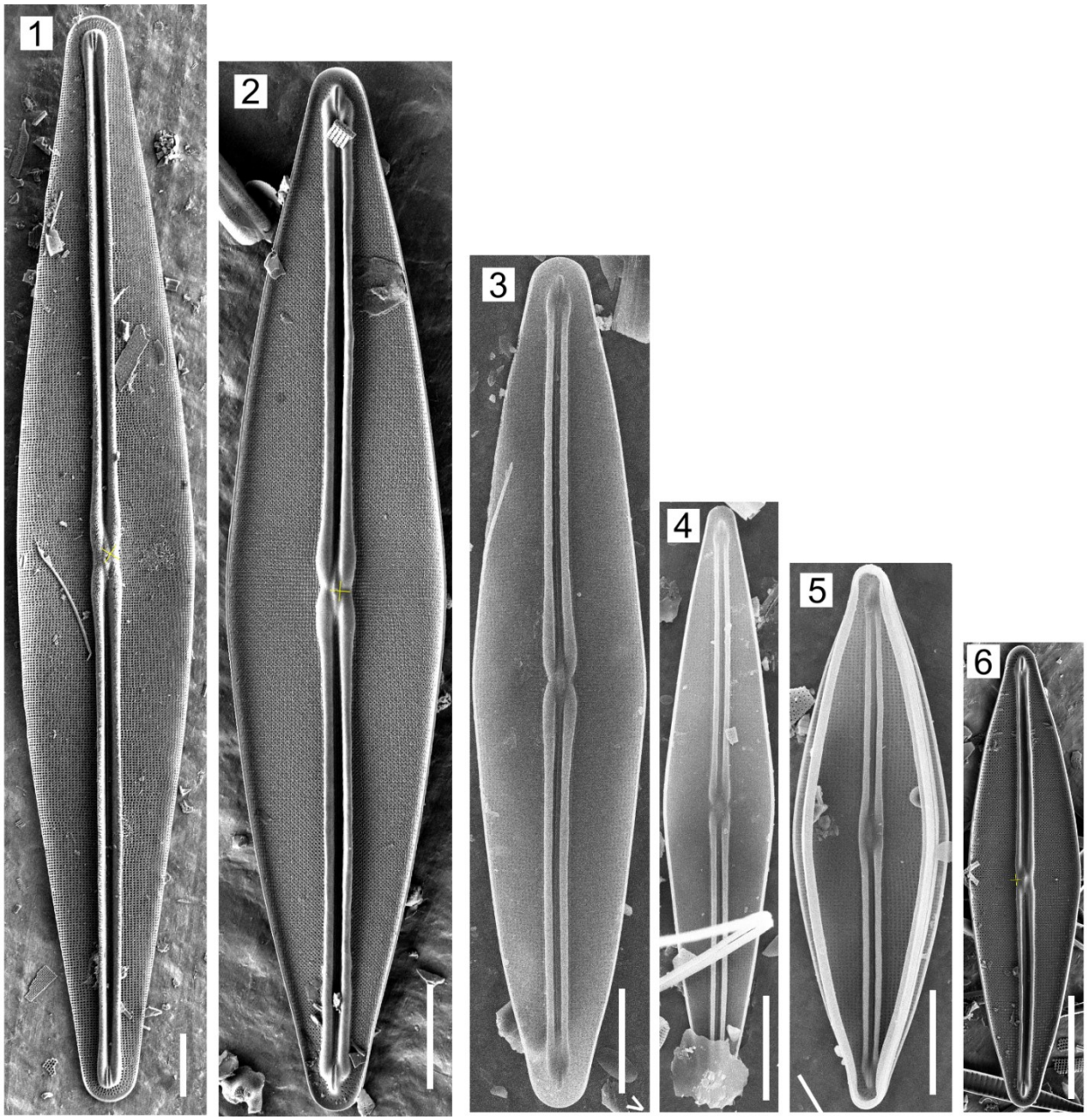
Source		This study					
	Species	*F. bahlsii*	*F. crassinervia*	*F. krammeri*	*F. gibsonea*	*F.* sp. (c)	*F.* cf. *saxonica* (b)
This study	*F. bahlsii*	–	19 – 23	17 – 21	17 – 21	16 – 20	20 – 24
	F. crassinervia	2.6 – 3.1	–	8	5 – 7	11	9
	F. krammeri	2.3 – 2.8	1.1	–	7 – 9	8	7
	F. gibsonea	2.3 – 2.8	0.7 – 0.9	0.9 – 1.2	–	1 – 3	8 – 10
	F. sp. (c)	2.2 – 2.7	1.5	1.1	0.1 – 0.4	–	9
	F. cf. *saxonica* (b)	2.7 – 3.2	1.2	0.9	1.1 – 1.3	1.2	–
GenBank	*F. erifuga*	2.0 – 2.6	1.6	1.2	0.3 – 0.5	0.4	1.3
	F. gondwana	2.3 – 2.8	1.9	1.5	0.5 – 0.8	0.7	1.6
	F. maoriana	2.6 – 3.1	1.5	1.2	1.2 – 1.5	1.3	1.1
	F. crassinervia x saxonica	2.0 – 2.6	1.1	0.4	0.9 – 1.2	1.1	1.1
	F. crassinervia	2.3 – 2.8	0.8	0.5	0.7 – 0.9	0.8	0.4
	F. saxonica	2.2 – 2.7	0.9	0.1	0.8 – 1.1	0.9	0.8
	F. sp.	2.3 – 2.8	1.9	1.5	0.5 – 0.8	0.7	1.6
	F. vulgaris	3.5 – 4.0	3.4	3.4	3.0 – 3.2	3.1	3.0

Table 2.5 Continued

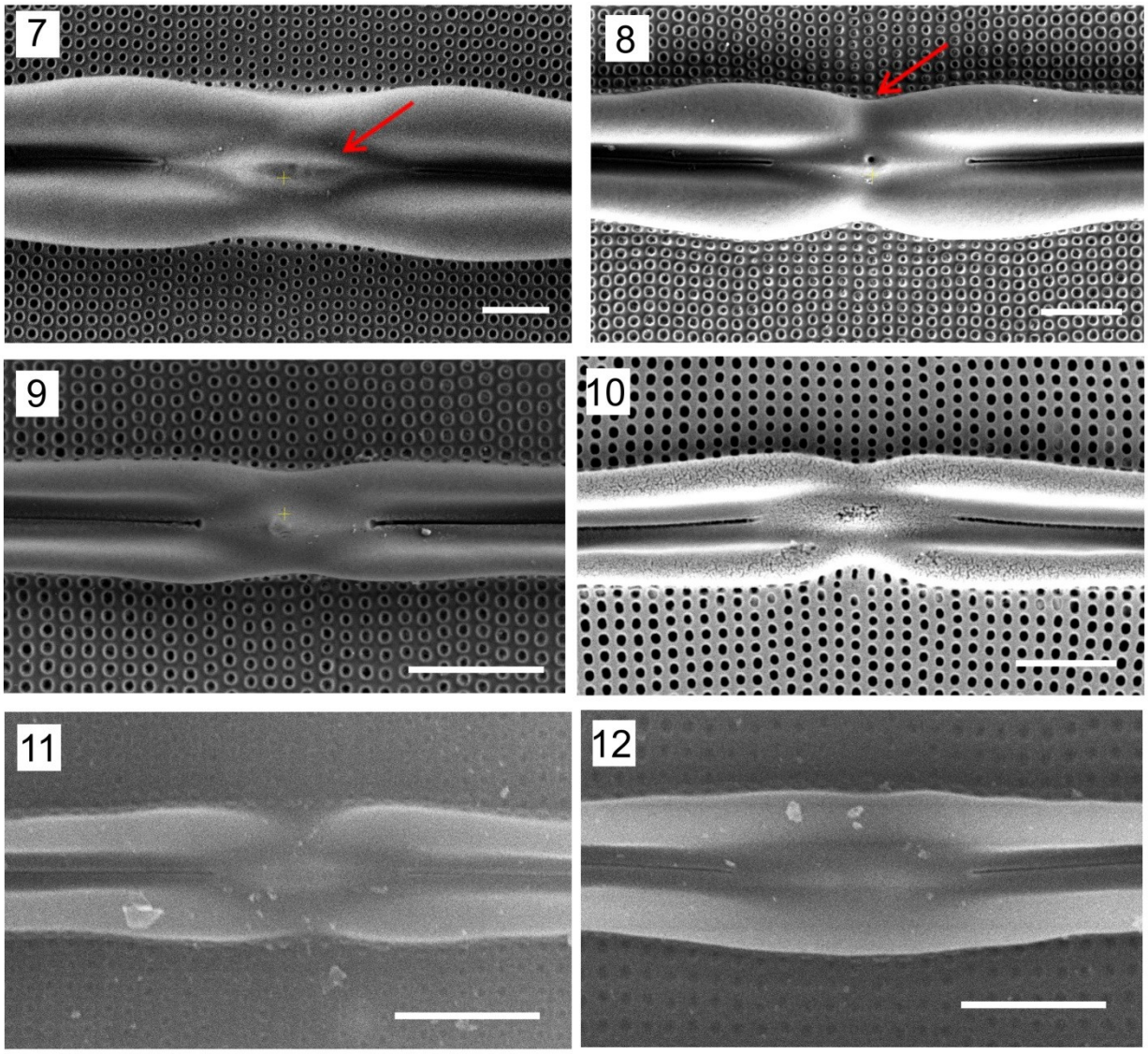
Source		*GenBank*							
	Species	*F. erifuga*	*F. gondwana*	*F. maoriana*	*F. crassinervia x saxonica*	*F. crassinervia*	*F. saxonica*	*F.* sp.	*F. vulgaris*
This study	*F. bahlsii*	15 – 19	17 – 21	19 – 23	15 – 19	17 – 21	16 – 20	17 – 21	26 – 30
	F. crassinervia	12	14	11	8	6	7	14	25
	F. krammeri	9	11	9	3	4	1	11	25
	F. gibsonea	2 – 4	4 – 6	9 – 11	7 – 9	5 – 7	6 – 8	4 – 6	22 – 24
	F. sp. (c)	3	5	10	8	6	7	5	23
	F. cf. *saxonica* (b)	10	12	8	8	3	6	12	22
GenBank	*F. erifuga*	–	2	7	9	7	8	2	23
	F. gondwana	0.3	–	9	11	9	10	2	25
	F. maoriana	0.9	1.2	–	10	5	8	9	24
	F. crassinervia x saxonica	1.2	1.5	1.3	–	5	2	11	26
	F. crassinervia	0.9	1.2	0.7	0.7	–	3	9	21
	F. saxonica	1.1	1.3	1.1	0.3	0.4	–	10	24
	F. sp.	0.3	0.3	1.2	1.5	1.2	1.3	–	25
	F. vulgaris	3.1	3.4	3.2	3.5	2.8	3.2	3.4	–

Table 2.6. Sequence divergence within species of *Frustulia.* NS: Nova Scotia, NY: New York, NL: Newfoundland, ON: Ontario.

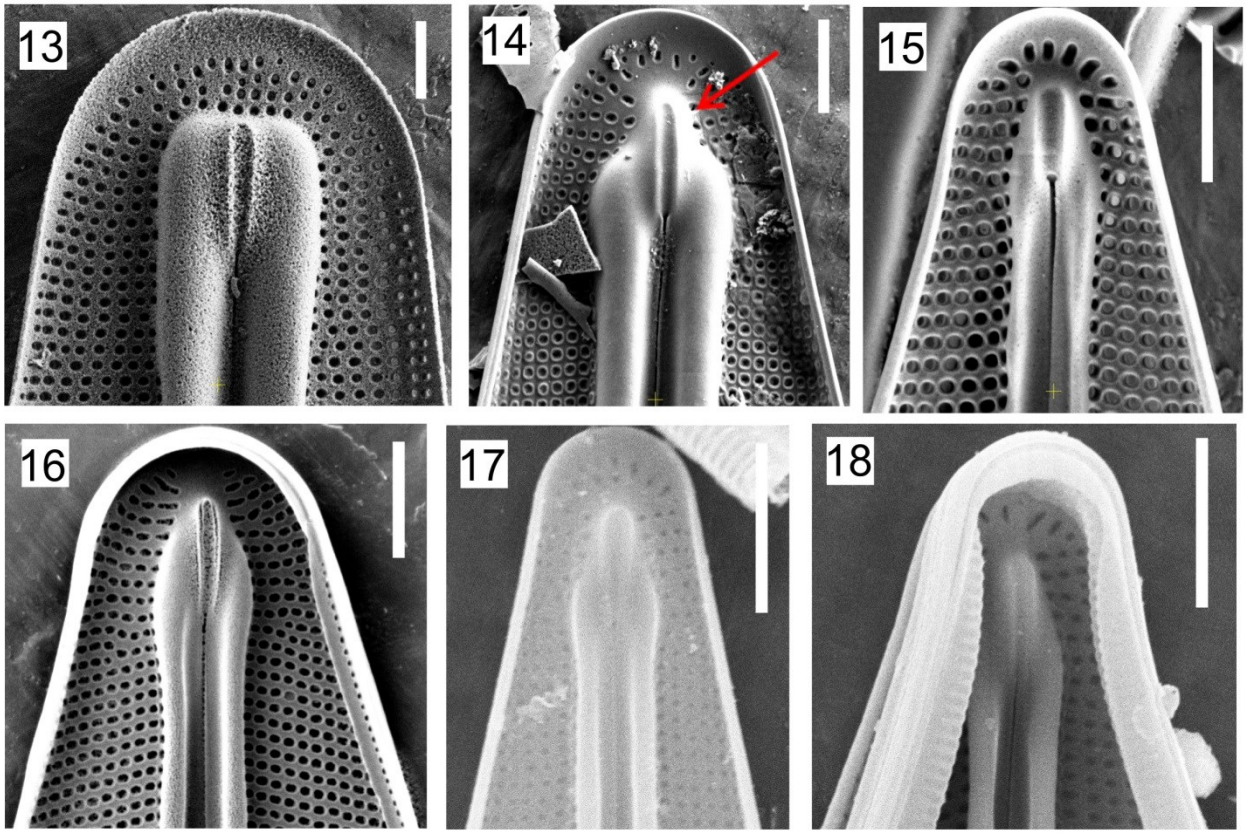
			rbcL		18s
Species	Locations sampled	*N*	Sequence divergence bp (%)	*N*	Sequence divergence bp (%)
F. bahlsii	NS, NY, NL, ON	23	0 – 2 (0.000 – 0.174)	9	0 – 8 (0.000 – 1.075)
F. crassinervia	NY	2	0 (0.000)	1	–
F. krammeri	ON, NS	10	0 – 2 (0.000 – 0.174)	2	0 (0.000)
F. gibsonea (a)	ON, NS, NL	16	0 (0.000)	9	0 – 2 (0.000 – 0.269)
F. sp. (c)	ON	2	2 (0.174)		
F. cf. *saxonica* (a)	ON	2	0 - 2 (0.000 – 0.174)	0	–
F. cf. *saxonica* (b)	ON	2	0 (0.000)	1	–



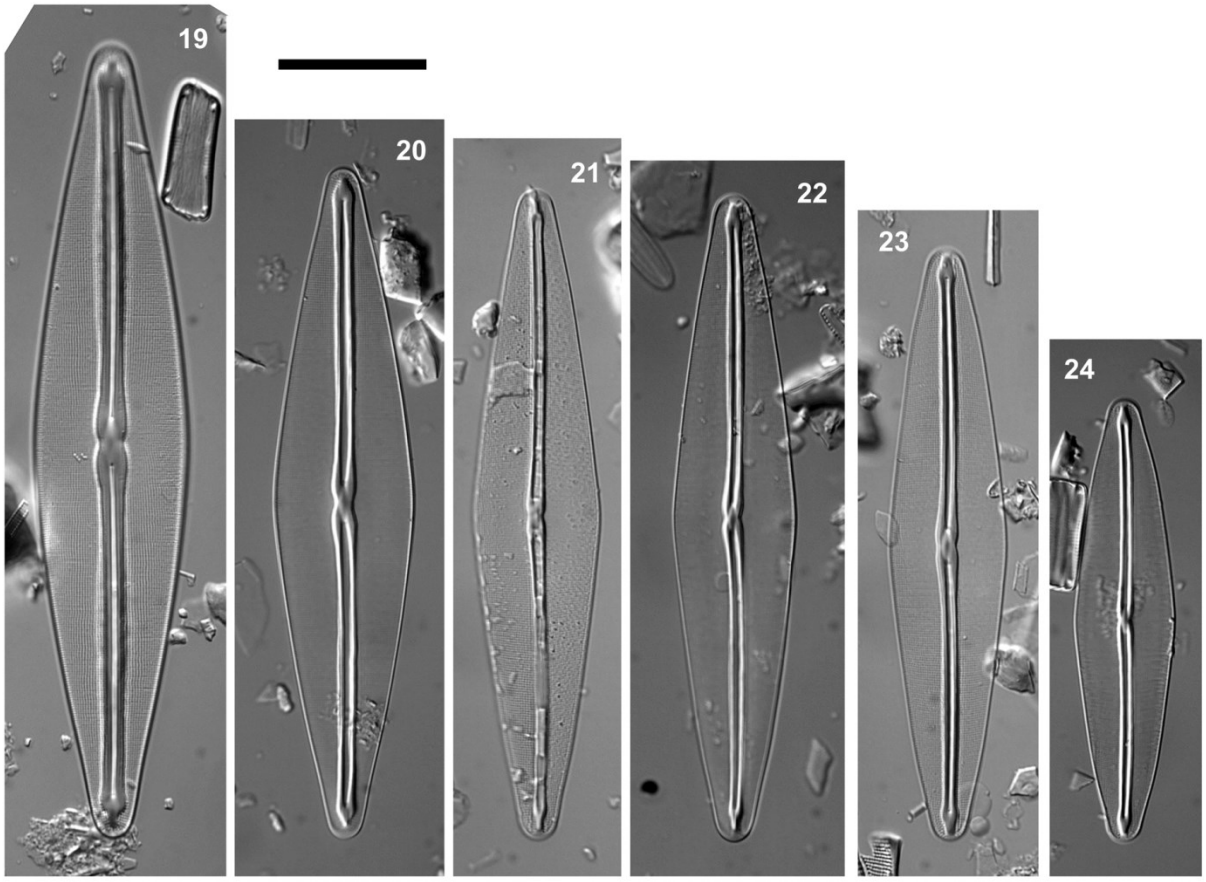
Figs 2.1–2.6. SEM micrographs of common North American taxa of *Frustulia*. Fig. 1. *F. bahlsii.* Fig. 2. *F. krammeri*. Fig. 3. *F. gibsonea.* Fig. 4. *F.* sp. (a) [*F.* cf. *saxonica*]. Fig. 5. *F.* sp. (b) [*F.* cf. *saxonica*]. Fig. 6. *F. crassinervia.* All figures show the internal views of frustules. Scale bars = 5 µm.



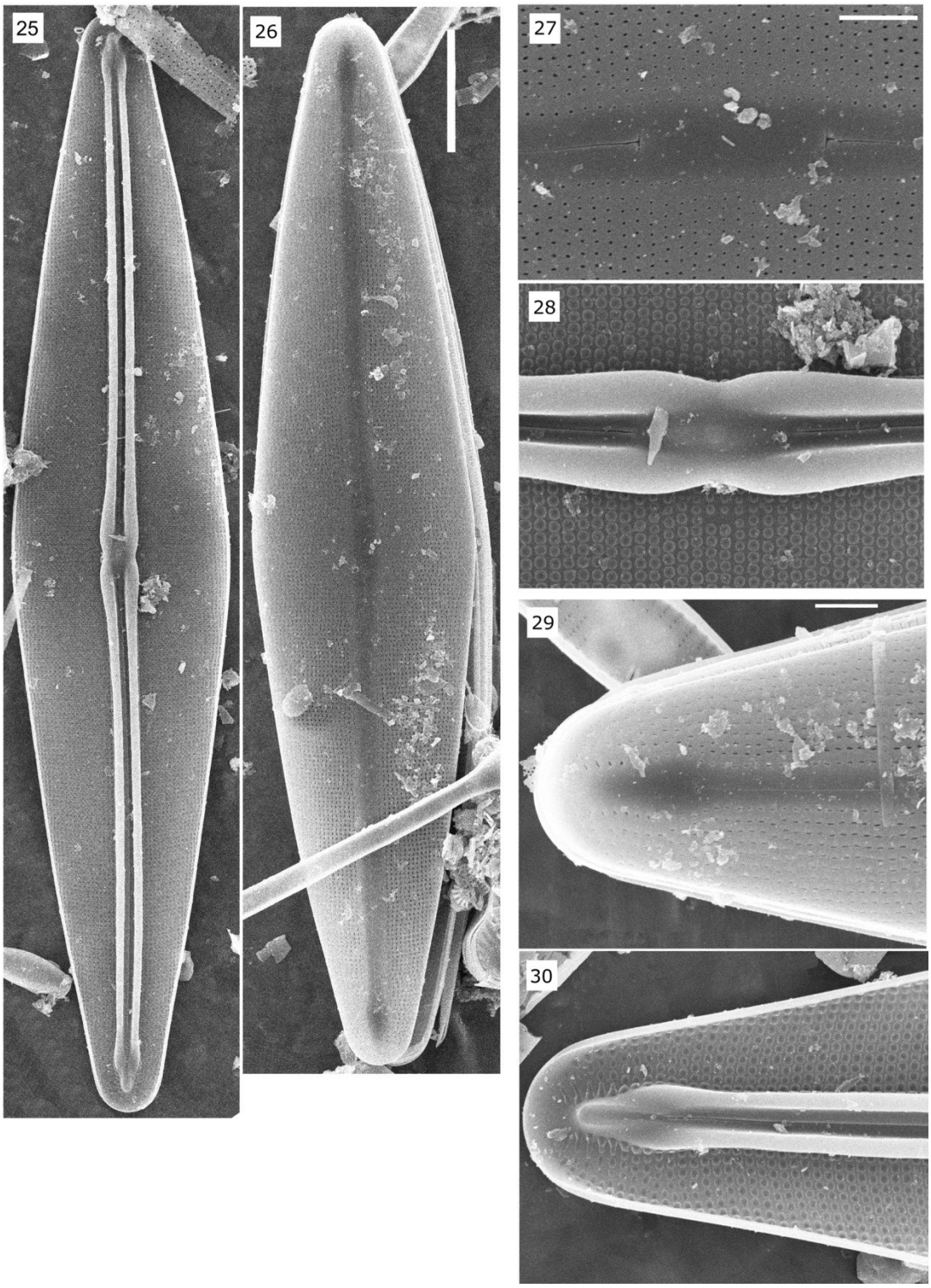
Figs 2.7–2.12. SEM micrographs of the internal central nodule and raphe ribs. Fig. 7. *F. bahlsii.* Arrow depicts central nodule. Fig. 8. *F. krammeri*. Arrow depicts the raphe ribs around the central nodule. Fig. 9. *F. crassinervia*. Fig. 10. *F. gibsonea.* Fig. 11. *F.* sp. (a) [*F.* cf. *saxonica*]. Fig. 12. *F.* sp. (b) [*F.* cf. *saxonica*]. Scale bars = 1 μm.



Figs 2.13–2.18. SEM micrographs of the internal apex and porte-crayon (ribs and helictoglossa) formation. Fig. 13. F. bahlsii. Fig. 14. F. krammeri. Arrow highlights the helictoglossa. Fig. 15. F. crassinervia. Fig. 16. F. gibsonea. Fig. 17. F. sp. (a) [F. cf. saxonica]. Fig. 18. F. sp. (b) [F. cf. saxonica]. Scale bars = 1 µm.



Figs 2.19–2.24. LM micrographs of *F. gibsonea*. Fig. 20. Holotype specimen for *F. gibsonea* from CANA 109952. Scale bar = 10 µm.



Figs 2.25–2. 30. SEM micrographs of *F. gibsonea* showing internal and external views. Fig 25, 26. Whole valve. Fig. 27. External central area. Fig. 28. Internal central area. Fig. 29. External apex. Fig. 30. Internal apex. Scale bars, Figs 25, 26 = 10 µm; Figs 27 – 30 = 2 µm.

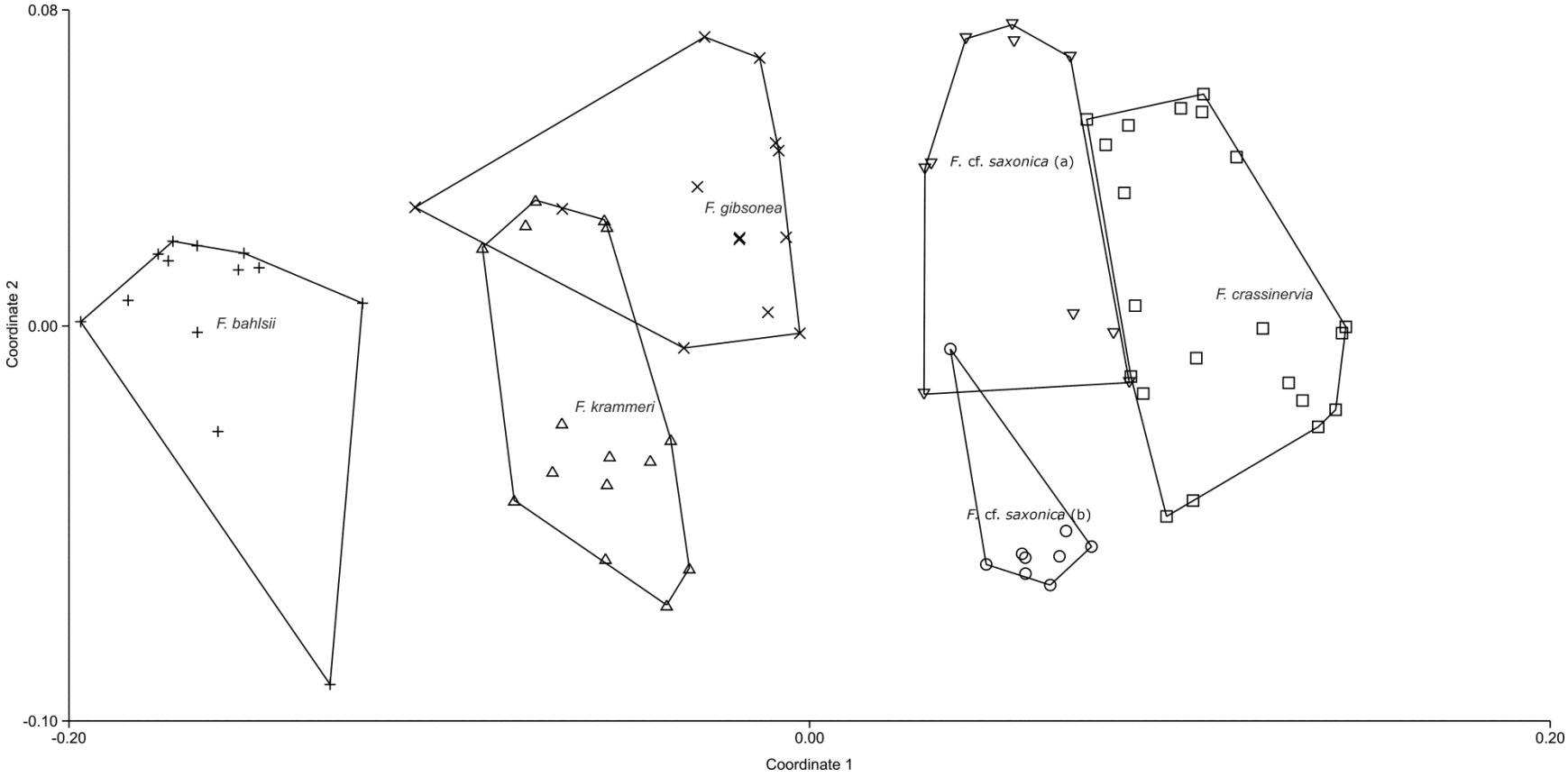


Fig. 2.31. Non metric multi-dimensional scaling (NMDS) plot showing the relative similarity of clades based on eleven morphological characters, as described in the Materials and Methods of Chapter 2. Stress = 0.082.

Fig. 2.32. Principal component analysis (PCA) results for North American *Frustulia* species' valve shapes. Type specimens, when available, were also included, and are represented by a star.

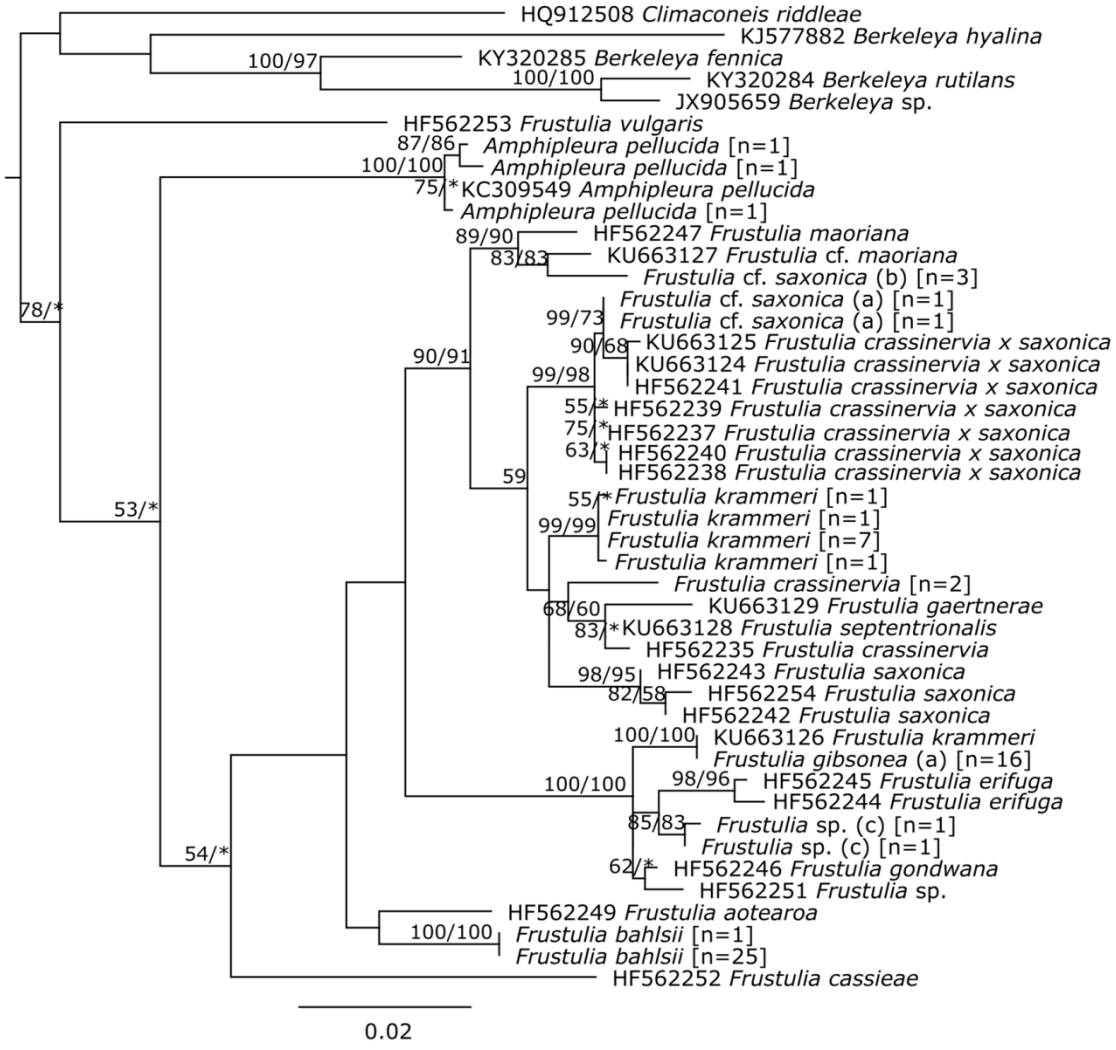


Fig. 2.33. Maximum likelihood tree of the genus *Frustulia*, using the *rbcL* gene. The GTRGAMMAI substitution model was used. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter. Bracketed numbers represent the number of specimens present within collapsed clades.

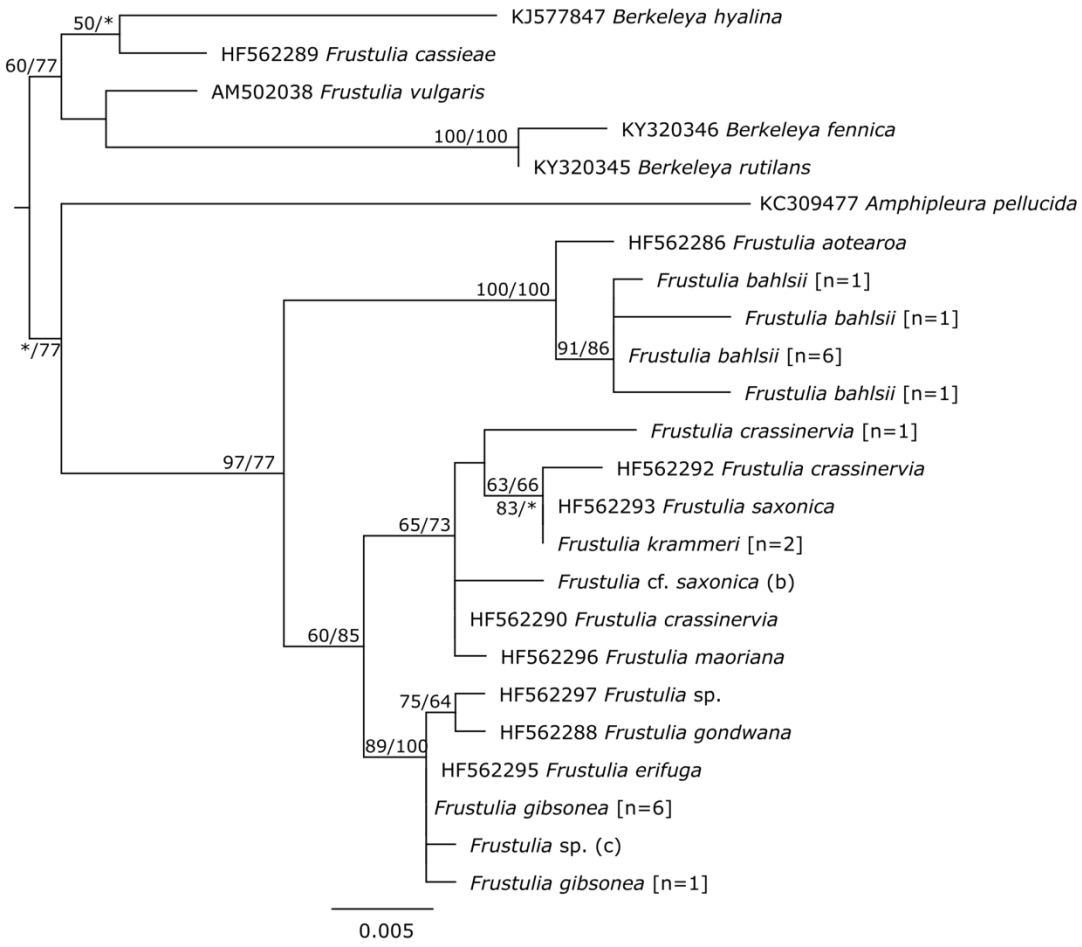


Fig. 2.34. Maximum likelihood tree of the genus *Frustulia*, using the 18S gene. The GTRGAMMAI substitution model was used. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter. Bracketed numbers represent the number of specimens present within collapsed clades.

Chapter 3 – Unrecognised Diversity in *Navicula* sensu stricto: A Morphological and Molecular Assessment

Disclaimer: new names presented here are not intended to be effectively published for nomenclatural purposes. These names will be published in a future scientific publication.

3.1 Abstract

Navicula Bory sensu stricto (s. str.) is a common and widespread genus that comprises over 200 species. Many of *Navicula*'s common species have been described as cosmopolitan. However, recent evidence suggests that these may be cryptic species complexes, comprised of multiple taxa with narrower distributions. This hidden diversity has enormous implications for our understanding of aquatic ecology worldwide. Using both molecular (*rbcL*, 18S rRNA, and *atpB* sequences) and morphological data (traditional, fine-scaled frustule characters, and shape analysis), I investigated widespread and common species from the genus. I recognize twenty-three taxa, including four unknowns. Two new species, *N. joguensa* sp. nov. and *N. cassensis* sp. nov., (provisional names) are described. These species are morphologically similar to *N. veneta*, and have been previously included within *N. veneta.* Plastid molecular data, from genes *rbcL* and *atpB*, shows that these taxa are distinct species. Furthermore, specimens that are morphologically identical to *N. tripunctata*, but that are more divergent from it than what is expected within a species, are observed. These specimens could represent the existence of cryptic diversity within *N. tripunctata.* This hidden diversity was uncovered due to the combination of molecular and morphology data – an approach that has a crucial role to play in the future of diatom systematics.

3.2 Introduction

Navicula Bory sensu stricto (s. str.) is one of the most commonly encountered diatom genera in the northern hemisphere (Underwood, 2001). It is a widespread genus, and species are found worldwide (Witkowski et al. 1998). The genus comprises more than 200 species, most of which can be found in freshwater environments (ca. 150), although some are also found in

brackish or marine habitats (Bruder & Medlin, 2008; Witkowski et al. 1998). With more research in marine ecosystems, it is anticipated that many more marine species will be identified.

Navicula presents a long history of taxonomic disagreements. The genus was created in 1822 by Bory de Saint-Vincent on the basis of its shape, which he described in French as resembling a "*navette de tisserand*" or weaver's shuttle. This object's resemblance to a small boat is the ultimate origin for both the French word "*navette*" and the Latin word "*navicula*". The number of species within the genus subsequently increased dramatically, as all diatoms with this characteristic elliptical-lanceolate valve shape were placed in *Navicula*, with little regard for other morphological features. In 1895, Cleve divided the genus into seven sections based on shared morphological features (Cleve, 1895). One of these sections, *Navicula* sect. *Lineolatae* Cleve, whose species share elliptical-lanceolate valves, but also characters like lineate areolae, straight raphes interrupted by central nodules, and short helictoglossae, is now treated as *Navicula* s. str. as it contains the type for the genus (Boyer, 1822; Bruder & Medlin, 2008). As for Cleve's other sections (*Navicula* sensu lato), most species have now been transferred to either newly erected genera, such as *Parlibellus* E.J. Cox (Cox, 1988), or to pre-existing ones, like *Sellaphora* Mereschowsky (Mann, 1989), although the limits between these closely related genera have never been clear. In addition, some species of *Navicula* s.l. still cannot be assigned to any existing genus or segregate of *Navicula*, a fact that prompted Lange-Bertalot (Lange-Bertalot & Moser, 1994) to erect the provisional genus *Naviculadicta* Lange-Bertalot. Although convenient, *Naviculadicta* has been heavily criticised because it has no defining characters (Kociolek, 1996).

Among modern morphological studies, it is agreed that the label *Navicula* should only be applied to species that fit the original diagnosis of *Navicula* sect. *Lineolatae* (Witkowski et al.

1998; Cox, 1999). However, no agreement exists on the minimal variation necessary for removal of a species from *Navicula* (Cox, 1999). Some species have been removed from the genus based on one or two features, such as species transferred to the genus *Fogedia* based on the presence of hyaline areas between striae and small differences in raphe structures (Witkowski et al. 1997). In her study of variation within *Navicula*, Cox (1999) demonstrated there was no true consensus on one morphological definition of the genus. While some authors interpret descriptions narrowly, and wish to circumscribe *Navicula* further, she argued that it would be more appropriate to modify the generic description in order to encompass more variation. However, morphological identifications can be problematic because variable environmental conditions can often mask key informative differences among taxa owing to phenotypic plasticity (Mann, 1999). In such cases, molecular data may prove to be useful in addressing these issues (Nakov et al. 2018).

Many of *Navicula*'s common species have been described as cosmopolitan, such as *N. cryptocephala* Kützing, *N. veneta* Kützing, *N. trivialis* Lange-Bertalot, *N. gregaria* Donkin, and *N. cryptotenella* Lange-Bertalot (Poulíčková et al. 2010). However, recent evidence suggests that these may be cryptic or pseudo-cryptic species complexes, comprising several taxa with restricted distributions. Species are considered to be cryptic if it is impossible to distinguish between them using morphology alone. In contrast, pseudocryptic species can be properly identified using fine-scale morphology, but they are so similar that there is a high probability of misidentification (Mann, 1999). Poulíčková et al. (2010) produced a phylogeny of four morphologically similar species within *Navicula* in order to look for cryptic and pseudo-cryptic diversity. Their phylogeny, derived from 28S sequences, demonstrated the existence of two pseudo-cryptic lineages within *N. cryptocephala.* These lineages could also be separated by the application of morphometrics on valve shape. Such hidden and poorly known diversity not only

has enormous implications for the conservation of rare species, but also to our understanding of aquatic ecology worldwide because such narrowly distributed taxa could highlight many unrecognised microhabitats.

Much uncertainty still exists around the phylogeny of *Navicula*. The few studies that have used molecular data to explore relationships within the genus have used only a small number of species that were exclusively sampled in Europe (Bruder & Medlin, 2008; Poulíčková et al. 2010). At present, no detailed genetic data are available for North American individuals, so the question as to how they may be related to European conspecifics remains unresolved. Fortunately, recent technical developments in DNA extraction and PCR amplification in single isolated cells or small groups of cells have now made it possible to reconstruct phylogenies and characterise genetic diversity on wide geographic scales without the need for established cultures (Hamilton et al. 2015).

The objective of this study was to reconstruct a phylogeny for major lineages of *Navicula* by sampling a wide range of morphological and taxonomic diversity within the genus. Geographically, a particular focus was given to North American individuals in order to address the current lack of data. A complete analysis of morphological variation was also conducted, and used jointly with the phylogeny to reinterpret the morphology of the genus.

3.3 Material and Methods

3.3.1 Sampling and isolations

Sediment samples were collected from 55 water bodies (Table 3.1). The samples were collected by macropipetting the sediment surface from the littoral zone of water bodies. From

this material, 89 living *Navicula* single-cells were isolated using a micropipette with contaminants removed by successive water isolations. These 89 *Navicula* cells were then divided into two groups in order to address some limitations of both of the methods used in this study: the single-cell amplification method, and the culturing method. The single-cell method allows for rapid processing of a high volume of specimens, but, as the cell is destroyed during DNA extraction, it does not allow for precise morphological observation or the possibility of maintaining a voucher specimen. The culturing method is more time-consuming, and does not work for all taxa, but it provides multiple cells which can then be processed for morphological observation and vouchering. The combination of these two methods provides a balanced approach.

Forty-five *Navicula* cells were transferred to PCR tubes containing 200 uL of 10% Chelex 100 solution (Bio-Rad Laboratories) (Hamilton et al. 2015) and 20 uL of silica beads.

The remaining 44 *Navicula* single-cells were placed in CHU-10 culture media with the addition of 10mg/L of Na_2SiO_3 and incubated in a growth chamber at 22ºC with a light intensity of 45 µmol m^{-2} $s^{-1}$and a 12:12 light:dark cycle. Cultures were allowed to grow from 3 weeks to 3 months, during which time they produced about 5 to >500 cells/mL. The cultures were verified for uniformity using an Olympus BH2 microscope. A portion of the culture material was then pipetted into PCR tubes containing the aforementioned Chelex solution.

Voucher information for all samples used in this study can be found in Table 3.1. GenBank accession numbers for all specimens used in this study can be found in Table 3.2. Although many short sequences for the 18S gene of *Navicula* species are found on GenBank, only sequences ≥500bp were included in this analysis because they were the only sequences to show significant overlap with the portion of the 18S amplified for this study.

3.3.2 DNA extraction, amplification, and sequencing

DNA extraction was performed on specimens as previously described (Hamilton et al. 2015). A nested PCR technique was used to amplify portions of the following molecular markers: the large subunit of the plastid gene ribulose-1,5-bisphosphate carboxylase/oxygenase (*rbcL*), the beta subunit of the plastid gene ATP synthase (*atpB*), and the 18S ribosomal RNA (18S rRNA) gene of nuclear rDNA. *RbcL* is recognized as a good gene sequence for species-level taxonomy of diatoms, whereas 18S (SSU) tends to be a more conserved barcode marker often used in higher-level analyses (Urbánková & Veselá 2013, and citations within). The a*tpB* is less commonly used, but was found to be a highly informative gene in a phylogeny of multiple orders of diatoms (Theriot et al. 2015).

PCR amplification of all targets was performed in two successive PCR reactions where the primers for the second amplification were nested within the target region of the first reaction. For the first amplification, reactions were carried out in a total volume of 15 µL, with 1 µL of DNA template, and final concentrations of: 1X DreamTaq Buffer (ThermoFisher Scientific), 2 mM dNTPs, 0.25 mM of each primer, 0.75 U of DreamTaq polymerase (ThermoFisher Scientific), and DNA-free water for the remaining volume. Amplification primers are listed in Table 3.3.

For the first PCR, thermocycling conditions were as follows (35 cycles): an initial step at 95ºC for 3 min, followed by a denaturing step at 95 ºC for 30 sec, an annealing step at 52 ºC for 30 sec, an extension step at 72 ºC for 1 min 30 sec, and a final step at 72 ºC for 10 min. The second PCR amplification was performed using the same procedure and cycling steps as the first

one, with the slight modification that 1 µL of product from the first reaction was used instead of 1 µL of DNA template and the amplification primers were nested within the PCR target of the first reaction. PCR products were visualized using 1.5 % agarose gels stained with Ethidium Bromide.

Depending on the strength of the band, the products of successful reactions were either sequenced directly or diluted by up to 60 µL of DNA-free water before sequencing. DNA sequencing followed the protocol outlined in Hamilton et al. (2015).

3.3.3 Phylogenetic analyses

Consensus sequences were assembled and aligned in Geneious Prime® 2020.0.3 (https://www.geneious.com), using MAFFT as a plugin (Kathoh & Standley, 2013). Because the 18S alignment could not be unambiguously aligned by MAFFT, 18S sequences were aligned using SSU-ALIGN (Nawrocki, 2009), a program that uses the conserved regions and secondary structure of SSU rRNA to assist with alignment. Columns that included a significant number of misaligned nucleotides were removed using the *ssu-mask* function. Out of 1881 columns, 267 were excluded from the final alignment.

Two trees were reconstructed for each molecular marker using maximum parsimony (MP) and maximum likelihood (ML) methods. Parsimony analyses were conducted in PAUP 4.0b10 (Swofford 2002) using heuristic searches and a random addition sequence of taxa for 1000 replicates with the MULTREES option on and with tree-bisection-reconnection (TBR) branch swapping conducted on a maximum of 1000 trees per replicate. Support for internal branches was assessed via a bootstrap (BS; Felsenstein 1985) using a simple addition sequence

for 1,000 replicates with the MULTREES option "off" which gives BS values similar to those seen when the option is "on" (DeBry and Olmstead 2000). Maximum likelihood trees were produced using the program RAxML Next Generation (RAxML-NG) (Kozlov et al. 2019). Partitioning schemes were determined using PartitionFinder version 2.1.1, with linked branch lengths, the Bayesian information criterion (BIC) for model selection, and the "all" search scheme (Lanfear 2012). For *rbcL,* two subsets were created: one for codon position 1, with K80+I+G as best fit model, and one for codon positions 2 and 3, with GTR+I+G as best fit model. For *atpB,* three subsets were created, one for each codon position. The best fit models were TVM+I+G, F81+I, and F81+I+G for the first, second, and third codon positions respectively. For the analysis combining both plastid genes, *rbcL* and *atpB*, four subsets created: one for *rbcL* codon positions 1 and 3 and *atpB* position 1, a second for *rbcL* position 2, a third for *atpB* position 2, and a fourth for *atpB* position 3, with HKY+I+G, TVM+G, F81+I+G and TVM+I+G as best fit models, respectively. Because it is a non-protein coding marker, 18S was not partitioned. The best fit model was TRN+I+G.

For all ML analyses, a total of 500 tree searches were performed for each marker, using 250 random and 250 parsimony-based starting trees. Branch support was assessed via bootstrap, with 1000 (standard/slow) replicates. An extended majority rule (MRE) based bootstopping test with the default cutoff value of 0.03 was performed post-hoc to evaluate convergence (Pattengale et al. 2010).

3.3.4 Morphological studies

Live cells were photographed from single-cell isolations at 300X magnification using an Olympus BH2 equipped with a Pixelink camera. To prepare for scanning electron microscopy (SEM), the sediment samples were digested with boiling H_2SO_4:HNO_3 in a 1:1 ratio for twenty minutes then washed through five dilutions of distilled water in order to remove all residual acid. Drops of digested samples were dried on aluminum pin stubs (Soquelec). These stubs were covered in ~500 Å Au-Pd, examined, and photographed using an APREO FEI SEM (Thermo-Fisher Scientific) with an acceleration voltage of 2 kV. Images were collected using T1 (backscatter) and secondary electron (SE) detectors.

The following data were recorded from the SEM micrographs: (1) valve length and (2) width, (3) stria density, (4) areola length, (5) thickening of the sternum, (6) shape of the proximal raphe endings, (7) distance between proximal pores, (8) width of the central area, (9) shape of the central area, (10) presence/absence of apical pores, (11) number of apical pores, (12) presence/absence of angled areolae around distal raphe ends, and (13) distance between the start of the axial hook and the valve margin.

Single-cells and SEM micrographs were linked together based on information gathered from the light microscopy (LM) pictures taken at isolation. In order to be matched to a valve, a photographed SEM specimen had to share a set of morphological characteristics with that LM micrograph: valve, valve shape, and other notable characteristics (central nodule or apex features) when they could be observed on the LM micrographs.

A non-metric multi-dimensional scaling (NMDS) analysis was performed using Past4.02 (Hammer et al. 2001). The thirteen previously described morphological characters recorded from SEM micrographs were used in this analysis. A Gower similarity index was used because the

dataset contained mixed variables (binary, ordinal, and measurement variables). Three dimensions were used to minimize stress (Kruskal, 1964).

Valve shapes were extracted from SEM images using DiaOutline (Wishkerman & Hamilton 2018), and a PCA analysis was performed in R 3.5.0 (R Core Team, 2018) with the following packages installed: MASS (Venables & Ripley 2002), ggplot2 (Wickham 2016), GGally (Schloerke et al. 2018), doBy (Højsgaard & Halekoh 2018), data.table (Dowle & Srinivasan 2018), plyr (Wickham 2011), gridExtra (Auguie 2017) and Momocs (Bonhomme et al. 2014).

3.4 Results

The DNA alignment datasets had a final length of 445 to 1545 characters for *rbcL,* 1028 to 1287 characters for *atpB*, and 500 to 2389 characters for 18S. The MRE-based bootstopping tests converged at 750, 950, and 550 trees for *rbcL*, *atpB*, and 18S respectively.

Nineteen major taxa are recognized within our analysis, and four other taxa are unidentified. Two new species, *N. joguensa* sp. nov. and *N. cassensis* sp. nov.*,* are described. SEM micrographs for each of the taxa observed in this study are presented in Figures 3.1 – 3.69. Phylogenetic trees are presented in Figures 3.70 – 3.80. Morphological analyses are shown in Figures 3.81 – 3.82.

3.4.1 Morphology of clades in this study

Navicula capitoradiata Germain 1981 Figs 3.14, 3.48, 3.63

Valves lanceolate, 34 – 37 µm long, 6 – 7 µm wide, with rostrate apices. Striae radiate around central nodule, strongly convergent at apices. Stria density 14 – 16 in 10 µm. Areolae slit-shaped, 0.31 – 0.46 µm in length. Proximal raphe fissures hook-shaped, separated by 1.05 – 1.10 µm. Central area slightly expanded, 2.27 – 2.74 µm wide. Sternum flat; distal raphe ends forming a hook – raphe curving 1.49 – 2.10 µm from valve margin. A single areola present above this structure. Angled areolae border distal raphe ends.

Navicula cassensis provisional name Figs 3.31, 3.34, 3.37

Valves lanceolate to linear-lanceolate, 27 – 37 µm in length, 5 – 7 µm in width, with rounded apices. Striae radiate around central nodule, and parallel to weakly convergent at apices. Stria density 12 – 14 in 10 µm. Areolae slit-shaped, 0.29 – 0.43 µm in length. Proximal raphe fissures teardrop-shaped and separated by 1.33 – 1.72 µm. Central area large, 2.25 – 3.92 µm wide, rectangular to rounded, and asymmetrical. Sternum flat to slightly thickened. Distal raphe ends forming a hook – raphe curving 1.61 – 2.21 µm from valve margin. A single areola sometimes visible above this structure. Angled areolae border distal raphe ends.

Differential diagnosis:

This species has distinctive features that are present in all valves: a large asymmetrical central area, and teardrop-shaped proximal raphe fissures that are relatively distant.

This species is morphologically similar to *N. veneta* and *N. joguensa,* however it has longer and more linear frustules, and a lower stria density.

Etymology: The species name ***cassensis*** is derived from a location where specimens were found: Casselman, Ontario.

Type slide & material: CANA 128733, in the Canadian Museum of Nature, Ottawa, Canada, collected by E. Huliganga, June 25th 2018.

Type habitat: Butternut Creek, 983 St Isidore Road (Route 650), Casselman Ontario (lat. 45.299122, long. -75.10783). Freshwater. Slightly alkaline (pH 8.19) and specific conductance of 56.0 µS cm^{-1}

Distribution: This species was also observed in eastern and northern Ontario, and in the Toronto area.

Comments:

Navicula cassensis, N. veneta and *N. joguensa* are pseudo-cryptic, and are difficult to distinguish using morphology alone.

Navicula cincta (Ehrenberg) Ralfs 1861 Figs 3.18, 3.41, 3.56

Valves lanceolate to linear-lanceolate, 23 µm in length, 5 µm in width, with rounded apices. Striae radiate around central nodule, weakly convergent at apices. Stria density 15 in 10 µm. Areolae slit-shaped, 0.38 µm in length. Proximal raphe fissures curved, 1.02 µm apart. Central area slightly expanded, 1.93 µm wide. External sternum flat; distal raphe ends forming a hook – raphe curving 1.30 µm from valve margin. A row of 3 areolae visible above this structure. Angled areolae border distal raphe ends.

Navicula cryptocephala Kützing 1844 Figs 3.16, 3.44, 3.59

Valves lanceolate to narrowly lanceolate, 26 –31 µm in length, 5 – 6 µm in width, with subcapitate apices. Striae strongly radiate around central nodule, and convergent at apices. Stria density 16 – 18 in 10 µm. Areolae slit-shaped, 0.32 – 0.36 µm in length. Proximal raphe fissures curved and separated by 1.12 – 1.35 µm. Central area large, 2.74 – 3.32 µm wide, and rounded. External sternum slightly thickened along raphe and becomes flat between the proximal fissures. Distal raphe ends forming a hook – raphe curving 1.20 – 1.63 µm from valve margin. A row of areolae visible around this structure. Angled areolae border distal raphe ends.

Navicula cryptotenella Lange-Bertalot 1985 Figs 3.20, 3.46, 3.61

Valves lanceolate to narrowly lanceolate, 22 – 24 µm in length, 5 – 6 µm in width, with rounded apices. Striae radiate around central nodule, and convergent at apices. Stria density 16 in 10 µm. Areolae slit-shaped, 0.26 – 0.40 µm in length. Proximal raphe fissures straight, teardrop-shaped 0.98 – 1.09 µm apart. Central area small, 1.59 – 2.26 µm wide. External sternum flat to thickened. Distal raphe ends forming a hook – raphe curving 1.17 – 1.53 µm from valve margin. Two to three areolae visible above this structure. Angled areolae border distal raphe ends in some specimens, but not all.

Navicula cryptotenelloides Lange-Bertalot 1985 Figs 3.24, 3.27, 3.30

Valves broadly lanceolate, 12 – 16 µm in length, 4 – 5 µm in width, with rounded apices. Striae radiate around central area, and convergent at the apices. Stria density 18 in 10 µm. Areolae slit-shaped, 0.33 – 0.38 µm in length. Proximal raphe fissures straight and separated by 0.74 – 0.89 µm. Central area moderately large, 1.45 – 2.05 µm wide. External sternum flat to slightly thickened. Distal raphe ends forming a hook – raphe curving 0.57 – 1.18 µm from valve margin. Areolae sometimes visible above this structure, numbering 2 – 3. Angled areolae border distal raphe ends in some specimens, but not all.

Navicula gregaria Donkin 1861 Figs 3.21, 3.45, 3.60

Valves lanceolate to broadly lanceolate, 21 – 29 µm in length, 5 – 6 µm in width, with rostrate apices. Striae weakly radiate around central nodule, and convergent at apices. Stria density 18 – 19 in 10 µm. Areolae slit-shaped, 0.30 – 0.35 µm in length. Proximal raphe fissures curved and 0.86 – 1.03 µm apart. Central area large and asymmetrical, 3.15 – 3.39 µm wide, with one side rounded, opposite side rectangular. Sternum flat; distal raphe ends forming a hook – raphe curving 1.37 – 1.75 µm from valve margin. Areolae sometimes visible above this structure, numbering 3 – 4. Angled areolae border distal raphe ends in some specimens, but not all.

Navicula joguensa provisional name Figs 3.32, 3.35, 3.38

Valves lanceolate to linear-lanceolate, 22 – 25 µm in length, 5 – 6 µm in width, with rounded apices. Striae radiate around central nodule, and convergent at apices. Stria

density 14 – 16 in 10 µm. Areolae slit-shaped, 0.19 – 0.35 µm in length. Proximal raphe fissures teardrop-shaped and separated by 1.18 – 1.45 µm. Central area large, rectangular to rounded, and asymmetrical, 3.42 – 3.94 µm wide. Sternum flat to slightly thickened. Distal raphe ends forming a hook – raphe curving 1.40 – 1.73 µm from valve margin. A single areola sometimes visible above this structure. Angled areolae border distal raphe ends.

Differential diagnosis:

This species has distinctive features that are present in all valves studied for this taxon: a large asymmetrical central area, and teardrop-shaped proximal raphe fissures that are relatively distant.

This species is morphologically similar to *N. veneta,* and *N. cassensis.* Compared to *N. veneta*, its frustules are longer and more linear, and its central areas are larger. Compared to *N. cassensis,* its frustules are smaller and its stria density is lower.

Etymology: The species name ***joguensa*** is derived from a location where specimens were found: Jogues, Ontario.

Type slide & material: CANA 128280, in the Canadian Museum of Nature, Ottawa, Canada, collected by A. Bouchard, May 19th 2018.

Type habitat: Creek on Chartrand Road, Jogues, Ontario, (lat. 49.6448, long. -83.6835). Freshwater. Acidic (pH 5.5) and specific conductance of 28.0 µS cm^{-1}.

Distribution: This species was also observed in eastern Ontario and Quebec.

Comments:

Navicula joguensa, N. veneta and *N. cassensis* are pseudo-cryptic, and are difficult to distinguish using morphology alone.

Navicula lundii E. Reichardt 1985 Figs 3.17, 3.43, 3.58

Valves lanceolate, 24 – 26 µm in length, 5 µm in width, with rounded apices. Striae radiate around central nodule, and convergent at the apices. Stria density is 14 – 16 in 10 µm. Areolae slit-shaped, 0.37 – 0.41 µm in length. Proximal raphe fissures curved and separated by 1.05 µm. Central area slightly widened, 1.95 – 2.21 µm wide. Sternum slightly thickened. Distal raphe ends forming a hook – raphe curving 1.25 – 1.34 µm from valve edge. Areolae sometimes visible above this structure. Angled areolae border distal raphe ends in some specimens, but not all.

Navicula menisculus Schumann 1867 Figs 3.23, 3.26, 3.29

Valves broadly lanceolate, 17 – 18 µm in length, 4 – 6 µm in width, with rounded apices slightly constricted in some specimens. Striae radiate around central nodule, and parallel at apices. Stria density 14 – 18 in 10 µm. Areolae slit-shaped, 0.25 – 0.41 µm in length. Proximal raphe fissures teardrop-shaped and 0.95 – 1.20 µm apart. Central area slightly to moderately widened, 1.84 – 2.43 µm wide. Sternum flat to thickened. Distal raphe ends forming a hook – raphe curving 1.18 – 1.46 µm from end of valve. Angled areolae border distal raphe ends.

Navicula peregrina (Ehrenberg) Kützing 1844 Figs 3.1, 3.3, 3.5

Valves lanceolate, 97 – 99 µm in length, 16 – 19 µm in width, with rounded and slightly constricted apices. Striae radiate around central nodule, and convergent at apices. Stria density 7 – 8 in 10 µm. Areolae slit-shaped, 0.53 – 0.57 µm in length. Proximal raphe fissures widened, hooked-shaped and separated by 2.40 – 2.90 µm. Central area widened into a rounded shape, 8.33 – 10.25 µm wide. Sternum thickened and very prominent. Distal raphe ends forming a hook – raphe curving 3.31 – 4.75 µm from valve margin. Areolae visible above this structure.

Navicula radiosa Kützing 1844 Figs 3.7, 3.52, 3.67

Valves narrowly lanceolate to linear-lanceolate, 50 – 84 µm in length, 9 – 11 µm in width, with rounded apices. Striae radiate around central nodule, and acutely convergent at apices. Stria density is 10 – 13 in 10 µm. Areolae slit-shaped, 0.38 – 0.77 µm in length. Proximal raphe fissures hook-shaped and separated by 1.12 – 1.85 µm. Central area expanded into a rounded shape, 3.53 – 5.94 µm wide. Sternum flat in most valves, but slightly thickened in some. Distal raphe ends forming a hook – raphe curving 2.19 – 3.91 µm from valve margin. Areolae visible above this structure in some valves. Angled areolae border distal raphe ends.

Navicula radiosafallax Kützing 1844 Figs 3.12, 3.47, 3.62

Valves narrowly lanceolate, 38 – 40 µm in length, 5 µm in width, with rounded apices. Striae strongly radiate around central nodule, and acutely convergent at apices. Stria density 16 – 18 in 10 µm. Areolae slit-shaped, 0.34 – 0.35 µm in length. Proximal raphe fissures curved and 0.77 – 1.02 µm apart. Central area size variable; shape ranges from elliptical to rounded. Sternum flat to thickened. Distal raphe ends forming a hook – raphe curving 1.03 – 1.46 µm valve margin. Angled areolae border distal raphe ends in some specimens, but not all.

Navicula reinhardtii (Grunow) Grunow 1880 Figs 3.9, 3.67, 3.68

Valves broadly lanceolate to rhomboidal, 55 µm in length, 18 µm in width, with rounded apices. Striae radiate around central nodule, and parallel at apices. Stria density 8 in 10 µm. Areolae slit-shaped, 0.47 µm in length. Proximal raphe fissures teardrop shaped and separated by 3.49 µm. Central area widened vertically and 8.59 µm in size. Sternum flat to slightly thickened. Distal raphe ends forming a hook – raphe curving 2.13 µm from valve margin. A row of 5 – 7 areolae visible above this structure.

Navicula rostellata Kützing 1844 Figs 3.13, 3.50, 3.65

Valves lanceolate, 39 µm in length and 9 µm in width, with rostrate apices. Striae radiate around central nodule, and convergent at apices. Stria density 14 in 10 µm. Areolae slit-shaped, 0.47 µm in length. Proximal raphe fissures curved and 1.44 µm apart. Central area large and asymmetrical, 4.80 µm wide, with one side rounded and other rectangular. Sternum flat to slightly thickened. Distal raphe ends forming a hook – raphe curving 2.39

µm from valve margin. A row of areolae visible above this structure. Angled areolae border distal raphe ends.

Navicula trivialis Lange-Bertalot 1980 Figs 3.11, 3.49, 3.64

Valves broadly lanceolate to rhomboid, 28 – 46 µm in length and 8 – 10 µm in width. Apices rostrate in some specimens and rounded in others. Striae radiate around central nodule and parallel to weakly convergent at apices. Stria density 12 – 14 in 10 µm. Areolae slit-shaped, 0.30 – 0.54 µm in length. Proximal raphe fissures curved and distant, 1.53 – 2.26 µm apart. Central area expanded into a rounded shape, 3.56 – 5.65 µm wide. Sternum thickened. Distal raphe ends forming a hook – raphe curving 1.37 – 1.96 µm from valve margin. A row of 3 – 5 areolae visible above this structure. Angled areolae border distal raphe ends.

Navicula veneta Kützing 1844 Figs 3.33, 3.36, 3.39

Valves broadly lanceolate to linear-lanceolate, 19 – 22 µm in length, 5 – 6 µm in width, with rostrate to rounded apices. Striae weakly radiate around central nodule, and convergent at apices. Stria density 14 – 16 in 10 µm. Areolae slit-shaped, 0.23 – 0.36 µm in length. Proximal raphe fissures teardrop-shaped and 0.94 – 1.41 µm apart. Central area large, 2.44 – 3.63 µm wide, rectangular, and often asymmetrical. Sternum flat to slightly thickened. Distal raphe ends forming a hook – raphe curving 1.12 – 1.63 µm from valve margin. A row of 1 – 5 areolae visible above this structure. Angled areolae border distal raphe ends.

Navicula viridulacalcis Lange-Bertalot 2000 Figs 3.8, 3.53, 3.69

Valves linear-lanceolate, 53 – 58 µm in length, 10 µm in width, with rostrate apices. Striae radiate around central nodule, and convergent at apices. Stria density 9 – 10 in 10 µm. Areolae slit-shaped, 0.62 – 0.78 µm in length. Proximal raphe fissures curved and 1.35 – 2.21 µm apart. Central area large and asymmetrical, 6.62 – 7.52 µm wide, with one side rounded, opposite side rectangular. Sternum flat; distal raphe ends forming a hook – raphe curving 1.96 – 3.73 µm from end of valve. A row of areolae visible above this structure. Angled areolae border distal raphe ends.

Navicula wildii Lange-Bertalot 1993 Figs 3.15, 3.42, 3.57

Valves lanceolate, 27 µm in length, 5 µm in width, with rounded and slightly constricted apices. Striae radiate around central nodule, and convergent at apices. Stria density 14 in 10 µm. Areolae slit-shaped, 0.56 µm in length. Proximal raphe fissures curved and 1.08 µm apart. Central area small, 2.46 µm wide. Sternum distinctly thickened. Distal raphe ends forming a hook – raphe curving 1.62 µm from valve margin. A row of areolae visible above this structure. Angled areolae border distal raphe ends.

Navicula cf. *tripunctata* Figs 3.10, 3.51, 3.66

Valves lanceolate to linear-lanceolate, 42 – 51 µm in length, 7 – 8 µm in width, with rounded apices. Striae weakly radiate around central nodule, and convergent or parallel at

apices. Stria density 11 – 12 in 10 µm. Areolae slit-shaped, 0.39 – 0.49 µm in length. Proximal raphe fissures straight and 1.69 – 1.97 µm apart. Central area large, 4.01 – 5.24 µm wide, rounded to rectangular, and often asymmetrical. Sternum flat to slightly thickened. Distal raphe ends forming a hook – raphe curving 1.59 – 2.46 µm from valve margin; a row of areolae visible above this structure.

Comments:

These specimens appear morphologically identical to *Navicula tripuncata*, however they are not identified as *N. tripunctata* because of molecular divergence.

Navicula cf. *cincta* Figs 3.22, 3.25, 3.28

Valves linear-lanceolate, 20 µm in length, 5 – 6 µm in width, with rounded apices. Striae radiate around central nodule, and parallel to weakly convergent at apices. Stria density 14 – 16 in 10 µm. Areolae slit-shaped, 0.41 – 0.43 µm in length. Proximal raphe fissures curved and separated by 0.96 – 1.00 µm. Central area of moderate size, 2.39 – 2.56 µm, and rounded. Sternum flat to slightly thickened. Distal raphe ends forming a hook – raphe curving 1.09 – 1.25 µm from valve margin; a row of 2 – 3 areolae visible above and around this structure. Angled areolae border distal raphe ends.

Comments:

These specimens are somewhat similar to *Navicula cincta*, however their central areas are larger and rounded in shape. The specimens are also slightly smaller and have longer areolae.

Navicula cf. *metareichardtiana* Figs 3.19, 3.40, 3.55

Valves lanceolate, 20 – 22 µm in length, 5 – 6 µm in width, with subcapitate apices. Striae strongly radiate around central nodule, and parallel to convergent at apices. Stria density is 16 – 18 in 10 µm. Areolae slit-shaped, 0.28 – 0.37 µm in length. Proximal raphe fissures straight to teardrop-shaped and 1.26 – 1.65 µm apart. Central area large, 2.22 – 3.34 µm, and rounded. Sternum flat to very slightly thickened. Distal raphe ends forming a hook – raphe curving 1.09 – 1.57 µm from valve margin; a row of 3 – 6 areolae visible above and around this structure. Angled areolae border distal raphe ends.

Comments:

These specimens are somewhat similar to the species' description of *Navicula metareichardtiana,* however the apices are more protracted, the central areas are larger, and the striae are slightly denser.

Navicula cf. *subwalkeri* Figs 3.2, 3.4, 3.6

Valves lanceolate, 85 – 100 µm in length, 14 – 18 µm in width, with rounded apices. Striae radiate around central nodule, and strongly convergent at apices. Stria density is 9 – 10 in 10 µm. Areolae slit-shaped, 0.66 – 0.88 µm in length. Proximal raphe fissures hook-shaped and separated by 2.34 – 2.63 µm. Central area is of moderate size, 6.32 – 8.04 µm, and irregular in shape. Sternum slightly to moderately thickened. Distal raphe ends forming a hook – raphe curving 3.11 – 4.41 µm from end of valve. A single areola sometimes visible above this structure. Angled areolae border distal raphe ends in some specimens, but not all.

Comments:

These specimens are somewhat similar to *Navicula subwalkeri,* however they are slightly smaller, the apices are more rounded, and the central areas are larger with irregular borders.

3.4.2 Phylogenetic analyses

The ML topologies were compatible with the strict consensus of parsimony analyses, as all differences between topologies were at poorly supported branches (BS < 70), except for one (BS, ML: 83, *rbcL* tree). As all differences between parsimony and ML analyses were only due to the lower resolution parsimony trees, parsimony results were summarised on ML trees as presented in Figs. 70 – 80.

Nineteen major lineages were recognized within our analysis, which correspond to the following taxa and groups: *N. capitoradiata, N. cassensis, N. cincta, N. cryptocephala, N. cryptotenella, N. cryptotenelloides, N. gregaria, N. jogensa, N. lundii, N. menisculus, N. peregrina, N. radiosa, N. radiosafallax, N. reinhardtii, N. rostellata, N. trivialis, N. veneta, N. viridulacalcis,* and *N. wildii.* Four other lineages were unidentified.

3.4.3 *rbcL* and *atpB*-based phylogenies

The individual gene tree for *rbcL* is presented in Figs. 3.70 – 3.72. In this tree, *Navicula* does not form a monophyletic group. Specimens *N. pseudoincerta* and three *N.* sp. (GenBank; accessions MN119676, MN119676, HQ33755 and HQ337555) are outside the *Haslea* outgroup.

The ingroup first separates into two major clades. The first major clade (Fig. 3.71, Clade A) further separates into three minor clades, which include *N. perminuta, N. vara, N. glacei*, and unidentified *Navicula* species, all from GenBank. The second major clade diverges further into three major clades, each of which separates further into multiple minor clades.

The second major clade (Fig. 3.71, Clade B) includes a well-supported minor clade with *N. tripunctata* (GenBank) and *N.* cf. *tripunctata* (this study) specimens (BS, ML/MP: 99/99), which is sister to a minor clade grouping *N. radiosafallax* (this study) and *N. lanceolata* (GenBank) (BS, ML/MP: 100/99). *Navicula capitoradiata* (GenBank and this study) (BS, ML/MP: 83/*) is sister to another minor clade including *N. radiosa* (GenBank and this study) and an unidentified *Navicula* species, which correspond to GenBank accession numbers KM999060, KM999057, KM999061, KM999058, and KM999062 (BS, ML/MP: 62/72). Both of these clades are closely related to *N.* cf. *subwalkerii* (this study). *Navicula viridulacalcis* (this study) (BS, ML/MP: 97/100) is sister to *N. rostellata* (Genbank and this study) (BS, ML/MP: 97/96).

The third major clade (Fig. 3.72, Clade C) includes *N. veneta* (GenBank and this study) (BS, ML/MP: 100/100), and two newly described species, *N. jogensis* and *N. cassensis* (BS, ML/MP: 99/99, 100/95). All of the *N. veneta* specimens from GenBank group with *N. veneta* from this study, and one *N.* sp. groups with *N. joguensa*.

The fourth major clade (Fig. 3.72, Clade D) features well-supported clades for *N. rhynchotella* (GenBank) (BS, ML/MP: 96/93), *N. peregrina* (this study) (BS, ML/MP: 100/100), *N. menisculus* (this study) (BS, ML/MP: 100/100), *N. cryptotenella* (GenBank and this study) (BS, ML/MP: 99/98), *N trivialis* (this study) (BS, ML/MP: 100/98) and *N. cryptocephala* (GenBank and this study) (BS, ML/MP: 99/95). *Navicula cryptotenelloides* (GenBank and this

study) specimens do not form a monophyletic group, but they are closely related to *N. cryptotenella*. *Navicula cryptocephala* is closely related to *N. wildii* (this study), *N. trivialis*, *N. cincta* (this study), and *N. lundii* (this study). *Navicula sclesviscensis* (GenBank) does not form a monophyletic group.

The individual gene tree for *atpB* is presented in Fig. 3.73. *Navicula* specimens do not form a monophyletic group. *Navicula cincta* and *N.* cf. *tripunctata* (from this study; Fig. 3.73, Clade A) are outside the *Haslea* outgroup. Other major trends are consistent with those from the *rbcL* tree, with the notable difference that *Navicula wildii* groups with *N. veneta*. New species *N. joguensa and N. cassensis* are again well-supported (Fig. 3.73, Clade B; BS, ML/MP: 100/100, 72/99).

The same close relationships are observed between *N. radiosa*, *N. capitoradiata*, and *N.* cf *subwalkerii* (Fig. 3.73, Clade C). *Navicula cryptocephala* is again close to *N. trivialis* and *N. lundii* (Fig. 3.73, Clade D).

The combined tree of both plastid genes is presented in Figs. 3.74 – 3.76. Here, *Haslea* specimens separate into two well-supported branches (BS, ML/MP: 99/*, 100/100). *Navicula* cf. *tripunctata* and *N. cincta* specimens (from this study) form a clade outside of a portion of the *Haslea* outgroup. *Navicula wildii* is sister to *N. veneta.* Other trends are in accordance to the individual gene trees.

3.4.4 18S-based phylogeny

The individual gene tree for 18S (Figs. 3.77 – 3.80) shows little to no support at most branches. *Navicula* does not form a monophyletic group, as multiple specimens group with

Haslea specimens. Supported monophyletic clades include *N. peregrina* (BS, ML/MP: 88/96) and *N. glacei* (BS, ML/MP: 88/71). *Navicula perminuta* and *N. vara* form a clade (BS, ML/MP: 62/68) that is closely related to *N. glacei,* which is also observed in the *rbcL* tree.

3.4.5 Genetic variability

Summary statistics for DNA sequences are given in Table 3.4. The *rbcL* marker showed the most variable and potentially informative sites of all three markers, and 18S showed the least. With all markers combined, sequence divergence was between 9.7 – 14.7% within the ingroup, and from 14.1 – 18.6% when the outgroup was included.

As suggested by phylogenetic trees, within the *N. veneta–N. jogensa–N. cassensis* species complex, *N. veneta* and *N. cassensis* are most closely related, with average divergences of 19 bp in *rbcL*, 15 bp in *atpB*, and < 1 bp in 18S. The *N.veneta* and *N. joguensa* clades are about equally as divergent as the *N. joguensa* and *N. cassensis* clades. These sister species respectively display average divergences of 30 and 28 bp in *rbcL*, 19 and 20 bp in *atpB*, and 1 bp in 18S. These values are comparable to those observed in other closely related species such as *N. radiosa* and *N. capitoradiata.*

3.4.6 Morphological analysis

The results of the NMDS are presented in Figure 3.81. In general, larger species (>50 μm) such as *N. perengrina, N. radiosa*, *N. virudulacalcis* and *N.* cf. *subwalkerii* show better separation than smaller species. *Navicula reinhardtii* and *N. trivialis* are also distant from other

species. Smaller *Navicula* taxa, such as *N. veneta, N. gregaria*, *N cryptocephala*, *N*. cf. *metareichardtiana*, *N. joguensa*, and *N. menisculus*, overlap almost entirely.

Navicula veneta and *N. joguensa* show significant overlap between each other. *Navicula cassensis* overlaps with *N.* cf. *tripunctata*.

In general, variability within each taxon is similar, with larger species such as *N.* cf. *subwalkerii*, *N. radiosa*, *N. trivialis*, and *N. radiosafallax* showing bigger variance clusters. There are no shape outliers within the taxa.

The results of the valve shape analysis of my *N. veneta* specimens are presented as a PCA in Figure 3.82. The type specimens for *N. veneta* are included based on measurements taken from the photographs provided by Lange-Bertalot (2001) (Plate 14, Fig. 23; Plate 65; Fig. 3). In total, 70.5% of the variation between specimens was explained by the first two PC axes (53.8%, axis 1; 16.7%, axis 2). The first axis represents valve width.

The range in shape for *N. veneta* is much larger than that for *N. jogensa* and *N. cassensis*. Some overlap is observed between *N. jogensa* and *N. cassensis*, while *N. veneta* is distinct. *Navicula veneta* is on the left side of axis 1, indicating that its frustules are broader, creating a different shape form than those of *N. jogensa* and *N. cassensis*. One of the type specimens for *N. veneta* falls into the shape range of *N. veneta* specimens from the present study, while the other is different than all three species.

3.5 Discussion

3.5.1 Monophyly of *Navicula*

The genus *Navicula* was created in 1822 by Bory de Saint-Vincent on the basis of its elliptical-lanceolate shape. Thereafter, all diatoms with this characteristic valve shape were placed in the genus. A subsequent treatment by Cleve (1895) divided the genus into seven sections. Today, it is agreed that the genus *Navicula* s. str. should only comprise species that fit Cleves's diagnosis of *Navicula* sect. *Lineolatae,* whose species share a number of characters, such as lineate areolae, straight raphes interrupted by central nodules, and short helictoglossae (Witkowski et al. 1998; Cox, 1999). In recent years, many species have been removed from the genus, but these removals have been criticised, especially when based on only one or two morphological features (Cox, 1999). For example, Witkowski (1997) transferred two species to *Fogedia* simply based on hyaline areas between striae and differences in raphe structures.

Navicula specimens do not form a monophyletic group in any of the trees presented in this study. In Nakov (2018), *Navicula* was also unnatural, with *Pseudogomphonema* and *Seminavis* specimens nesting among *Navicula* specimens. Bruder & Medlin (2008) obtained a similar result, with a *Neidium* specimen nesting within *Navicula* specimens in their LSU rDNA tree, and *Pseudogomphonema* and *Seminavis* specimens nesting in *Navicula* in their *rbcL* tree.

In our trees, *Halsea* specimens are nested within *Navicula*. It is therefore not surprising that species within *Haslea* were initially identified in the genus *Navicula* (Hustedt, 1961–1966). These two genera were separated when the genus *Haslea* was erected in 1974, on the basis of its spindle-shaped valves and striae organization (Simonsen, 1974). However, in recent times, transfers are still being made between the two genera. Li (2017) used three criteria to transfer two species to *Navicula* from *Haslea*: the presence of hook-shaped terminal fissures, the short length of the helictoglossae, and the thickening of valves at axial areas. Furthermore, Li (2017) proposed that true *Haslea* specimens always feature a distinctive "sandwich-like" valve structure

with two layers held together by saepes, an external surface featuring multiple longitudinal grooves with recessed areolae, long helictoglossae, apical raphe endings that end or bend before the apex of the valve, and proximal raphe endings that may bend but are never hooked (like the ones seen in *Navicula radiosa*; Fig. 3.52).

A re-examination of the micrographs of *Halsea nusantara* (MH681881; NC_044491) and *Haslea* sp. (KT257734) showed that they displayed the typical characters of true *Haslea* specimens as described above (Prasetiya et al. 2019; Gastineau et al. 2016). However, the presence of the "sandwich-like" structure could not be determined because broken valves are necessary to see this character. Broken valves were not depicted in the available plates. The other *Haslea* specimens in our trees were not re-examined as high resolution SEM micrographs were unavailable.

This suggests that, although the two genera appear to be morphologically distinct, they may be quite closely related on a molecular level. Alternatively, taxonomic identifications of specimens maybe in error or the genes used here may not be well suited to distinguish between the two.

The plastid genes *rbcL* and *atpB* generally performed well in our phylogenic analyses. However, Guo et al. (2015) found that less conserved genes, such as *rbcL*, tend to perform poorly at resolving higher level phylogenies. Although *atpB* was not included in Guo et al. (2015), our results demonstrate that it evolves at a rate similar to *rbcL*, as base pair (bp) differences within and between taxa are very similar. For example, within *N. radiosa*, *rbcL* shows bp differences from 0 to 17, while *atpB* shows bp differences between 0 and 18. Likewise, at the species level, *N. radiosa* and *N. capitoradiata*, show bp differences for *rbcL* between 15 and 18, while *atpB* shows bp differences from 6 to 14.

Guo et al. (2015) also noted that 18S is more conserved and is likely to perform better at higher taxonomic levels. In our 18S tree, two *Navicula* specimens, belonging to the species *N. minima* sensu Grunow (1880) and *N. pulchripora* I.Kaczmarska & A.M.Chan, group with our *Haslea* specimens. However, these species are no longer believed to be positioned within *Navicula sensu stricto.* At present, *N. minina* has been transferred to the genus *Sellaphora* since Wetzel et al. (2015) demonstrated that it matched the morphological concept of *Sellaphora seminulum.* Furthermore, they cited several studies which showed that *N. minima* and *Sellaphora* species formed a monophyletic group (Behnke et al. 2004; Bruder & Medlin 2008; Zimmermann et al. 2015). As for *N. pulchripora,* its position within *Navicula* has been questioned, since Kaczmarska & Chan (1995) found that it did not conform morphologically to *Navicula* section *Lineolatae*, which now forms *Navicula* sensu stricto. It is instead more similar to *Navicula* section *Microstigmaticae*, which is a highly heterogeneous group (Cox, 1988). The taxonomic status of the species placed in this section is uncertain. Other new genera, such as *Parlibellus*, have been erected from some of the species of section *Microstigmaticae* (Cox, 1989). It is expected that with more research, more genera will be defined from this section.

By contrast, the *Navicula* specimens found outside of *Halsea* in our *rbcL* and *atpB* trees are from *N. pseudoincerta* and *N. cincta*, species consistently found to match the concept of *Navicula* sensu stricto (Cox, 1999).

I therefore propose that the placement of these specimens outside of *Haslea* may be due to the poor performance of *rbcL* and *atpB* at higher taxonomic levels.

3.5.2 Genetic and morphological variability

Four major clades can be recognised in our plastid analyses. These clades are identical in the *rbcL*, *atpB*, and combined plastid trees, with the following exceptions: in the *atpB* and combined plastid trees, *N.* cf. *tripunctata* and *N. cincta* are in a clade outside of the *Haslea* outgroup, and *N. wildii* is in a clade sister to *N. veneta.* In contrast, in the *rbcL* tree, *N.* cf. *tripunctata* is in a different clade with *N. tripunctata* specimens, *N. cincta* is in a clade sister to a clade comprised of *N. trivialis* and *N. cryptocephala* specimens, and *N. wildii* is in a clade sister to *N. trivialis.*

The topology seen in plastid trees was not observed in our nuclear 18S tree, which was poorly resolved and received little branch support. Previous studies have shown that 18S performs better at higher taxonomic levels, such as genus-level, and does not appear to be well-suited for species-level phylogenies as seen here and in Guo et al. (2015).

Base pair differences within our major clades were comparable to those observed in *Frustulia* (Bouchard et al. 2019; Chapter 2). For example, the major clades in *Frustulia rbcL* trees typically displayed base pair divergences between 0 – 7.2%, whereas the major *rbcL* clades in our *Navicula* analyses showed divergence values ranging from 0 – 8.0%. Our *atpB* clades showed similar values – clades within the ingroup had divergence values ranging from 0 – 7.3%, while Clade A had a divergence value of 8.7%. This demonstrates that *rbcL* and *atpB* appear to evolve at similar rates.

The clades observed in our plastid trees were largely consistent with the results from Bruder & Medlin (2008). Using LSU (26S), SSU (18S), and *rbcL* data, they found three major clades in their analyses. These groups were entirely compatible with the clades observed in our trees, and any differences, such as the presence of a fourth major clade in this analysis, appear to

be due to our taxonomic sampling being more extensive. An updated overview of phylogenetic relationships within *Navicula* as compared to Bruder & Medlin (2008) is provided in Table 3.5.

Although molecular data were able to distinguish four major clades, no significant morphological differences were found between them. This is consistent with what Bruder & Medlin (2008) observed, as they also were unable to find morphological differences that characterized major *Navicula* lineages, even noting that shape differences were greater within each group than between groups. It is possible that no such characters exist, and that the molecular variation observed here does not translate into perceptible differences in external morphology.

3.5.3 Two new, pseudo-cryptic species, *N. joguensa* and *N. cassensis* (provisional names)

Navicula veneta, N. joguensa, and *N. cassensis* are very similar morphologically, and difficult to distinguish using morphology alone. An analysis of *N. cassensis* shows that it is distinct from *N. veneta* and *N. joguensa* by only a few fine morphological differences. In particular, *N. cassensis* specimens have slightly larger frustules with lower stria densities and more rounded central areas as compared to *N. veneta* and *N. joguensis* which have asymmetrical central areas, with one rectangular and one rounded side. Although *N. veneta* and *N. joguensis* could not be separated by discrete morphological characters, they could be effectively separated using frustule shape data. In our shape analysis, *N. joguensa* and *N. cassensis* overlap, but *N. veneta* is clearly separated. Its specimens have broader frustules, with apices that are more rostrate as opposed to the rounded apices seen in *N. joguensis* and *N. cassensis* specimens.

Navicula joguensis and *N. cassensis* diverged by 16 – 41 bp in *rbcL*, and by 13 – 21 bp in *atpB*. This is comparable to the divergence observed between other closely related *Navicula* species in this study. For example, our *N. radiosa* and *N. capitoradiata* specimens, diverged by 6 – 14 bp in *rbcL* and from 15 – 18 bp in *atpB*. Comparable divergences have also been found in other diatom genera. In Bouchard et al. (2019) and Chapter 2, the average divergence between the closely related species *Frustulia krammeri* and *F. crassinervia* was 20 bp. This further suggests that the taxa are distinct species.

These results demonstrate the usefulness of combining multiple data types in species-level taxonomic analyses of diatoms. Here, the combination of molecular data, fine-scale character morphology and frustule shape allowed us to uncover hidden diversity which might otherwise have been overlooked.

This was also observed in Bouchard et al. 2019 and Chapter 2, where specimens previously described as *F.* cf. *krammeri* based on morphology were found to be a distinct species due to molecular data. The newly described species, *F. gibsonea,* was not closely related to *F. krammeri,* despite morphological similarities.

It appears that *N. joguensa* and *N. cassensis* are pseudo-cryptic, and have previously been treated under *N. veneta*. In Lange-Bertalot (2001), for example, specimens in Figures 23 – 30 of Plate 14 are described as *N. veneta*; however, it is likely that some of these specimens represent our pseudo-cryptic species. In particular, the specimen depicted in their Figure 23 measures 25 µm in length, has a linear-lanceolate valve shape, a wide rectangular central area, and a stria density of 16 in 10 µm. These characters point to this specimen belonging to *N. joguensa.* However, in our shape analysis, the type material for *N. veneta* falls within the variation seen for my *N. veneta* specimens, suggesting that my *N. veneta* specimens are correctly identified.

Based on the facts outlined above, we describe *N. joguena* and *N. cassensis* as new species.

3.5.4 Possible unrecognised diversity in *Navicula*

My analyses identified four entities that may warrant recognition as species after further study. Three of these, which I identify as *N.* cf. *subwalkeri*, *N.* cf. *metareichardtiana*, and *N.* cf. *cincta*, were not described as new species here because only one or two specimens were found. Further work, with increased sampling, is required before these potentially new taxa could be confidently positioned within *Navicula*. More specimens are available for the fourth taxon, *N.* cf. *tripunctata*, but molecular data does not clearly position this taxon within *Navicula*. In our analyses, individuals of *N.* cf. *tripunctata* are placed between two outgroup clades in both our *atpB* and combined plastid trees, but they form a clade with *N. tripunctata* specimens from GenBank in our *rbcL* tree. Because of this contradiction, further work is required.

Navicula cf. *subwalkeri* is morphologically distinct from all other species in this study. Molecular data indicates that it is most closely related to *N. radiosa* and *N. capitoradiata*. The divergence between *N.* cf. *subwalkeri* and *N. radiosa* is from 10 – 38 bp in *rbcL* sequences, and from 15 – 18 bp in *atpB* sequences. Between *N.* cf. *subwalkeri* and *N. capitoradiata,* sequences diverge by 11 – 55 bp for *rbcL,* and by 21 – 22 bp for *atpB*. This is a range of divergence that is typical for closely related species in diatoms (Bouchard, 2019).

Navicula cf. *metareichardtiana* is morphologically similar to *N. gregaria* and to *N. veneta* and N. *joguensa,* which are species that are found in Clade B of the combined plastid tree. However, molecular data suggests that *N.* cf. *metareichardtiana* is closest to *N. cryptotenelloides*

in Clade C, and that the two differ by only 9 – 16 bp in their *rbcL* sequences. This low level of divergence suggests they are very closely related species. Nonetheless, there are significant morphological differences between the two with *N. cryptotenelloides* possessing significantly smaller frustules with small central areas than what is seen in *N.* cf. *metareichardtiana*. Further research into this taxon is required in order to determine whether it warrants recognition at the species level and to clarify its relationship to other *Navicula*.

Navicula cf. *cincta* is morphologically similar to *N. cincta* and to *N. radiosafallax,* although there is an important difference in valve size with the latter. In contrast, molecular data suggests *N.* cf. *cincta* is most closely related to *N. wildii* (divergence of 17 bp), *N. trivialis* (divergence of 12 – 32 bp), and *N. cryptocephala* (divergence of 14 – 18 bp). Unfortunately, only a single specimen was found for this taxon, and its *atpB* sequence could not be successfully amplified and sequenced. More work is required to determine its taxonomic status within *Navicula*.

N. cf. *tripunctata* is morphologically identical to *N. tripunctata.* However, *N. tripunctata* is a marine species, and our *N.* cf. *tripunctata* specimens were collected in freshwater locations. Furthermore, the *rbcL* sequences of our *N.* cf. *tripunctata* specimens diverge from those of *N. tripunctata* specimens obtained from GenBank by 49 – 87 bp, which is significantly higher than what is expected within a species. This indicates that there may be cryptic diversity within *N. tripunctata.* As *N. tripunctata* is the type species for *Navicula* s.s., this hidden diversity could have serious ramifications on our understanding of how the genus *Navicula* is defined. Just as it was observed that pseudo-cryptic species *N. joguensa* and *N. cassensis* were previously treated within *N. veneta,* it is possible that analyses comprising *N. tripunctata* have included or will include the cryptic specimens observed here. This inclusion certainly skews the species concept

of *N. tripunctata*, which could lead to an important redefinition of *Navicula* itself since this species is the type upon which *Navicula* is based.

Although *Navicula* was not monophyletic in our trees, this appeared to be due to the poor performance of molecular markers and misidentified GenBank specimens. Caution should be exercised when using specimens from third-party sources. Furthermore, the four-clade structure identified within *Navicula* using phylogenetic trees could not be reproduced using morphological data alone. The molecular variation observed between these clades does not appear to translate into significant differences in external morphology. Two new species, *N. joguensa* and *N. cassensis*, were described. Their specimens had been previously treated under *N. veneta*, but the combination of multiple data types used here demonstrates that they are distinct species. Four potentially cryptic taxa, for which further study is required, were observed. It is likely that there is further hidden diversity within *Navicula,* which could be uncovered using methods that allow for the combination of DNA sequence data and traditional morphological data such as the protocol used here.

Table 3.1 Location and collection date for sediment samples used in this study. Species names followed by an asterisk "*" were cultured as described in the Materials and Methods of Chapter 3.

CANA Accession Number	Taxon	Isolation Number	Date Collected	Latitude	Longitude	Country	Province/ State	Description of habitat
93607	*N. peregrina* *N. radiosa*	1C1 1B1	2013-06-01	45.28719	-75.81302	Canada	Ontario	Ditch – alongside road
93655	*N. wildii* *N. radiosa* *N. radiosa*	6B2 6B5 6B7	2013-09-07	45.37658	-75.36088	Canada	Quebec	Lake
109928	*N. tripunctata* *N. tripunctata*	27E1 27E5	2015-06-14	45.24938	-75.79074	Canada	Ontario	Jock river – alongside road
109930	*N. radiosa*	27F6	2015-06-14	45.44628	-75.81279	Canada	Ontario	Ditch – draining a quarry
109967	*N. radiosa* *N. radiosa*	38E4 38E6	2015-07-11	44.111978	-77.35316	Canada	Ontario	Trent River – out from small swamp area
125938	*N. radiosafallax*	60E3	2016-07-24	43.844996	-74.40231	United States	New York	Lake Durant – in Adirondack Park
125948	*N. radiosa* *N. radiosa* *N. radiosa* *N. radiosa* *N. radiosa*	56C5 56C6 56C7 56D4 56D1	2016-09-04	62.444128	-114.3501	Canada	Northwest Territories	Great Slave Lake – trail from Rat Lake
125957	*N. lundii*	57C8	2016-09-05	62.54877	-113.9314	Canada	Northwest Territories	Wetland – alongside road
125992	*N. radiosa*	59C7	2016-09-07	62.54525	-114.4021	Canada	Northwest Territories	Pond – at the end of Martin Lake
126310	*N.* cf. *subwalkeri*	66B5	2017-05-05	49.3126	-83.4919	Canada	Ontario	Copell Lake – at docking station

Table 3.1 - Continued

CANA Accession Number	Taxon	Isolation Number	Date Collected	Latitude	Longitude	Country	Province/ State	Description of habitat
126313	*N. radiosa* *N.* cf. *subwalkeri*	66C6 66C4	2017-05-05	49.7315	-84.298	Canada	Ontario	Rabbit Lake
127091	*N. radiosa**	103E7	2017-08-08	45.443	-75.8104	Canada	Ontario	Pond
127814	*N. trivialis*	71D3	2017-06-26	45.36056	-75.13569	Canada	Ontario	Larose Forest
127816	*N. tripunctata*	71E7	2017-07-04	45.272544	-75.21831	Canada	Ontario	Little Castor River
127817	*N. trivialis*	71E8	2017-07-04	45.272544	-75.21831	Canada	Ontario	Payne River
127832	*N. viridulacalcis*	71F7	2017-07-10	45.360566	-75.13569	Canada	Ontario	Larose Forest
128280	*N. joguensa** *N. joguensa**	73B6 73B7	2018-05-19	49.6448	-83.6835	Canada	Ontario	Creek
128281	*N.* cf. *cincta** *N. cincta**	73D8 73C2	2018-05-19	49.3833	-83.4234	Canada	Ontario	Pond
128283	*N. radiosa** *N.* cf. *metareichardtiana**	73C7 73C4	2018-05-19	49.4352	-83.4553	Canada	Ontario	Pond
128286	*N. cassensis**	73C3	2018-05-19	49.445	-84.1145	Canada	Ontario	Samuel Lake
128291	*N. veneta**	73D2	2018-05-20	47.4858	-80.1613	Canada	Ontario	Swan Lake – at boat launch

Table 3.1 - Continued

CANA Accession Number	Taxon	Isolation Number	Date Collected	Latitude	Longitude	Country	Province/ State	Description of habitat
128292	*N. veneta**	73F8	2018-05-28	45.4139	-75.6684	Canada	Ontario	Ottawa River – trail near Hurdamn Station
	*N. gregaria**	73E7						
	*N. cryptotenelloides**	73F5						
	*N. cryptotenella**	73F2						
	*N. menisculus**	73F1						
	*N. menisculus**	73E8						
128293	*N. radiosa*	101F7	2018-05-28	45.4027	-75.6396	Canada	Ontario	Storm Water Pond
	*N. radiosa**	103A5						
	N. peregrina	101F6						
128298	*N. rostellata**	103E6	2018-08-27	45.452468	-75.72616	Canada	Quebec	Leamy Lake
128299	*N. lundii**	103E5	2018-08-19	45.639	-75.603	Canada	Quebec	Pond
128339	*N. capitoradiata**	103A8	2018-06-19	44.689036	-75.49243	United States	New York	Oswegatchie River – at boat launch
	*N. capitoradiata**	103B1						
128450	*N. trivialis*	62D7	2016-10-24	45.22011	-75.46442	Canada	Ontario	South Nation River –at watershed
128451	*N. peregrina*	70B5	2018-06-03	44.849897	-75.53375	Canada	Ontario	Ditch
128518	*N.* cf. *subwalkeri*	49B2	unkown	unkown	unkown	Canada	Northwest territories	unkown
128596	*N. radiosa*	56A3	2016-06-27	44.753855	-76.51201	Canada	Ontario	Farren Lake

Table 3.1 - Continued

CANA Accession Number	Taxon	Isolation Number	Date Collected	Latitude	Longitude	Country	Province/ State	Description of habitat
128597	*N. tripunctata*	100F5		45.147785	-75.648	Canada	Ontario	Rideau River – at boat launch
	N. radiosa	100F7						
	N. trivialis	104C7						
	N. peregrina	101B1						
128598	*N. joguensa*	101B7	2018-06-03	45.98255	-75.67504	Canada	Ontario	Wetland – alongside ride
128599	*N. cassensis**	101F3	2018-06-03	44.919252	-75.63699	Canada	Ontario	Creek – alongside road
	*N. cassensis**	101E8						
	*N. radiosa**	101F1						
128711	*N. viridulacalcis*	83A8	2017-08-03	45.339130	-75.16570			Wolf Creek
128716	*N. viridulacalcis*	83C4	2017-08-03	45.486481	-75.08106			Harris Creek
128726	*N. viridulacalcis*	83D7	2017-08-21	45.339130	-75.16570			Wolf Creek
	N. cryptocephala	83F5						
128761	*N. trivialis**	103B8	2018-07-23	45.272544	-75.21831	Canada	Ontario	Payne River
128732	*N. viridulacalcis*	102A8	2018-06-25	45.360566	-75.13569	Canada	Ontario	Larose Forest
128733	*N. cassensis**	103B4	2018-06-25	45.299122	-75.10783	Canada	Ontario	Butternut Creek
	*N. cassensis**	103D2						
	N. trivialis	102B6						

Table 3.1 - Continued

CANA Accession Number	Taxon	Isolation Number	Date Collected	Latitude	Longitude	Country	Province/ State	Description of habitat
128734	*N. veneta**	105B1	2018-06-25	45.278784	-75.15094	Canada	Ontario	Little Castor River - drainage
	*N. veneta**	105B2						
128737	*N. cassensis**	105B5	2018-06-25	45.265119	-75.15761	Canada	Ontario	Whissel Creek - drainage
	*N. trivialis**	105B4						
128738	*N. veneta**	105C2	2018-06-25	45.272544	-75.21831	Canada	Ontario	Payne River
	*N. radiosa**	102D5						
	*N. trivialis**	105C1						
128739	*N. joguensa**	105B8	2018-06-25	45.272544	-75.21831	Canada	Ontario	Little Castor River
128759	*N. trivialis**	103E4	2018-07-23	45.360566	-75.13569	Canada	Ontario	Larose Forest
128765	*N. veneta**	105F2	2018-07-27	45.36443	-75.66362	Canada	Ontario	Sawmill Creek Wetalnds
128773	*N. veneta**	103E1	2018-08-07	45.278784	-75.15094	Canada	Ontario	Little Castor River - drainage
128792	*N. trivialis*	105F8	2018-08-22	45.26464	-75.73172	Canada	Ontario	Storm Water Pond – Chapman Mills
128809	*N. trivialis**	103C6	2018-07-26	43.685958	-79.3667	Canada	Ontario	Pond - Don Valley Brick Works Park
128810	*N. veneta**	103C4	2018-07-27	43.667659	-79.35905	Canada	Ontario	Pond – Riverdale Farm
128811	*N. cassensis**	103C7	2018-07-27	43.667386	-79.36036	Canada	Ontario	Duck Pond – Riverdale Farm

Table 3.1 - Continued

CANA Accession Number	Taxon	Isolation Number	Date Collected	Latitude	Longitude	Country	Province/ State	Description of habitat
128813	*N. veneta**	73F7	2018-06-03	45.249417	-75.7906	Canada	Ontario	Jock River – alongside road
129302	*N. joguensa**	103A3	2018-05-31	45.443785	-75.81287	Canada	Quebec	Stream – draining quarry
129303	*N. cryptocephala**	101F2	2018-06-03	44.964901	-75.67504	Canada	Ontario	Kemptville Creek
Not assigned	*N. reinhardtii*	77F7	2018-07-07	40.94749	20.90391	Macedonia		Lake Prespa
	N. reinhardtii	77F8						

Table 3.2. GenBank accession numbers for specimens used in this study.

Taxon	*rbcL*	*atpB*	18S	Source
Gyrosigma acuminatum	HQ912462			Theriot et al. 2010
Gyrosigma acuminatum		KM009662		Theriot et al. 2015
Gyrosigma acuminatum			HQ912598	Theriot et al. 2010
Gyrosigma limosum	KY320287			An et al. 2017
Gyrosigma limosum			KY320348	An et al. 2017
Haslea nusantara		NC_044491		Prasetiya et al. 2019
Haslea nusantara		MH681881		Prasetiya et al. 2019
Haslea ostrearia	HE663062			Gastineau et al. 2013
Haslea ostrearia			AY485523	Damste et al. 2004
Haslea sp.	KT257734			Gastineau et al. 2016
Haslea spicula			HM805034	Pniewski et al. 2010
Navicula agatkae	KY320292			An et al. 2017
Navicula agatkae			KY320353	An et al. 2017
Navicula arenaria			KJ961668	Medlin, 2014
Navicula arenaria			KU561174	Yang et al. Unpublished
Navicula atomus var. *permitis*			AJ867024	Rimet et al. Unpublished
Navicula bottnica	JX905666			Hamsher & Saunders, 2014
Navicula brockmannii			AM502020	Bruder & Medlin, 2007
Navicula bruneli	JX905667			Hamsher & Saunders, 2014
Navicula capitatoradiata	KT072920			Keck et al. 2016
Navicula capitatoradiata	AM710479			Bruder & Medlin, 2007
Navicula capitatoradiata			AM502012	Bruder & Medlin, 2007
Navicula capitatoradiata			KT072982	Keck et al. 2016
Navicula cari	AM710457			Bruder & Medlin, 2007
Navicula cari			AM501991	Bruder & Medlin, 2007
Navicula cari			GU295220	Dang et al. Unpublished
Navicula cari			KY054972	Huang et al. 2017
Navicula cf. *cryptocephala*	KM999112			Hamilton et al. 2015
Navicula cf. *duerrenbergiana*	MG968533			De Decker et al. 2018
Navicula cf. *duerrenbergiana*	AY571749			Jones et al. 2005
Navicula cf. *lundii*	KT072906			Keck et al. 2016
Navicula cf. *lundii*			KT072956	Keck et al. 2016
Navicula cf. *normaloides*			KU867915	Schlie & Karsten, 2016
Navicula cf. *perminuta*	HQ337542			Hamsher et al. 2011
Navicula cf. *perminuta*	HQ337541			Hamsher et al. 2011
Navicula cf. *salinarum*	KY320293			An et al. 2017
Navicula cf. *salinarum*			KY320354	An et al. 2017
Navicula cincta	KT072934			Keck et al. 2016
Navicula cincta			KT072988	Keck et al. 2016
Navicula clavata			KT757317	Eswaran et al. 2016

Table 3.2 Continued

Taxon	*rbcL*	*atpB*	18S	Source
Navicula cryptocephala	KC736601			Kermarrec et al. 2013
Navicula cryptocephala	HQ912467			Theriot et al. 2010
Navicula cryptocephala	KM084946			Zimmermann et al. 2014
Navicula cryptocephala	KM084944			Zimmermann et al. 2014
Navicula cryptocephala	AM710463			Bruder & Medlin, 2007
Navicula cryptocephala	AM710439			Bruder & Medlin, 2007
Navicula cryptocephala	HQ337543			Hamsher et al. 2011
Navicula cryptocephala		KM009667		Theriot et al. 2015
Navicula cryptocephala			AJ639850	Sánchez et al. 2006
Navicula cryptocephala			AM501973	Bruder & Medlin, 2007
Navicula cryptocephala			AM501996	Bruder & Medlin, 2007
Navicula cryptocephala			HQ912603	Theriot et al. 2010
Navicula cryptocephala			KC736631	Kermarrec et al. 2013
Navicula cryptocephala			KU561108	Yang et al. Unpublished
Navicula cryptocephala			KU561183	Yang et al. Unpublished
Navicula cryptocephala			KY054973	Huang et al. 2017
Navicula cryptocephala var. *veneta*			AJ297724	Beszteri et al. 2001
Navicula cryptocephala var. *veneta*			JQ610162	Novis et al. 2012
Navicula cryptocephala var. *veneta*			KF482040	Brinkmann et al. Unpublished
Navicula cryptocephala var. *veneta*			KU561116	Yang et al. Unpublished
Navicula cryptocephala var. *veneta*			KU561199	Yang et al. Unpublished
Navicula cryptocephala var. *veneta*			KX257364	Saba et al. 2016
Navicula cryptocephala var. *veneta*			KY054975	Huang et al. 2017
Navicula cryptotenella	KM084994			Zimmermann et al. 2014
Navicula cryptotenella	AM710496			Bruder & Medlin, 2007
Navicula cryptotenella	AM710478			Bruder & Medlin, 2007
Navicula cryptotenella	AM710482			Bruder & Medlin, 2007
Navicula cryptotenella			AM502011	Bruder, Unpublished
Navicula cryptotenella			AM502015	Bruder & Medlin, 2007
Navicula cryptotenella			AM502029	Bruder & Medlin, 2007
Navicula cryptotenella			KF417683	Brinkmann et al. Unpublished
Navicula cryptotenella			KU561109	Brinkmann et al. Unpublished
Navicula cryptotenelloides	MK639353			Miao & Kim, Unpublished
Navicula cryptotenelloides	KT072927			Keck et al. 2016
Navicula cryptotenelloides			KT072980	Keck et al. 2016
Navicula diserta			AJ535159	Medlin & Kaczmarska, 2004
Navicula flagellifera	KY320296			An et al. 2017

Table 3.2 Continued

Taxon	*rbcL*	*atpB*	18S	Source
Navicula flagellifera			KY320357	An et al. 2017
Navicula glaciei	MH670902			Liu, Unpublished
Navicula glaciei	HQ337545			Hamsher et al. 2011
Navicula glaciei	HQ337544			Hamsher et al. 2011
Navicula glaciei			EF106788	Lee et al. Unpublished
Navicula glaciei			EF106789	Lee et al. Unpublished
Navicula gregaria	KY320297			An et al. 2017
Navicula gregaria	AM710440			Bruder & Medlin, 2007
Navicula gregaria	KM084960			Zimmermann et al. 2014
Navicula gregaria	KM084954			Zimmermann et al. 2014
Navicula gregaria	KM084949			Zimmermann et al. 2014
Navicula gregaria			AM501974	Bruder & Medlin, 2007
Navicula gregaria			HM805037	Pniewski et al. 2010
Navicula gregaria			KF482041	Brinkmann et al. Unpublished
Navicula gregaria			KU341754	Akagha, Unpublished
Navicula gregaria			KU561150	Yang et al. Unpublished
Navicula gregaria			KU561154	Yang et al. Unpublished
Navicula gregaria			KY054974	Huang et al. 2017
Navicula gregaria			KY320358	An et al. 2017
Navicula hippodontofallax			KT943636	Witkowski et al. 2016
Navicula incertata	KY320298			An et al. 2017
Navicula incertata			KY320359	An et al. 2017
Navicula lanceolata	FJ002148			Rampen et al. Unpublished
Navicula lanceolata			AY485484	Damste et al. 2004
Navicula lanceolata			KT861066	Le Gall et al. Unpublished
Navicula lanceolata			KT861071	Le Gall et al. Unpublished
Navicula microdigitoradiata	MG968534			De Decker et al. 2018
Navicula minima			HM449712	Kim Tiam et al. 2012
Navicula pelliculosa			AJ544657	Behnke et al. 2015
Navicula pelliculosa			AY485454	Damste et al. 2004
Navicula pelliculosa			EU260468	Ki et al. 2009
Navicula perminuta	KY320300			An et al. 2017
Navicula perminuta	KY320299			An et al. 2017
Navicula perminuta	KP400287			Stachura-Suchoples et al. 2015
Navicula perminuta	KP400283			Stachura-Suchoples et al. 2015
Navicula perminuta	KP400282			Stachura-Suchoples et al. 2015
Navicula perminuta	KP400281			Stachura-Suchoples et al. 2015
Navicula perminuta	KP400284			Stachura-Suchoples et al. 2015

Table 3.2 Continued

Taxon	*rbcL*	*atpB*	18S	Source
Navicula perminuta	JQ432375			Zhao, Unpublished
Navicula perminuta	MN097772			Prelle et al. 2019
Navicula perminuta	KM592938			Zimmermann et al. 2015
Navicula perminuta			HM805043	Pniewski et al. 2010
Navicula perminuta			HM805044	Pniewski et al. 2010
Navicula perminuta			JQ045340	Zhao, Unpublished
Navicula perminuta			KJ961666	Medlin, 2014
Navicula perminuta			KU561132	Yang et al. Unpublished
Navicula perminuta			KY320360	An et al. 2017
Navicula perminuta			KY320361	An et al. 2017
Navicula phyllepta	EU938320			Vanelslander et al. 2009
Navicula phyllepta	EU938319			Vanelslander et al. 2009
Navicula phyllepta	EU938318			Vanelslander et al. 2009
Navicula phyllepta	EU938317			Vanelslander et al. 2009
Navicula phyllepta	EU938316			Vanelslander et al. 2009
Navicula phyllepta	EU938315			Vanelslander et al. 2009
Navicula phyllepta	EU938314			Vanelslander et al. 2009
Navicula phyllepta	EU938312			Vanelslander et al. 2009
Navicula phyllepta	EU938313			Vanelslander et al. 2009
Navicula phyllepta	FJ624251			Vanelslander et al. 2009
Navicula phyllepta	FJ624250			Vanelslander et al. 2009
Navicula phyllepta	FJ624249			Vanelslander et al. 2009
Navicula phyllepta	FJ624248			Vanelslander et al. 2009
Navicula phyllepta	FJ624247			Vanelslander et al. 2009
Navicula phyllepta	FJ624246			Vanelslander et al. 2009
Navicula phyllepta	FJ624245			Vanelslander et al. 2009
Navicula phyllepta	FJ624244			Vanelslander et al. 2009
Navicula phyllepta	FJ624243			Vanelslander et al. 2009
Navicula phyllepta	FJ624242			Vanelslander et al. 2009
Navicula phyllepta			AY485456	Damste et al. 2004
Navicula phyllepta			EU938307	Vanelslander et al. 2009
Navicula phyllepta			EU938308	Vanelslander et al. 2009
Navicula phyllepta			EU938309	Vanelslander et al. 2009
Navicula phyllepta			EU938310	Vanelslander et al. 2009
Navicula phyllepta			EU938311	Vanelslander et al. 2009
Navicula phyllepta			FJ624231	Vanelslander et al. 2009
Navicula phyllepta			FJ624232	Vanelslander et al. 2009
Navicula phyllepta			FJ624233	Vanelslander et al. 2009
Navicula phyllepta			FJ624234	Vanelslander et al. 2009
Navicula phyllepta			FJ624235	Vanelslander et al. 2009
Navicula phyllepta			FJ624236	Vanelslander et al. 2009

Table 3.2 Continued

Taxon	*rbcL*	*atpB*	18S	Source
Navicula phyllepta			FJ624237	Vanelslander et al. 2009
Navicula phyllepta			FJ624238	Vanelslander et al. 2009
Navicula phyllepta			FJ624239	Vanelslander et al. 2009
Navicula phyllepta			FJ624240	Vanelslander et al. 2009
Navicula phyllepta			FJ624241	Vanelslander et al. 2009
Navicula phyllepta			FJ624253	Vanelslander et al. 2009
Navicula phyllepta			FJ640068	Vanelslander et al. 2009
Navicula phyllepta			KC178569	Sanyo et al. Unpublished
Navicula phyllepta			KJ961667	Medlin, 2014
Navicula phyllepta			KU341750	Akagha, Unpublished
Navicula phyllepta			KU341752	Akagha, Unpublished
Navicula phyllepta			KU561104	Yang et al. Unpublished
Navicula pseudacceptata	JQ432376			Zhao, Unpublished
Navicula pseudacceptata			JN674064	Zhao, Unpublished
Navicula pseudoincerta	MN119676			Dong & Zhang, Unpublished
Navicula pulchripora			KF177774	Kirkwood et al. Unpublished
Navicula radiosa	KM084955			Zimmermann et al. 2014
Navicula radiosa	AM710438			Bruder & Medlin, 2007
Navicula radiosa	AM710501			Bruder & Medlin, 2007
Navicula radiosa	AM710494			Bruder & Medlin, 2007
Navicula radiosa			AM501972	Bruder & Medlin, 2007
Navicula radiosa			AM502027	Bruder & Medlin, 2007
Navicula radiosa			AM502034	Bruder & Medlin, 2007
Navicula ramosissima	HQ337548			Hamsher et al. 2011
Navicula ramosissima	JX905668			Hamsher & Saunders, 2014
Navicula ramosissima	KY320302			An et al. 2017
Navicula ramosissima	KY320301			An et al. 2017
Navicula ramosissima	FJ002097			Rampen et al. Unpublished
Navicula ramosissima			AY485512	Damste et al. 2004
Navicula ramosissima			KY320362	An et al. 2017
Navicula ramosissima			KY320363	An et al. 2017
Navicula reinhardtii	AM710442			Bruder & Medlin, 2007
Navicula reinhardtii			AM501976	Bruder & Medlin, 2007
Navicula reinhardtii			KU561136	Yang et al. Unpublished
Navicula rhynchocephala var. *hankensis*	JQ432374			Zhao, Unpublished
Navicula rhynchocephala var. *hankensis*			JQ045339	Zhao, Unpublished
Navicula rhynchotella	KM084953			Zimmermann et al. 2014
Navicula rhynchotella	KM084952			Zimmermann et al. 2014
Navicula rhynchotella	KM084950			Zimmermann et al. 2014
Navicula rhynchotella	KM084951			Zimmermann et al. 2014

Table 3.2 Continued

Taxon	*rbcL*	*atpB*	18S	Source
Navicula salinarum	KY320303			An et al. 2017
Navicula salinarum			KY320364	An et al. 2017
Navicula salinarum f. minima	KY320304			An et al. 2017
Navicula salinarum f. minima			KY320365	An et al. 2017
Navicula salinicola	KY320305			An et al. 2017
Navicula salinicola	HQ337550			Hamsher et al. 2011
Navicula salinicola	HQ337549			Hamsher et al. 2011
Navicula salinicola	AY604699			Fox & Sorhannus, 2003
Navicula salinicola	HQ337546			Hamsher et al. 2011
Navicula salinicola			GQ219689	Sorhannus & Fox, 2012
Navicula salinicola			KY320366	An et al. 2017
Navicula salinicola	JQ003571			Stoof-Leichsenring et al. 2012
Navicula saprophila			AJ867025	Rimet et al. Unpublished
Navicula sclesviscensis	FJ002098			Rampen et al. Unpublished
Navicula sclesviscensis			AY485483	Damste et al. 2004
Navicula slesvicensis	KM084942			Zimmermann et al. 2014
Navicula sp.	HQ337552			Hamsher et al. 2011
Navicula sp.	MH125182			Rachmayanti et al. Unpublished
Navicula sp.	KY320295			An et al. 2017
Navicula sp.	KY320294			An et al. 2017
Navicula sp.	KY320306			An et al. 2017
Navicula sp.	KY320309			An et al. 2017
Navicula sp.	KY320308			An et al. 2017
Navicula sp.	KM999105			Hamilton et al. 2015
Navicula sp.	KY320307			An et al. 2017
Navicula sp.	KY320310			An et al. 2017
Navicula sp.	AB430695			Sato et al. Unpublished
Navicula sp.	EU090048			Choi et al. 2008
Navicula sp.	EU090047			Choi et al. 2008
Navicula sp.	KP400305			Stachura-Suchoples et al. 2015
Navicula sp.	KM999062			Hamilton et al. 2015
Navicula sp.	KM999061			Hamilton et al. 2015
Navicula sp.	KM999060			Hamilton et al. 2015
Navicula sp.	KM999058			Hamilton et al. 2015
Navicula sp.	KM999057			Hamilton et al. 2015
Navicula sp.	KM999059			Hamilton et al. 2015
Navicula sp.	MH040272			Lobban et al. 2018
Navicula sp.	MH064106			Sabir et al. 2018
Navicula sp.	MH064107			Sabir et al. 2018

Table 3.2 Continued

Taxon	*rbcL*	*atpB*	18S	Source
Navicula sp.	MH064105			Sabir et al. 2018
Navicula sp.	MH064104			Sabir et al. 2018
Navicula sp.	MH064108			Sabir et al. 2018
Navicula sp.	MH040271			Lobban et al. 2018
Navicula sp.	KT943661			Witkowski et al. 2016
Navicula sp.	MH064103			Sabir et al. 2018
Navicula sp.	KM999111			Hamilton et al. 2015
Navicula sp.	KM999064			Hamilton et al. 2015
Navicula sp.	KM999110			Hamilton et al. 2015
Navicula sp.	KM999109			Hamilton et al. 2015
Navicula sp.	KM999108			Hamilton et al. 2015
Navicula sp.	KM999107			Hamilton et al. 2015
Navicula sp.	KM999106			Hamilton et al. 2015
Navicula sp.	KM999063			Hamilton et al. 2015
Navicula sp.	MH064109			Sabir et al. 2018
Navicula sp.	FJ002133			Rampen et al. Unpublished
Navicula sp.	MK454985			Kryk et al. Unpublished
Navicula sp.	FN392686			Moro et al. 2010
Navicula sp.	AM710474			Bruder & Medlin, 2007
Navicula sp.	KT943657			Witkowski et al. 2016
Navicula sp.	KT943681			Witkowski et al. 2016
Navicula sp.	KT943660			Witkowski et al. 2016
Navicula sp.	KT943658			Witkowski et al. 2016
Navicula sp.	KT943659			Witkowski et al. 2016
Navicula sp.	AM710466			Bruder & Medlin, 2007
Navicula sp.	KJ577899			Nakov et al. 2015
Navicula sp.	HQ337553			Hamsher et al. 2011
Navicula sp.	MH125179			Rachmayanti et al. Unpublished
Navicula sp.	HQ337554			Hamsher et al. 2011
Navicula sp.	HQ337555			Hamsher et al. 2011
Navicula sp.	HQ337551			Hamsher et al. 2011
Navicula sp.			AB430615	Sato et al. Unpublished
Navicula sp.			AM501999	Bruder & Medlin, 2007
Navicula sp.			AM502007	Bruder & Medlin, 2007
Navicula sp.			AY485460	Damste et al. 2004
Navicula sp.			AY485502	Damste et al. 2004
Navicula sp.			AY485513	Damste et al. 2004
Navicula sp.			EF106790	Lee et al. Unpublished
Navicula sp.			FJ546706	Beer et al. Unpublished
Navicula sp.			FN398345	Moro et al. 2010

Table 3.2 Continued

Taxon	*rbcL*	*atpB*	18S	Source
Navicula sp.			JF708144	Subin et al. Unpulished
Navicula sp.			KC840690	Xu & Wang, Unpublished
Navicula sp.			KF417679	Brinkmann et al. Unpublished
Navicula sp.			KF417680	Brinkmann et al. Unpublished
Navicula sp.			KF417682	Brinkmann et al. Unpublished
Navicula sp.			KJ577862	Nakov et al. 2015
Navicula sp.			KJ961665	Medlin, 2014
Navicula sp.			KM998987	Hamilton et al. 2015
Navicula sp.			KM998988	Hamilton et al. 2015
Navicula sp.			KM998989	Hamilton et al. 2015
Navicula sp.			KM998990	Hamilton et al. 2015
Navicula sp.			KM998991	Hamilton et al. 2015
Navicula sp.			KM998992	Hamilton et al. 2015
Navicula sp.			KM998993	Hamilton et al. 2015
Navicula sp.			KM998994	Hamilton et al. 2015
Navicula sp.			KT861019	Le Gall et al. Unpublished
Navicula sp.			KT943631	Witkowski et al. 2016
Navicula sp.			KX253954	Moreno, 2015
Navicula sp.			KY094989	Guillou, Unpublished
Navicula sp.			KY320355	An et al. 2017
Navicula sp.			KY320356	An et al. 2017
Navicula sp.			KY320367	An et al. 2017
Navicula sp.			KY320368	An et al. 2017
Navicula sp.			KY320369	An et al. 2017
Navicula sp.			KY320370	An et al. 2017
Navicula sp.			KY320371	An et al. 2017
Navicula sp.			LC189083	Kataoka et al. Unpublished
Navicula sp.			MK310104	Loftus & Johnson, 2019
Navicula sp.	EF143288			Evans et al. 2008
Navicula subminuscula			AJ867026	Rimet et al. Unpublished
Navicula symmetrica	KT072930			Keck et al. 2016
Navicula symmetrica			KT072983	Keck et al. 2016
Navicula tripunctata	KT072925			Keck et al. 2016
Navicula tripunctata	KM084986			Zimmermann et al. 2014
Navicula tripunctata	KM084936			Zimmermann et al. 2014
Navicula tripunctata	KM084935			Zimmermann et al. 2014
Navicula tripunctata	AM710495			Bruder & Medlin, 2007
Navicula tripunctata			AM502028	Bruder & Medlin, 2007
Navicula tripunctata			KF417681	Brinkmann et al. Unpublished
Navicula tripunctata			KT072979	Keck et al. 2016
Navicula tripunctata			KU561179	Yang et al. Unpublished

Table 3.2 Continued

Taxon	*rbcL*	*atpB*	18S	Source
Navicula tripunctata			MF992125	Hawk & Geller, 2019
Navicula trivialis	KY320311			An et al. 2017
Navicula trivialis			KY320372	An et al. 2017
Navicula vara	JX853711			Xiao, Unpublished
Navicula vara			JQ045338	Zhao, Unpublished
Navicula vara	JX863893			Bruder & Medlin, 2007
Navicula veneta	AM710441			Bruder & Medlin, 2007
Navicula veneta	AM710437			Bruder & Medlin, 2007
Navicula veneta	AM710436			Bruder & Medlin, 2007
Navicula veneta			AM501970	Bruder & Medlin, 2007
Navicula veneta			AM501971	Bruder & Medlin, 2007
Navicula veneta			AM501975	Bruder & Medlin, 2007
Navicula viridula var. *rostellata*	KT072916			Keck et al. 2016
Navicula viridula var. *rostellata*			KT072966	Keck et al. 2016
Navicula zhengii			KT943632	Witkowski et al. 2016
Navicula zhengii			KT943633	Witkowski et al. 2016
Navicula zhengii			KT943634	Witkowski et al. 2016
Navicula zhengii			KT943635	Witkowski et al. 2016
Uncultured *Navicula*			GQ324976	Arp et al. 2010
Uncultured *Navicula*			LC510613	Warqaa & Fikrat, Unpublished
Uncultured *Navicula* clone			GQ324977	Arp et al. 2010
Uncultured *Navicula* sp.			FR681847	Ghosh, Unpublished
Uncultured *Navicula* sp.			FR681848	Ghosh, Unpublished
Uncultured *Navicula* sp.			FR681849	Ghosh, Unpublished
Uncultured *Navicula* sp.			FR681850	Ghosh, Unpublished
Uncultured *Navicula* sp.			FR681851	Ghosh, Unpublished
Uncultured *Navicula* sp.			FR681852	Ghosh, Unpublished

Table 3.3. Primers used in nested PCR amplifications.

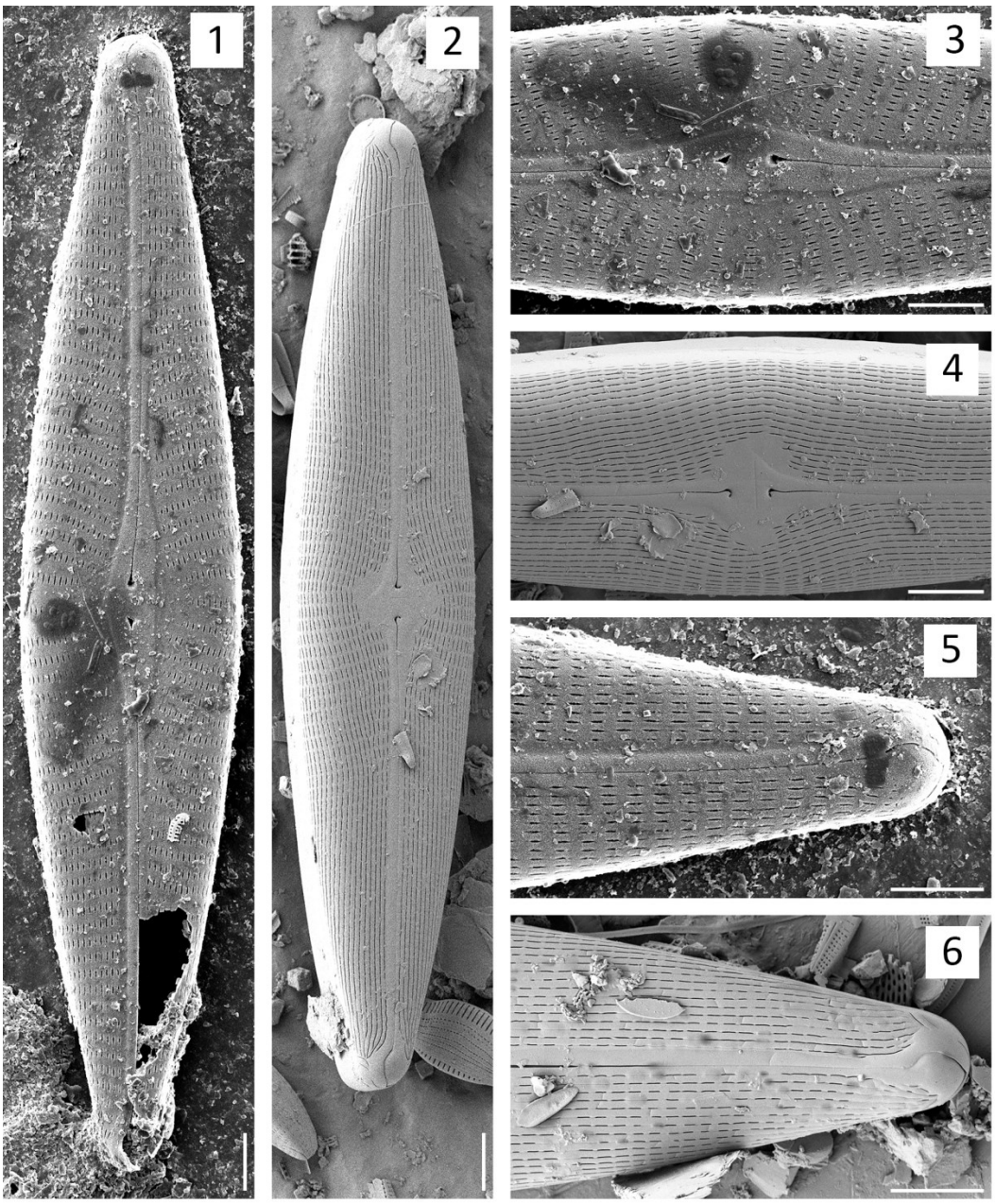
Primer type	Name	Sequence (5'-3')	Primer direction	Source
External	rbcL66+	TTTAAGGAGAAATAAATGTCTCAATCTG	F	Alverson et al., 2007; Daugbjerg and Anderson, 1997
	rbcLdp7	AAASHDCCTTGTGTWAGTVTC	R	
	18SP2F	CTGGTTGATTCTGCCAGT	F	Ruck and Theriot, 2011
	18SP4R	TGATCCTTCYGCAGGTTCAC	R	
	atpB20f	AAGGTTACGTTTGTCAAATTAT	F	Theriot et al. 2015
	atoB1422r	ATTTTAGAGTTTCTGCTTTTGC	R	
Internal	rbcL40+	GGACTCGAATVAAAAGTGAACG	F	Ruck and Theriot, 2011
	rbcL1444	GCGAAATCAGCTGTATCTGTWG	R	
	18S 434F	ACAATAAATAACAATGCCGGGC	F	This study
	18S1449R	GVRTRCATCAGTGTAGCGCG	R	
	atpB77f	CGTTWVTCAAATTATTGGTCC	F	Theriot et al. 2015
	atpB1413r	GTTCTGGTAAWTCATCTAATTCAC	R	

Table 3.4. Summary statistics for DNA sequences used in this study.

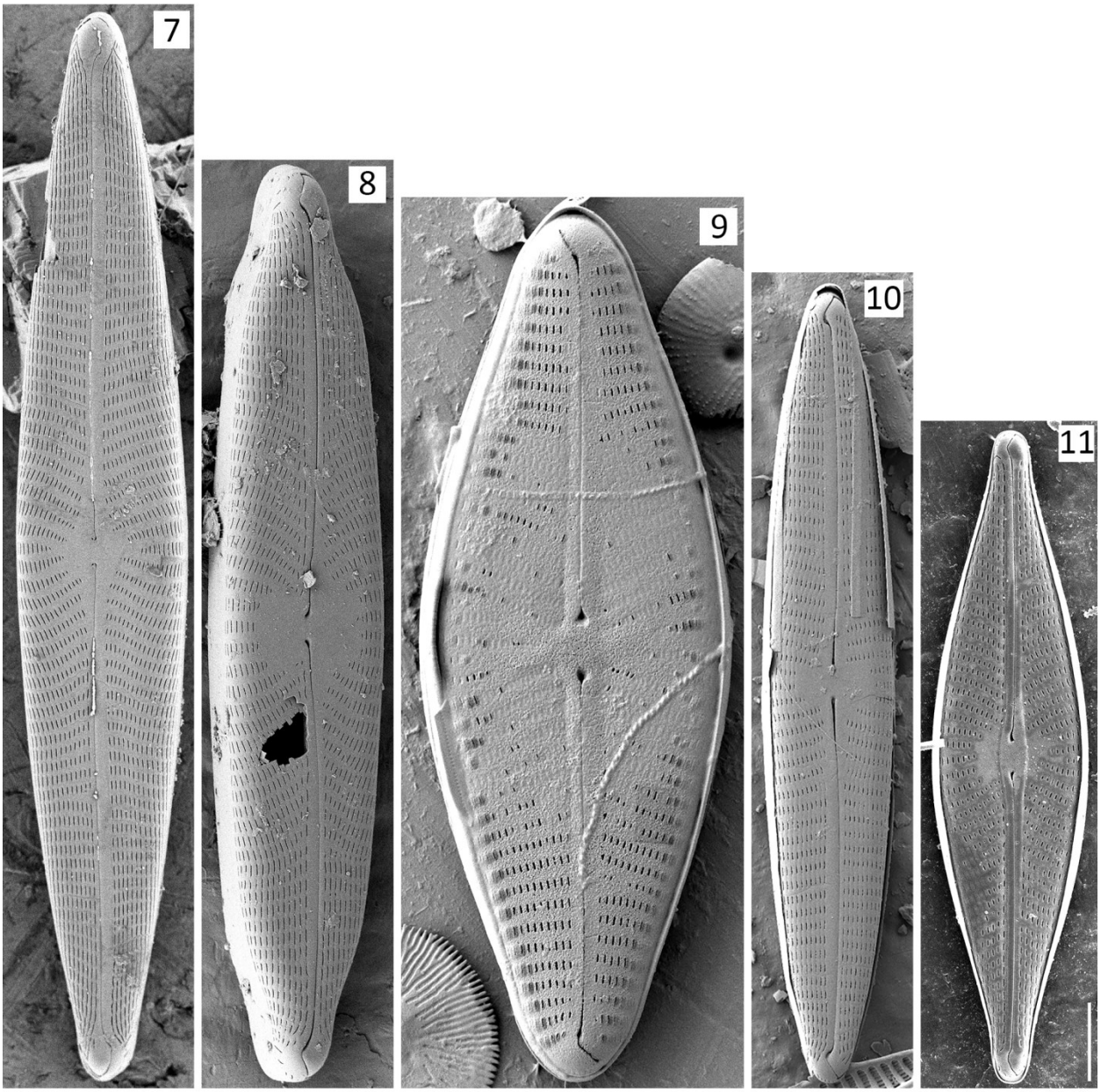
	rbcL	*atpB*	18S
Sequence length range (bp)	445 –1545	1028 –1287	389 – 1605
Sequence divergence # bp (%) – ingroup	0–129 (0 – 14.706%)	0 – 111 (0 – 9.689%)	0–96 (0 – 14.118%)
outgroup included	0–158 (0 – 18.627%)	0 – 152 (0 – 12.167%)	0–104 (0 – 14.118%)
Number of variable sites – ingroup	443	299	297
outgroup included	458	335	312
Potentially informative sites – ingroup	322	213	143
outgroup included	332	249	160
Constant sites – ingroup	1108	988	1053
outgroup included	1093	952	1038
Autapomorphic sites – ingroup	121	86	154
outgroup included	126	86	152

Table 3.5. Updated overview of phylogenetic relationships within *Navicula,* as compared to Bruder & Medlin (2008).

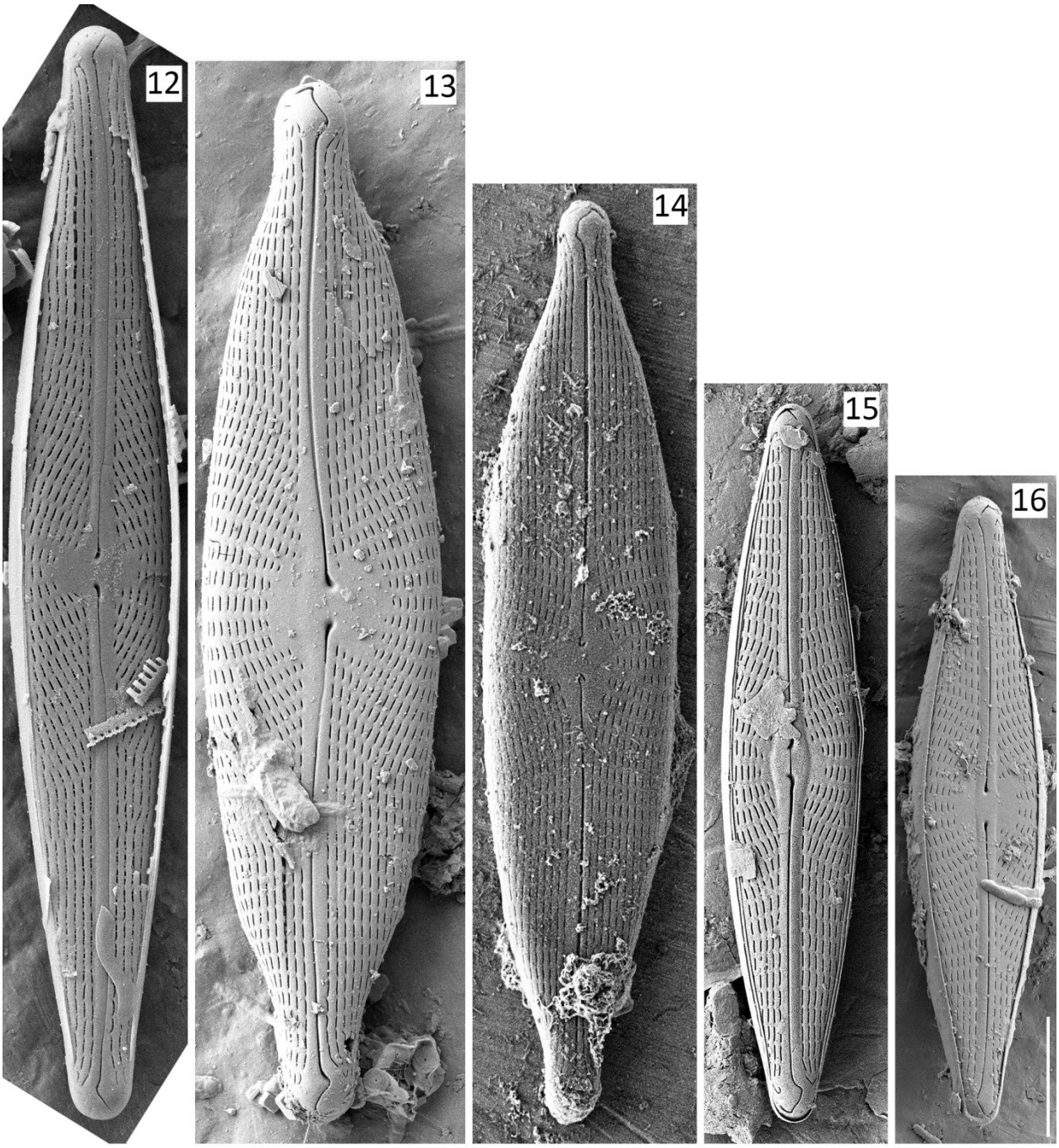
Study	Group 1	Group 2	Group 3
Bruder & Medlin, 2008	*N. capitoradiata* *N. cari* *N. cf. duerrenbergiana* *N. cf. erifuga* *N. lanceolata* *N. radiosa* *N. ramosissima* *N. tripunctata*	*N. cryptocephala var. veneta* *N. gregaria* *N. sclesviscensis* *N. veneta*	*N. cryptocephala* *N. cryptotenella* *N. phyllepta* *N. reinhardtii*
Additions from this study	*N. bottnica* *N. bruneli* *N. cf. lundii* *N. cf. perminuta* *N. cf. subwalkeri* *N. flagelifera* *N. incertata* *N. rhynchocephala var. harkensis* *N. rostellata* *N. symmetrica* *N. viridula var. rostellata* *N. viridulacalcis*	*N. agatkae* *N. cassensis* *N. joguensa* *N. pseudoacceptata*	*N. cf. cryptocephala* *N. cf. salinarum* *N. cryptotenelloides* *N. lundii* *N. menisculus* *N. microdigitoradiata* *N. rhychotella* *N. salinarum* *N. salinarum var. minima* *N. sleviscensis* *N. trivialis* *N. vara*



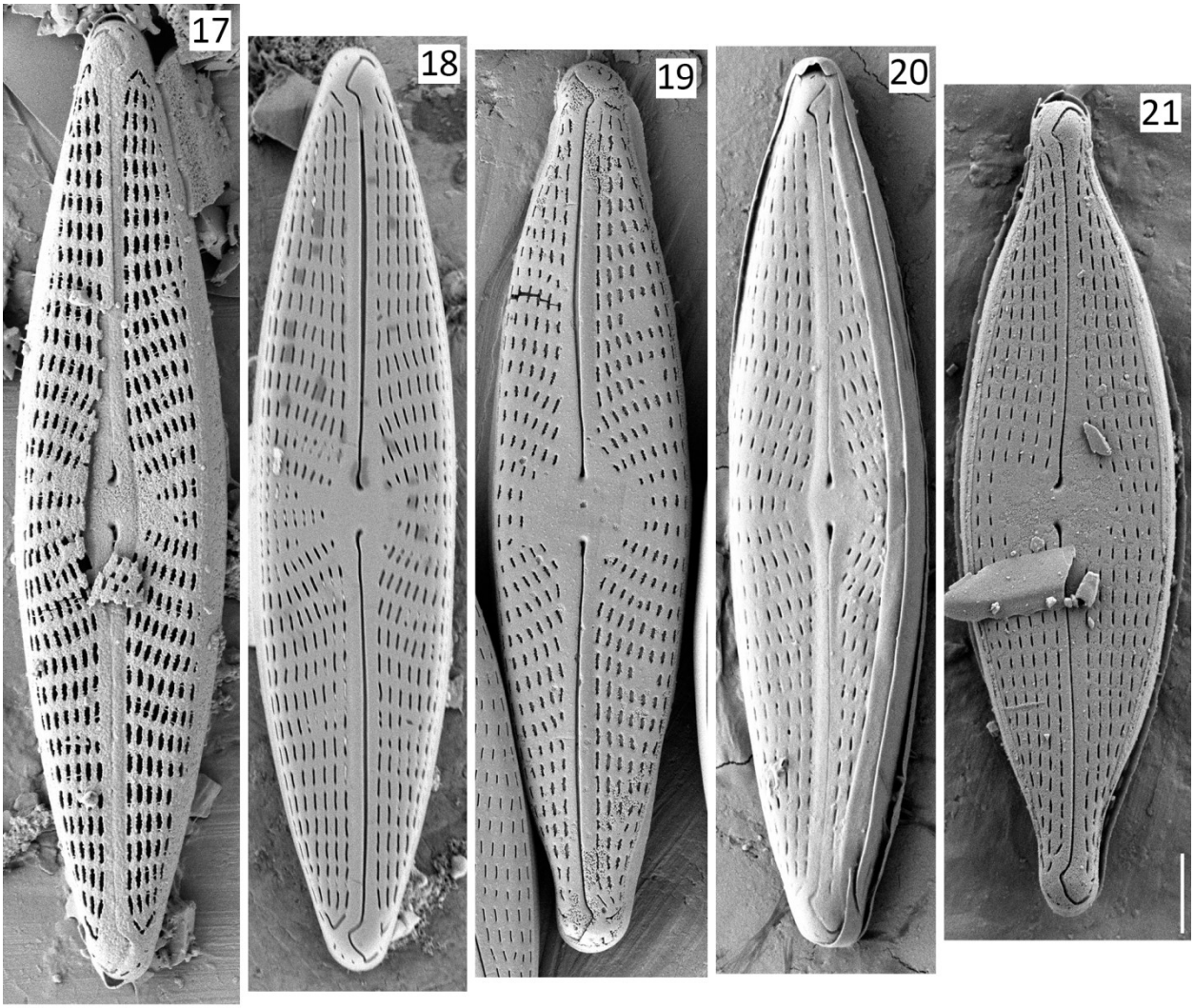
Figs 3.1 – 3.6. SEM micrographs of common taxa of *Navicula*. Figs. 1, 3, 5. *N. peregrina.* Figs. 2, 4, 6. *N.* cf. *subwalkeri*. Figs. 1 – 2. Whole valve. Figs. 3 – 4. Central area. Figs. 5 – 6. Apex. All figures show the external views of frustules. Scale bars = 5 µm.



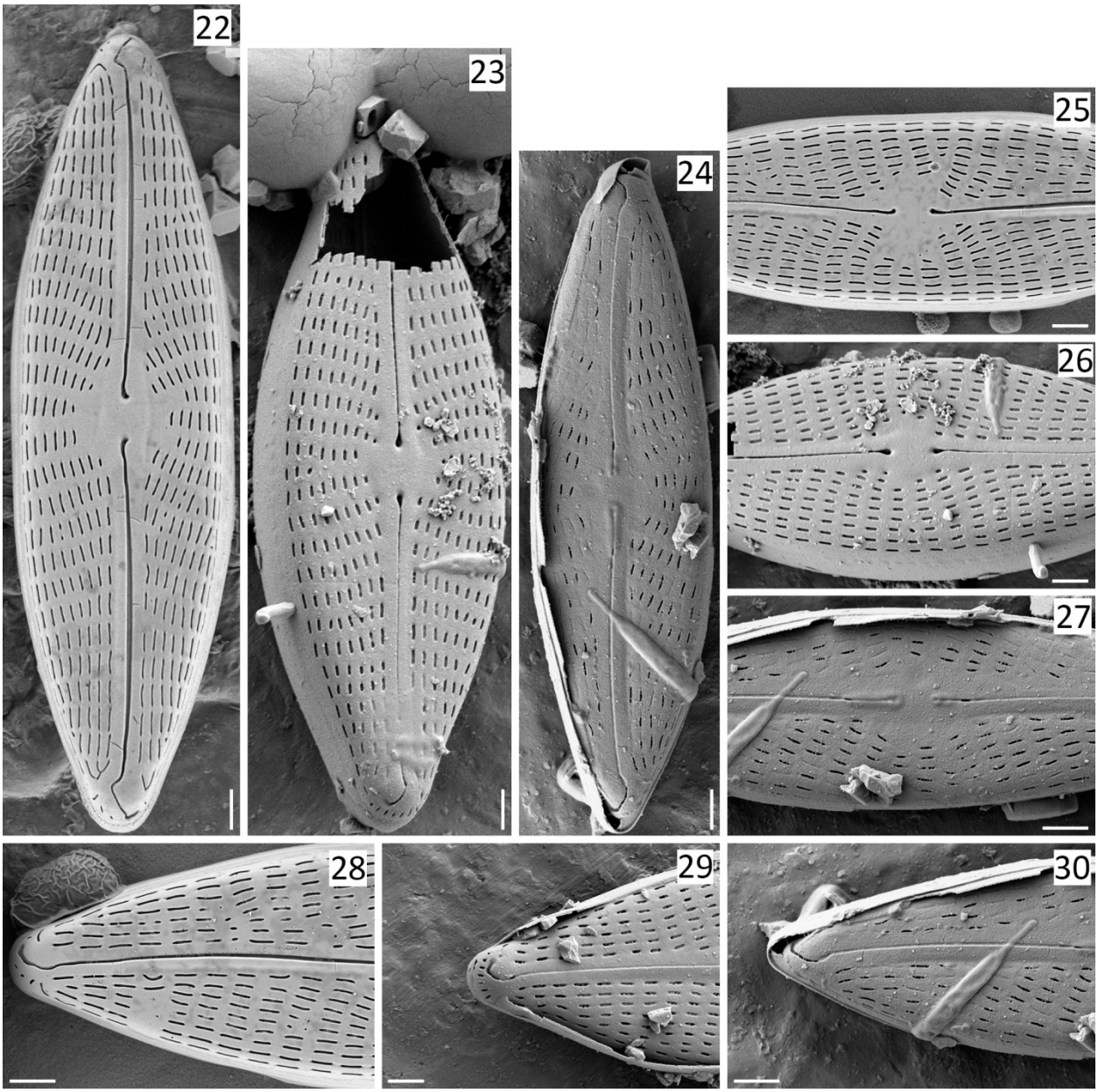
Figs 3.7 – 3.11. SEM micrographs of common taxa of *Navicula*. Fig. 7. *N. radiosa.* Fig. 8. *N. viridulacalcis.* Fig. 9. *N. reinhardtii.* Fig. 10. *N.* cf. *tripunctata.* Fig. 11. *N. trivialis*. All figures show the external views of frustules. Scale bars = 5 µm



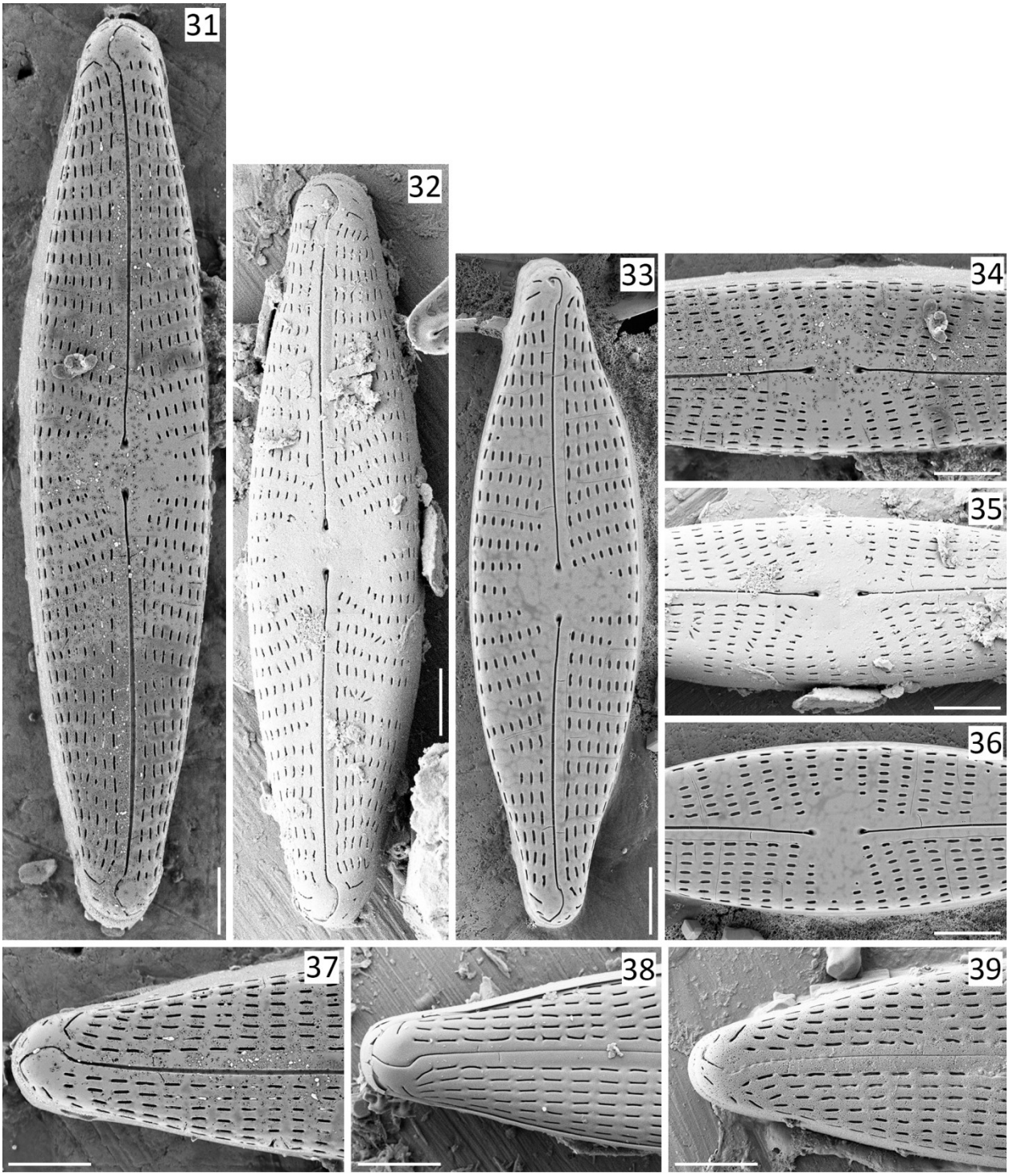
Figs 3.12 – 3.16. SEM micrographs of common taxa of *Navicula*. Fig. 12. *N. radiosafallax.* Fig. 13. *N. rostellata.* Fig. 14. *N. capitoradiata.* Fig. 15. *N. wildii.* Fig. 16. *N. cryptocephala.* All figures show the external views of frustules. Scale bars = 5 µm



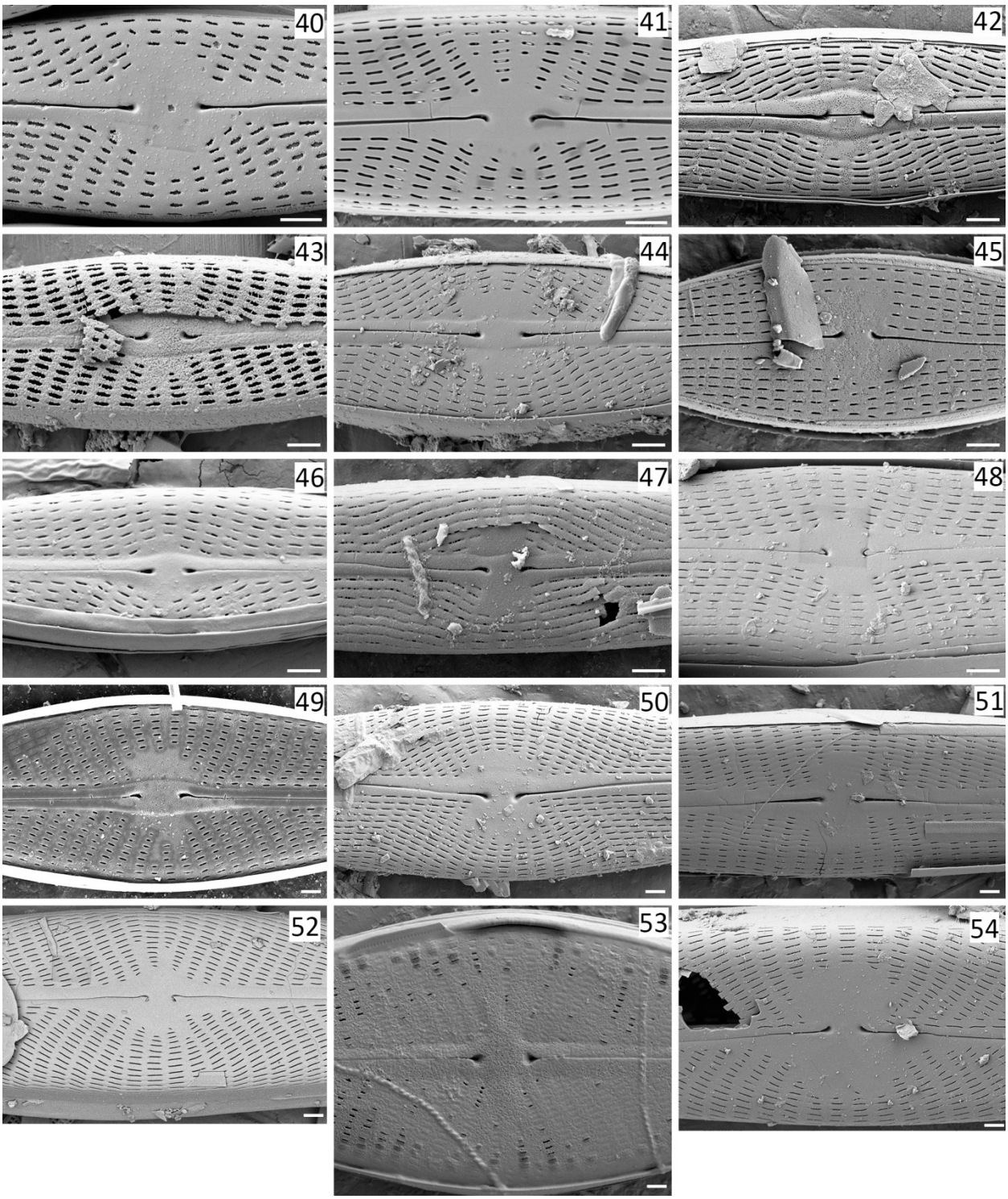
Figs 3.17 – 3.21. SEM micrographs of common taxa of *Navicula*. Fig. 17. *N. lundii.* Fig. 18. *N. cincta.* Fig. 19. *N.* cf. *metareichardtiana.* Fig. 20. *N. cryptotenella.* Fig. 21. *N. gregaria.* All figures show the external views of frustules. Scale bars = 2 µm



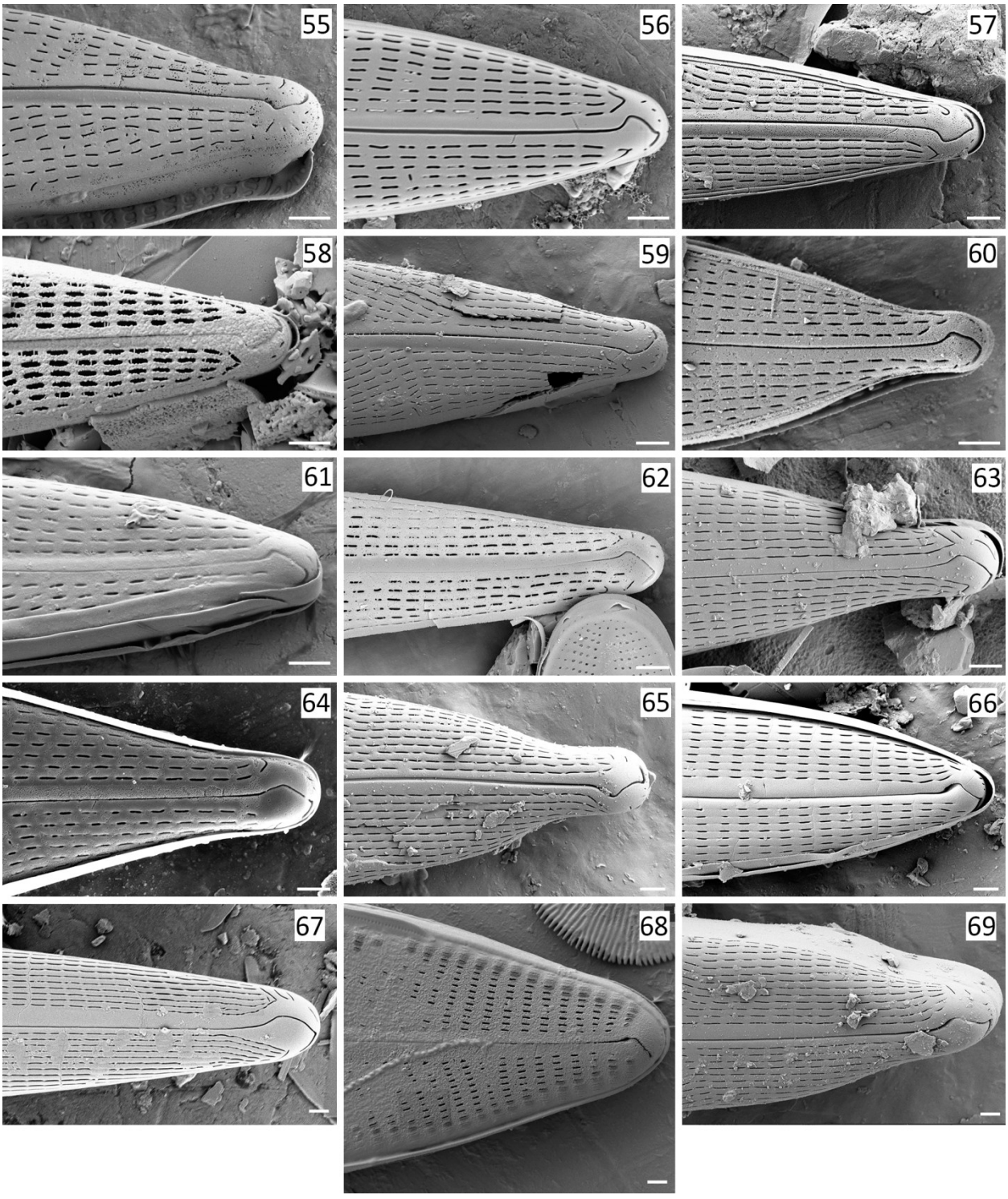
Figs 3.22 – 3.30. SEM micrographs of common taxa of *Navicula*. Figs. 22, 25, 28. *N.* cf. *cincta.* Figs. 23, 26, 29. *N. menisculus.* Fig. 24, 27, 30. *N. cryptotenelloides*. Figs. 22 – 24. Whole valve. Figs. 25 – 27. Central area. Figs. 28 – 30. Apex. All figures show the external views of frustules. Scale bars = 1 µm.



Figs 3.31 – 3.39. SEM micrographs of common taxa of *Navicula*. Figs. 31, 34, 37. *N. cassensis.* Figs. 32, 35, 38. *N. joguensa.* Figs. 33, 36, 39. *N. veneta.* Figs. 31 – 33. Whole valve. Figs. 34 – 36. Central area. Figs. 37 – 39. Apex. All figures show the external views of frustules. Scale bars = 2 µm.



Figs 3.40 – 3.54. SEM micrographs of the external central nodule. Fig. 40. *N.* cf. *metareichardtiana*. Fig. 41. *N. cincta.* Fig. 42. *N. wildii*. Fig. 43. *N. lundii.* Fig. 44. *N. cryptocephala.* Fig. 45. *N. gregaria.* Fig. 46. *N. cryptotenella.* Fig. 47. *N. radiosafallax.* Fig. 48. *N. capitoradiata.* Fig. 49. *N. trivialis.* Fig. 50. *N. rostellata.* Fig. 51. *N.* cf. *tripunctata.* Fig. 52. *N. radiosa*. Fig. 53. *N. reinhardtii.* Fig. 54. *N. viridulacalcis.* Scale bars = 1 µm.



Figs 3.55 – 3.69. SEM micrographs of the external apex. Fig. 55. *N.* cf. *metareichardtiana*. Fig. 56. *N. cincta.* Fig. 57. *N. wildii*. Fig. 58. *N. lundii.* Fig. 59. *N. cryptocephala.* Fig. 60. *N. gregaria.* Fig. 61. *N. cryptotenella.* Fig. 62. *N. radiosafallax.* Fig. 63. *N. capitoradiata.* Fig. 64. *N. trivialis.* Fig. 65. *N. rostellata.* Fig. 66. *N.* cf. *tripunctata.* Fig. 67. *N. radiosa.* Fig. 68. *N. reinhardtii.* Fig. 69. *N. viridulacalcis.* Scale bars = 1 μm.

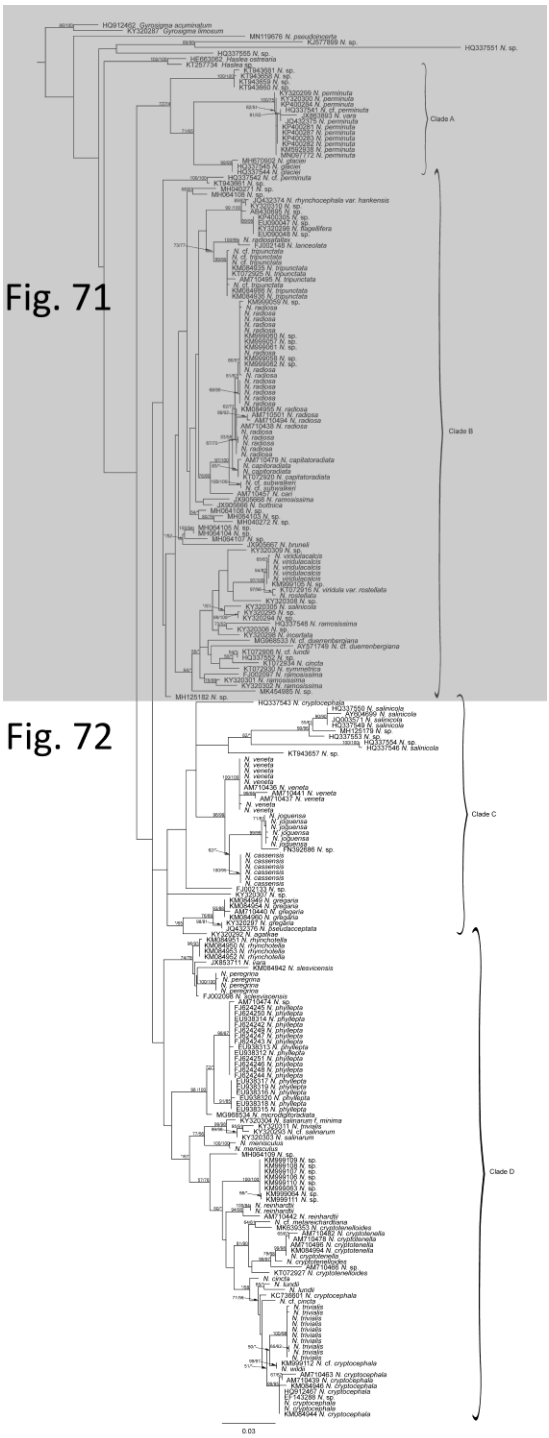


Fig. 3.70. Overview of the combined maximum likelihood and parsimony tree of the genus *Navicula* using the *rbcL* gene. The detailed view of each portion of the tree is shown in Figs. 3.71 – 3.72, as illustrated here.

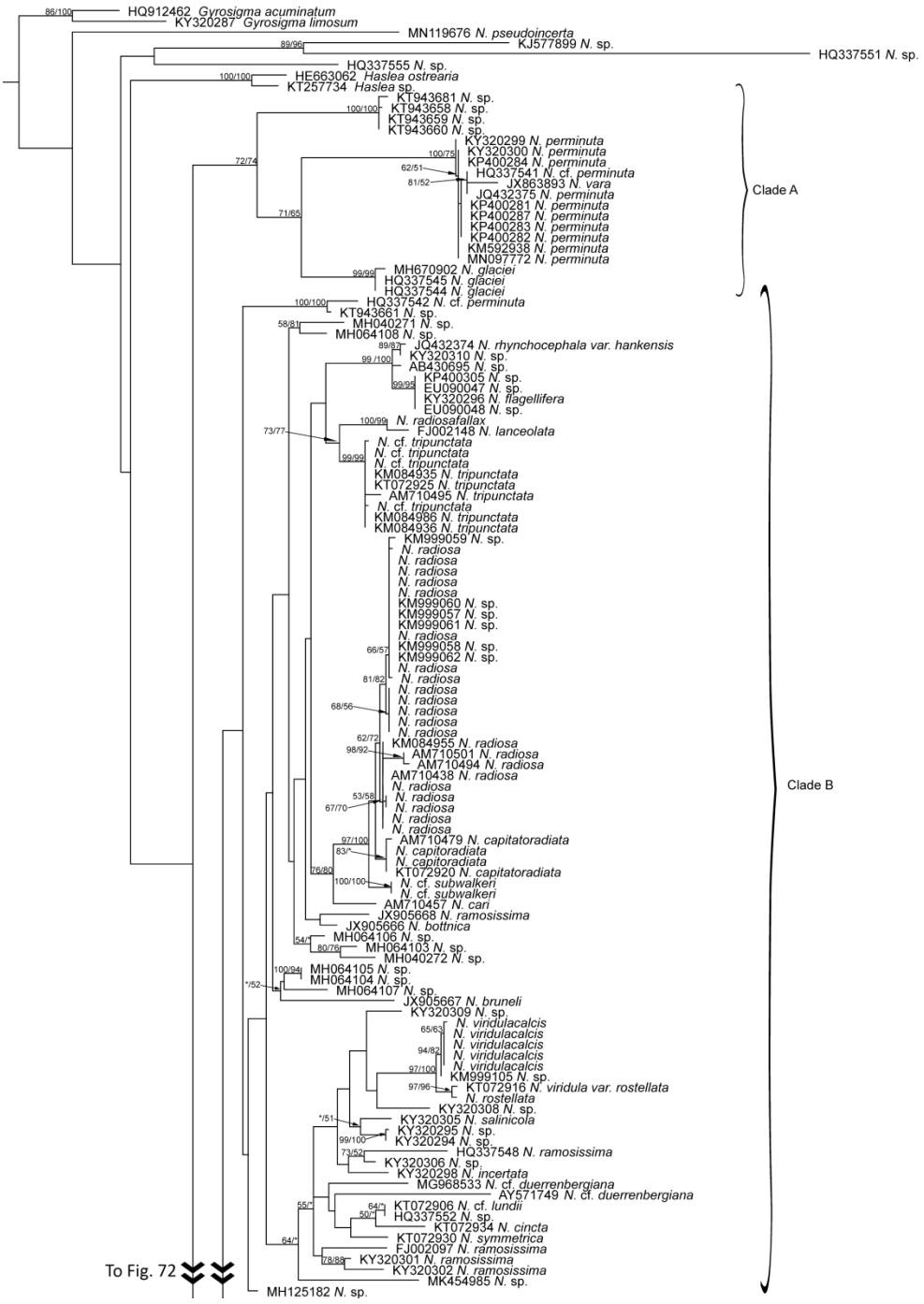


Fig. 3.71. Combined maximum likelihood and parsimony tree of the genus *Navicula* using the *rbcL* gene. The analysis was partitioned, with two subsets: one for codon position 1, with K80+I+G as best fit model, and one for codon positions 2 and 3, with GTR+I+G as best fit model. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter.

Fig. 3.72. Combined maximum likelihood and parsimony tree of the genus *Navicula* using the *rbcL* gene. The analysis was partitioned, with two subsets: one for codon position 1, with K80+I+G as best fit model, and one for codon positions 2 and 3, with GTR+I+G as best fit model. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter.

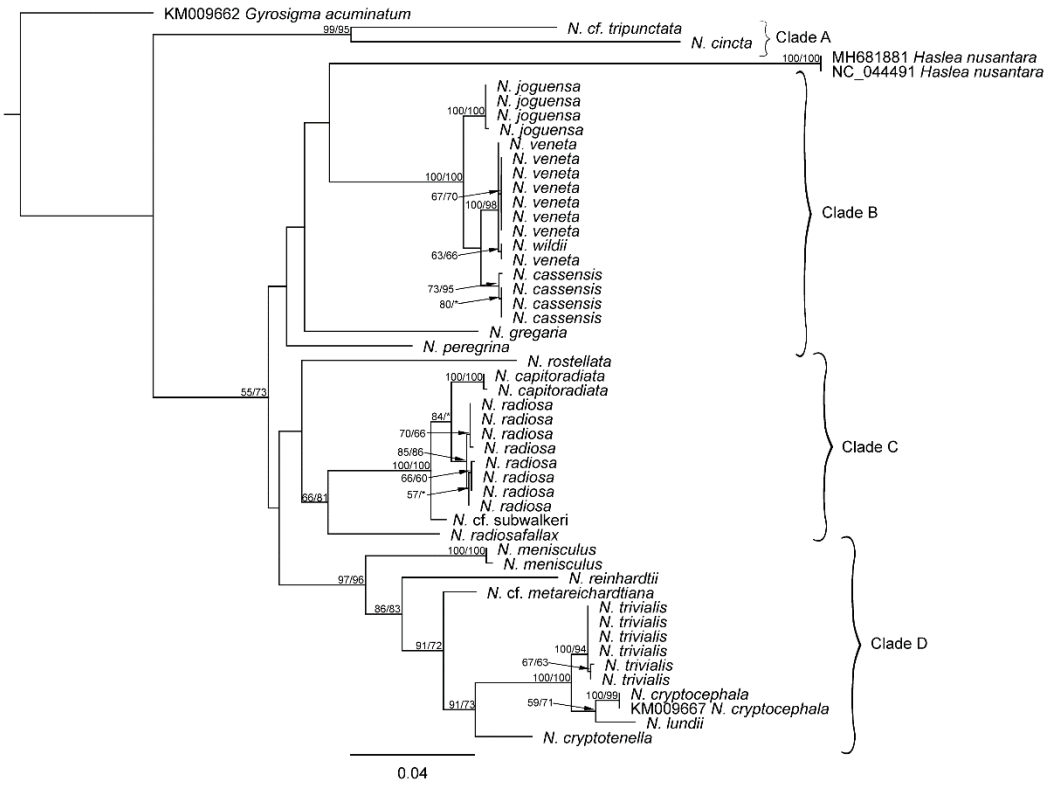


Fig. 3.73. Combined maximum likelihood and parsimony tree of the genus *Navicula* using the *atpB* gene. The analysis was partitioned, with a subset created for each codon position. The best fit models were TVM+I+G, F81+I, and F81+I+G for the first, second, and third codon positions respectively. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter.

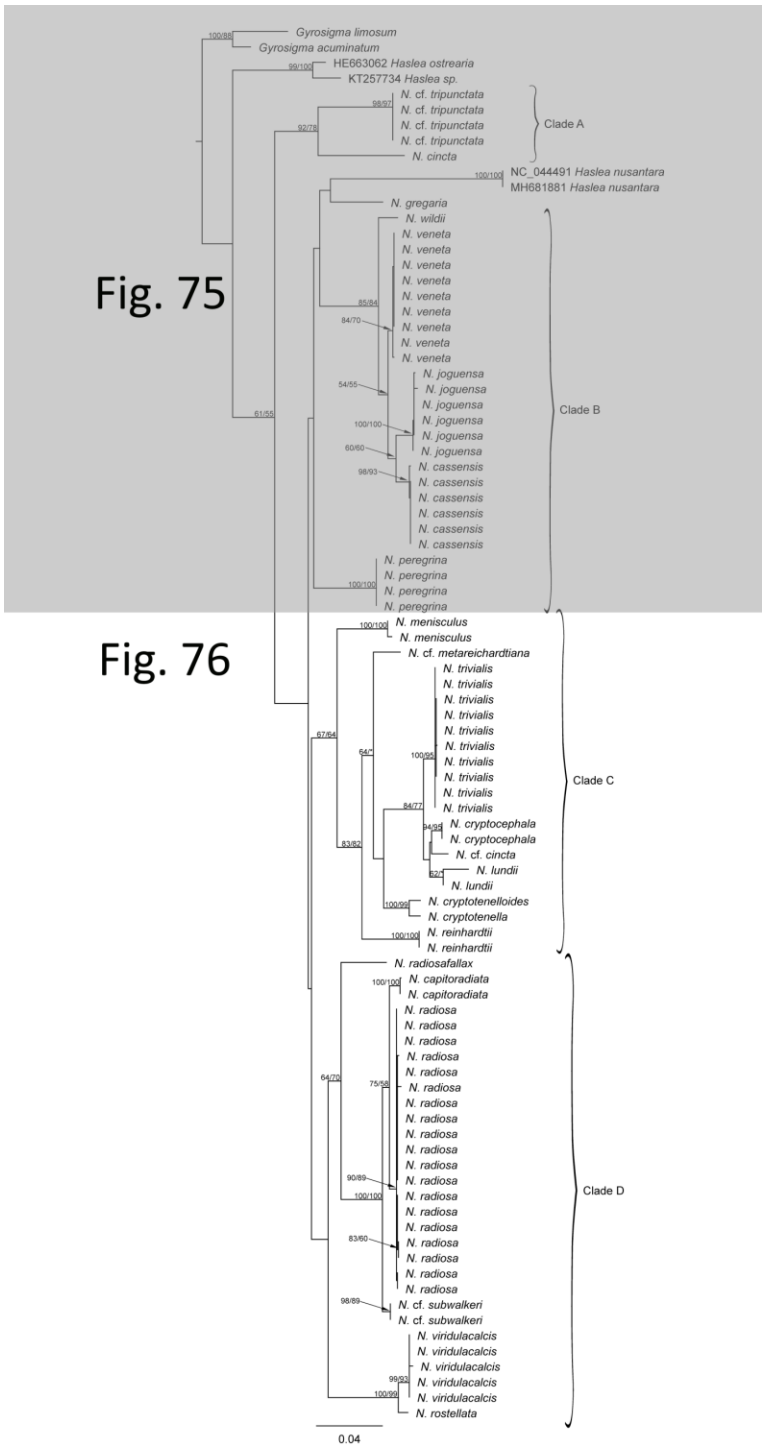


Fig. 3.74. Overview of the combined maximum likelihood and parsimony tree of the genus *Navicula* using the combination of two plastid genes, *rbcL* and *atpB*. The detailed view of each portion of the tree is shown in Figs. 3.75 – 3.76, as illustrated here.

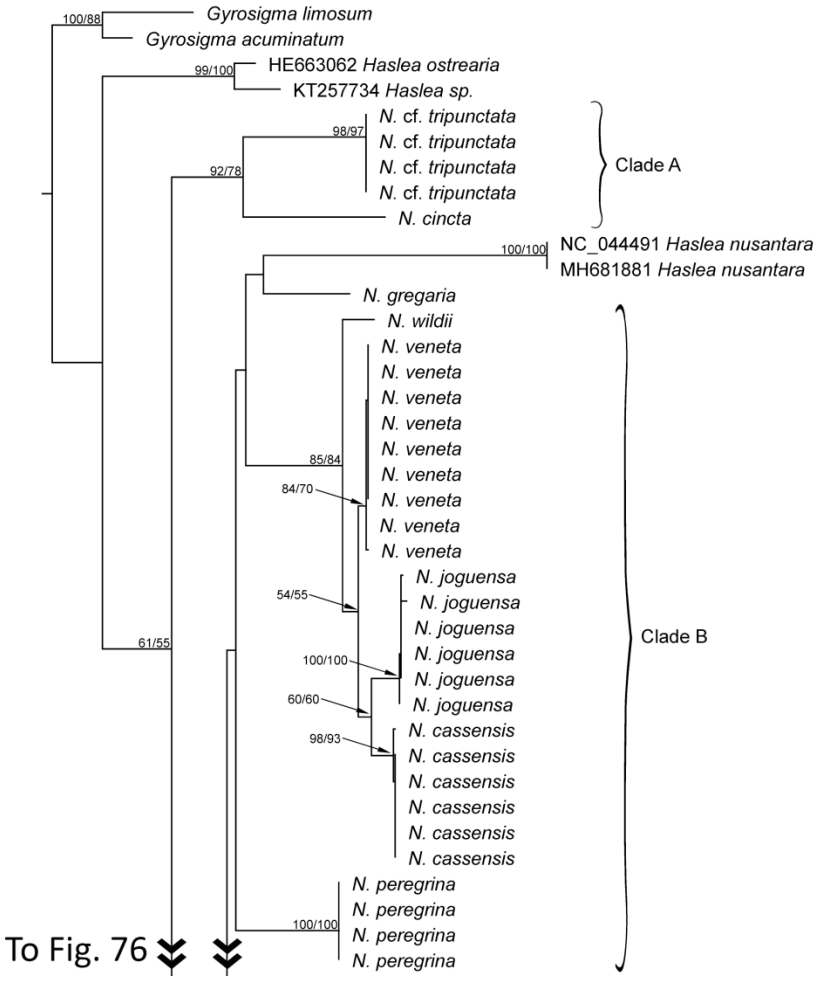


Fig. 3.75. Combined maximum likelihood and parsimony tree of the genus *Navicula* using the combination of two plastid genes, *rbcL* and *atpB*. The analysis was partitioned, with four subsets created: one for *rbcL* codon positions 1 and 3 and *atpB* position 1, a second for *rbcL* position 2, a third for *atpB* position 2, and a fourth for *atpB* position 3, with HKY+I+G, TVM+G, F81+I+G and TVM+I+G as best fit models, respectively. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter.

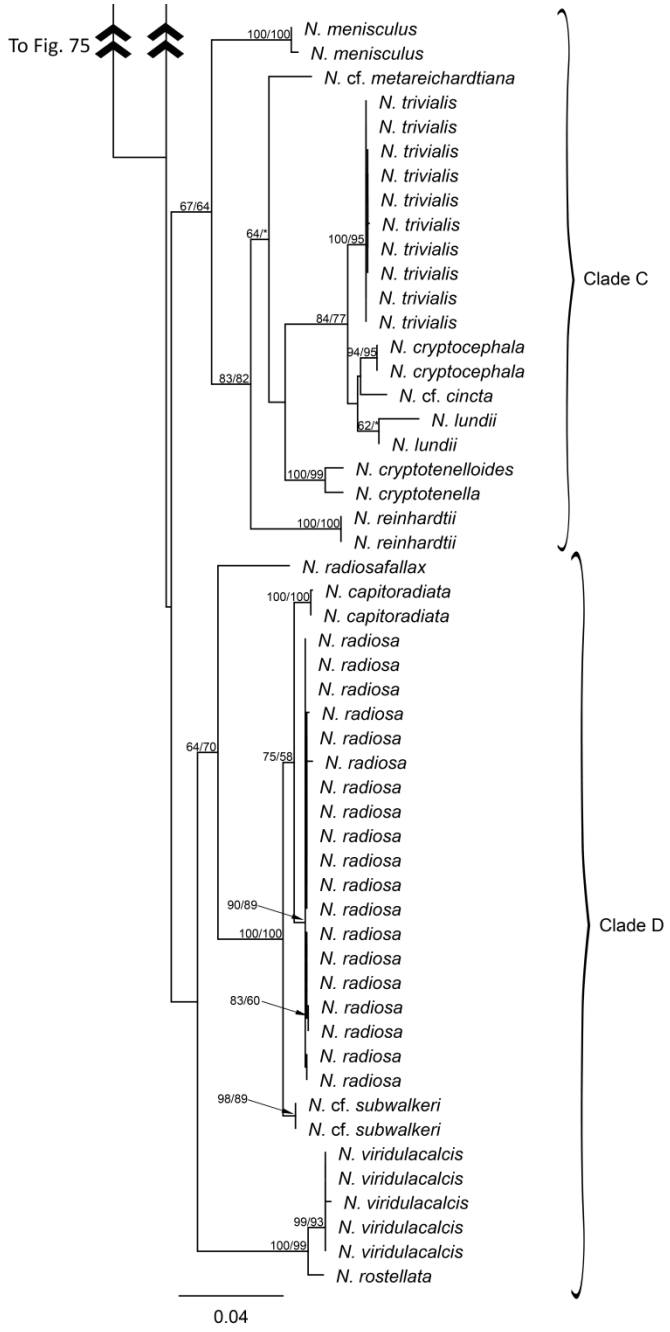


Fig. 3.76. Combined maximum likelihood and parsimony tree of the genus *Navicula* using the combination of two plastid genes, *rbcL* and *atpB*. The analysis was partitioned, with four subsets created: one for *rbcL* codon positions 1 and 3 and *atpB* position 1, a second for *rbcL* position 2, a third for *atpB* position 2, and a fourth for *atpB* position 3, with HKY+I+G, TVM+G, F81+I+G and TVM+I+G as best fit models, respectively. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter.

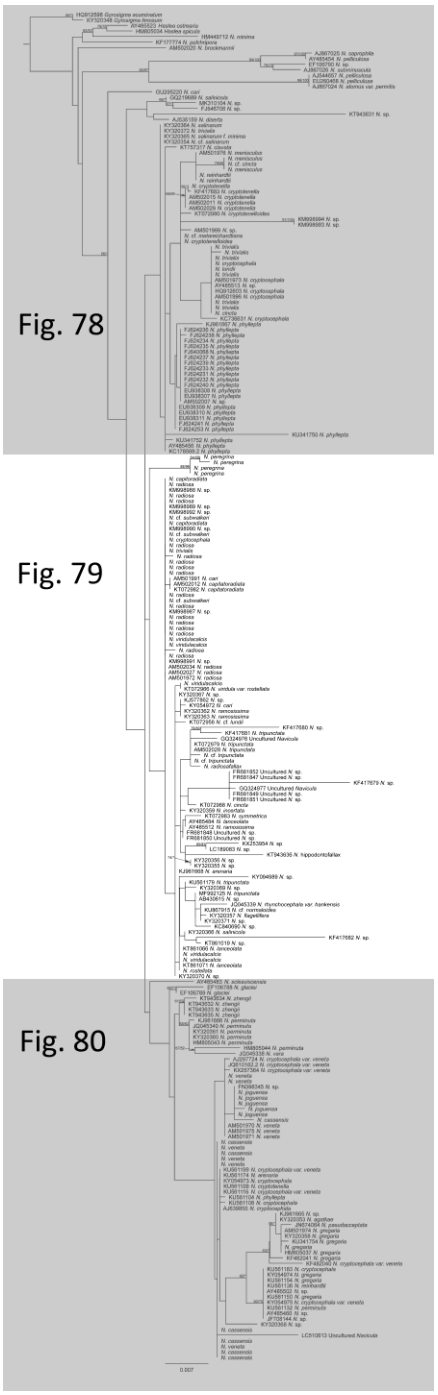


Fig. 3.77. Overview of the combined maximum likelihood and parsimony tree of the genus *Navicula* using the 18S gene. The detailed view of each portion of the tree is shown in Figs. 3.78 – 3.80, as illustrated here.

Fig. 3.78. Combined maximum likelihood and parsimony tree of the genus *Navicula* using the 18S gene. The best fit model used is TRN+I+G. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter.

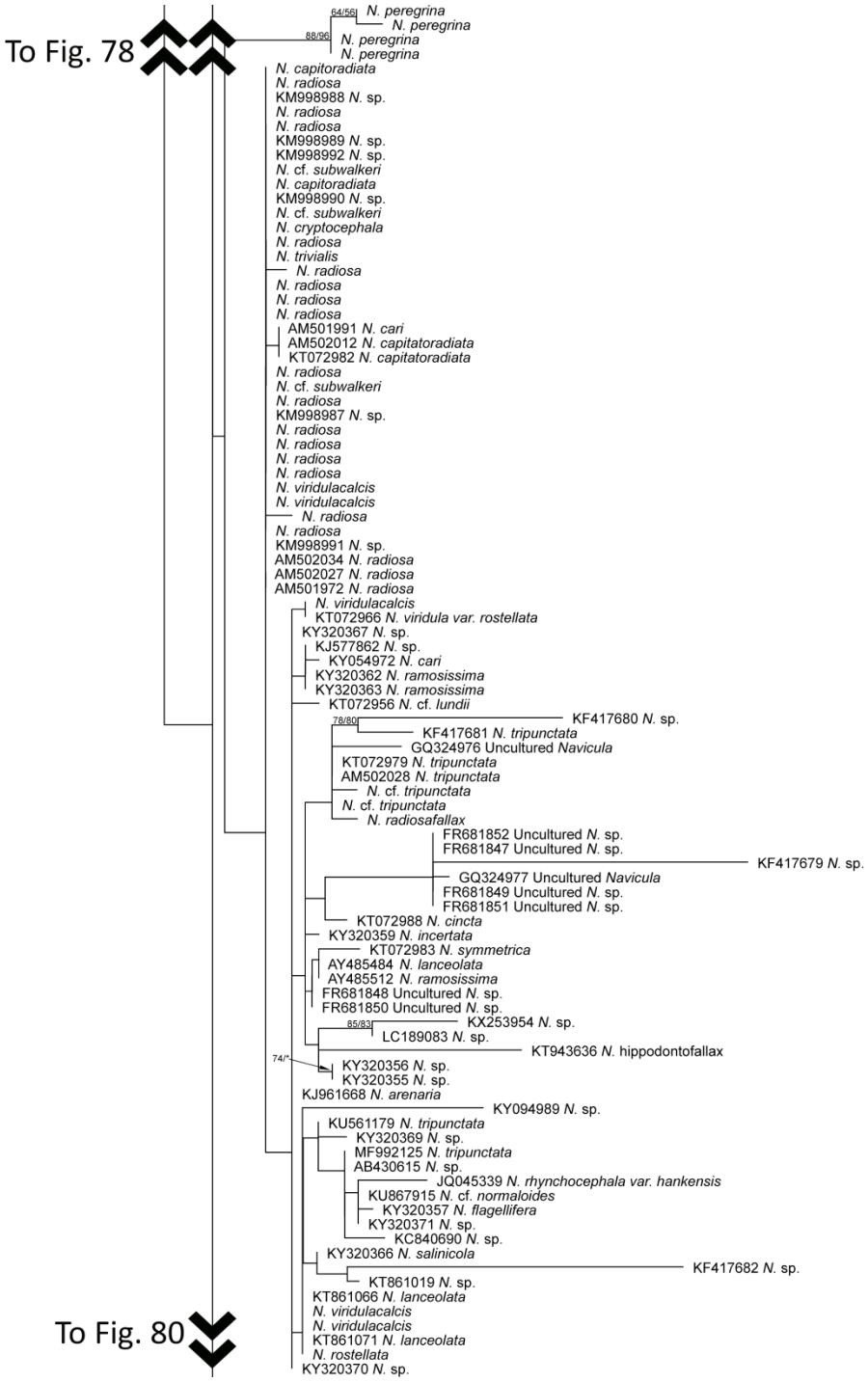


Fig. 3.79. Combined maximum likelihood and parsimony tree of the genus *Navicula* using the 18S gene. The best fit model used is TRN+I+G. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter.

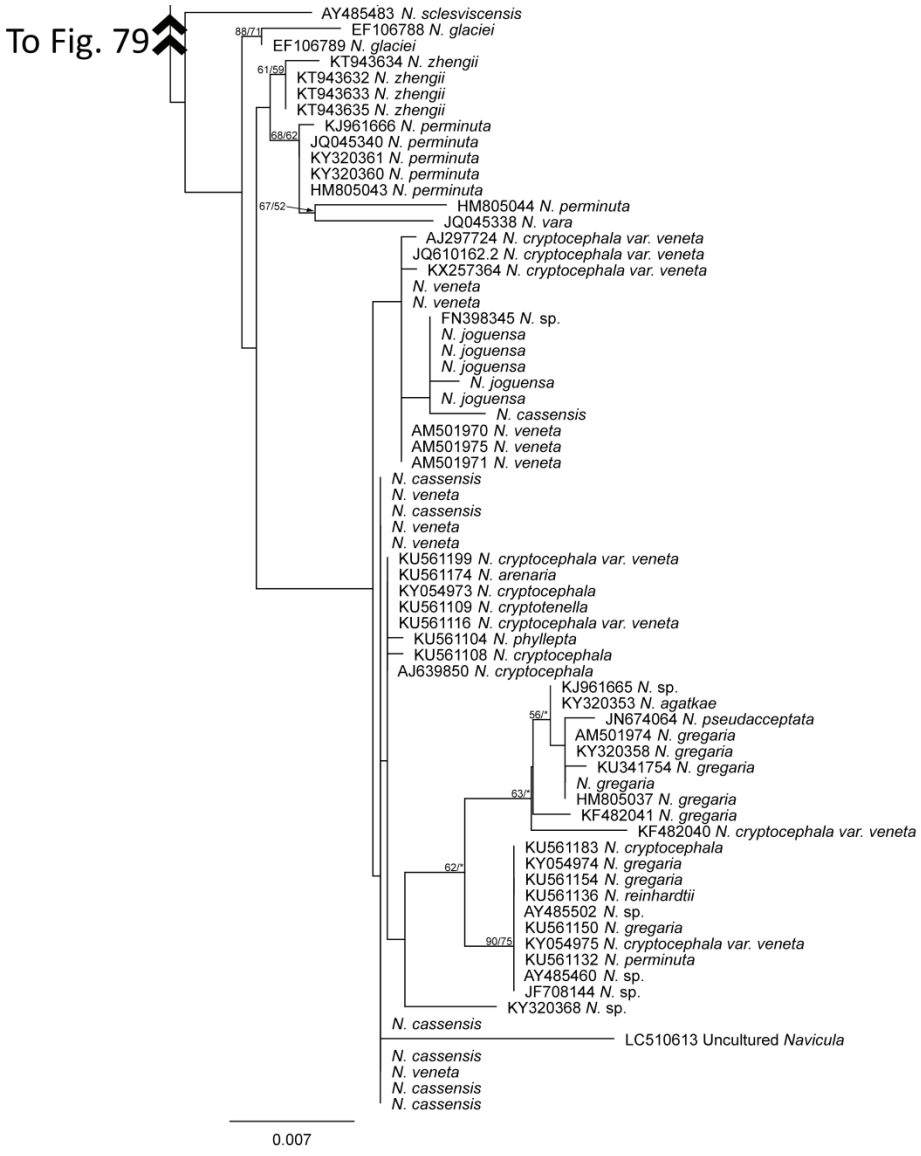


Fig. 3.80. Combined maximum likelihood and parsimony tree of the genus *Navicula* using the 18S gene. The best fit model used is TRN+I+G. Branch labels indicate bootstrap support values as percentages for maximum likelihood and maximum parsimony analyses (ML/MP). Values below 50% BS are not presented. When the support value for one analysis is above that threshold, but the other is not, an asterisk (*) is used in place of the latter.

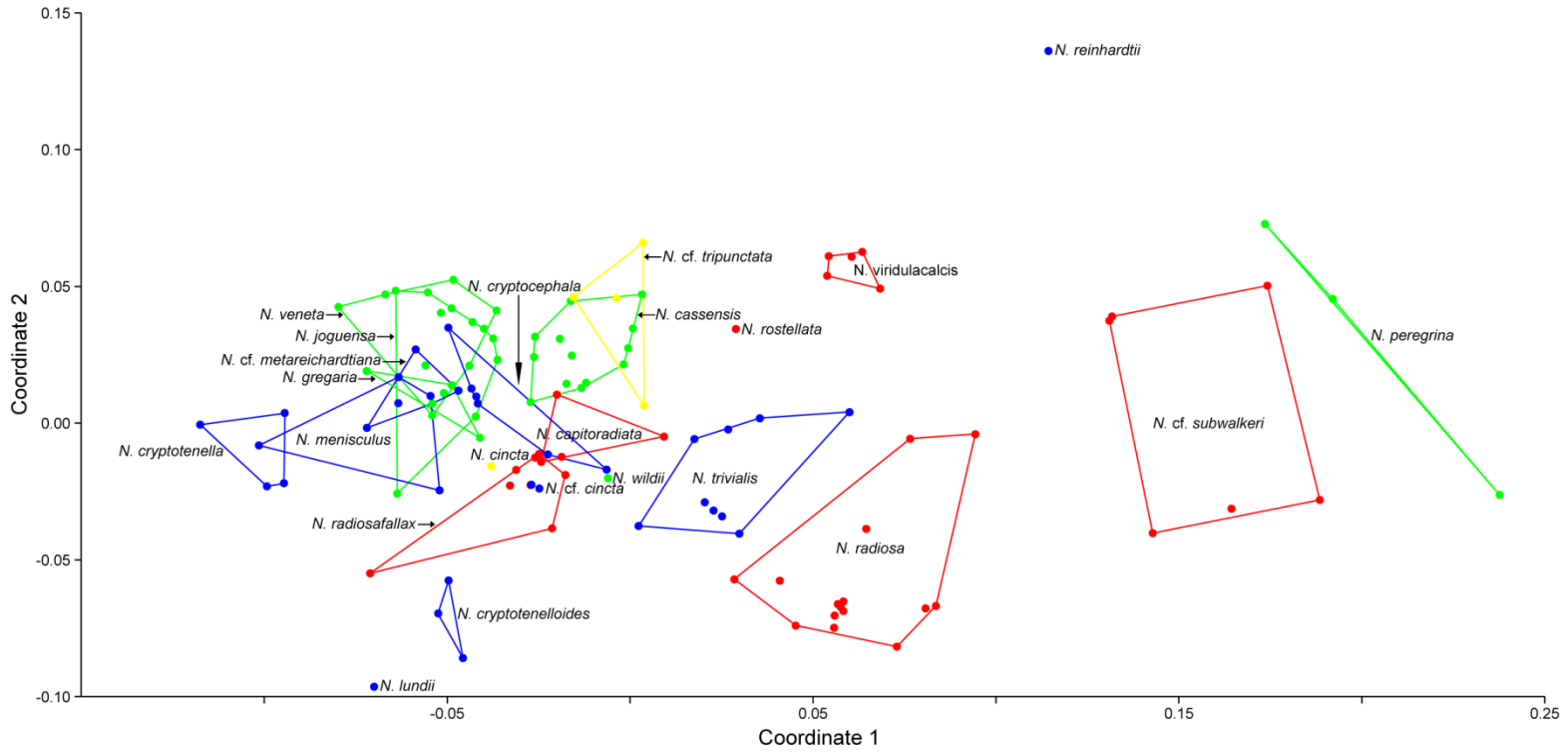


Fig. 3.81. Non metric multi-dimensional scaling (NMDS) plot showing the relative similarity of clades based on thirteen morphological characters, as described in the Material and Methods of Chapter 3. Stress = 0.1412. Species are color-coded based on the major clade they belong to in the combined plastid phylogenetic tree (Figs 3.70 – 3.72). Clade A = Yellow; Clade B = Green; Clade C = Blue; Clade D = Red. Note that species do not group by major clade.

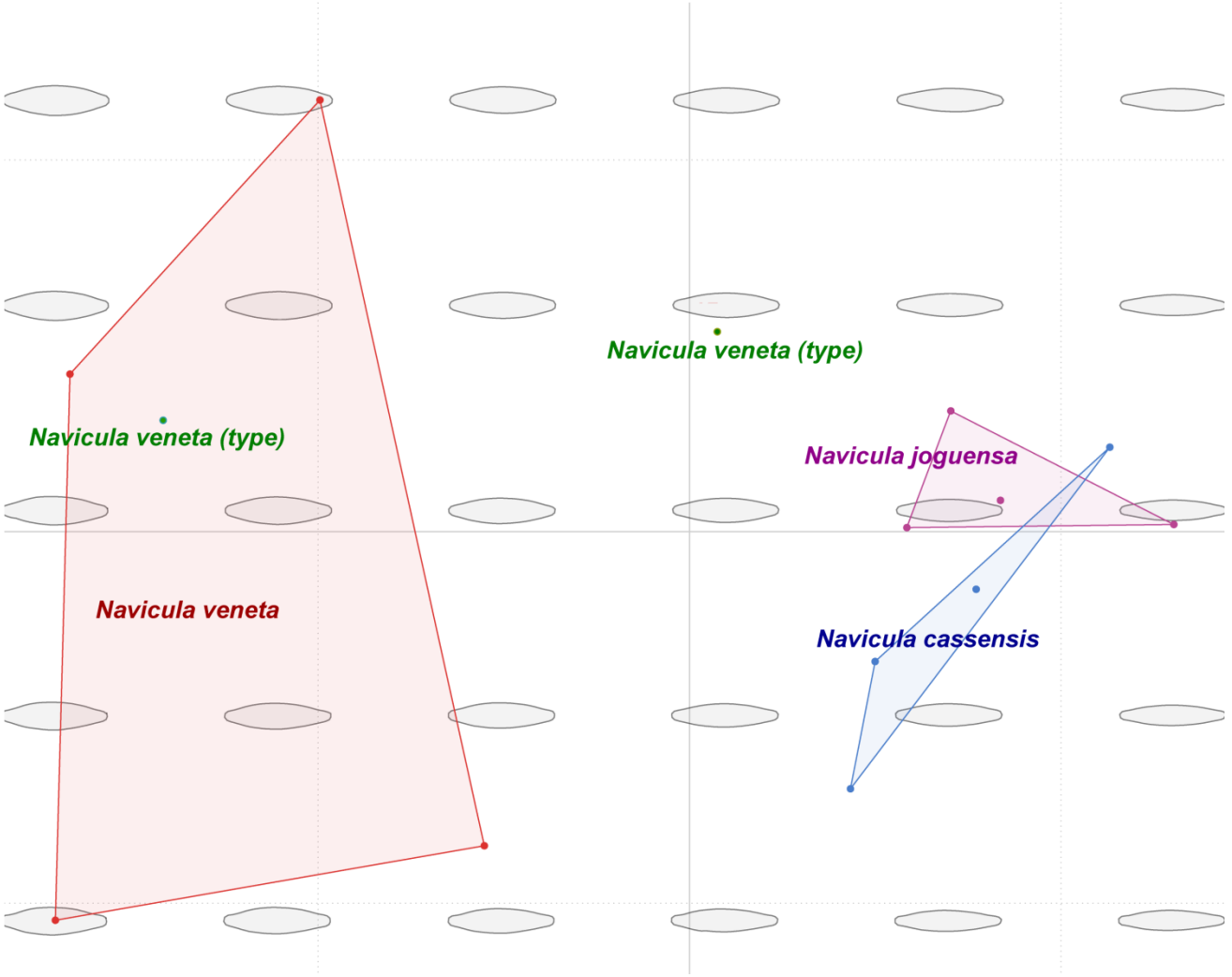


Fig. 3.82. Principal component analysis (PCA) results for valve shape of pseudo-cryptic species *N. veneta, N. joguensa*, *and N. cassensis*. Type specimens for *N. veneta,* as based on measurements taken from the photographs provided by Lange-Bertalot (2001) (Plate 14, Fig. 23; Plate 65; Fig. 3), were also included.

Chapter 4 – General Conclusions

The overall objective of this thesis was to use molecular and morphological methods to investigate major lineages in two selected diatom genera. The genus *Frustulia* was studied using a single-cell DNA extraction technique and two molecular markers, namely, the plastid gene *rbcL* and the nuclear ribosomal gene 18S (Chapter 2). The genus *Navicula* s. str. was then studied using a modified version of this single-cell protocol, which incorporated the use of small cultures, and three molecular markers, *rbcL*, *atpB* (plastid), and 18S (Chapter 3).

Pseudo-cryptic diversity was found within both *Frustulia* and *Navicula*. Within *Frustulia*, a new species, *F. gibsonea,* was described. This taxon was previously known as *F.* cf. *krammeri* based on morphology alone – however, the combination of molecular and morphological data used in this study shows that it is a distinct species. Similarly, two new species, *N. joguensa* and *N. cassensis,* were described within *Navicula*. These species appear to have been previously included within *N. veneta* on the basis of their morphological similarities. Molecular data shows that they are distinct species, albeit closely related to *N. veneta.*

Possible cryptic diversity was also detected within *N. tripunctata.* Several specimens that were morphologically identical to *N. tripunctata* were found to be more molecularly divergent than what is expected within a diatom species. Further research on these specimens will be necessary to confirm this finding. As *N. tripunctata* is the type species for *Navicula,* this hidden diversity could have significant consequences for our understanding of the genus as it could lead to an entire recircumscription of the genus.

Both *Frustulia* and *Navicula* are not monophyletic. It is possible that certain species could be more closely related to other genera than to species within *Frustulia* and *Navicula*, or

that the genes used here are not well-suited to distinguish between genera. In Chapter 3, two *Navicula* specimens from GenBank formed a clade with the outgroup in 18S trees. These specimens did not belong to *Navicula* sensu stricto, but to species that have been transferred to other genera. However, caution should be employed when drawing conclusions from GenBank specimens as the identification of the material used for sequencing cannot be easily confirmed.

Similar interspecific divergence levels were observed for *Frustulia* and *Navicula*. In both genera, close molecular relationships between species were not necessarily reflected in similar morphological structures. In *Navicula,* I was unable to find common morphological characters for the major clades in my trees. Conversely, molecularly distant species such as *F. gibsonea* and *F. krammeri* were very similar in morphology. This illustrates the importance of using multiple data types when making taxonomic decisions in diatom systematics.

The molecular markers used within this thesis showed a variety of divergence levels. The plastid markers *rbcL* and *atpB* showed similar levels of divergence both within and between species. Despite a few discrepancies in the positioning of certain species of *Navicula* between *rbcL* and *atpB* trees, most of the data were concordant. In my study of *Navicula*, the combination of *rbcL* and *atpB* sequences in a partitioned analysis produced a well-supported tree. While these markers generally performed well, *rbcL* was too conserved to delimit the *F. saxonica-crassinervia-krammeri* species complex. Further work, using less conserved markers, may be useful in interpreting this complex and others. The other marker used, 18S, appears to be too conserved for species-level phylogenies in diatoms as the resolution among closely related taxa was poor.

The single-cell DNA extraction technique used in Chapter 2 presented some limitations. A large working distance was required in order to isolate cells, which resulted in low imaging

resolution. This made it difficult to perceive fine morphological characteristics and identify species. As the cells were subsequently destroyed during the DNA extraction process, it was not possible to examine them directly with scanning electron microscopy. Instead, isolated cells were linked to SEM micrographs taken from the sediment samples using the photographs taken at the time of isolation; a sometimes laborious process.

In order to address these limitations, this protocol was modified in Chapter 3 in the following ways: small cultures, of about 5 to 500 cells/mL, were grown from isolated cells. DNA was extracted from a portion of this culture material, while the rest was examined under SEM. This allowed for a thorough direct examination of the morphology of my specimens, and easier species identification. My amplification success was also higher using this technique, which can likely be attributed to the presence of more DNA in the pre-amplification samples. Culturing success was high (> 80%) for small (< 50 μm) Navicula species, and lower (≈ 50%) for large species (> 50 μm).

By exploring methods that allow for the combination of DNA sequence data and traditional morphological data, this thesis can serve as a model to better understand the diversity and evolution of diatoms. The single-cell sequencing and small culture methods used here allow for a relatively rapid and easy study of large numbers of specimens compared to traditional culturing methods. Because these techniques allow for such large-scale sampling of diatoms, they could play a major future role in addressing key questions in diatom systematics and evolution.

References

Aaronson B., de Angelis, Frank, O. & Baker, H. 1971. Secretion of vitamins and amino acids into the environment by *Ochromonas danica*. *Journal of Phycology* 7: 215–218.

Ajani, P., Murray, S., Hallegraeff, G., Lundholm, N., Gillings, M., Brett, S. & Armand, L. 2013. The diatom genus *Pseudo-nitzschia* (Bacillariophyceae) in New South Wales, Australia: Morphotaxonomy, molecular phylogeny, toxicity, and distribution. *Journal of Phycology* 49(4): 765–785.

Alverson, A. J., Jansen, R. K. & Theriot, E. C. 2007. Bridging the Rubicon: Phylogenetic analysis reveals repeated colonizations of marine and fresh waters by thalassiosiroid diatoms. *Molecular Phylogenetics and Evolution* 45(1): 193–210.

Amato A., Kooistra H.C.F., Ghiron J.H.L., Mann D.G., Pröschold T. & Montresor M. 2007. Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist* 158: 193–207.

An, S. M., Choi, D. H., Lee, J. H., Lee, H. & Noh, J. H. 2017. Identification of benthic diatoms isolated from the eastern tidal flats of the Yellow Sea: Comparison between morphological and molecular approaches. *PLOS ONE*, 12(6): e0179422.

Antoine, D., Andre, J-M. & Morel, A. 1995. Oceanic primary production: 2. Estimate at global scale from satellite (Coastal Zone Colour Scanner) chlorophyll. *Global Biogeochemical Cycles* 10: 57–69.

Arp, G., Bissett, A., Brinkmann, N., Cousin, S., De Beer, D., Friedl, T., Mohr, K. I., Neu, T. R., Reimer, A., Shiraishi, F., Stackebrandt, E. & Zippel, B. 2010. Tufa-forming biofilms of German karstwater streams: Microorganisms, exopolymers, hydrochemistry and calcification. *Geological Society, London, Special Publications* 336(1): 83–118.

Auguie B. 2017. gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra

Behnke, A., Friedl, T., Chepurnov, V. A. & Mann, D. G. 2004. Reproductive compatibility and rDNA sequence analyses in the *Sellaphora pupula* species complex (Bacillariophyta): rDNA and sexual compatibility in *Sellaphora*. *Journal of Phycology* 40(1): 193–208.

Beszteri, B., Acs, E., Makk, J., Kovács, G., Márialigeti, K. & Kiss, K. T. 2001. Phylogeny of six naviculoid diatoms based on 18S rDNA sequences. *International Journal of Systematic and Evolutionary Microbiology* 51(4): 1581–1586.

Beszteri B., Ács É. & Medlin L.K. 2005. Ribosomal DNA sequence variation among sympatric strains of the *Cyclotella meneghiniana* complex (Bacillariophyceae) reveals cryptic diversity. *Protist* 156: 317–333.

Bold, H. C. & Wynne, M. J. 1978. *Introduction to the algae*. Prentice-Hall, Hoboken. 706 pp.

Bonhomme V., Picq, S., Gaucherel C. & Claude, J. 2014. Momocs: Outline Analysis Using R. *Journal of Statistical Software* 56(13): 1–24.

Bouchard, A. J., Hamilton, P. B., Savoie, A. M., & Starr, J. R. 2019. Molecular and morphological data reveal hidden diversity in common North American *Frustulia* species (Amphipleuraceae). *Diatom Research* 34(4): 205-223.

Bruder, K. & Medlin, L. K. 2007. Molecular assessment of phylogenetic relationships in selected species/genera in the naviculoid diatoms (Bacillariophyta). I. The genus *Placoneis*. *Nova Hedwigia* 85(3): 331–352.

Bruder, K. & Medlin, L. K. 2008. Morphological and Molecular Investigations of Naviculoid Diatoms. Iii. *Hippodonta* and *Navicula* S. S. *Diatom Research* 23(2): 331–347.

Cantonati, M., Leira, M., Angeli, N. & Lopez Rodriguez, C. 2012. *Naviculadicta langebertalotii* sp. nov (Bacillariophyta) from streams in Galicia (N-W Spain). *Nova Hedwigia* 141: 71–79.

Chapman, R. L. 2013. Algae: The world's most important "plants" – an introduction. *Mitigation and Adaptation Strategies for Global Change* 18(1): 5–12.

Choi, H.-G., Joo, H. M., Jung, W., Hong, S. S., Kang, S. & Kang, S.-H. 2008. Morphology and phylogenetic relationships of some psy-chrophilic polar diatoms (Bacillariophyta). *Nova Hedwigia* 133: 7–30.

Cleve, P. T. 1895. Synopsis of the Naviculoid Diatoms, Part II. *Kongliga Svenska-Vetenskaps Akademiens Handlingar* 27(3): 1–219.

Cox, E. J. 1988. Taxonomic studies on the Diatom Genus *Navicula* V. The establishment of *Parlibellus* gen. nov. for some members of *Navicula* sect. Microstigmaticae. *Diatom Research* 3(1) :9–38.

Cox E.J. 1996. *Identification of Freshwater Diatoms from Live Material*. Chapman and Hall, London. 158 pp.

Cox, E. 1999. Studies on the genus *Navicula* Bory VIII. Variation in valve morphology in relation to the generic diagnosis based on *Navicula tripunctata* (O.F. Müller) Bory. *Diatom Research* 14(2): 207–237.

Cox, E. & Williams, D. 2000. Systematics of naviculoid diatoms: The interrelationships of some taxa with a stauros. *European Journal of Phycology* 35(3): 273–282.

Damste, J. S., Muyzer, G., Abbas, B., Rampen, S.W., Masse, G., Allard, W. G., Belt, S. T., Robert, J. M., Rowland, S. J., Moldowan, J. M., Barbanti, S. M., Fago, F. J., Denisevich, P., Dahl, J., Trindade, L. A. & Schouten, S. 2004. The Rise of the Rhizosolenid Diatoms. *Science* 304(5670): 584–587.

Daugbjerg, N. & Andersen, R. A. 1997. A Molecular Phylogeny of the Heterokont Algae Based on Analyses of Chloroplast-Encoded *rbcL* Sequence Data. *Journal of Phycology* 33(6): 1031–1041.

Daugbjerg, N., & Guillou, L. 2001. Phylogenetic analyses of Bolidophyceae (Heterokontophyta) using *rbc* L gene sequences support their sister group relationship to diatoms. *Phycologia* 40(2): 153–161.

Debry, R., & Olmstead, R. 2000. A Simulation Study of Reduced Tree-Search Effort in Bootstrap Resampling Analysis. *Systematic Biology* 49(1): 171–179.

De Decker, S., Vanormelingen, P., Pinseel, E., Sefbom, J., Audoor, S., Sabbe, K. & Vyverman, W. 2018. Incomplete Reproductive Isolation Between Genetically Distinct Sympatric Clades of the Pennate Model Diatom *Seminavis robusta*. *Protist* 169(4): 569–583.

Dowle M. & Srinivasan A. 2018. data.table: Extension of "data.frame". R package version 1.11.4. https://CRAN.R-project.org/package=data.table

Edlund M.B. & Brant L.A. 1997. *Frustulia bahlsii* sp. nov., A Freshwater Diatom from the Eastern U.S.A. *Diatom Research* 12(2): 207–216.

Eswaran, K., Sivagama Sundari, M., Somasundaram, S. T. and Anantharaman, P. 2016. A simple approach combining microscopic and molecular techniques to identify diatoms isolated from Vellar Estuary, South East Coast of India. *Journal of Algal Biomass Utilization* 7(1): 71–77.

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). *Journal of Phycology* 44(1): 215–231.

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

Fox, M. G. & Sorhannus, U. M. 2003. RpoA: A Useful Gene for Phylogenetic Analysis in Diatoms. *Journal of Eukaryotic Microbiology* 50(6): 471–475.

Furnas, M. 1990. In situ growth rates of marine phytoplankton: approaches to measurement, community, species growth rates. *Journal of Plankton Research* 12(6): 1117–1151.

Gaiser E.E. & Johansen J. 2000. Freshwater diatoms from Carolina bays. *Diatom Research* 15: 75–130.

Gastineau, R., Leignel, V., Jacquette, B., Hardivillier, Y., Wulff, A., Gaudin, P., Bendahmane, D., Davidovich, N. A., Kaczmarska, I. & Mouget, J.-L. 2013. Inheritance of Mitochondrial DNA in the Pennate Diatom Haslea ostrearia (Naviculaceae) during Auxosporulation Suggests a Uniparental Transmission. *Protist* 164(3): 340–351.

Gastineau, R., Hansen, G., Davidovich, N. A., Davidovich, O., Bardeau, J.-F., Kaczmarska, I., Ehrman, J. M., Leignel, V., Hardivillier, Y., Jacquette, B., Poulin, M., Morançais, M., Fleurence, J. & Mouget, J.-L. 2016. A new blue-pigmented hasleoid diatom, *Haslea provincialis*, from the Mediterranean Sea. *European Journal of Phycology* 51(2): 156–170.

Gift, N. Stevens, P. F. 1997. Vagaries in the delimitation of character states in quantitative variation – an experimental study. *Systematic Biology* 46: 112–125.

Guiry, M. C. 2012. How many species of algae are there? *Journal of Phycology* 48: 1057–1063.

Guiry M.D. & Guiry G.M. 2019. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. http://www.algaebase.org; searched on 11 July 2019.

Guo L., Ren Y., Sui Z., Liu Y. & Zhang S. 2015. Comparison of potential diatom 'barcode' genes (the 18S rRNA gene and ITS, COI, *rbcL*) and their effectiveness in discriminating and determining species taxonomy in the Bacillariophyta. *International Journal of Systematic and Evolutionary Microbiology* 65(4): 1369–1380.

Hamilton, P. B., Lefebvre, K. E. & Bull, R. D. 2015. Single cell PCR amplification of diatoms using fresh and preserved samples. *Frontiers in Microbiology* 6(1084).

Hammer, Ø., Harper, D.A.T. & Ryan, P.D. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* 4(1): 9pp.

Hamsher, S. E., Evans, K. M., Mann, D. G., Poulíčková, A. & Saunders, G. W. 2011. Barcoding Diatoms: Exploring Alternatives to COI-5P. *Protist* 162(3): 405–422.

Hamsher, S. E. & Saunders, G. W. 2014. A floristic survey of marine tube-forming diatoms reveals unexpected diversity and extensive co-habitation among genetic lines of the *Berkeleya rutilans* complex (Bacillariophyceae). *European Journal of Phycology* 49(1): 47–59.

Hawk, H. L. & Geller, J. B. 2019. DNA entombed in archival seashells reveals low historical mitochondrial genetic diversity of endangered white abalone *Haliotis sorenseni. Marine and Freshwater Research* 70(3): 359–370.

Højsgaard S. & Halekoh U. 2018. doBy: Groupwise Statistics, LSmeans, Linear Contrasts, Utilities. R package version 4.6-1. https://CRAN.R-project.org/package=doBy

Hustedt F. 1930. Heft 10. *Bacillariophyta (Diatomeae). In: A. Pascher (ed.), Die Süsswasser-Flora Mitteleuropas, Zweite Auflage*. Verlag von Gustav Fischer, Jena. 466pp.

Hustedt, F. 1961–1966. Die Kieselalgen Deutschlands, Österreichs und der Schweiz mit Berücksichtigung der übrigen Länder Europas sowie der angrenzenden meeresgebeite. In: *Kryptogamen-Flora von Deutschland, Österreich und der Schweiz* (Ed. by L. Rabenhorst), Vol. Band VII. Teil 3. Akademische Verlagsgesellschaft, Leipzig. 816 pp.

John, J. 2012. *A Beginner's Guide to Diatoms*. Gantner Verlag, Liechtenstein, 150pp.

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. *European Journal of Phycology* 40(1): 61–87.

Kaczmarska, I., & Chan, A. M. 1995. *Navicula pulchripora*, A New Species From The Coastal Waters Of Texas, USA. *Diatom Research* 10(1): 131–143.

Katoh, K. & D. M. Standley. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 30(4): 772–80.

Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Mentjies P. & Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12): 1647–1649.

Keck, F., Rimet, F., Franc, A. & Bouchez, A. 2015. Phylogenetic signal in diatom ecology: perspectives for aquatic ecosystems biomonitoring. *Ecological Applications* 26(3): 861–872.

Kermarrec, L., Franc, A., Rimet, F., Chaumeil, P., Humbert, J. F. & Bouchez, A. 2013. Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: A test for freshwater diatoms. *Molecular Ecology Resources* 13(4): 607–619.

Ki, J.-S., Cho, S.-Y., Katano, T., Jung, S. W., Lee, J., Park, B. S., Kang, S.-H. & Han, M.-S. 2009. Comprehensive comparisons of three pennate diatoms, *Diatoma tenuae*, *Fragilaria vaucheriae,* and *Navicula pelliculosa*, isolated from summer Arctic reservoirs (Svalbard 79°N), by fine-scale morphology and nuclear 18S ribosomal DNA. *Polar Biology* 32(2): 147–159.

Kim Tiam, S., Feurtet-Mazel, A., Delmas, F., Mazzella, N., Morin, S., Daffe, G. & Gonzalez, P. 2012. Development of q-PCR approaches to assess water quality: Effects of cadmium on gene expression of the diatom *Eolimna minima*. *Water Research* 46(4): 934–942.

Kociolek, J. P., Theriot, E. C. & Williams, D. M. 1989. Inferring diatoms phylogeny: a cladistics perspective. *Diatom Research* 4(2): 289–300.

Kociolek, J. P. 1996. Comment: taxonomic instability and the creation of Naviculadicta Lange-Bertalot in Lange-Bertalot & Moser, a new catch-all genus of diatoms. *Diatom Research* 11(1): 219–222.

Kooistra, W. H. C. F., Gersonde, R., Medlin, L. K., & Mann, D. G. 2007. The Origin and Evolution of the Diatoms: Their Adaptation to a Planktonic Existence. In*: Evolution of Primary Producers in the Sea*. Elsevier, Amsterdam. 476pp.

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. 2019. RAxML-NG: A Fast, Scalable and User-Friendly Tool for Maximum Likelihood Phylogenetic Inference. *Bioinformatics* 35(21): 4453–55.

Kruskal, J. B. 1964. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29(2): 115–129.

Kulikovskiy, M. S., Andreeva, S. A., Gusev, E. S., Kuznetsova, I. V., & Annenkova, N. V. 2016. Molecular phylogeny of monoraphid diatoms and raphe significance in evolution and taxonomy. *Biology Bulletin* 43(5): 398–407.

Lanfear R. 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* 29(6): 1695–1701.

Lange-Bertalot, H. & Moser, G. 1994. *Brachysira*. Monographie der Gattung und *Naviculadicta* nov. gen. *Bibliotheca Diatomologica* 29: 1–212.

Lange-Bertalot H. & Jahn R. 2000. On the Identity of *Navicula* (*Frustulia*) *rhomboides* and *Frustulia saxonica* (Bacillariophyceae). *Systematics and Geography of Plants* 70(2): 255–261.

Lange-Bertalot, H. 2001. *Navicula* sensu stricto, 10 genera separated from *Navicula* sensu lato, *Frustulia*. Diatoms of Europe. Vol. 2. Edited by H. Lange-Bertalot. A.R.G. Gantner Verlag, Königstein, Germany. 526pp.

Lee, R. E. 2008. *Phycology*. Cambridge University Press.

Li, Y., Chen, X., Sun, Z., & Xu, K. 2017. Taxonomy and molecular phylogeny of three marine benthic species of *Haslea* (Bacillariophyceae), with transfer of two species to *Navicula*. *Diatom Research* 32(4): 451–463.

Lobban, C. S., Tharngan, B. G. & Ashworth, M. P. 2018. Four new *Licmophora* species (Licmophorales), with a review of valve characters and exploration of cingulum characters, including a new septum type. *Diatom Research* 33(2): 187–217.

Loftus, S. E. & Johnson, Z. I. 2019. Reused Cultivation Water Accumulates Dissolved Organic Carbon and Uniquely Influences Different Marine Microalgae. *Frontiers in Bioengineering and Biotechnology* 7(101).

Lundholm, N., Bates, S.S., Baugh, K.A., Bill, B.D., Connell, L.B., Léger, C. & Trainer, V.L. 2012. Cryptic and pseudo-cryptic diversity in diatoms—with descriptions of *Pseudo-nitzschia hasleana* sp. nov. and *P. fryxelliana*. *Journal of Phycology* 48: 436–454.

M.D. Guiry in Guiry, M.D. & Guiry, G.M. 2020. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. http://www.algaebase.org.

Mann, D. G. 1989. The diatom genus *Sellaphora*: Separation from *Navicula*. *British Phycological Journal* 24(1): 1–20.

Mann D.G. 1999. The species concept in diatoms. *Phycologia* 38: 437–495.

Mayr, E. 2000. The Biological Species Concept. In: *Species concepts and phylogenetic theory: A debate.* (Ed. by: Wheeler, Q. D., & Meier, R.). Columbia University Press, New York City. 245pp.

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(3): 245–270.

Medlin, L. K. 2011. The Permian–Triassic mass extinction forces the radiation of the modern marine phytoplankton. *Phycologia* 50(6): 684–693.

Medlin, L. K. 2014. Evolution of the Diatoms: VIII. Re-Examination of the SSU-Rrna Gene Using Multiple Outgroups and a Cladistic Analysis of Valve Features. *Journal of Biodiversity, Bioprospecting and Development* 1(3): 16pp.

Moreno, C. 2015. *A transcriptomic comparison of physiological responses to iron and light in southern ocean diatoms.* M.Sc. thesis, University of North Carolina, Chapel Hill, United States of America.

Moro, I., Maistro, S., Cassaro, L., Rascio, N. & Andreoli, C. 2010. Morphology, 18S rDNA sequence and *rbcL* phylogeny of *Navicula veneta* (Bacillariophyceae) from thermal muds in Italy. *Cryptogamie Algologie* 31(2): 209–219.

Nabors, W. 2004. *Introduction to Botany*. Pearson Education, London. 656 pp

Nakov, T., Ashworth, M. & Theriot, E. C. 2015. Comparative analysis of the interaction between habitat and growth form in diatoms. *ISME J.* 9(1): 246–255.

Nakov T., Beaulieu J.M. & Alverson A.J. 2018. Accelerated diversification is related to life history and locomotion in a hyperdiverse lineage of microbial eukaryotes (Diatoms, Bacillariophyta). *New Phytologist* 219(1): 462–473.

National Oceanic and Atmospheric Administration (NOAA). (2014). Ocean Explorer. Retrieved Nov. 10, 2017, from http://oceanexplorer.noaa.gov/facts/oceanproduction.html

Nawrocki, E. P. 2009. *Structural RNA Homology Search and Alignment using Covariance Models.* Ph.D. thesis, Washington University in Saint Louis, School of Medicine, Saint Louis, United States of America.

Novis, P. M., Braidwood, J. & Kilroy, C. 2012. Small diatoms (Bacillariophyta) in cultures from the Styx River, New Zealand, including descriptions of three new species. *Phytotaxa* 64: 11–45.

Nonoyama T, Kazamia E, Nawaly H, Gao X, Tsuji Y, Matsuda Y, Bowler C, Tanaka T, G. Dorrell R. 2019. Metabolic Innovations Underpinning the Origin and Diversification of the Diatom Chloroplast. *Biomolecules* 9(8):322.

Pniewski, F.F., Friedl, T. & Latala, A. 2010. Identification of diatom isolates from the Gulf of Gdańsk: testing of species identifications using morphology, 18S rDNA sequencing and DNA barcodes of strains from the Culture Collection of Baltic Algae (CCBA). *Oceanological and Hydrobiological Studies* 39(3): 3–20.

Poulíčková, A., Veselá, J., Neustupa, J., & Škaloud, P. 2010. Pseudocryptic diversity versus cosmopolitanism in diatoms: a case study on *Navicula cryptocephala* Kütz. (Bacillariophyceae) and morphologically similar taxa. *Protist* 161(3): 353–369.

Prasetiya, F. S., Gastineau, R., Poulin, M., Lemieux, C., Turmel, M., Syakti, A. D., Hardivillier, Y., Widowati, I., Risjani, Y., Iskandar, I., Subroto, T., Falaise, C., Arsad, S., Safitri, I., Mouget, J.-L. & Leignel, V. 2019. *Haslea nusantara* (Bacillariophyceae), a new blue diatom from the Java Sea, Indonesia: Morphology, biometry and molecular characterization. *Plant Ecology and Evolution* 152(2): 188–202.

Prelle, L. R., Graiff, A., Gründling-Pfaff, S., Sommer, V., Kuriyama, K. & Karsten, U. 2019. Photosynthesis and Respiration of Baltic Sea Benthic Diatoms to Changing Environmental Conditions and Growth Responses of Selected Species as Affected by an Adjacent Peatland (Hütelmoor). *Frontiers in Microbiology* 10(1500).

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reavie, E.D. & Andresen, N.A. 2020. *Navicula* from the coastal Laurentian Great Lakes. *Diatom Monographs* 19(2): [1]-202, 45 figs, 15 pls.

Riisberg, I., Orr, R. J. S., Kluge, R., Shalchian-Tabrizi, K., Bowers, H. A., Patil, V., Edvardsen, B., & Jakobsen, K. S. 2009. Seven Gene Phylogeny of Heterokonts. *Protist* 160(2): 191–204.

Round F.E., Crawford R.M. & Mann D.G. 1990. *The Diatoms, Biology & Morphology of the Genera.* Cambridge University Press, Cambridge. 751pp.

Ruck, E., Theriot, E. 2011. Origin and Evolution of the canal raphe system. *Protist* 162(5): 723–737.

Ryther, J.H. 1969. Photosynthesis and fish production in the sea. *Science* 166: 72–76.

Saba, F., Noroozi, M., Ghahremaninejad, F., Amoozegar, M. A., Sedghi, M., Shahzadeh Fazeli, S. A. & Papizadeh, M. 2016. Isolation, Purification, and Identification of Three Diatom species (Bacillariophyceae) from Gomishan Wetland using Phylogeny and Silica-Wall Ultra-Structure Analysis. *Rostaniha* 17(1): 28–39.

Sabir, J. S. M., Theriot, E. C., Manning, S. R., Al-Malki, A. L., Khiyami, M. A., Al-Ghamdi, A. K., Sabir, M. J., Romanovicz, D. K., Hajrah, N. H., El Omri, A., Jansen, R. K. & Ashworth, M. P. 2018. Phylogenetic analysis and a review of the history of the accidental phytoplankter, *Phaeodactylum tricornutum* Bohlin (Bacillariophyta). *PLOS ONE* 13(6): e0196744.

Sánchez, O., Ferrera, I., Vigués, N., Oteyza, T. G. de, Grimalt, J. & Mas, J. 2006. Role of cyanobacteria in oil biodegradation by microbial mats. *International Biodeterioration & Biodegradation* 58(3–4): 186–195.

Sato, S., Matsumoto, S. & Medlin, L. K. 2009. Fine structure and 18S rDNA phylogeny of a marine araphid pennate diatom *Plagiostriata goreensis* gen. et sp. nov. (Bacillariophyta). *Phycological Research* 57(1): 25–35.

Schlie, C. & Karsten, U. 2016. Growth of the Antarctic sea ice diatom *Navicula* cf. *normaloides* Cholnoky at different temperatures and salinities. *Algological Studies,* 151-152(1): 39–49.

Schloerke B., Crowley J., Cook D., Briatte F., Marbach M., Thoen E., Elberg A. & Larmarange J. 2018. GGally: Extension to 'ggplot2'. R package version 1.4.0 https://CRAN.R-project.org/package=GGally

Scotland, R. W., Olmstead, R. G., Bennett, J. R. 2003. Phylogeny reconstruction: the role of morphology. *Systematic Biology* 52: 539–548.

Simonsen, R. 1974. The diatom plankton of the Indian Ocean expedition of R/V ‘‘Meteor’’1964–1965. ‘‘Meteor’’ Forschungsergebnisse, Reihe D, 19: 1–107.

Siver P.A. & Baskette G. 2004. A morphological examination of *Frustulia* (Bacillariophyceae) from the Ocala National Forest, Florida, USA. *Canadian Journal of Botany* 82(5): 629–644.

Siver P.A. & Hamilton P.B. 2011. *Diatoms of North America: The Freshwater Flora of Waterbodies on the Atlantic Coastal Plain*. Iconographia Diatomologica 22. A.R.G. Gantner Verlag KG, Ruggell. 916pp.

Smetacek, V. S. 1985. Role of sinking in diatom life-history cycles: Ecological, evolutionary and geological significance. *Marine Biology* 84(3): 239–251.

Smol, J. 1988. Paleoclimate proxy data from freshwater arctic diatoms. *SIL Proceedings, 1922–2010* 23(2): 837–844.

Smol, J., & Stoermer, E. 2010. *The Diatoms: Applications for the Environmental and Earth Sciences*. Cambridge University Press, Cambridge. 686 pp.

Sorhannus, U. & Fox, M. G. 2012. Phylogenetic Analyses of a Combined Data Set Suggest that the Attheya Lineage is the Closest Living Relative of the Pennate Diatoms (Bacillariophyceae). *Protist* 163(2): 252–262.

Souffreau, C., Verbruggen, H., Wolfe, A. P., Vanormelingen, P., Siver, P. A., Cox, E. J. & Vyverman, W. 2011. A time-calibrated multi-gene phylogeny of the diatom genus *Pinnularia*. *Molecular Phylogenetics and Evolution* 61(3): 866–879.

Stachura-Suchoples, K., Enke, N., Schlie, C., Schaub, I., Karsten, U. & Jahn, R. 2015. Contribution towards a morphological and molecular taxonomic reference library of benthic marine diatoms from two Arctic fjords on Svalbard (Norway). *Polar Biology* 39(11): 1933–1956.

Stamatakis A. 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* 30(9): 1312–1313.

Starr J.R., Harris S.A. & Simpson D.A. 2004. Phylogeny of the Unispicate Taxa in Cyperaceae Tribe Cariceae I: Generic Relationships and Evolutionary Scenarios. *Systematic Botany* 29(3): 528–544.

Stevens, P. F. 2000. On characters and character states: do overlapping and non-overlapping variation, morphology and molecules all yield data of the same value? *Homology and systematics* 81–105.

Stoof-Leichsenring, K. R., Epp, L. S., Trauth, M. H. & Tiedemann, R. 2012. Hidden diversity in diatoms of Kenyan Lake Naivasha: A genetic approach detects temporal variation. *Molecular Ecology* 21(8): 1918–1930.

Swofford D.L. 2002. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods).* Version 4. Sinauer Associates, Sunderland, Massachusetts.

Theriot, E. C., Ashworth, M. P., Nakov, T., Ruck, E. & Jansen, R. K. 2015. Dissecting signal and noise in diatom chloroplast protein encoding genes with phylogenetic information profiling. *Molecular Phylogenetics and Evolution* 89: 28–36.

Theriot, E. C., Ashworth, M., Ruck, E., Nakov, T. & Jansen, R. K. 2010. A preliminary multigene phylogeny of the diatoms (Bacillariophyta): Challenges for future research. *Plant Ecology and Evolution* 143(3): 278–296.

Tiffany, L. H. 1985. *Algae: The Grass of Many Waters*. Springfield, Irvine. 199pp.

Underwood, G. J. C. (2001). *Encyclopedia of Ocean Sciences*. Elsevier Ltd, Amsterdam. 4306 pp.

Urbánková P. & Veselá J. 2013. DNA-barcoding: a case study in the diatom genus *Frustulia* (Bacillariophyceae). *Nova Hedwigia* 142: 147–162.

Urbánková P., Scharfen V. & Kulichová J. 2016. Molecular and automated identification of the diatom genus *Frustulia* in northern Europe. *Diatom Research* 31(3): 217–229.

Van de Vijver, B., & Kopalová, K. 2014. Four *Achnanthidium* species (Bacillariophyta) formerly identified as *Achnanthidium minutissimum* from the Antarctic Region. *European Journal of Taxonomy*, 79: 1–19.

Vanelslander, B., Créach, V., Vanormelingen, P., Ernst, A., Chepurnov, V. A., Sahan, E., Muyzer, G., Stal, L. J., Vyverman, W. & Sabbe, K. 2009. Ecological differentiation between sympatric pseudocryptic species in the estuarine benthic diatom *Navicula phyllepta* (Bacillariophyceae)1: ecology of pseudocryptic species. *Journal of Phycology* 45(6): 1278–1289.

Vanormelingen P., Evans K.M., Chepurnov V., Vyverman W. & Mann D.G. 2013. Molecular species discovery in the diatom *Sellaphora* and its congruence with mating trials. *Fottea* 13(2): 133–148.

Venables W.N. & Ripley, B.D. 2002. *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. 498pp.

Veselá J., Urbánková P., Cerná K. & Neustupa J. 2012. Ecological variation within traditional diatom morphospecies: diversity of *Frustulia rhomboides* sensu lato (Bacillariophyceae) in European freshwater habitats. *Phycologia* 51: 552–561.

Wetzel, C.E., Ector, L., Van de Vijver, B., Compère, P. & Mann, D.G. 2015 Morphology, typification and critical analysis of some ecologically important small naviculoid species (Bacillariophyta) *Fottea* 15(2): 203–234.

Wiens, J. 2001. Character analysis in morphological phylogenetics: problems and solutions. *Systematic Biology* 50(5): 689–699.

Wickham H. 2009. plyr: Tools for splitting, applying and combining data. R package version 0.1.9 http://CRAN.R-project.org/package=plyr.

Wickham H. 2011. The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical Software*, 40(1): 1–29.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York. 260pp.

Witkowski, A., Metzeltin, D., Lange-Bertalot, H., & Bafana, G. 1997. *Fogedia* gen. nov. (Bacillariophyceae), a new naviculoid genus from the marine littoral. *Nova Hedwigia* 65: 79–98.

Witkowski, A., Lange-Bertalot, H, & Stachura, K. 1998. New and confused species in the genus *Navicula* (Bacillariophyceae) and the consequences of restrictive generic circumscription. *Cryptogamie Algol*. 19(1): 83–108.

Witkowski, A., Li, C., Zgłobicka, I., Yu, S., Ashworth, M., Dąbek, P., Qin, S., Tang, C., Krzywda, M., Ruppel, M., Theriot, E. C., Jansen, R. K., Car, A., Płociński, T., Wang, Y., Sabir, J. S. M., Daniszewska-Kowalczyk, G., Kierzek, A. & Hajrah, N. H. 2016. Multigene Assessment of Biodiversity of Diatom (Bacillariophyceae) Assemblages from the Littoral Zone of the Bohai and Yellow Seas in Yantai Region of Northeast China with some Remarks on Ubiquitous Taxa. *Journal of Coastal Research* 74: 166–195.

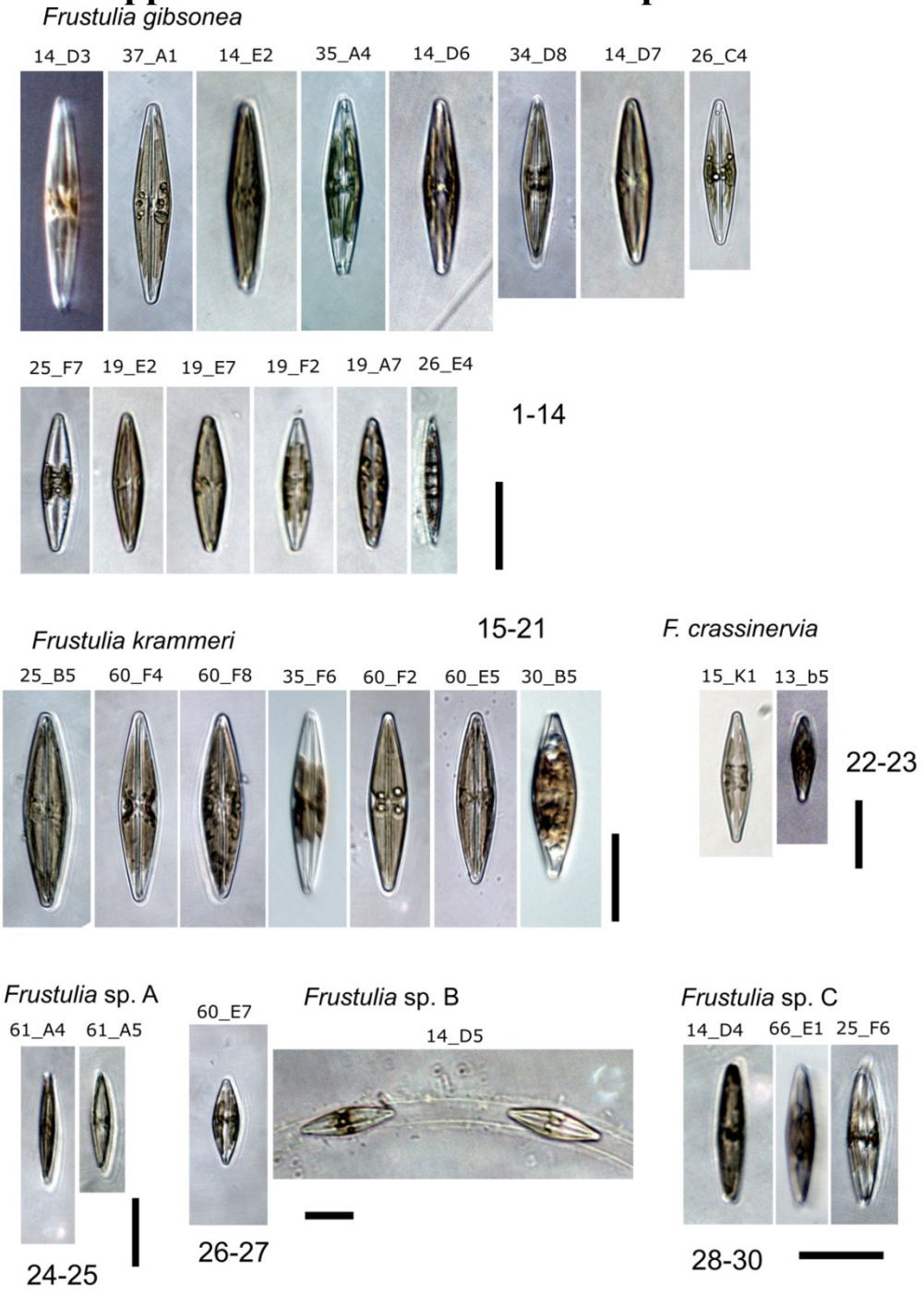
Witkowski, A., Ashworth, M., Li, C., Sagna, I., Yatte, D., Górecka, E., Franco, A. O. R., Kusber, W. H., Klein, G., Lange-Bertalot, H., Dąbek, P., Theriot, E. C. & Manning, S. R. (2020). Exploring Diversity, Taxonomy and Phylogeny of Diatoms (Bacillariophyta) from Marine Habitats. Novel Taxa with Internal Costae. *Protist* 171(2): 125713.

Yool, A., & Tyrrell, T. 2003. The role of diatoms in regulating the ocean's silicate cycle. *Global Biogeochemical Cycles* 17: 1–24.

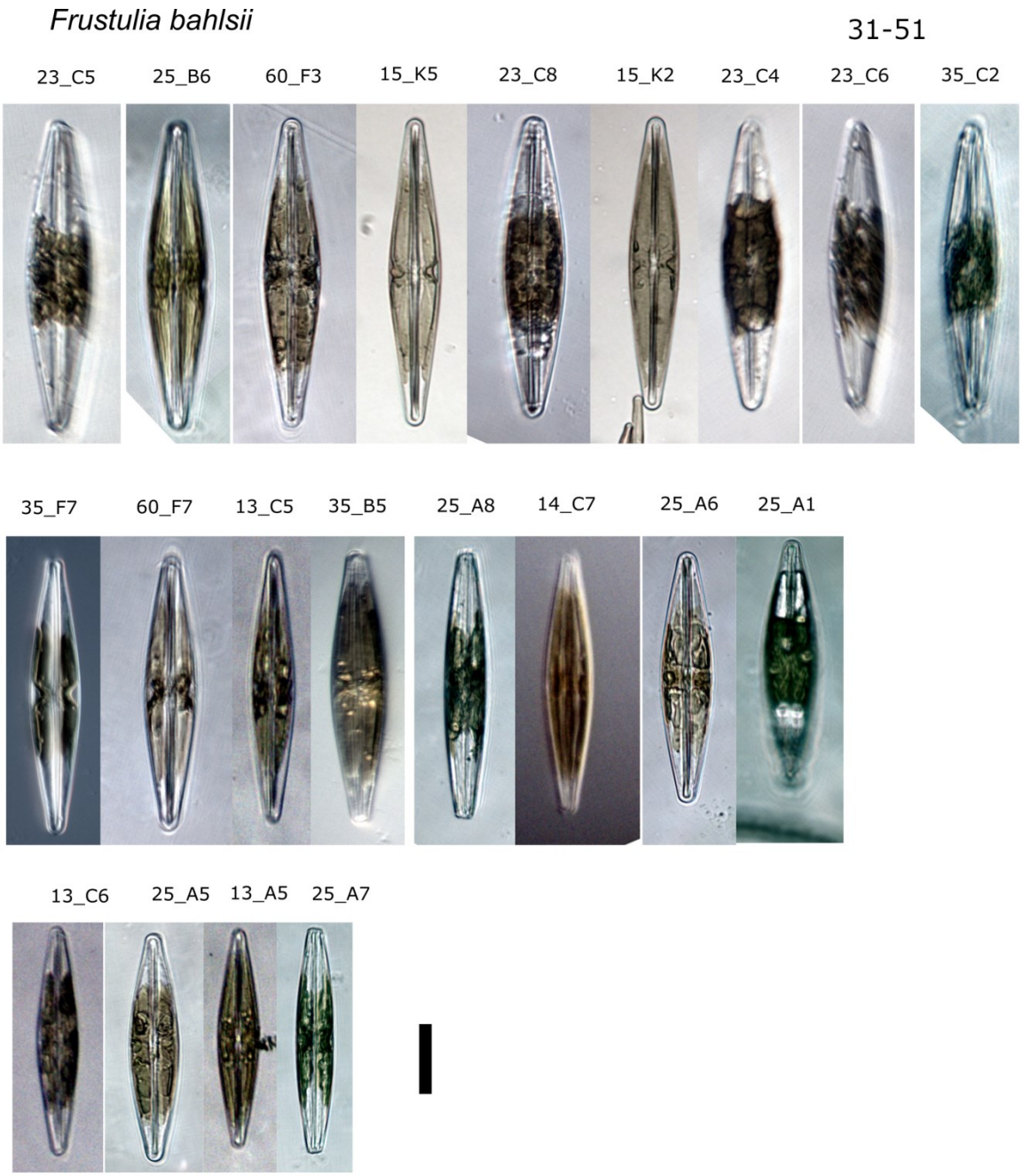
Zimmermann, J., Abarca, N., Enk, N., Skibbe, O., Kusber, W.-H. & Jahn, R. 2014. Taxonomic Reference Libraries for Environmental Barcoding: A Best Practice Example from Diatom Research. *PLOS ONE* 9(9): e108793.

Zimmermann, J., Glöckner, G., Jahn, R., Enke, N. & Gemeinholzer, B. 2015. Metabarcoding vs. Morphological identification to assess diatom diversity in environmental studies. *Molecular Ecology Resources* 15(3): 526–542.

Appendix A – Supplemental Material from Chapter 2



Figs A.1–A.30. LM images of live specimens used in the DNA analyses. Figs 1–14. *F. gibsonea*. Figs 15–21. *Frustulia krammeri*. Figs. 22–23. *Frustulia crassinervia*. Figs 24–25. *Frustulia* sp. (a) [*F.* cf. *saxonica*]. Figs 26–27. *Frustulia* sp. (b) [*F.* cf. *saxonica*]. Figs 28–30. *Frustulia* sp. (c). Scale bars, Figs 1-21 = 50 µm. Figs 22-25 = 30 µm. Figs 26-27 = 25 µm. Figs 28-30 = 50 µm.



Figs A.31–A.49. LM images of live specimens for *Frustulia bahlsii* used in the DNA analyses. Scale bar = 50 µm.