# An integrative approach to epilithic diatom diversity analysis in tropical streams from the Lerma-Chapala Basin, Central Mexico

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy

of Freie Universität Berlin

by

**Luis Demetrio Mora Hernández**

from El Sauz, Mexico

Berlin, November 2017

The work on this doctoral thesis was conducted from October 2013 to November 2017 at the Botanic Garden and Botanical Museum Berlin-Dahlem of Freie Universität Berlin under the supervision of Prof. Dr. Thomas Borsch.

1st Reviewer: **Prof. Dr. Thomas Borsch**

2nd Reviewer: **PD Dr. Franz Hölker**

Date of defence: 23.04.2018

*a mi mamá y mi teli*

## Acknowledgements

I thank Prof. Dr. Thomas Borsch for allowing me to conduct my research activities at the Botanic Garden and Botanical Museum Berlin and for mentoring my thesis. I also thank PD Dr. Franz Hölker for his time to be the second reviewer of this thesis.

Accomplishing this dissertation was made possible through the financial support from the following research and educational organizations from the Mexican government: Consejo Nacional de Ciencia y Tecnología (CONACyT) granted me a doctoral scholarship. Dirección General de Relaciones Internacionales, Secretaría de Educación Pública (DGRI-SEP) granted a complementary scholarship for postgraduate studies. Consejo de Ciencia y Tecnología del Estado de Querétaro (CONCyTEQ) financed the transportation back to Mexico to conduct the second and third field trips. Comisión Nacional de Acuacultura y Pesca (CONAPESCA) authorized the permission to collect samples.

I would like to express my deep gratitude to my supervisor Dr. Regine Jahn for allowing me to conduct my research over the last four years in the research group led by her. For her supervision, support, thoroughness, for giving me the freedom to conduct my research but also helping me keep focused. I am deeply thankful to my three supervisors Dr. Nélida Abarca, Dr. Neela Enke and Dr. Jonas Zimmermann. I really appreciate all your advice, help, support, cheerfulness and for keeping me going. Thank you Neela for the initiation to this project and all the interesting workshops. Thank you Nélida for your listening and for all the talks and advice on diatom identification as well as non diatom topics during our lunch time. Thank you Jonas for always being so supportive and for your critical thinking that helped me get new insights into molecular studies and for inspiration. I thank you all for your promptness in reviewing my writings, particularly at the very end phase of this dissertation.

My special gratitude goes to Dr. Javier Carmona Jiménez for co-supervising my research. For supporting my field trip campaigns as well as my stays at his laboratory while doing the water chemistry analyses. For his wisdom about rivers and for his wise advice while observing macroalgae.

I thank Wolf-Henning Kusber for his help in all aspects of diatom nomenclature and for always being a helping hand.

I am indebted to Jana Bansemer for introducing me to molecular biology techniques and diatom cultivation, for her cheerfulness and for her never ending patience while listening to my very first words in German. I thank Dr. Oliver Skibbe for his tips and tricks on diatom cultivation.

THANKS TO:

Monika Lüchow, Kim Govers and Juliane Bettig for guiding and accompanying me while working at the Scanning Electron Microscope. Without your advice it would not have been possible to obtain the beautiful images presented in this work.

Susan Mbedi and Kirsten Richter (BEGENDIV) for their advice while conducting the library preparation for metabarcoding, for their endurance and guidance. Verena Deutschmeyer for her help in the laboratory during her stay in the BGBM.

Sebastian Proft for lots of help with bioinformatics and his promptness in replying to my questions. Lina Ahn Van for your joy, some English clarifications and for sharing your expertise on cymbelloid diatoms. To Dr. Katarzyna Stachura-Suchoples for sharing your knowledge on fossil diatoms.

Petra Werner for the moments shared at the orange lab while making diatom preparations.

Calvinna Caswara and Dr. José Grau for their help with several bioinformatics tasks.

Matthias Simon for IT support and for his long standing patience in helping me recover my accounts the very many times I forgot my passwords.

Verónica Aguilar Zamora for creating the maps contained in this thesis and the creation of many more maps.

Dr. Enrique Cantoral Uriza for introducing me to the diatom world back in 2008.

My friends and colleagues from Laboratorio de Ecosistemas de Ribera, for their help and advice while preparing material for the field trips and while conducting the water chemistry analyses: Mariana Cartagena Alcántara, Rogelio Rodríguez Flores, Victor Hugo Salinas and María Elena Valeriano.

My friends and colleagues, current and former members of the BGBM: Dr. Patricia Hernández-Ledesma, Dr. Sascha Ismail, Dubán Canal Gallego, Sarah Bollendorf, Dr. Elmira Maharramova, Teresa Ortuño, Nana Silakadze, Dr. Pedro González Gutierrez and Dr. Virginia Duwe.

Herr and Frau Wörner for their help and support at home and for all the logistics when having visitors.

Dr. Zeina El Habre, Dr. Han Cheng Lie and soon (Dr.) Neta Tuvia for your advice during our peer mentoring sessions.

My Berlin friends Ulli, Martin R., Uwe, Jose, Robert, Sora, Tim and Sven L.

My CO-Gruppe-Jungs for all the stories shared.

Current and former members I have met in the Lausi-WG for their cheerfulness and support.

Andrea García Servín and Jimena Pérez Sánchez.

My very good friend Dr. Maja Grubišić for sharing those good times at the laboratory and outside the Botanic Garden. For your support, help and encouragement.

My dear friend Martin Jagodzinski for your support during the last four years, for never doubting me. For your help while preparing the figure plates contained in this work.

My family:

Boris König for your support in all scenarios, for those adventurous trips, our bike trips, for your critical thinking, for your never ending biology questions and everyting else.

To my parents for always supporting me, at home or away, for always being there. For your help while collecting the samples that form the basis of this work. Gracias.

**Summary**

The study and application of diatoms as biological indicators in Mexico is infrequent, despite the multiple advantages offered by these organisms in the assessment and monitoring of freshwaters. In the few diatom studies of Mexican freshwaters, there seems to be low species diversity and an inherent cosmopolitanism of taxa, contrary to what would have been expected of a megadiverse country. This was mainly because of 1) force-fitting identifications based on monographs from temperate regions and 2) the lone use of light microscopy for identifications, which not always differentiates between closely related species. In order to better assess the diatom diversity of the country and to set an identification baseline for future studies using diatoms as biological indicators, this dissertation presents an integrative approach to epilithic diatom diversity analysis in tropical streams from the Lerma-Chapala Basin, Central Mexico.

This study is based on a representative sampling of diatoms from small, mostly undisturbed, mountain streams from the north and east sections of the basin, as well as samples from the west of the basin, which include the heavily polluted Lerma River and some of its major tributaries. Samples were collected during the most contrasting periods of the year, the rainy and dry seasons.

In the first part of this work, the diatom samples of the small mountain streams from the north and east sections of the basin were cultivated for morphological, ecological, molecular and phylogenetic analysis. The morphological evaluation resulted in the largest diatom diversity reported for Mexico to date, 274 infrageneric taxa, including the description of two new species, *Brachysira altepetlensis* and *Sellaphora queretana*. The ecological analysis revealed that the community composition observed was mainly driven by the ionic composition of the water, with indicator taxa identified for the varying conditions in pH, conductivity and nutrients. Under the premise that diatom identifications at species level in environmental DNA (eDNA) metabarcoding studies rely heavily on the completeness of reference databases, a regional morphological and molecular taxonomic reference library was assembled by diatom cultivation. The eDNA metabarcoding approach tested here, which integrates molecular and tree-based phylogenetic methods, revealed a larger diversity than the one recorded by morphological analysis. One quarter of the taxa assigned to species level in the eDNA metabarcoding approach was attributable to the herein assembled taxonomic reference library, supporting the aforementioned premise. The use of a regional sequence reference database is to increase the identification success, particularly in poorly studied regions such as the tropics. By comparing the diversity retrieved by morphology and eDNA metabarcoding, it was found that neither morphology nor eDNA metabarcoding were a better method than the other in catching the entire

diversity; they were rather complementary. The cultivation of diatoms revealed a concealed diversity not detected by morphology or eDNA metabarcoding, suggesting cultivation as a further method to unravel species diversity from environmental samples. The relative abundances recorded by morphology (diatom valves) and eDNA metabarcoding (sequence reads) showed large disparities, even after the application of correction factors. This suggests that further methodological improvements are needed in order to establish eDNA metabarcoding as a standard method for water analysis. Furthermore, the results presented here support the retrieval of DNA reference barcodes from High-Throughput Sequencing data.

In the second part of this study all the samples, including those from the Lerma River and its major tributaries, were studied by microscopy (light and scanning electron microscopy) in order to prepare a detailed illustrated identification guide. The analysis resulted in 307 infrageneric taxa, with the description of ten new species, belonging to *Cocconeis*, *Craticula*, *Gomphonema* and *Sellaphora*. This identification guide represents the baseline for diatom identification in future monitoring studies and programs in the region.

In the third part of this dissertation, the taxonomy and systematics of the *Planothidium lanceolatum*/*P. frequentissimum* species complexes were explored by morphological and molecular data obtained from clonal cultures and from sequences deposited at the INSDC databases. Besides Mexico, the analyzed strains came from France, Germany, the Faroe Islands, Korea, Lake Baikal in Russia, New Zealand and the USA. The analysis resulted in the recognition of eight species, with three described as new to science. Both molecular and morphology-based phylogenetic analyses led to postulate the sinus and cavum as important stable taxonomic characters. The fine-grained taxonomy applied in this study allowed to revisit the distribution of *Planothidium frequentissimum* and *P. lanceolatum*, taxa previously considered to be cosmopolitan.

Morphological and molecular based approaches for the study of diversity are often seen as antagonistic, with morphology seen as methodologically outdated and time consuming while molecular methods are seen as a means of choice. However, the results of this study rather underline complementarity of both methods. The integrative approach to the study of diatoms presented in this dissertation allowed an improved assessment of the diversity of epilithic diatoms from the Lerma-Chapala Basin, Central Mexico, setting the baseline for future monitoring studies in this biodiversity rich but threatened region of Mexico.

**Zusammenfassung**

Untersuchungen zu Anwendungen von Diatomeen als Bioindikatoren sind trotz der vielfältigen Vorteile, die diese Organismen bei der Beobachtung und Begutachtung von Süßwasser bieten, in Mexiko selten. In den wenigen Studien zu Mexikos Süßwassersystemen wird von einer geringen Artenvielfalt und überwiegend kosmopolitischen Taxa ausgegangen. Dies rührt vor allem aus (1) der Identifikation der Taxa ausschließlich nach Monographien der gemäßigten Zonen und (2) der alleinigen Nutzung der Lichtmikroskopie zur Identifikation, die nicht immer eine Unterscheidung nah verwandter Arten erlaubt. Zur genaueren Untersuchung der Diatomeenvielfalt des Landes und zur Schaffung einer Grundlage für zukünftige Arbeiten, die Diatomeen als Bioindikatoren behandeln, wird mit dieser Dissertation ein integrativer Ansatz zur Analyse der Vielfalt epilithischer Diatomeen in tropischen Bächen des Lerma-Chapala Beckens in Zentralmexiko vorgelegt.

In dieser Arbeit wurden repräsentative Proben von Diatomeen aus kleinen, weitgehend ungestörten Bergbächen aus den nördlichen und östlichen Teilen sowie dem westlichen Teil des Beckens verwendet. Diese Region schließt auch den stark verschmutzten Fluss Lerma und einige seiner wichtigsten Zuflüsse ein. Es wurden Proben während der beiden kontrastreichsten Saisons des Jahres gesammelt, der Trocken- und der Regenzeit.

Im ersten Teil dieser Arbeit wurden die Diatomeen aus den Proben der kleinen Bergbäche aus dem Nord- und Ostteil des Beckens für morphologische, ökologische, molekulare und phylogenetische Analysen kultiviert. Die morphologischen Untersuchungen ergaben die größte, bisher dokumentierte Vielfalt an beschriebenen Diatomeenarten Mexikos: 274 infragenerische Arten, darunter die Beschreibung zweier neuer Arten, *Brachysira altepetlensis* und *Sellaphora queretana*. Die ökologische Untersuchung zeigte, dass die beobachtete Zusammensetzung der Gemeinschaft vor allem durch die Ionenkonzentration des Wassers bestimmt wird. In Abhängigkeit der weiteren Parameter pH, Leitfähigkeit und Nährstoffgehalt konnten relevante Indikatorarten identifiziert und zugeordnet werden. In dem Wissen, dass die Identifikation von Diatomeen auf Artenebene mittels Environmental DNA (eDNA) Metabarcoding stark von der Vollständigkeit von Referenzdatenbanken abhängt, wurde eine regionale morphologische und molekulartaxonomische Referenzbibliothek mittels Diatomeenkultivierung erstellt. Der hier getestete eDNA-Metabarcoding-Ansatz, der molekulare und phylogenetische Stammbaum-Methoden zur Taxazuordnung einsetzt, offenbarte eine größere Vielfalt als die morphologischen Untersuchungen. Ein Viertel der Taxa, die durch die eDNA-Metabarcoding-Analyse auf Artenebene zugeordnet wurden, konnte nur auf Grundlage der eigens erstellten Referenzbibliothek zugeordnet werden. Der Nutzen einer regionalen

Sequenzdatenbank als Referenz liegt in der Steigerung des Identifikationserfolges, speziell in weniger gut untersuchten Regionen wie den Tropen. Im Vergleich der gefundenen Artenvielfalten durch morphologische Untersuchungen einerseits und eDNA-Metabarcoding-Analysen andererseits, zeigte sich, dass keine der beiden Methoden der anderen überlegen ist; die Methoden sind vielmehr als komplementär zu betrachten. Die Kultivierung und Analyse zusätzlicher Diatomeen brachte noch weitere Arten zum Vorschein, die durch reine Evaluierung der Umweltproben nicht detektiert werden konnten. Dies unterstreicht die enorme Bedeutung der Diatomeenkultivierung als ergänzende Methode zur Erschließung der Artenvielfalt. Die relativen Häufigkeiten, die durch die morphologische Untersuchung (Diatomeenschalen) oder die eDNA-Metabarcoding-Analyse (Sequenzen) verzeichnet werden konnten, zeigten im Vergleich der beiden Ansätze große Unterschiede, sogar nach Anwendung von Korrekturfaktoren. Dies weist darauf hin, dass weitere methodische Verbesserungen notwendig sind um eDNA-Metabarcoding als Routineverfahren zur Wasseranalyse einzusetzen. Desweiteren zeigen die Ergebnisse, dass DNA-Barcodes für die Referenzbibliothek durch Hochdurchsatzsequenzierung (HTS) gewonnen werden können.

Im zweiten Teil der Arbeit wurden alle Proben, auch diejenigen des Flusses Lerma und seinen größeren Zuflüssen durch Mikroskopie (Licht - und Rasterelektronenmikroskopie) untersucht, um detaillierte und bebilderte Bestimmungsreferenzen zu erhalten. Es konnten 307 infragenerische Taxa identifiziert werden, wobei zehn neue Arten beschrieben werden konnten, die zu den Gattungen *Cocconeis*, *Craticula*, *Gomphonema* und *Sellaphora* gehören. Diese Identifikationsdaten dienen als Grundlage für zukünftige Forschungsarbeiten in der Region.

Im dritten Teil der Dissertation wurde die Taxonomie und Systematik des Artenkomplexes *Planothidium lanceolatum* / *P. frequentissimum* mittels morphologischer und molekularer Daten, die aus Klonkulturen und aus INSDC-Sequenzen gewonnen wurden, untersucht. Neben Mexiko stammen die untersuchten Kulturen aus Frankreich, Deutschland, den Färöer Inseln, Korea, dem russischen Baikalsee, Neuseeland und den USA. Es konnten acht Arten identifiziert werden, darunter drei für die Wissenschaft neue. Sowohl die molekularen als auch die morphologisch-phylogenetischen Untersuchungen führten zur Erkenntnis, dass Sinus und Cavum wichtige stabile, taxonomische Merkmale darstellen. Der detaillierte und integrative Ansatz der Taxonomie, der in dieser Arbeit angewendet wurde, erlaubte eine Neubewertung der Zuordnung der Taxa *Planothidium frequentissimum* und *P. lanceolatum*, die bis dato als kosmopolitisch eingestuft wurden.

Morphologische und molekularbasierte Ansätze zur Untersuchung der Artenvielfalt werden oft als antagonistisch behandelt, wobei die morphologische Analyse als methodisch überholt und zeitaufwändig gilt und die molekulare Methode als Mittel der Wahl erachtet wird. Die Ergebnisse dieser Arbeit zeigen jedoch gerade die Komplementarität beider Ansätze auf. Der integrative Ansatz

zur Untersuchung von Diatomeen, der in dieser Dissertation vorgelegt wird, führt zu einer besseren Beschreibung der Artenvielfalt epilithischer Diatomeen des Lerma-Chapala Beckens in Zentralmexiko. Es wurde somit die Grundlage für zukünftige Forschungsarbeiten in dieser artenreichen, aber gefährdeten Region Mexikos geschaffen.

## Dissertation outline and collaboration statement

This dissertation is a cumulative work composed of a rationale and six chapters. The rationale states the fundamental reasons of my research. The general introduction (Chapter 1) provides the background and research objectives of the dissertation. The following four chapters are manuscripts either published (Chapters 2 and 5), submitted (Chapter 4) in peer-reviewed journals, or in preparation for submission (Chapter 3) to a peer-reviewed monographic series. Chapters 2-5 contain an abstract, introduction, material and methods, results, discussion, conclusion and references (Chapter 3 contains no discussion because it is an identification guide). In the general conclusions and outlook (Chapter 6) I provide the major findings of the dissertation and future research prospects.

**Rationale**

**Chapter 1**. General Introduction

**Chapter 2**. **Mora D**, Carmona J, Jahn R, Zimmermann J, Abarca N (2017) Epilithic diatom communities of selected streams from the Lerma-Chapala Basin, Central Mexico, with the description of two new species. PhytoKeys 88: 39-69. https://doi.org/10.3897/phytokeys.88.14612

Own contribution: lead authorship, performed the field, laboratory and microscopy work, data analysis and wrote the manuscript.

**Chapter 3**. Abarca N, **Mora D**, Israde-Alcántara I, Jahn R (*in preparation*) Diatoms from the Lerma-Chapala River Basin, Central Mexico: an identification guide.

Own contribution: shared lead authorship, performed the field, laboratory and microscopy work as well as the data analysis of the samples collected during the 2013-2014 campaigns and wrote the manuscript.

**Chapter 4**. **Mora D**, Abarca N, Proft S, Grau J, Enke N, Carmona J, Skibbe O, Jahn R, Zimmermann J (*submitted to Freshwater Science*) Morphology and metabarcoding! A test with stream diatoms from Mexico highlights the complementarity of methods.

Own contribution: lead authorship, performed the field, laboratory and microscopy work, data analysis and wrote the manuscript.

**Chapter 5**. Jahn R, Abarca N, Gemeinholzer B, **Mora D**, Skibbe O, Kulikovskiy M, Gusev E, Kusber WH, Zimmermann J (2017) *Planothidium lanceolatum* and *Planothidium frequentissimum* reinvestigated with molecular methods and morphology: four new species and the taxonomic importance of the sinus and cavum. Diatom Research 32(1): 75-107. https://doi.org/10.1080/0269249x.2017.1312548

Own contribution: performed the field, laboratory and microscopy work for the Mexican strains of 2014, conducted the morphometric analysis and contributed to the manuscript.

**Chapter 6**. General conclusions and outlook

## Rationale

Although freshwaters represent only a tiny fraction (0.01%) of the water on Earth, they contain almost 6% of the global biodiversity. Freshwater ecosystems provide invaluable ecosystem services essential to humanity such as supply of clean water. Despite their importance, freshwater ecosystems face serious challenges such as pollution, overexploitation, habitat degradation, flow modification and the invasion of alien species. In order to counteract degradation, legislations have been adopted at national and continental scale in several parts of the world, e.g. the Water Framework Directive in the European Union and the Clean Water Act in the United States. Those legislations establish the use of biological indicators in addition to the traditionally used physical and chemical indicators.

The benefit of using biological indicators over physical and chemical indicators relies on the fact that biological communities provide an overall assessment of environmental quality including a historical record. Among the organisms used as indicators, diatoms are one of the most widely used. But their identification remains a major impediment in their use and the conclusion that can be drawn from them. Diatom identification, based on morphological features of their cell walls, is challenging due to their microscopic size, enormous diversity and plasticity to environmental conditions, requiring highly skilled taxonomists.

In spite of the advantages of using diatoms as biological indicators, their use in Mexico is very rare. In the few diatom studies of Mexican freshwaters, there seems to be a low diversity and an inherent cosmopolitanism, contrary to what would be expected of a megadiverse country. This is mainly because of force-fitting identifications based on monographs from temperate regions.

This dissertation contributes to filling this gap by performing an integrative analysis of the epilithic diatom diversity in tropical streams of the Lerma-Chapala Basin, Central Mexico. Diversity was analyzed using morphological, ecological, molecular and phylogenetic methods, as well as by the cultivation of diatoms. In this study, the diversity identified by all the employed methods is critically assessed to determine if one method performs better than the others, as several studies have suggested. Abundance data retrieved from both morphology and metabarcoding was also critically assessed because abundances are essential in monitoring studies. Furthermore, the importance of the sinus and the cavum as morphological characters in *Planothidium* is discussed, as well as the distribution of some of its species considered to have wide distributions. Overall, this dissertation highlights the importance of integrative approaches to better assess diatom diversity.

# 1 General introduction

## 1.1 A primer to the diatoms

Diatoms (Bacillariophyta) are unicellular eukaryotes, predominantly photoautotrophic, characterized by silica cell walls (Round et al. 1990). Diatoms are key global players in the biogeochemical cycles of carbon and silicon, being responsible for about one-fifth of the net primary production through photosynthesis on Earth, making diatoms more productive than the tropical rainforests (Falkowski et al. 1998; Field et al. 1998; Nelson et al. 1995; Sarthou et al. 2005).

### 1.1.1 Discovery and early classifications of diatoms

The invention of the microscope led to the discovery of microorganisms including diatoms. Diatoms colonizing the roots of the pond-weed *Lemna*, observed by an anonymous person in 1703 (Anonymous 1703), represent the first record of diatoms. His findings, probably corresponding to *Tabellaria flocculosa*, were published in the Philosophical Transactions of the Royal Society of London (Round et al. 1990; Williams and Kociolek 2011). Even though the renowned microscopist Antoni van Leeuwenhoek probably also observed diatoms (Leeuwenhoek 1703), his illustrations are not deemed completely verifiable as such (Round et al. 1990). Several diatoms were described with Latin binomials already in the second half of the 18th century, with the contributions of Otto Friedrik Müller of significant importance (Müller 1783; 1786), because what he described as *Vibrio paxillifer* later became the type of *Bacillaria*, the first diatom genus (Gmelin 1791). Müller considered diatoms to be animals, calling them "animacula infusoria" along with dinoflagellates, amoebae, ciliates and other protists. Diatoms continued to be treated as animals by several authors e.g. Bory de Saint-Vincent (1822) until Kützing (1844) treated them as plants; the sole exception before Kützing being Ehrenberg (1838), who treated diatoms as Infusoria (Jahn 1995). Over time, improvements in microscopy translated into progress in the knowledge of diatom biology and taxonomy, diatoms being indeed a preferred test for microscope lenses (Round et al. 1990). During the second half of the 19th century, microscopy was very popular and diatoms were sought after because of their beauty and intricacies, marveling observers alike including Charles Darwin, who dedicated a few lines to the diatoms in his seminal work, On the Origin of Species: *"Few objects are more beautiful than the minute siliceous cases of the diatomaceæ: were these created that they might be examined and admired under the higher powers of the microscope? The beauty in this latter case, and in many others, is apparently wholly due to symmetry of growth"* (Darwin 1866).

### 1.1.2 Origin

According to the fossil record which is extensive in diatoms, the origin of diatoms dates back to the early Jurassic, around 190 million years ago. But molecular clock data suggest an earlier origin during the Triassic no earlier than 240 million years ago (Kooistra and Medlin 1996; Sims et al. 2006; Sorhannus 2007). This 60 million years difference between the oldest fossils and the molecular clock data suggest that the early diatoms were not silicified (Raven and Waite 2004). The earliest diatoms are thought to be marine and the first major invasion of terrestrial habitats occurred later in the Cenozoic, though minor terrestrial incursions may have occurred in the Mesozoic (Sims et al. 2006).

Diatoms originated from a secondary endosymbiotic event. The first endosymbiotic event occurred about 1.5 billion years ago, after a eukaryotic heterotroph engulfed a cyanobacterium, giving rise to the ancestor cell of glaucophytes, chlorophytes and rhodophytes, as well as land plants. The secondary endosymbiotic event occurred around 500 million years ago, after another eukaryotic heterotroph engulfed a red alga, giving rise to the plastids of the cryptophytes, haptophytes and stramenopiles (group including diatoms and brown algae) (Armbrust 2009; Armbrust et al. 2004; Yoon et al. 2004). There are several particularities of diatoms that serve as evidence of their origin from a secondary endosymbiotic event, such as: having a complete urea cycle; chloroplasts having four bounding membranes; and the ability to produce and oxidize fatty acids to generate metabolic intermediates, this later feature probably allowing diatoms to survive the long darkness in the poles (Allen et al. 2011; Armbrust 2009; Armbrust et al. 2004; Cox 2011). There are two significant events in the evolution of diatoms shaping their genomes. Horizontal gene transfer from Chlamydiae (obligate intracellular bacteria) occurred during the first endosymbiotic event. And the discovery of large numbers of green algal genes in diatoms and other stramenopiles has shed light into a putative green algal endosymbiont having been acquired parallel to the red algal endosymbiont in the second endosymbiotic event (Becker et al. 2008; Dorrell et al. 2017; Dorrell and Smith 2011).

### 1.1.3 Diversity

Diatoms are the most species diverse group of algae; algae understood as a polyphyletic group of taxa with different evolutionary histories (e.g. prokaryotes and eukaryotes) and organizational levels (e.g. unicellular, filamentous, colonial or simple tissues), but being predominantly aquatic and photoautotrophic.

The number of described diatom species goes from 10,000 to 12,000 species (Guiry 2012; Mann and Droop 1996), with more than 25,000 verified species names at the California Academy of Sciences, giving an idea of the large synonymy in diatoms (Alverson 2008).

Diversity estimates of extant diatoms are discordant. Mann and Droop (1996) put their estimate at 200,000 species, 20 times larger than the described species; these authors based their estimate on *Sellaphora pupula* and other species complexes with sympatric morphodemes isolated reproductively, arguing that taxonomy was too coarse at the time, hiding a significant share of the diversity; they also argue that large regions of the planet are poorly explored e.g. the tropics and the marine benthos.

Guiry (2012) puts diatom diversity at 20,000 species, with only 8,000 species to be discovered from the current 12,000 known. In his conservative estimation, the author goes on to claim that larger estimates e.g Mann and Droop (1996) might be the result of "microspecies" (i.e. apomictically reproducing linages), therefore not worth of been considered in diversity estimations.

In the most recent species richness estimation, Mann and Vanormelingen (2013) follow the same approach as in Mann and Droop (1996) in order to find a multiplier to estimate extant diversity. The 11 examples of well defined, actively researched and consistently treated taxa for which they provide multipliers, range from marine to freshwater taxa. Those taxa have been studied either morphologically, molecularly or by mating experiments, or through a combination of the three approaches, including the well-studied genera *Pseudo-nitzschia*, *Skeletonema* and *Sellaphora*. Mann and Vanormelingen (2013) state that their multipliers, going from x2.1 for *Aneumastus* to x14 for *Brachysira* should be regarded as minimal. As in Mann and Droop (1996), the authors claim that large regions of the planet where diatoms are abundant still remain unexplored. Mann and Vanormelingen (2013) assert that epipsammic diatoms (inhabiting sand grains) of both freshwater and marine environments have been poorly studied as well as most freshwaters of the tropics; if the study of the previously mentioned environments already represents a big challenge, the authors claim that challenge would be minor if compared to the ultimate challenge posed by the study of littoral and sublittoral diatoms not only from the tropics, but also from temperate and polar regions. Mann and Vanormelingen (2013) conclude that there are no less than 30,000 species but probably as many as 100,000 species (Mann and Vanormelingen 2013). They suggest that the 200,000 species estimation of Mann and Droop (1996) remains possible due to the large share of species diversity that the unexplored areas of the planet might reveal as well as the cryptic diversity that is being unmasked by molecular studies.

### 1.1.4 Distribution

Diatoms are found in all waters of the planet, from marine to fresh waters, from the tropics to the poles. They also inhabit soils and aerosols (Johansen 2010; Kawecka and Olech 1993; Round et al. 1990). According to their life form, diatoms can be divided into planktonic (suspended in the water

column), benthic (attached to surfaces) and subaerial. The benthic communities can be further divided into epilithic (growing on stones), epiphytic (growing on other photoautotrophic organisms such as other algae and plants), epipsammic (growing on sand grains) and epizoic (growing on animals) (Lowe 2011; Round et al. 1990). Diatoms possessing a raphe (raphids) on either one or both valves of the frustule dominate in benthic habitats because the raphe is a structure that allows motility and colonization on benthic habitats (Sims et al. 2006). On the other hand, centrics (round) dominate in the water column because of their form allowing them to remain suspended longer in the water, in contrast to pennates (bilateral symmetry) which would rapidly sink, unless they link together through spines or mucilage binding.

According to the "ubiquity hypothesis" (Finlay 2002), microbial free-living eukaryotes smaller than 2 mm are distributed worldwide due to their small size, large population sizes and because their dispersal is not restricted by geographical barriers. This author further claims that ecosystem functions mediated by microbes will never be compromised (e.g. nutrient cycling) due to large pool "seedbanks" of microbes. If size is one of the drivers of distribution, then diatom species should be considered cosmopolitan, because the size of most diatoms ranges from 10-100 µm, though there are species smaller than 2 µm and as large as 4 mm (Harwood 2010; Round et al. 1990). Following the "ubiquity hypothesis", Finlay et al. (2002) claim that assertions of diatom species having restricted distributions (endemics) to a particular environment or region might be the result of undersampling, taxonomic synonymy, poor identifications, the uniqueness of the habitat and because of the difficulty of detecting rare species.

The ubiquity of diatoms proposed by Finlay et al. (2002) seems to be reinforced by the apparent wide distribution of the European freshwater species contained in the seminal monographs of Krammer and Lange-Bertalot (1991a; 1991b; 1997a; 1997b); this is because these monographs are globally used, even in tropical regions, leading to misidentifications due to "force-fitting". Another issue in using those monographs is that wide species boundaries were set, leading to lumping of several species into species complexes, e.g. *Achnanthidium minutissimum*, *Gomphonema parvulum* and *Sellaphora pupula* just to name a few. Tropical America is no exception when it comes to the use of the before mentioned identification monographs, with a seemingly cosmopolitanism of its regional floras, though detailed studies have found a larger diversity than previous reports with also a large number of unknown taxa previously lumped into species complexes (Mora et al. 2017; Morales et al. 2014). Nowadays it seems unlikely to find large amounts of shared species between the tropics with north temperate regions due to the evidence that diatoms have biogeography or even restricted distributions to a particular region (Abarca et al. 2014; Vanormelingen et al. 2008) therefore putting into question the ubiquity hypothesis.

### 1.1.5 Morphology

Diatoms are typically brown in color (Fig. 1), because their chloroplasts are covered by carotenoid pigments (β-carotene, diatoxanthin, diadinoxanthin and fucoxanthin) (Goodwin 1974), masking the green color of their chlorophylls, *a*, $c_2$, and $c_1$ or $c_3$ (Stauber and Jeffrey 1988). The number and shape of chloroplasts varies across species. Most of the non-raphid diatoms have many chloroplasts of small size. On the other hand, raphid diatoms have fewer chloroplasts, typically one but larger in comparison to the ones of non-raphid diatoms (Cox 2011).

The hallmark of diatoms are their symmetrically perforated and ornamented cell walls, known as the frustule (Fig. 2) though there are a few exceptions to this e.g. *Phaeodactylum tricornutum* and endosymbiotic species. Although *P. tricornutum* does not normally produce silica cell walls, it has been observed to produce a siliceous wall on only one side of the cell when the cell form is oval, with the other side formed of organic material (Round et al. 1990). Regarding endosymbiotic species, it has been found that endosymbiotic diatoms of Foraminifera have the ability to create silica walls after expulsion or isolation (Lee et al. 1982; Reimer and Lee 1988).

The diatom frustule is mainly constituted of hydrated silica ($SiO_2 \cdot nH_2O$) but also organic components. A frustule consists of two valves which are joined by bands known as cingular bands or copulae (constituting the cingulum). The older valve (epivalve) and its cingulum (epicingulum) are called the epitheca; whereas the younger valve (hypovalve) and its cingulum (hypocingulum) are called the hypotheca (Fig. 2), with the hypotheca always underlapping the epitheca, like a petri dish (Cox 2011; Round et al. 1990).

Among the characteristic perforations and structure of the diatom valves are the pores (e.g. areolae, stigmata, pore fields, rotae), processes (fultoportulae or “strutted process” and rimoportulae or “labiate process”), raphe (longitudinal slit through the valve), ridges and spines (Cox 2011; Round et al. 1990).

Even though the copulae of the cingulum are less complex than the valves, in some taxa they can be robust and intricate. The cingulum is made of two to several copulae (Fig. 2), which are normally perforated but there are taxa with no perforations (Cox 2011).

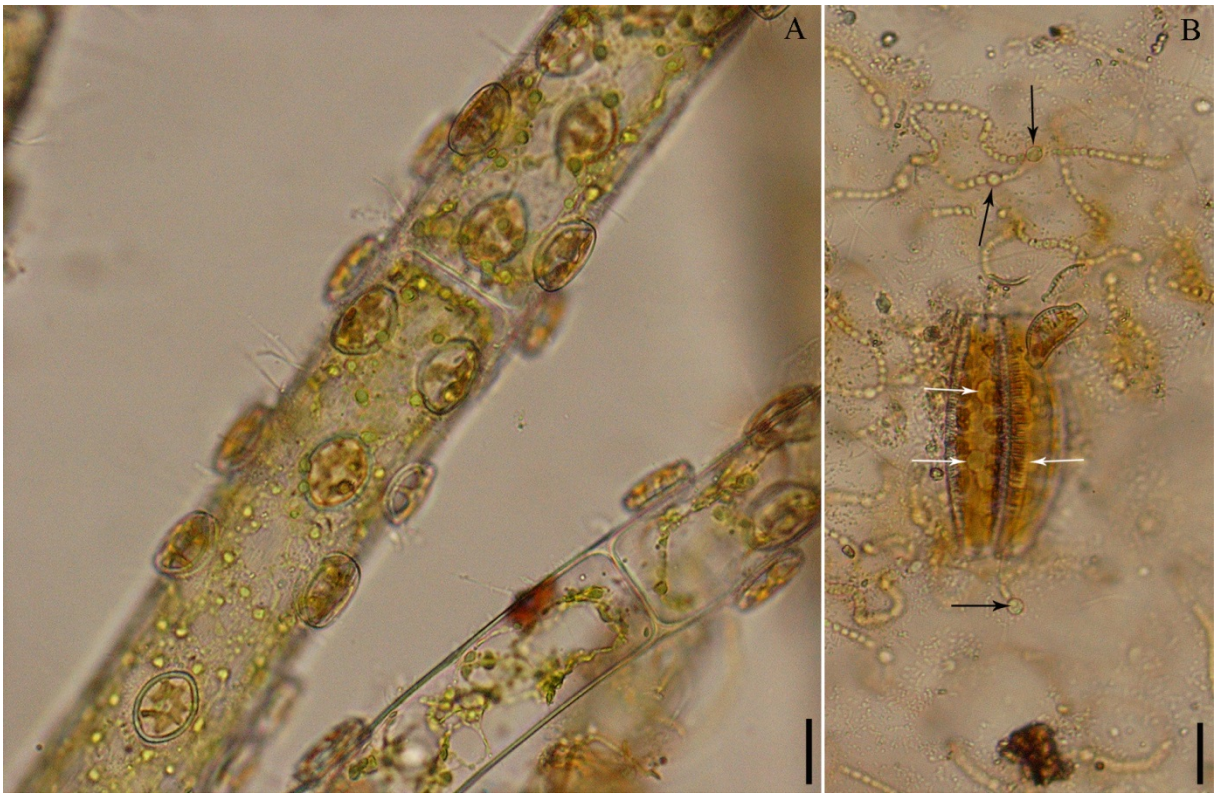



**Fig. 1.** Live cells of diatoms in tropical streams from Central Mexico showing their characteristic brown color. **A** – *Cocconeis pediculus* attached to filaments of the green alga *Cladophora* sp., collected from Calvillo, Guanajuato. **B** – *Epithemia sorex* and *E. turgida* associated to a colonial growth of *Nostoc* sp. collected from La Mesa, Guanajuato. The cells of *E. turgida* contain endosymbiotic cyanobacteria (white arrows). The presence of colonial growths of *Nostoc* sp. is an indicator of poor-nutrient freshwaters, since this cyanobacterium can actively fix atmospheric nitrogen in anaerobic conditions within heterocysts (black arrows) (Sand-Jensen 2014). Endosymbiotic cyanobacterial cells within *E. turgida* allow this diatom to become abundant in waters with low nutrient availability (Stancheva et al. 2013). Scale bars 20 µm.

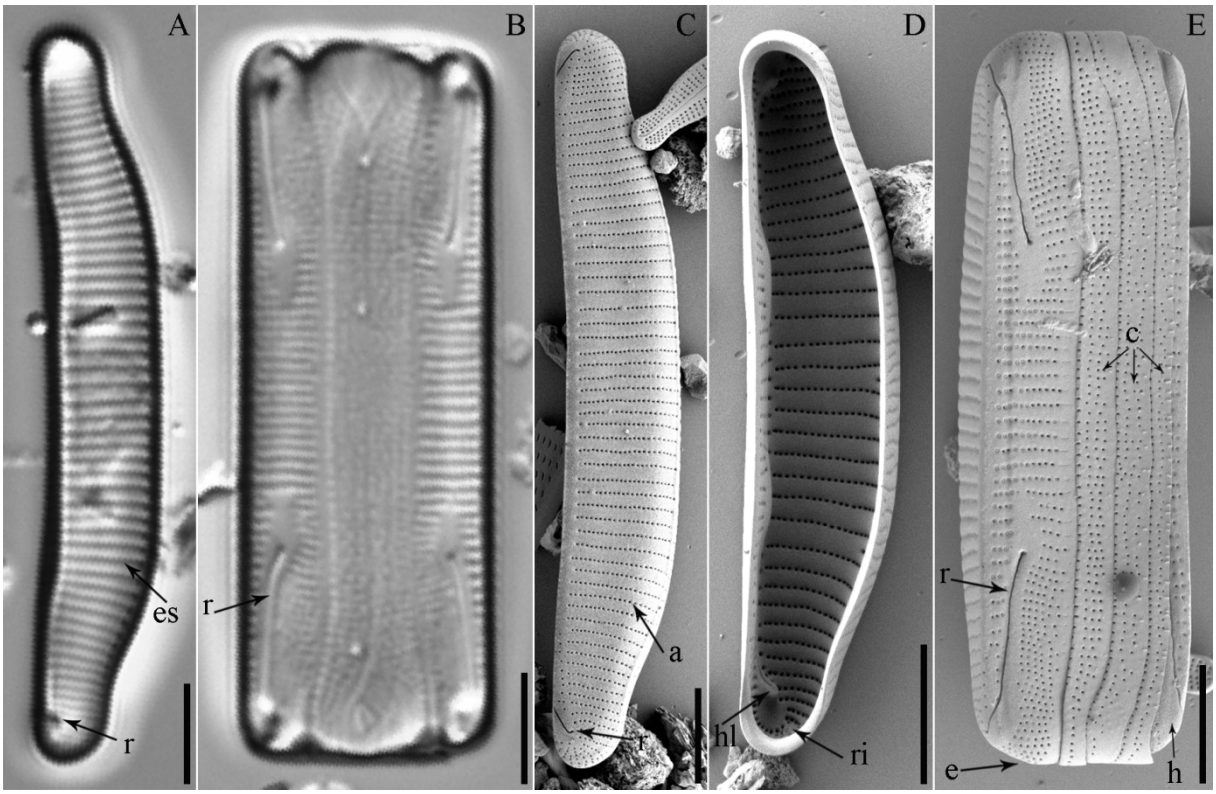



**Fig. 2.** Morphology of the diatom frustule exemplified by *Eunotia minor* collected from Laguna de Servín 1, Querétaro, Mexico. **A, B** – light microscope pictures of a valve in external view (A) and of a frustule in girdle view (B) showing the raphe (r) and the striae. **C-E** – scanning electron microscope pictures of valves in external (C) and interval view (D), as well as of a frustule in girdle view (E), showing the epivalve (e), hypovalve (h), three copulae (c) of the epicingulum, raphe (r), helictoglossa (hl), rimoportula (ri) and the areolae (a) that constitute the striae. Scale bars 5 µm.

### 1.1.6 Reproduction

Diatoms mainly reproduce asexually although in rare occasions sexual reproduction occurs (Round et al. 1990). During cell division, each of the progeny cells inherits one of the valves of the mother cell and synthesizes a new one, therefore the wall formation is semiconservative unlike most algae (Cox 2011). Even though diatom cells are normally solitary, the close proximity of the valves during cell division allows the formation of chains and filaments through the production of spines (Fig. 3) and projections, but also by mucilage binding.

Since the new valves are formed within the mother cells, diatoms undergo size reduction after the formation of the new hypovalves. But there are diatoms that do not reduce their size because of flexibility in their cingulum and valves (Cox 2011; Round et al. 1990).

With respect to gamete fusion, sexual reproduction can be oogamous (female gametes are non-motile and larger than males gametes, which are motile) isogamous (gametes of similar shape and

size) and anysogamous (similar morphology but one motile and the other sesile) (Mann 1993; Round et al. 1990). Sexual reproduction results in the production of an auxospore, linked to the restoration of size because a cell of maximum size is formed. But there are species that are able to produce auxospores without sexual reproduction, which might be a mechanism for restoring cell size if no sexual mechanisms are triggered (Julius and Theriot 2010).

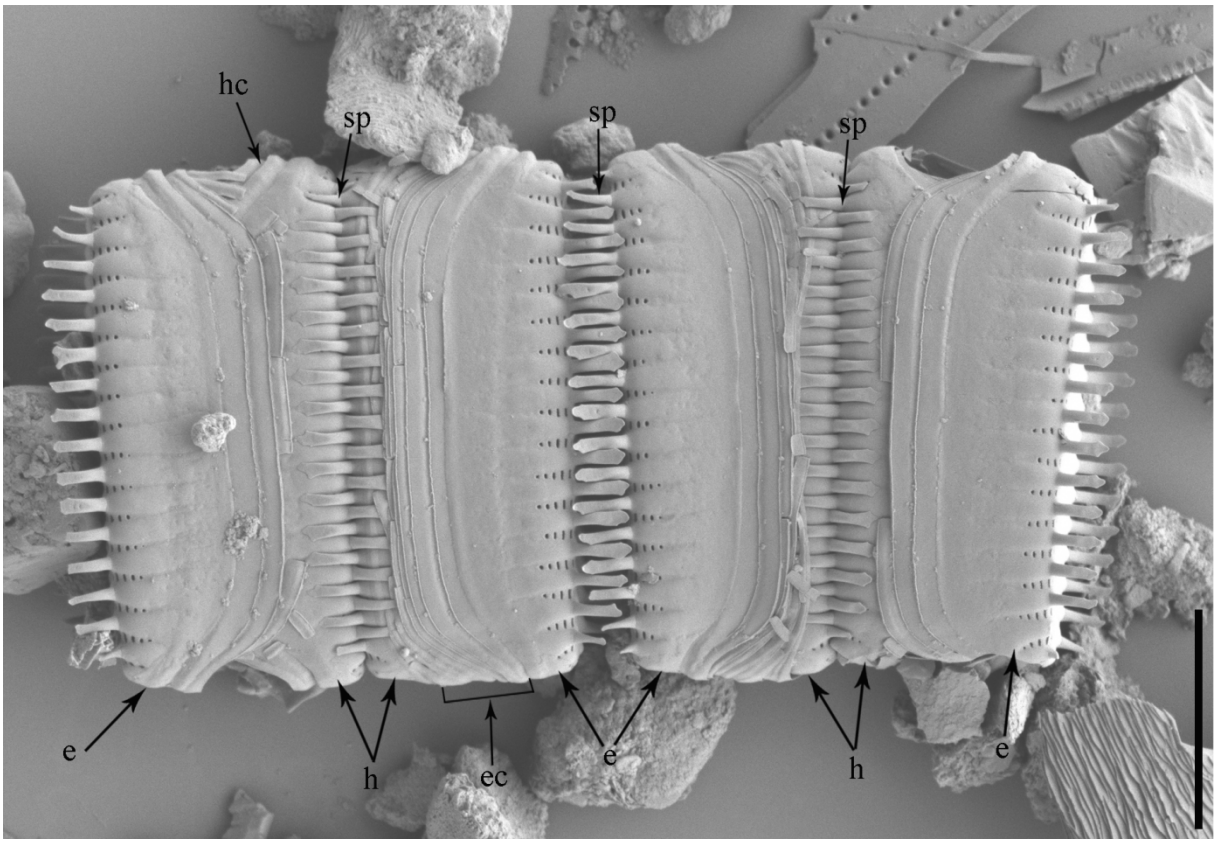



**Fig. 3.** Cell division of chain-forming *Staurosira venter* from Laguna de Servín 1, Querétaro, Mexico. Scanning electron microscope picture of two mother cells, seen in girdle view, showing the start of separation of the daughter cells, resulting in four cells, joined by interlinking spines; epivalve (e), hypovalve (h), epicingulum (ec), hypocingulum (hc) and spines (sp). Scale bar 5 µm.

### 1.1.7 Classification

The large morphological variation in shape, structure and symmetry of the diatom valves have been the basis of diatom identification and classification for a long time, implying a phenetic or morphological species concept (Alverson 2008; Cox 2011).

The first diatom classifications date back to the nineteenth century, by Agardh (1824) and Kützing (1844), based on the shape of valves and on their growth form. A century later, Hustedt (1930) based

his classification on symmetry, proposing the order Pennales (bilateral) and Centrales (radial) within the phylum Bacillariophyta. In the pursuit of a more natural classification, Simonsen (1979) retained the orders Centrales and Pennales, belonging to the class Bacillariophyceae, basing his classification on both symmetry and type of reproduction. A decade later Round et al. (1990) proposed one of the most comprehensive and widely used classifications to date, basing it not only on symmetry and morphology of the valves, but also on observations of the protoplast, nuclei and sexual reproduction. Diatoms were reinstated within the phylum Bacillariophyta, recognizing three classes: Coscinodiscophyceae (centrics), Fragilariophyceae (pennate araphids) and Bacillariophyceae (pennate raphids) (Fig. 4).

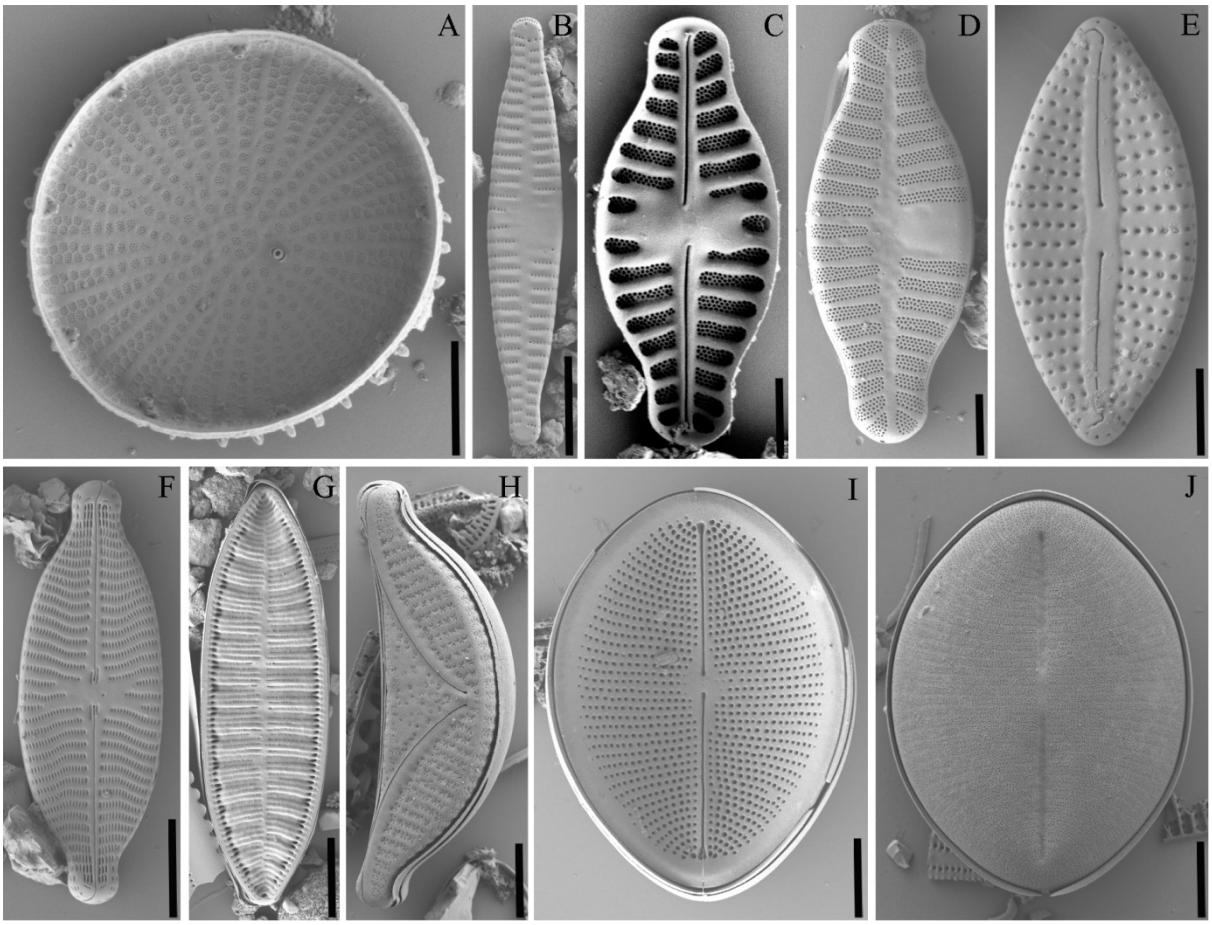



**Fig. 4.** Morphological diversity of diatoms in tropical streams from Central Mexico. **A** – *Cyclostephanos invisitatus*. **B** – *Fragilaria* sp. **C, D** – *Planothidium rostratum*, internal view of the raphe valve (C) and external view of the sternum valve (D). **E** – *Craticula subminuscula*. **F** – *Geissleria decussis*. **G** – *Surirella angusta*. **H** – *Epithemia sorex*. **I, J** – *Cocconeis pediculus*, external views of the raphe valve (I) and of the sternum valve (J). Scale bars 2 µm (A, C-E) and 5 µm (B, F-J). In accordance with the classification of Round et al. (1990), the following diatoms correspond to their three classes: Coscinodiscophyceae (A), Fragilariophyceae (B) and Bacillariophyceae (C-J). Raphid diatoms (Bacillariophyceae) dominate the composition in streams because the raphe allows mobility and therefore successful colonization of benthic habitats e.g. stony river beds.

With the use of molecular methods, previous classifications have been challenged. The centric diatoms of Round et al. (1990) have been found to be paraphyletic, whereas the pennates, both raphid and araphid are monophyletic. Based on cytological, morphological and molecular data, Medlin and Kaczmarska (2004) proposed amendments to the classes Bacillariophyceae (raphid diatoms) and Coscinodiscophyceae (radial centrics), and proposed a new class, the Mediophyceae (bipolar centrics); but this classification has been heavily challenged since its inception (Alverson and Theriot 2005; Theriot et al. 2010; Theriot et al. 2009; Williams and Kociolek 2007). In the last classification presented to date, Cox (2015) integrates previous classifications (Medlin and Kaczmarska 2004; Round et al. 1990) and recent outcomes on some of the classes (Ashworth et al. 2013; Theriot et al. 2011), resulting in four classes: Coscinodiscophyceae, Mediophyceae, Fragilariophyceae and Bacillariophyceae.

Regarding their position within the tree of life, the most recent classification based on a phylogenomic approach places the diatoms within the SAR supergroup, including Stramenopiles, Alveolates and Rhizaria. Diatoms are placed within the Stramenopiles (heterokonts) along with brown (Phaeophyta) and golden algae (Chrysophyta), but also heterotrophic protists (Burki 2014).

## 1.2 Diatoms as biological indicators

Despite their importance for mankind, freshwater ecosystems are being degraded at more accelerated rates than their terrestrial counterparts (Dudgeon et al. 2006). In order to assess and monitor environmental change, the use of biological indicators as a complimentary tool to physical and chemical measurements has been established. Biological indicators are species or assemblages whose presence or absence, and changes in abundance and morphology reflect the characteristics of a habitat (Stevenson et al. 2010). Biological indicators have the advantage over physical and chemical indicators that they record the historical ecological conditions, e.g. in lakes and streams from weeks to millennia (Mccormick and Cairns 1994; Stevenson et al. 2010; Williamson et al. 2008).

According to Bellinger and Sigee (2010), the ideal characteristics of a biological indicator are: narrow tolerance to the environmental parameter of interest, wide distribution, rapid response to environmental change, well resolved taxonomy and ease of identification by non-specialists. Among several groups of organisms used as indicators (e.g. algae including diatoms, fish, macroinvertebrate and macrophytes), diatoms are widely used because they better assess changes in water quality in a promptly manner (Hering et al. 2006; Mccormick and Cairns 1994).

The following attributes make diatoms attractive for their use as biological indicators in freshwaters: high species diversity, widespread distribution, short life cycles, relative ease of sampling, processing, quantification and storage, as well as their position at the base of aquatic food webs (Dixit et al. 1992; Hering et al. 2006; Kelly et al. 2009; Licursi and Gomez 2009; Lowe 2011). But their use has also disadvantages, the main being their complex taxonomy, requiring in-depth knowledge by qualified personnel for identification. Unlike other biological indicators such as macroinvertebrates where identifications at the genus or family level are sufficient, species level identification is required for diatoms.

## 1.3 Methods for the identification of diatoms in biomonitoring

### 1.3.1 Morphology

The large morphological variation in shape, structure and symmetry of the diatom valves are the foundations of diatom identification (Cox 2011). Because of the tiny size of diatoms, typically 10-100 µm, observations for their identification are conducted by light microscopy (LM) and scanning electron microscopy (SEM). However, identifications in ecological and monitoring studies are mainly conducted by LM (Morales et al. 2001). Nevertheless, the limitations of LM in resolving the fine structure of the diatom valves are evident and well documented (Fig. 2) (Cox 1975; Siver and Kling 1997). The low resolution offered by LM, due to the limited wave lengths of visible light, has led to clustering of taxa in all groups of algae including diatoms, resulting in underestimations of diversity, and overestimation of distribution and tolerance to environmental parameters (Mora et al. 2017; Morales et al. 2001; Rimet and Bouchez 2012; Siver 1995). As exemplified with small fragilarioid diatoms, Morales et al. (2001) perfectly illustrate the limitations of LM in comparison to SEM, which in turn leads to wrong conclusions in bioindication; this is because different taxa are clustered in the same species due to their similarity under LM even though they have different ecological tolerances. In conclusion, reliable identifications are crucial when using diatoms as bioindicators in order to obtain precise evaluations, with SEM observations improving resolution to a great extent (Cox 1998; Morales et al. 2001; Rimet and Bouchez 2012). However, there are cases when even detailed SEM observations are not able to resolve the taxonomical identity of diatoms, e.g. cryptic and semi-cryptic species. In such cases, molecular studies have proven to be a valuable identification tool (Abarca et al. 2014; Kermarrec et al. 2013a; Poulickova et al. 2008; Trobajo et al. 2009).

### 1.3.2 DNA barcoding and eDNA metabarcoding

Even though the advantages of the use of diatoms as biological indicators are well-known, one of the major impediments to their usage is their taxonomic identification, which is crucial in recording meaningful results (e.g. calculation of trophic indices). Identification can be problematic because of the subtle boundaries among species that have been traditionally based on the morphology of the valves. Morphological identification of diatoms can be time consuming, requiring in-depth knowledge of the diatom diversity of the area of analysis and awareness of the morphological plasticity a species displays under varying environmental conditions and during the life cycle.

To overcome the identification impediment, DNA barcoding (Hebert et al. 2003) and environmental DNA metabarcoding (Taberlet et al. 2012b) are being adopted in taxonomic identification of diatoms as an alternative to the traditional morphology-based identification approaches (Kermarrec et al. 2014; Zimmermann et al. 2015; Zimmermann et al. 2011).

A DNA barcode is a tool for the correlation of unknown sequences representing organisms to known sequences in a reference database (Hebert et al. 2003; Ratnasingham and Hebert 2007). A barcode consists of a molecular marker that can be easily sequenced in one Sanger run, unambiguously identifying a taxon independent of its life cycle (Hebert et al. 2003; Mann et al. 2010; Moritz and Cicero 2004).

Environmental DNA (eDNA) metabarcoding refers to the use of universal primers for the amplification of a barcode region from DNA of bulk samples of water, soil or air (Taberlet et al. 2012a; Yu et al. 2012) and sequencing by High-Throughput Sequencing (HTS). After sequencing, a pipeline involving several bioinformatics procedures is followed (Zimmermann et al. 2015), resulting in a list of Molecular Operational Taxonomic Units (MOTUs). The list of MOTUs can then be correlated to morphologically identified taxa from DNA reference libraries (Zimmermann et al. 2014), resulting in a species list that can be used in water quality assessments (Vasselon et al. 2017). Current developments focus on eliminating the need of a DNA reference library (Apotheloz-Perret-Gentil et al. 2017).

The choice of a barcoding marker that has the resolution power desired is of critical importance, diatom studies needing species level resolution. Impediments for DNA based taxonomical approaches could be incomplete lineage sorting and introgression (Hudson et al. 2002; Meyer and Paulay 2005).

Another concern in determining species is the natural intraspecific and intragenomic variability and interspecific divergence of the barcoding marker. This is particularly problematic when a single

traditionally recognized species or bioindicator taxon comprises a variety of different genotypes; sequences corresponding to different genotypes within the same taxon may cluster into different MOTUs, and thus artificially inflate taxonomic richness genotypes (Balint et al. 2016; Brown et al. 2015). The clustering of MOTUs relies on thresholds that depend on the level of overlap between intraspecific and interspecific variation across the organisms in question. Thresholds of 3% are of common use, with sequences having less than 97% similarity considered as belonging to a different species, whereas sequences with more than 97% similarity pooled together into MOTUs. But the use of thresholds is particularly problematic in delineating closely related species in taxonomically understudied groups (Meyer and Paulay 2005). Since threshold based approaches have produced mixed results, tree-based approaches implementing evolutionary models are used in the delimitation of species boundaries, i.e. shifts in tree branching rates (Monaghan et al. 2009), have been successfully implemented in diatoms (Visco et al. 2015; Zimmermann et al. 2015)

#### 1.3.2.1 The barcoding marker

As the quest for the ideal barcoding marker continues, mitochondrial, plastid and nuclear regions have been explored, including *cox*1, *rbc*L, 5.8S+ITS2, ITS, 18S and 18S V4 (Evans et al. 2007; Moniz and Kaczmarska 2009; Zimmermann et al. 2011). With mixed results depending on the groups of diatoms to be evaluated, the search for the ideal marker continues, although the 18S V4 region has been declared pre-barcode for protists (Pawlowski et al. 2012). Important to mention is the impossibility of designing universal diatom primers (Moniz and Kaczmarska 2009; Zimmermann et al. 2011). For monitoring of streams, the preferred barcodes are *rbc*L and 18S V4 (Apotheloz-Perret-Gentil et al. 2017; Kermarrec et al. 2014; Vasselon et al. 2017; Visco et al. 2015; Zimmermann et al. 2015).

#### 1.3.2.2 Sources of reference barcodes in diatoms

Under current methodologies, the translation of the list of MOTUs into species lists requires a barcode reference library whose taxonomy has been validated by morphological identifications (Kermarrec et al. 2014; Vasselon et al. 2017; Zimmermann et al. 2015).

The completeness of the used reference library correlates with the identification success of the MOTUs. Library completeness in diatoms can prove a major challenge because building a library implies cultivation of diatoms. Even though cultivation remains the best source for diatom barcodes, it is time-consuming and never yields 100% success because of taxa that are recalcitrant to culturing conditions (Chen et al. 2013). That is why single cell PCR amplifications have been proposed as an alternative to cultivation for obtaining barcode libraries (Chen et al. 2013; Hamilton et al. 2015; Lang and Kaczmarska 2011). But single cell isolations are disadvantageous because it is almost impossible

to check the morphological identity of the isolated cell because the cell is destroyed during DNA extraction. Another disadvantage is that no DNA is available for further amplification (e.g. if different markers are to be tested) because the amplifications are done directly from the isolated cell without a previous extraction. In order to increase the sources of reference barcodes, HTS data has been proposed as a source of barcodes if the targeted diatom can be unambiguously linked to its morphological identification; this can only be done for a relative small number of barcodes (Rimet et al. 2018).

#### 1.3.2.3 Quantification of abundance from HTS data

Discrepancies in abundances retrieved by morphological and metabarcoding approaches is a topic of debate when sequence numbers are used as abundance proxy in the calculation of indices for biological monitoring studies (Apotheloz-Perret-Gentil et al. 2017; Vasselon et al. 2017; Visco et al. 2015). The barcoding marker and its ability to discriminate between closely related species, as well as primer specificity are important factors to take into account before using sequence reads to calculate abundances (Elbrecht et al. 2017).

Cell size is a determinant factor of disparities observed when comparing abundance data retrieved from microscopy and metabarcoding. It has been proposed that there is a correlation between biovolume and gene copies of the SSU rDNA (Godhe et al. 2008; Zhu et al. 2005), which explains why small-sized species are underrepresented in metabarcoding data compared to microscopy, whereas large species might be overrepresented in read abundances compared to morphology-based abundances. This hypothesis has also been proven by Vasselon et al. (submitted) in their study of both mock and natural communities using the *rbc*L gene; these authors achieved a reduction in abundance disparities from both methods by calculating correction factors for cell biovolume.

Methodological procedures can be another source of disparity in abundance data. During PCR amplification, differential primer efficiency, specificity and template competition are determinants in the final number of sequences (Elbrecht and Leese 2015; Kermarrec et al. 2013b). This is of particular importance because primer efficiency varies among species and environmental samples contain several species.

## 1.4 Central Mexico – a rich but threatened center of biodiversity

### 1.4.1 Biodiversity

Central Mexico, where the Lerma-Chapala Basin is located (Fig. 5), is an environmentally heterogeneous region (Figs. 6-8), rich in biodiversity, where two biodiversity hotspots converge, the Madrean Pine-Oak Woodlands and Mesoamerica (Mittermeier et al. 2011; Myers et al. 2000). From the 17,000 plant species occurring in Mesoamerica, 17% are endemic to the hotspot; in the Madrean Pine-Oak Woodland, 75% of its plant diversity (5,300 species) is endemic. Even though there are no endemism records per hotspot for microorganisms such as diatoms, there are records for other freshwater organisms such as amphibians and fish. In Mesoamerica, the endemism levels are as high as 61% and 67% for amphibians and fish respectively; in the Madrean Pine-Oak Woodlands these figures are 23% and 21% respectively (Mittermeier et al. 2004; Mittermeier et al. 2011).

The rich biological diversity of the Lerma-Chapala Basin in Central Mexico is explained in part by its complex geology. The basin is situated in the Central Plateau of Mexico, surrounded by the largest mountain chains of the country to the west (Sierra Madre Occidental) and to the east (Sierra Madre Oriental). To the south of the Lerma-Chapala Basin is the Trans-Mexican Volcanic Belt, which is the tallest mountain chain in the country, running from west to east across Central Mexico, with elevations up to 5,700 m asl (Cotler et al. 2006; Mittermeier et al. 2004; Sedeño-Díaz and López-López 2007). The rich biodiversity of Central Mexico is also defined by its position within the Mexican Transition Zone, an area where biotic elements from the Neotropical and Neartic biogeographic regions converge (Huidobro et al. 2006; Olson et al. 2001).

### 1.4.2 Threats

From the original extent of both hotspots, Mesoamerica and the Madrean Pine-Oak Woodlands, only 20% remains. Despite their outstanding biodiversity, the protected areas in categories I-IV of the International Union for Conservation of Nature (IUCN) account for only 5.7% of the extent of Mesoamerica and a low 1.9% of the Madrean Pine-Oak Woodlands. The major threats to both hotspots are deforestation, clearance for agriculture, overgrazing and soil erosion, Mesoamerica having one of the highest deforestation rates in the world (Mittermeier et al. 2004).

Concerning the threats to freshwater ecosystems in general, Dudgeon et al. (2006) name five major threats: flow-modification, overexploitation, water pollution, habitat-degradation and invasion of alien species. All these five threats are taking place at varying degrees of magnitude in the two hotspots.

When specifically looking at the major threats to freshwater ecosystems in the Lerma-Chapala Basin, the degree at which perturbations are taking place is distressing. The flow of the Lerma River and its tributaries have been regulated for decades through reservoirs and water-diversion dams for agricultural irrigation, with sections of the river and tributaries left without any flow during the dry season (Mercado-Silva et al. 2009; Mercado-Silva et al. 2006). Due to flow modification and overexploitation, Lake Chapala is drying up; this lake is the largest in Mexico and one of the main water sources for Guadalajara, the second largest city of the country (Bertrab 2003). Overexploitation of groundwater is particularly critical in the semi-arid north of the basin, with 400% more water extracted than the natural recharge (Mahlknecht et al. 2004). Pollution is also taking its toll in the basin, with large sections of the Lerma River and its tributaries, as well as Chapala Lake heavily contaminated by wastewater discharges from industry, agriculture, farming along with domestic sewage; heavy metals and pesticides are among the most dangerous pollutants (Rosales-Hoz et al. 2000; Sedeño-Díaz and López-López 2007; Tejeda et al. 2010; Zarazua et al. 2006). The arrival of alien species such as several species of fish (Mercado-Silva et al. 2009; Mercado-Silva et al. 2006) and the invasive water hyacinth (*Eichhornia crassipes*) is changing the food web structure of both lotic and lentic environments and having negative impacts in migratory species. To add to those already negative effects, population control of the water hyacinth in Chapala Lake is precarious and dangerous. The plants are directly sprayed with pesticides to control their invasiveness, to no real decrease of their populations but only to the detriment to local organisms and human health since both fishermen and fish catches are contaminated by the sprayed pesticides (Stong et al. 2013; Villamagna et al. 2010). On the whole, the aquatic ecosystems of the basin face colossal threats that require equally enormous assessment, monitoring, management and conservation tasks to safeguard the rich biodiversity of the Lerma-Chapala Basin and the ecosystem services this biodiversity provides.

### 1.4.3 Diatoms in the assessment and monitoring of treats to freshwaters in Central Mexico

Even though diatoms have proved to be a powerful bioindication tool, their study in Mexico is rare. The study of stream diatoms from Mexico has mainly concentrated in the center of the country, with a focus on species composition, which is in fact the baseline for bioindication. Those studies have found a seemingly low diversity due to the application of broad species boundaries and because of force-fitting into already described species from temperate regions (Mora et al. 2017). From those studies, only four were aimed at bioindication, which is a fairly new research area in the country.

Two studies were conducted within the Lerma-Chapala Basin, both reassuring the polluted nature of the Lerma-River (Abarca-Mejía 2010; Segura-Garcia et al. 2012) as it has been previously revealed by physical and chemical evaluations (Sedeño-Díaz and López-López 2007). The other two studies were conducted within the Basin of Mexico, identifying a gradient of perturbation as the studied river approaches Mexico City (Carmona Jiménez et al. 2016; Jujnovsky et al. 2010).

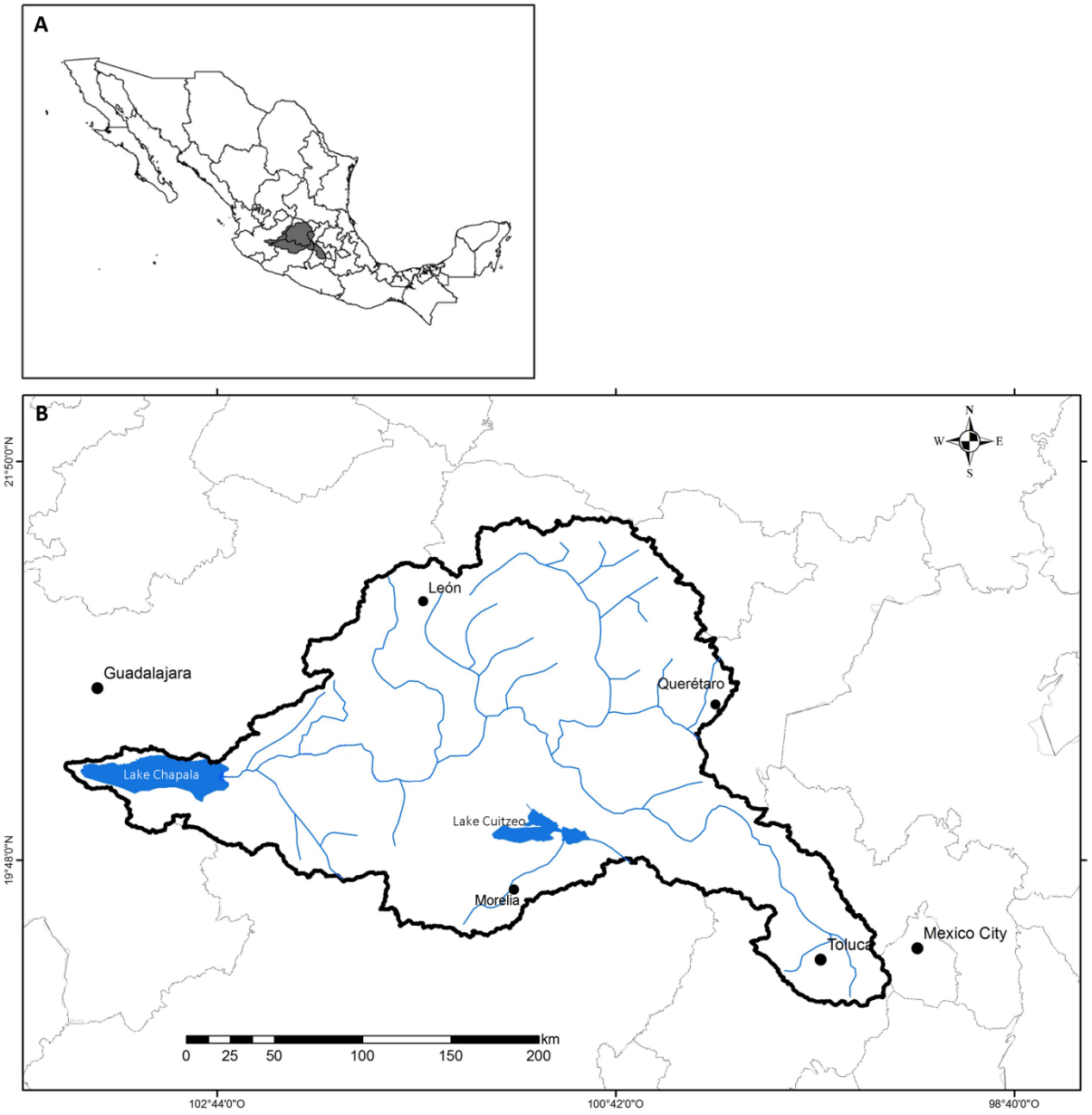



Figure 5. Location of the Lerma-Chapala Basin in Central Mexico (A). Detail of the Lerma-Chapala Basin with its major river (Lerma) and tributaries, as well as lakes, including Chapala, the largest lake of Mexico, an important source of water for Guadalajara, the second most populated metropolitan area in the country.


**Fig. 6.** Seasonal variation in tropical streams from Central Mexico. **A** – rainy season (October 2013) and **B** – dry season (February 2014) in San Martín, Guanajuato.


**Fig. 7.** Seasonal variation in tropical streams from Central Mexico. **A** – Calvillo, Guanajuato, during the rainy season (October 2013). **B** – Ojo de Agua de Calvillo, Guanajuato, during the rainy season (October, 2013).


**Fig. 8.** Seasonal variation in tropical streams from Central Mexico. **A** – La Laborcilla 2, Guanajuato, during the rainy season (September 2013). **B** – Los Ailes 2, Querétaro, during the dry season (February 2014).

## 1.5 Objectives

The main goal of this thesis was to conduct an integrative approach to the diversity analysis of epilithic diatoms in tropical streams from the Lerma-Chapala Basin, Central Mexico. The particular objectives per approach are stated below:

Morphology

- Assessment of the diversity by morphological analyses through light and scanning electron microscopy.
- Quantification of taxa abundances.
- Description and illustration of the morphological diversity.

Ecology

- Identification of the environmental factors that determine the species composition observed and quantified by microscopy.

Cultivation

- Establishment of a regional morphological and molecular taxonomic reference library, unambiguously linking morphological and molecular data.

Metabarcoding

- Assessment of the diversity by environmental DNA metabarcoding and comparison to the diversity retrieved from morphological analysis.
- Determination of taxa abundances and critical comparison to the abundances quantified by microscopy.
- Evaluation of the potential of High-Throughput Sequencing data as a source of reference DNA barcodes.

Phylogeny

- Elucidation of the taxonomy and systematics of the *Planothidium lanceolatum*/*P. frequentissimum* species complexes using morphological and molecular data.
- Determination if the sinus and cavum are stable taxonomic characters.
- Investigation of biogeographic distribution patterns.

## 1.6 References


Abarca-Mejía N (2010) Diatom community analysis and quality assessment of the polluted tropical Lerma River (Mexico). Dissertation, Freie Universität Berlin, Germany.

Abarca N, Jahn R, Zimmermann J, Enke N (2014) Does the cosmopolitan diatom *Gomphonema parvulum* (Kutzing) Kutzing have a biogeography? PLoS One 9(1): e86885. https://doi.org/10.1371/journal.pone.0086885.

Agardh C (1824) *Systema algarum*. Litteris Berlingianus, Lundae, 312 pp.

Allen AE, Dupont CL, Obornik M, Horak A, Nunes-Nesi A, McCrow JP, Zheng H, Johnson DA, Hu H, Fernie AR, Bowler C (2011) Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. Nature 473(7346): 203-207. https://doi.org/10.1038/nature10074.

Alverson AJ (2008) Molecular systematics and the diatom species. Protist 159(3): 339-353. https://doi.org/10.1016/j.protis.2008.04.001.

Alverson AJ, Theriot EC (2005) Comments on recent progress toward reconstructing the diatom phylogeny. Journal of Nanoscience and Nanotechnology 5(1): 57-62. https://doi.org/10.1166/jnn.2005.007.

Anonymous (1703) Two letters from a gentleman in the country, relating to Mr. Leeuwenhoeck's letter in Transaction, No. 283. Philosophical Transactions of the Royal Society of London 23(1494-1498.

Apotheloz-Perret-Gentil L, Cordonier A, Straub F, Iseli J, Esling P, Pawlowski J (2017) Taxonomy-free molecular diatom index for high-throughput eDNA biomonitoring. Mol Ecol Resour. https://doi.org/10.1111/1755-0998.12668.

Armbrust EV (2009) The life of diatoms in the world's oceans. Nature 459(7244): 185-192. https://doi.org/10.1038/nature08057.

Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kröger N, Lau WWY, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science 306(5693): 79-86. https://doi.org/10.1126/science.1101156.

Ashworth MP, Nakov T, Theriot EC (2013) Revisiting Ross and Sims (1971): toward a molecular phylogeny of the Biddulphiaceae and Eupodiscaceae (Bacillariophyceae). Journal of Phycology 49(6): 1207-1222. https://doi.org/10.1111/jpy.12131.

Balint M, Bahram M, Eren AM, Faust K, Fuhrman JA, Lindahl B, O'Hara RB, Opik M, Sogin ML, Unterseher M, Tedersoo L (2016) Millions of reads, thousands of taxa: microbial community structure and associations analyzed via marker genes. FEMS Microbiol Rev 40(5): 686-700. https://doi.org/10.1093/femsre/fuw017.

Becker B, Hoef-Emden K, Melkonian M (2008) Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes. BMC Evolutionary Biology 8(1): 203. https://doi.org/10.1186/1471-2148-8-203.

Bellinger EG, Sigee DC (2010) Freshwater algae: identification and use as bioindicators. John Wiley & Sons, Ltd, 271 pp.

Bertrab Ev (2003) Guadalajara's water crisis and the fate of Lake Chapala: a reflection of poor water management in Mexico. Environment and Urbanization 15(2): 127-140. https://doi.org/10.1177/095624780301500204.

Bory de Saint-Vincent J (1822) Dictionnaire Classique d'Histoire Naturelle, 1. Paris.

Brown EA, Chain FJ, Crease TJ, MacIsaac HJ, Cristescu ME (2015) Divergence thresholds and divergent biodiversity estimates: can metabarcoding reliably describe zooplankton communities? Ecol Evol 5(11): 2234-2251. https://doi.org/10.1002/ece3.1485.

Burki F (2014) The eukaryotic tree of life from a global phylogenomic perspective. Cold Spring Harbor perspectives in biology 6(5): a016147. https://doi.org/10.1101/cshperspect.a016147.

Carmona Jiménez J, Ramírez Rodríguez R, Bojorge-García MG, González Hidalgo B, Cantoral-Uriza EA (2016) Estudio del valor indicador de las comunidades de algas bentónicas: una propuesta de evaluación y aplicación en el río Magdalena, Ciudad de México. Revista internacional de contaminación ambiental 32(139-152.

Chen XG, Zhang J, Huang Y, Hou YP (2013) Diatom taxa identification based on single-cell isolation and rDNA sequencing. Forensic Science International: Genetics Supplement Series 4(1): e308-e309. https://doi.org/10.1016/j.fsigss.2013.10.157.

Cotler H, Mazari M, De Anda J (2006) Atlas de la cuenca Lerma-Chapala, construyendo una visión conjunta. INE-UNAM, 191 pp.

Cox EJ (1975) A reappraisal of the diatom genus Amphipleura Kütz. Using light and electron microscopy. British Phycological Journal 10(1): 1-12. https://doi.org/10.1080/00071617500650011.

Cox EJ (1998) The identity and typification of some naviculoid diatoms (Bacillariophyta) from freshwater or brackish habitats. Phycologia 37(3): 162-175. https://doi.org/10.2216/i0031-8884-37-3-162.1.

Cox EJ (2011) Morphology, cell wall, cytology, ultrastructure and morphogenetic studies. In: Seckbach J, Kociolek P (Eds) The diatom world. Springer Netherlands, Dordrecht, 21-45.

Cox EJ (2015) Coscinodiscophyceae, Mediophyceae, Fragilariophyceae, Bacillariophyceae (Diatoms). In: Frey W (Ed) Syllabus of Plant Families, A Engler's Syllabus der Pflanzenfamilien, Part 2/1. 13 ed. Borntraeger Science Publishers, Stuttgart, 61-103 pp.

Darwin CR (1866) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London.

Dixit SS, Smol JP, Kingston JC, Charles DF (1992) Diatoms - Powerful Indicators of Environmental-Change. Environmental science & technology 26(1): 22-33. https://doi.org/10.1021/es00025a002.

Dorrell RG, Gile G, McCallum G, Méheust R, Bapteste EP, Klinger CM, Brillet-Guéguen L, Freeman KD, Richter DJ, Bowler C (2017) Chimeric origins of ochrophytes and

haptophytes revealed through an ancient plastid proteome. eLife 6(e23717. https://doi.org/10.7554/eLife.23717.

Dorrell RG, Smith AG (2011) Do red and green make brown?: perspectives on plastid acquisitions within chromalveolates. Eukaryotic Cell 10(7): 856-868. https://doi.org/10.1128/EC.00326-10.

Dudgeon D, Arthington AH, Gessner MO, Kawabata Z, Knowler DJ, Leveque C, Naiman RJ, Prieur-Richard AH, Soto D, Stiassny ML, Sullivan CA (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. Biological reviews 81(2): 163-182. https://doi.org/10.1017/S1464793105006950.

Ehrenberg CG (1838) Die Infusionsthierchen als vollkommene Organismen. Leopold Voss, Leipzig, i-xviii +548 pp.

Elbrecht V, Leese F (2015) Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass--Sequence Relationships with an Innovative Metabarcoding Protocol. PLoS One 10(7): e0130324. https://doi.org/10.1371/journal.pone.0130324.

Elbrecht V, Vamos EE, Meissner K, Aroviita J, Leese F, Yu D (2017) Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. Methods in Ecology and Evolution 8(10): 1265-1275. https://doi.org/10.1111/2041-210x.12789.

Evans KM, Wortley AH, Mann DG (2007) An assessment of potential diatom “barcode” genes (*cox*1, *rbc*L, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). Protist 158(3): 349-364. https://doi.org/10.1016/j.protis.2007.04.001.

Falkowski PG, Barber RT, Smetacek V (1998) Biogeochemical controls and feedbacks on ocean primary production. Science 281(5374): 200-206. https://doi.org/10.1126/science.281.5374.200.

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281(5374): 237-240. https://doi.org/10.1126/science.281.5374.237

Finlay BJ (2002) Global dispersal of free-living microbial eukaryote species. Science 296(5570): 1061-1063. https://doi.org/10.1126/science.1070710

Finlay BJ, Monaghan EB, Maberly SC (2002) Hypothesis: the rate and scale of dispersal of freshwater diatom species is a function of their global abundance. Protist 153(3): 261-273. https://doi.org/10.1078/1434-4610-00103.

Gmelin J (1791) In *Systema Naturae* Edition 13 1(6). Lipsiae.

Godhe A, Asplund ME, Harnstrom K, Saravanan V, Tyagi A, Karunasagar I (2008) Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. Appl Environ Microbiol 74(23): 7174-7182. https://doi.org/10.1128/AEM.01298-08.

Goodwin T (1974) Carotenoids and biliproteins. In: Stewart WDP (Ed) Algal physiology and biochemistry, Botanical Monographs 10. Oxford, 176-205.

Guiry MD (2012) How many species of algae are there? Journal of Phycology 48(5): 1057-1063. https://doi.org/10.1111/j.1529-8817.2012.01222.x.

Hamilton PB, Lefebvre KE, Bull RD (2015) Single cell PCR amplification of diatoms using fresh and preserved samples. Front Microbiol 6(1084. https://doi.org/10.3389/fmicb.2015.01084.

Harwood DM (2010) Diatomite. In: Smol JP, Stoermer EF (Eds) The diatoms: applications for the environmental and earth sciences. Cambridge University Press, Cambridge, 570-574.

Hebert PD, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc Biol Sci 270(1512): 313-321. https://doi.org/10.1098/rspb.2002.2218.

Hering D, Johnson RK, Kramm S, Schmutz S, Szoszkiewicz K, Verdonschot PFM (2006) Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. Freshwater Biology 51(9): 1757-1785. https://doi.org/10.1111/j.1365-2427.2006.01610.x.

Hudson RR, Coyne JA, Huelsenbeck J (2002) Mathematical consequences of the genealogical species concept. Evolution 56(8): 1557-1565. https://doi.org/10.1554/0014-3820(2002)056[1557:MCOTGS]2.0.CO;2.

Huidobro L, Morrone JJ, Villalobos JL, Alvarez F (2006) Distributional patterns of freshwater taxa (fishes, crustaceans and plants) from the Mexican Transition Zone. Journal of biogeography 33(4): 731-741. https://doi.org/10.1111/j.1365-2699.2005.01400.x.

Hustedt F (1930) Bacillariophyta (Diatomeae). In: Pascher A (Ed) Die Süsswasser-Flora Mitteleuropas. Gustav Fischer, Jena.

Jahn R (1995) C. G. Ehrenberg's concept of the diatoms. Archiv für Protistenkunde 146(2): 109-116. https://doi.org/10.1016/S0003-9365(11)80104-1.

Johansen JR (2010) Diatoms of aerial habitats. In: Smol JP, Stoermer EF (Eds) The diatoms: applications for the environmental and earth sciences. Cambridge University Press, Cambridge, 465-472.

Jujnovsky J, Almeida-Leñero L, Bojorge-García M, Monges YL, Cantoral-Uriza E, Mazari-Hiriar M (2010) Hydrologic ecosystem services: water quality and quantity in the Magdalena River, Mexico City. Hidrobiológica 20(2): 113-126.

Julius ML, Theriot EC (2010) The diatoms: a primer. In: Smol JP, Stoermer EF (Eds) The diatoms: applications for the environmental and earth sciences. Cambridge University Press, Cambridge, 8-22.

Kawecka B, Olech M (1993) Diatom communities in the Vanishing and Ornithologist Creek, King George Island, South Shetlands, Antarctica. In: van Dam H (Ed) Twelfth International Diatom Symposium: Proceedings of the Twelfth International Diatom Symposium, Renesse, The Netherlands, 30 August – 5 September 1992. Springer Netherlands, Dordrecht, 327-333.

Kelly M, Bennett C, Coste M, Delgado C, Delmas F, Denys L, Ector L, Fauville C, Ferreol M, Golub M, Jarlman A, Kahlert M, Lucey J, Ni Chathain B, Pardo I, Pfister P, Picinska-Faltynowicz J, Rosebery J, Schranz C, Schaumburg J, van Dam H, Vilbaste S (2009) A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. Hydrobiologia 621(1): 169-182. https://doi.org/10.1007/s10750-008-9641-4.

Kermarrec L, Bouchez A, Rimet F, Humbert JF (2013a) First evidence of the existence of semi-cryptic species and of a phylogeographic structure in the *Gomphonema parvulum* (Kützing) Kützing complex (Bacillariophyta). Protist 164(5): 686-705. https://doi.org/10.1016/j.protis.2013.07.005.

Kermarrec L, Franc A, Rimet F, Chaumeil P, Frigerio JM, Humbert JF, Bouchez A (2014) A next-generation sequencing approach to river biomonitoring using benthic diatoms. Freshwater Science 33(1): 349-363. https://doi.org/10.1086/675079.

Kermarrec L, Franc A, Rimet F, Chaumeil P, Humbert JF, Bouchez A (2013b) Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: a test for freshwater diatoms. Mol Ecol Resour 13(4): 607-619. https://doi.org/10.1111/1755-0998.12105.

Kooistra WH, Medlin LK (1996) Evolution of the diatoms (Bacillariophyta): IV. A reconstruction of their age from small subunit rRNA coding regions and the fossil record. Molecular phylogenetics and evolution 6(3): 391-407. https://doi.org/10.1006/mpev.1996.0088.

Krammer K, Lange-Bertalot H (1991a) Bacillariophyceae. 3. Teil: Centrales, Fragilariaceae, Eunotiaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (Eds) Süßwasserflora von Mitteleuropa. Spektrum, Heidelberg–Berlin, 598 pp.

Krammer K, Lange-Bertalot H (1991b) Bacillariophyceae. 4. Teil: Achnanthaceae. Kritische Ergänzungen zu Navicula (Lineolatae) und Gomphonema Gesamtliteraturverzeichnis Teil 1–4. In: Ettl H, Gärtner G, Gerloff J, Heynig H, Mollenhauer D (Eds) Süßwasserflora von Mitteleuropa. Gustav Fischer, Stuttgart–Jena, 437 pp.

Krammer K, Lange-Bertalot H (1997a) Bacillariophyceae. 1. Teil: Naviculaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (Eds) Süßwasserflora von Mitteleuropa. Gustav Fischer, Jena, 610 pp.

Krammer K, Lange-Bertalot H (1997b) Bacillariophyceae. 2. Teil: Bacillariaceae, Epithemiaceae, Surirellaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (Eds) Süßwasserflora von Mitteleuropa. Gustav Fischer, Jena, 610 pp.

Kützing FT (1844) Die kieselschaligen Bacillarien oder Diatomeen. W. Köhne, Nordhausen, 152 pp.

Lang I, Kaczmarska I (2011) A protocol for a single-cell PCR of diatoms from fixed samples: method validation using *Ditylum brightwellii* (T. West) Grunow. Diatom Research 26(1): 43-49. https://doi.org/10.1080/0269249x.2011.573703.

Lee MJ, Ellis R, Lee JJ (1982) A comparative study of photoadaptation in four diatoms isolated as endosymbionts from larger foraminifera. Marine Biology 68(2): 193-197. https://doi.org/10.1007/bf00397606.

Leeuwenhoek A (1703) Part of a letter from Mr Antony van Leeuwenhoek, F.R.S. concerning green weeds growing in water, and some animalcula found about them. Philosophical Transactions 23(277-288): 1304-1311.

Licursi M, Gomez N (2009) Effects of dredging on benthic diatom assemblages in a lowland stream. J Environ Manage 90(2): 973-982. https://doi.org/10.1016/j.jenvman.2008.03.004.

Lowe RL (2011) The importance of scale in understanding the natural history of diatom communities. The Diatom World. Springer, 293-311.

Mahlknecht J, Schneider JF, Merkel BJ, Navarro de León I, Bernasconi SM (2004) Groundwater recharge in a sedimentary basin in semi-arid Mexico. Hydrogeology Journal 12(5): 511-530. https://doi.org/10.1007/s10040-004-0332-6.

Mann DG (1993) Patterns of sexual reproduction in diatoms. In: van Dam H (Ed) Twelfth International Diatom Symposium: Proceedings of the Twelfth International Diatom Symposium, Renesse, The Netherlands, 30 August – 5 September 1992. Springer Netherlands, Dordrecht, 11-20.

Mann DG, Droop SJM (1996) 3. Biodiversity, biogeography and conservation of diatoms. Hydrobiologia 336(1): 19-32. https://doi.org/10.1007/bf00010816.

Mann DG, Sato S, Trobajo R, Vanormelingen P, Souffreau C (2010) DNA barcoding for species identification and discovery in diatoms. Cryptogamie, Algologie 31(4): 557-577.

Mann DG, Vanormelingen P (2013) An Inordinate Fondness? The Number, Distributions, and Origins of Diatom Species. Journal of Eukaryotic Microbiology 60(4): 414-420. https://doi.org/10.1111/jeu.12047.

Mccormick PV, Cairns J (1994) Algae as Indicators of Environmental-Change. Journal of Applied Phycology 6(5-6): 509-526. https://doi.org/10.1007/Bf02182405.

Medlin LK, 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. https://doi.org/10.2216/i0031-8884-43-3-245.1.

Mercado-Silva N, Helmus MR, Zanden MJV (2009) The effects of impoundment and non-native species on a river food web in Mexico's central plateau. River Research and Applications 25(9): 1090-1108. https://doi.org/10.1002/rra.1205.

Mercado-Silva N, Lyons J, Díaz-Pardo E, Gutiérrez-Hernández A, Ornelas-García CP, Pedraza-Lara C, Zanden MJV (2006) Long-term changes in the fish assemblage of the Laja River, Guanajuato, central Mexico. Aquatic Conservation: Marine and Freshwater Ecosystems 16(5): 533-546. https://doi.org/10.1002/aqc.737.

Meyer CP, Paulay G (2005) DNA Barcoding: Error Rates Based on Comprehensive Sampling. PLOS Biology 3(12): e422. https://doi.org/10.1371/journal.pbio.0030422.

Mittermeier R, Robles Gil P, Hoffman M, Pilgrim J, Brooks T, Mittermeier C, Lamoreux J, Da Fonseca G (2004) Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions. CEMEX, Mexico City, 390 pp.

Mittermeier RA, Turner WR, Larsen FW, Brooks TM, Gascon C (2011) Global Biodiversity Conservation: The Critical Role of Hotspots. In: Zachos FE, Habel JC (Eds) Biodiversity Hotspots: Distribution and Protection of Conservation Priority Areas. Springer Berlin Heidelberg, Berlin, Heidelberg, 3-22.

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJ, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst Biol 58(3): 298-311. https://doi.org/10.1093/sysbio/syp027.

Moniz MBJ, Kaczmarska I (2009) Barcoding diatoms: Is there a good marker? Molecular Ecology Resources 9(65-74. https://doi.org/10.1111/j.1755-0998.2009.02633.x.

Mora D, Carmona J, Jahn R, Zimmermann J, Abarca N (2017) Epilithic diatom communities of selected streams from the Lerma-Chapala Basin, Central Mexico, with the description of two new species. PhytoKeys 88(39-69. https://doi.org/10.3897/phytokeys.88.14612.

Morales EA, Siver PA, Trainor FR (2001) Identification of diatoms (Bacillariophyceae) during ecological assessments: Comparison between Light Microscopy and Scanning Electron Microscopy techniques. Proceedings of the Academy of Natural Sciences of Philadelphia 151(1): 95-103. https://doi.org/10.1635/0097-3157(2001)151[0095:Iodbde]2.0.Co;2.

Morales EA, Wetzel CE, Rivera SF, Van de Vijver B, Ector L (2014) Current taxonomic studies on the diatom flora (Bacillariophyceae) of the Bolivian Altiplano, South America, with possible consequences for palaeoecological assessments. Journal of Micropalaeontology 33(2): 121-129. https://doi.org/10.1144/jmpaleo2014-007.

Moritz C, Cicero C (2004) DNA barcoding: promise and pitfalls. PLoS Biol 2(10): e354. https://doi.org/10.1371/journal.pbio.0020354.

Müller OF (1783) Strand-Parlebandet och Armbandet, tvanne microscopiska Strandvaxter. Kongl Vetenskaps Academiens, Nya Handlingar 4(80-85 and Tab. III.

Müller OF (1786) *Animalcula infusoria fluviatilia et marina.* Lvi + 367 pp. Havniae: N. Moller.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature 403(6772): 853-858. https://doi.org/10.1038/35002501.

Nelson DM, Tréguer P, Brzezinski MA, Leynaert A, Quéguiner B (1995) Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. Global Biogeochemical Cycles 9(3): 359-372. https://doi.org/10.1029/95GB01070.

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GVN, Underwood EC, D'Amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR (2001) Terrestrial ecoregions of the world: a new map of life on earth: a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. BioScience 51(11): 933-938. https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, Bowser SS, Cepicka I, Decelle J, Dunthorn M, Fiore-Donno AM, Gile GH, Holzmann M, Jahn R, Jirku M, Keeling PJ, Kostka M, Kudryavtsev A, Lara E, Lukes J, Mann DG, Mitchell EA, Nitsche F, Romeralo M, Saunders GW, Simpson AG, Smirnov AV, Spouge JL, Stern RF, Stoeck T, Zimmermann J, Schindel D, de Vargas C (2012) CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. PLoS Biol 10(11): e1001419. https://doi.org/10.1371/journal.pbio.1001419.

Poulickova A, Spackov J, Kelly MG, Duchoslav M, Mann DG (2008) Ecological variation within Sellaphora species complexes (Bacillariophyceae): specialists or generalists? Hydrobiologia 614(1): 373-386. https://doi.org/10.1007/s10750-008-9521-y.

Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7(3): 355-364. https://doi.org/10.1111/j.1471-8286.2007.01678.x.

Raven JA, Waite AM (2004) The evolution of silicification in diatoms: inescapable sinking and sinking as escape? New Phytologist 162(1): 45-61. https://doi.org/10.1111/j.1469-8137.2004.01022.x.

Reimer CW, Lee JJ (1988) New species of endosymbiotic diatoms (Bacillariophyceae) inhabiting larger Forminifera in the Gulf of Elat (Red Sea), Israel. Proceedings of the Academy of Natural Sciences of Philadelphia 140(2): 339-351.

Rimet F, Abarca N, Bouchez A, Kusber WH, Jahn R, Kahlert M, Keck F, Kelly MG, Mann DG, Piuz A, Trobajo R, Tapolczai K, Vasselon V, Zimmermann J (2018) The potential of high throughput sequencing (HTS) of natural samples as a source of primary taxonomic information for reference libraries of diatom barcodes. Fottea 18(1): *in press*. https://doi.org/10.5507/fot.2017.013.

Rimet F, Bouchez A (2012) Biomonitoring river diatoms: Implications of taxonomic resolution. Ecological Indicators 15(1): 92-99. https://doi.org/10.1016/j.ecolind.2011.09.014.

Rosales-Hoz L, Carranza-Edwards A, Lopez-Hernandez M (2000) Heavy metals in sediments of a large, turbid tropical lake affected by anthropogenic discharges. Environmental Geology 39(3): 378-383. https://doi.org/10.1007/s002540050017.

Round FE, Crawford RM, Mann DG (1990) Diatoms: biology and morphology of the genera. Cambridge University Press, 747 pp.

Sand-Jensen K (2014) Ecophysiology of gelatinous *Nostoc* colonies: unprecedented slow growth and survival in resource-poor and harsh environments. Annals of Botany 114(1): 17-33. https://doi.org/10.1093/aob/mcu085.

Sarthou G, Timmermans KR, Blain S, Tréguer P (2005) Growth physiology and fate of diatoms in the ocean: a review. Journal of Sea Research 53(1): 25-42. https://doi.org/10.1016/j.seares.2004.01.007.

Sedeño-Díaz JE, López-López E (2007) Water Quality in the Río Lerma, Mexico: An Overview of the Last Quarter of the Twentieth Century. Water Resources Management 21(10): 1797-1812. https://doi.org/10.1007/s11269-006-9128-x.

Segura-Garcia V, Cantoral-Uriza EA, Israde I, Maidana N (2012) Epilithic diatoms (Bacillariophyceae) as indicators of water quality in the Upper Lerma River, Mexico. Hidrobiológica 22(1): 16-27.

Simonsen R (1979) The diatom system: ideas on phylogeny. Bacillaria 2(9-71.

Sims PA, Mann DG, Medlin LK (2006) Evolution of the diatoms: insights from fossil, biological and molecular data. Phycologia 45(4): 361-402. https://doi.org/10.2216/05-22.1.

Siver PA (1995) The distribution of chrysophytes along environmental gradients: their use as biological indicators. In: Sandgren CD, Smol JP, Kristiansen J (Eds) Chrysophyte algae: ecology, phylogeny and development. Cambridge University Press, Cambridge, 232-268.

Siver PA, Kling H (1997) Morphological observations of Aulacoseira using scanning electron microscopy. Canadian Journal of Botany 75(11): 1807-1835. https://doi.org/10.1139/b97-894.

Sorhannus U (2007) A nuclear-encoded small-subunit ribosomal RNA timescale for diatom evolution. Marine Micropaleontology 65(1): 1-12. https://doi.org/10.1016/j.marmicro.2007.05.002.

Stancheva R, Sheath RG, Read BA, McArthur KD, Schroepfer C, Kociolek JP, Fetscher AE (2013) Nitrogen-fixing cyanobacteria (free-living and diatom endosymbionts): their use in southern California stream bioassessment. Hydrobiologia 720(1): 111-127. https://doi.org/10.1007/s10750-013-1630-6.

Stauber JL, Jeffrey SW (1988) Photosynthetic pigments in fifty-one species of marine diatoms. Journal of Phycology 24(2): 158-172. https://doi.org/10.1111/j.1529-8817.1988.tb04230.x.

Stevenson R, Pan Y, van Dam H (2010) Assessing environmental conditions in rivers and streams with diatoms. The Diatoms: Applications for the Environmental and Earth Sciences, 2nd ed Cambridge University Press, Cambridge.

Stong T, Osuna CA, Shear H, de Anda Sanchez J, Ramírez G, de Jesús Díaz Torres J (2013) Mercury concentrations in common carp (Cyprinus carpio) in Lake Chapala, Mexico: A lakewide survey. Journal of Environmental Science and Health, Part A 48(14): 1835-1841. https://doi.org/10.1080/10934529.2013.823340.

Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012a) Environmental DNA. Mol Ecol 21(8): 1789-1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012b) Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol 21(8): 2045-2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x.

Tejeda S, Zarazúa G, Ávila-Pérez P, Carapia-Morales L, Martínez T (2010) Total reflection X-ray fluorescence spectrometric determination of elements in water hyacinth from

the Lerma River. Spectrochimica Acta Part B: Atomic Spectroscopy 65(6): 483-488. https://doi.org/10.1016/j.sab.2010.04.005.

Theriot EC, Ashworth M, Ruck E, Nakov T, Jansen RK (2010) A preliminary multigene phylogeny of the diatoms (Bacillariophyta): challenges for future research. Plant Ecology and Evolution 143(3): 278-296. https://doi.org/10.5091/plecevo.2010.418.

Theriot EC, Cannone JJ, Gutell RR, Alverson AJ (2009) The limits of nuclear-encoded SSU rDNA for resolving the diatom phylogeny. European Journal of Phycology 44(3): 277-290. https://doi.org/10.1080/09670260902749159.

Theriot EC, Ruck E, Ashworth M, Nakov T, Jansen RK (2011) Status of the pursuit of the diatom phylogeny: are traditional views and new molecular paradigms really that different? In: Seckbach J, Kociolek P (Eds) The diatom world. Springer Netherlands, Dordrecht, 119-142.

Trobajo R, Clavero E, Chepurnov VA, Sabbe K, Mann DG, Ishihara S, Cox EJ (2009) Morphological, genetic and mating diversity within the widespread bioindicator *Nitzschia palea* (Bacillariophyceae). Phycologia 48(6): 443-459. https://doi.org/10.2216/08-69.1.

Vanormelingen P, Verleyen E, Vyverman W (2008) The diversity and distribution of diatoms: from cosmopolitanism to narrow endemism. Biodiversity and Conservation 17(2): 393-405. https://doi.org/10.1007/s10531-007-9257-4.

Vasselon V, Bouchez A, Rimet F, Jacquet S, Trobajo R, Corniquel M, Tapolczai K, Domaizon I (submitted) A correction factor inferred from cell biovolume improves quantification in diatom metabarcoding. Molecular Ecology Resources.

Vasselon V, Rimet F, Tapolczai K, Bouchez A (2017) Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). Ecological Indicators 82(1-12. https://doi.org/10.1016/j.ecolind.2017.06.024.

Villamagna AM, Murphy BR, Trauger DL (2010) Behavioral response of American Coots (Fulica americana) to water hyacinth (Eichhornia crassipes) in Lake Chapala, Mexico. Waterbirds 33(4): 550-555. https://doi.org/10.1675/063.033.0416.

Visco JA, Apotheloz-Perret-Gentil L, Cordonier A, Esling P, Pillet L, Pawlowski J (2015) Environmental Monitoring: Inferring the Diatom Index from Next-Generation Sequencing Data. Environ Sci Technol 49(13): 7597-7605. https://doi.org/10.1021/es506158m.

Williams DM, Kociolek JP (2007) Pursuit of a natural classification of diatoms: History, monophyly and the rejection of paraphyletic taxa. European Journal of Phycology 42(3): 313-319. https://doi.org/10.1080/09670260701419921.

Williams DM, Kociolek JP (2011) An overview of diatom classification with some prospects for the future. In: Seckbach J, Kociolek P (Eds) The Diatom world. Springer, 47-91.

Williamson CE, Dodds W, Kratz TK, Palmer MA (2008) Lakes and streams as sentinels of environmental change in terrestrial and atmospheric processes. Frontiers in Ecology and the Environment 6(5): 247-254. https://doi.org/10.1890/070140.

Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. Molecular biology and evolution 21(5): 809-818. https://doi.org/10.1093/molbev/msh075.

Yu DW, Ji YQ, Emerson BC, Wang XY, Ye CX, Yang CY, Ding ZL (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods in Ecology and Evolution 3(4): 613-623. https://doi.org/10.1111/j.2041-210X.2012.00198.x.

Zarazua G, Ávila-Pérez P, Tejeda S, Barcelo-Quintal I, Martínez T (2006) Analysis of total and dissolved heavy metals in surface water of a Mexican polluted river by total reflection X-ray fluorescence spectrometry. Spectrochimica Acta Part B: Atomic Spectroscopy 61(10): 1180-1184. https://doi.org/10.1016/j.sab.2006.06.010.

Zhu F, Massana R, Not F, Marie D, Vaulot D (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. FEMS Microbiol Ecol 52(1): 79-92. https://doi.org/10.1016/j.femsec.2004.10.006.

Zimmermann J, Abarca N, Enke N, Skibbe O, Kusber WH, Jahn R (2014) Taxonomic reference libraries for environmental barcoding: a best practice example from diatom research. PLoS One 9(9): e108793. https://doi.org/10.1371/journal.pone.0108793.

Zimmermann J, Glöckner G, Jahn R, Enke N, Gemeinholzer B (2015) Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Mol Ecol Resour 15(3): 526-542. https://doi.org/10.1111/1755-0998.12336.

Zimmermann J, Jahn R, Gemeinholzer B (2011) Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. Organisms Diversity & Evolution 11(3): 173-192. https://doi.org/10.1007/s13127-011-0050-6.

# 2 Epilithic diatom communities of selected streams from the Lerma–Chapala Basin, Central Mexico, with the description of two new species

**Demetrio Mora**[1*], Javier Carmona[2], Regine Jahn[1], Jonas Zimmermann[1] and Nélida Abarca[1]

[1]Botanischer Garten und Botanisches Museum Berlin–Dahlem, Freie Universität Berlin, Königin–Luise–Straße 6–8, 14195 Berlin, Germany

[2]Facultad de Ciencias, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, Coyoacán, 04510 México, D.F., México

PhytoKeys 88: 39-69. https://doi.org/10.3897/phytokeys.88.14612

## 2.1 Abstract

The Lerma–Chapala Basin, in Central Mexico, is geologically heterogeneous, climatically diverse and boasts high biodiversity, lying within two Biodiversity Hotspots, namely Mesoamerica and the Madrean Pine–Oak Woodlands. Epilithon and water samples were collected in the basin from 14 sampling sites three times each, two sampling campaigns during the rainy season and one in the dry season. A total of 274 infrageneric taxa in 48 genera were recorded. The taxonomic composition observed was dominated by taxa from the genera *Nitzschia*, *Gomphonema*, *Pinnularia*, *Navicula*, *Sellaphora* and *Eunotia*. About a third of the taxa found could not be identified to the species level. From those unidentified morphodemes, two are described as new species, namely *Brachysira altepetlensis* and *Sellaphora queretana*. Furthermore, *Eolimna rhombica* is transferred to *Sellaphora*. Canonical Correspondence Analysis (CCA) revealed that specific conductivity and pH were the main environmental factors driving the community composition observed. Three groups of samples were identified after the CCA: 1) characterized by acidic waters and low conductivity; 2) with circumneutral waters, low specific conductivity and high temperature and phosphorous concentrations; and 3) characterized by circumneutral waters, high conductivity and low nitrogen concentrations. The indicator value method (IndVal), based on the relative abundance and relative frequency of the most abundant taxa was calculated based on the groups observed in the CCA, identifying the characteristic taxa for each of the three groups.

**Key words**: Central Mexico, diatom communities, epilithon, indicator species, Lerma–Chapala Basin, mountain streams, new species

## 2.2 Introduction

Lotic environments, i.e. streams, are unidirectional flows of water. They are characterized by a broad spatial (i.e. substrate, slope, vegetation) and temporal (i.e. water velocity, light) heterogeneity, which determines the specialized biota that inhabit them (Giller and Malmqvist 1998, Allan and Castillo 2007). Stream diatoms have features that allow them to thrive in flowing waters, such as the morphological and physiological ability to adhere directly or by means of stalks or mucilage pads to different substrate types to avoid being dragged away by water. Apart from water velocity, physical and chemical variables of the water such as temperature, pH, specific conductivity and nutrient concentrations are determining factors for diatom composition and community structure (Bellinger and Sigee 2010, Stevenson et al. 2010).

Even though there is mounting evidence of the applied use of diatoms as indicators of environmental change in lotic environments (Kelly 1998, Potapova and Charles 2002, 2007, Smol and Stoermer 2010), diatom studies from Mexican streams are relatively scarce, despite the increasing pressure these environments are facing to satisfy human demand for clean water.

Diatom studies of lotic environments from Mexico have been mostly focused on the center of the country: Antigua River Basin (Vázquez et al. 2011); Balsas River Basin (Valadez-Cruz et al. 1996, Bojorge-García et al. 2010, 2014), Lerma–Chapala Basin (Abarca-Mejía 2010, Segura-García et al. 2010, 2012, 2016, Mora et al. 2015), Mexico Basin (Ramírez-Vázquez et al. 2001, Ramírez-Vázquez and Cantoral-Uriza 2003, Bojorge-García and Cantoral-Uriza 2007, Carmona-Jiménez et al. 2016); Pánuco River Basin (Cantoral-Uriza et al. 1997) and Papaloapan River Basin (Tavera et al. 1994). Most of these studies focused on the flora *per se* but also on community structure and bioindication. Despite the research done, the diatom diversity of the region seems to be low due to clustering of taxa into species complexes and force–fitting into already described taxa.

The studies conducted in the Lerma–Chapala Basin have been focused on the polluted Lerma River and some of its main tributaries (Abarca-Mejía 2010, Segura–García et al. 2010, 2012, 2016, Mora et al. 2015). But no study has been conducted so far on the headwater streams of the basin, which are important in the establishment of reference conditions for biological integrity evaluations based on regional characteristics of the streams and its associated diatom flora (Stoddard et al. 2006, Tornés et al. 2007).

In order to contribute to the studies done in the Lerma–Chapala Basin, one of the most important basins of the country regarding population and trade, the aims of this study are: to document the epilithic diatom diversity from selected headwater and midland streams from the Lerma–Chapala Basin, Central Mexico; to illustrate the most abundant taxa; and to identify the environmental factors that determine the variation observed in diatom composition.

## 2.3 Methods

### 2.3.1 Study area

The Lerma–Chapala Basin is located in Central Mexico, covering an area of 53,591.3 km$^2$ (Fig. 1). It is geologically heterogeneous, has a strong elevational gradient, is climatically diverse, has well defined rainy (June to October) and dry seasons (November to May) and boasts high biodiversity. It lies within two Biodiversity Hotspots, namely Mesoamerica and the Madrean Pine–Oak Woodlands (Cotler et al. 2006, CEPFa, b 2017).

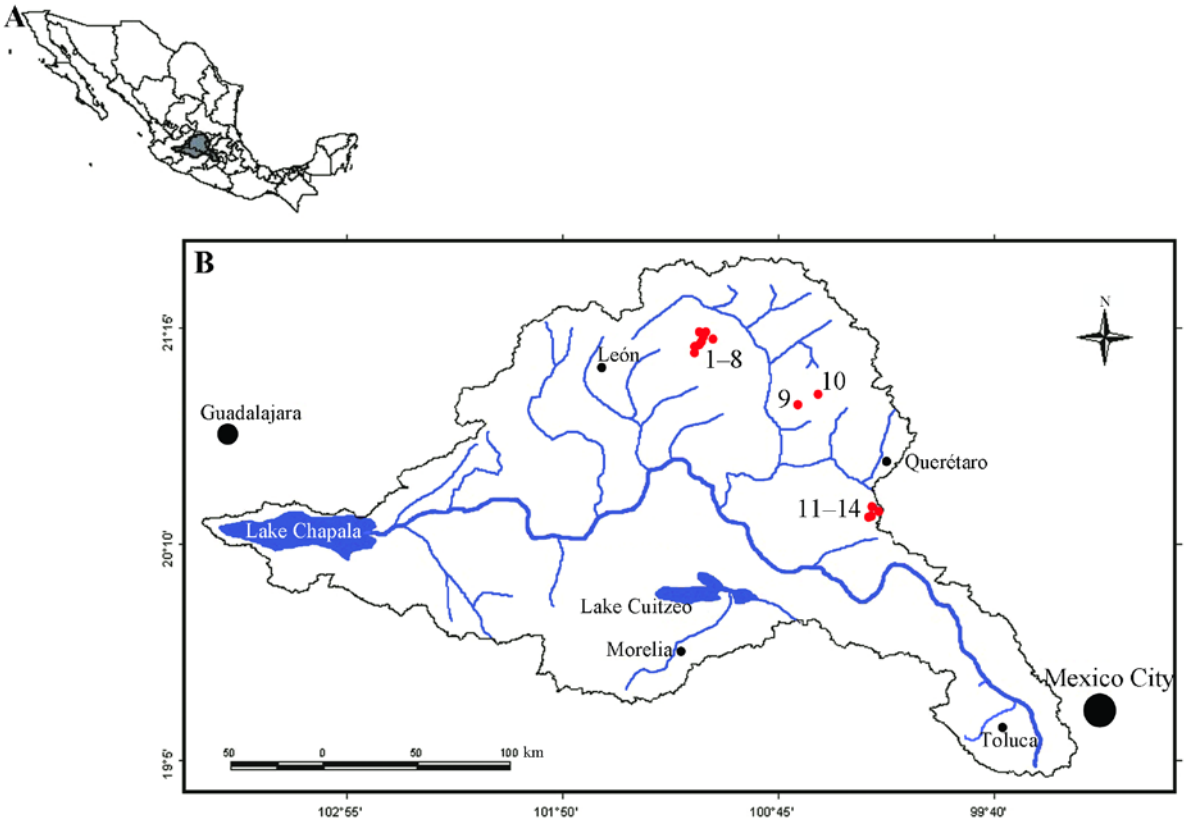



Fig. 1. Location of the area of study. A Map of Mexico, showing the location of the Lerma–Chapala Basin in the center of the country. B Location of the 14 sampling sites in the Lerma–Chapala Basin, indicated by red dots. The numbers next to the red dots refer to the name of the sampling site in Table 1.

This basin is one of the most important centers in the country for agriculture and industry, and has a population of more than 15 million inhabitants (Wester et al. 2005, Cotler et al. 2006). But the Lerma–Chapala Basin is also one of the most environmentally degraded basins in the country, facing serious water related issues because of overexploitation and pollution of surface and underground waters (Aparicio 2001, Wester et al. 2005).

Table 1: Sampling sites in the Lerma-Chapala Basin, including site number and name, type of water body, geographical coordinates and elevation.

| Site | | Water body | Latitude (N) | Longitude (W) | Elevation (m a.s.l) |
|---|---|---|---|---|---|
| 1 | La Mesa | Stream | 21° 05' 28.69" | 101° 08' 18.98" | 2215 |
| 2 | Calvillo | Stream | 21° 06' 50.40" | 101° 08' 04.10" | 2138 |
| 3 | Ojo de Agua de Calvillo | Stream | 21° 07' 41.80" | 101° 07' 04.50" | 2102 |
| 4 | Peña Colorada | Stream | 21° 09' 03.84" | 101° 05' 58.96" | 2110 |
| 5 | San Martín | Stream | 21° 09' 24.50" | 101° 03' 11.30" | 2017 |
| 6 | Paredones | Stream | 21° 11' 20.60" | 101° 06' 53.40" | 2089 |
| 7 | La Laborcilla 1 | Stream | 21° 11' 04.70" | 101° 06' 14.60" | 2076 |
| 8 | La Laborcilla 2 | Stream | 21° 11' 20.10" | 101° 05' 37.90" | 2065 |
| 9 | El Membrillo | Stream | 20° 50' 21.22" | 100° 38' 43.46" | 2114 |
| 10 | Guanajuatito | Spring fed-creek | 20° 53' 23.98" | 100° 32' 30.72" | 2120 |
| 11 | Los Ailes 1 | Stream | 20° 19' 58.72" | 100° 15' 17.09" | 2358 |
| 12 | Laguna de Servín 1 | Stream | 20° 18' 18.10" | 100° 17' 38.10" | 2409 |
| 13 | Laguna de Servín 2 | Stream | 20° 18' 45.20" | 100° 17' 25.60" | 2409 |
| 14 | Los Ailes 2 | Stream | 20° 20' 50.20" | 100° 16' 45.50" | 2317 |

The 14 sampling sites selected for this study are located in the north and central–east sections of the Lerma–Chapala Basin at elevations ranging from 2,000 to 2,400 meters above sea level. Of those 14 sites, one is a perennial spring–fed creek and 13 correspond to streams that have water during most part of the year (Fig. 1, Table 1). Sampling sites 1–8 are located at the foothills of the Sierra de Santa

Rosa, an oak–forested mountain range of priority for the conservation of biodiversity in Mexico (Arriaga et al. 2000); the mean temperature of the area is 16.1 °C and the average rainfall is 642 mm (CNA 2017a). Sites 9 and 10 are in a flat area dominated by shrub land and subsistence agriculture; the mean air temperature is 18.8 °C and the average rainfall is 566 mm (CNA 2017b). Sites 11–14 are located at the foothills of a small mountainous area dominated by pine–oak forests and subsistence agriculture; the mean air temperature is 15.6 °C and the average rainfall is 774 mm (CNA 2017c).

### 2.3.2 Sampling

Water and epilithon samples were collected three times from each sampling site in: September/October 2013, rainy season (sampling campaign I); February 2014, dry season (sampling campaign II); and September 2014, rainy season (sampling campaign III); resulting in 42 water and epilithon samples. Each epilithon sample was collected from five cobbles across a transversal section of the stream, brushing with a disposable toothbrush ten square centimeters of epilithic growth from each of the five cobbles to make a composite sample, fixed in 70% alcohol. *In situ* measurements of pH, water temperature, specific conductivity corrected to 25°C and total dissolved solids were recorded using a Hanna multi–sensor (HI 991300, California, USA). Dissolved oxygen was recorded with an YSI–85 oxygen meter (YSI, Ohio, USA). Dissolved oxygen saturation percentages were calculated from dissolved oxygen data according to correcting factors for elevation and water temperature. Water velocity was recorded with a Global Water FP111 velocity meter (Texas, USA). At each sampling site, a 500 ml sample of water was filtered through 0.22 µm and 0.45 µm filter membranes (Millipore, Massachusetts, USA) and collected in sterile polypropylene bottles for chemical analysis. Samples were kept cold and in the dark before laboratory analysis. The subsequent chemical laboratory analyses were adapted from Standard Methods for the Examination of Water and Wastewater and analyzed using a DR 3900 laboratory Spectrophotometer (Hach Company, Loveland, Colorado) (HACH 2003, APHA 2005): nitrite nitrogen ($NO_2^-$–N), nitrate nitrogen ($NO_3^-$–N), ammonium nitrogen ($NH_4^+$–N), soluble reactive phosphorus (SRP, in theory, mostly in the form of orthophosphate, $PO_4^{3-}$–P) and total alkalinity (as $CaCO_3$). Dissolved inorganic nitrogen (DIN) was calculated as the sum of the three inorganic nitrogen forms in water (nitrites, nitrates and ammonium).

The Riparian Forest Quality index (QBR from its Catalan abbreviation) was calculated in order to evaluate the riparian habitat quality (Munné et al. 2003). This index evaluates quantitatively four components of the riparian habitat: 1) Total riparian vegetation cover, evaluates the vegetation cover of all plants except for annuals and also taking into account the connectivity between the riparian area and surrounding terrestrial vegetation. 2) Vegetation cover structure, it assesses the

structural complexity of the riparian habitat, which is determined by the percent coverage and patch distribution of trees, shrubs and aquatic plants. 3) Cover quality, takes into account the number of native tree and shrub species (dependent of the river type) and also evaluates if the river has alterations such as man-made structures, presence of alien species and garbage. To determine the river type, the following geomorphological criteria are evaluated: slope and form of the riparian zone, presence of islands in the river and percentage of hard substrata. 4) River channel alterations, evaluates how pristine or altered is the river, considering if the river has been permanently channelized, if there are rigid structures or fluvial terraces constraining the flow. Each component of the index scores between 0 and 25, therefore the index score go from 0 to 100. The index has five classes: natural condition, good quality, fair quality, poor quality and bad quality. The native vegetation, needed to calculate this index, was identified following Zamudio et al. (1992), Carranza-González (1995), Carranza-González and Madrigal-Sánchez (1995), Calderón de Rzedowski and Rzedowski (2001), Rzedowski and Calderón de Rzedowski (2004).

### 2.3.3 Diatom analysis

Fractions of the diatom samples were cleaned by adding aliquots of 35% hydrogen peroxide and heating at 80°C until no bubbling was observed. After the digestion was completed, peroxide remnants were removed by rinsing at least three times with distilled water. Samples were finally diluted with distilled water in order to avoid high concentrations of valves and sediment. Three permanent slides per sample were made using the high refraction index mounting medium Naphrax®. The slides were scanned and the diatoms photographed under the light microscope (LM) in order to account for diatom diversity, using a Zeiss Axioscope microscope with Differential Interference Contrast equipped with an AXIOAM MRc camera. In order to estimate the relative abundance of the taxa, a minimum of 500 valves per sample were counted and identified with the 100x immersion oil objective. Aliquots of cleaned sample material for scanning electron microscopy observations were mounted on stubs, sputter-coated with gold-palladium and observed under a Hitachi FE 8010 scanning electron microscope (SEM) operated at 1.0 kV. Samples and slides are stored at the Diatom Collection of the Botanical Garden and Botanical Museum Berlin–Dahlem, Freie Universität Berlin. Diatoms were identified to the lowest taxonomical level possible using monographs as well as papers for particular taxa (Supplementary material 1 – Appendix 1). Taxa identified with ‘cf.’ (*confer*) before the epithet indicate that it could be that taxon but the taxonomic identity is still uncertain, ‘aff.’ (*affinis*) that it has some similarity to the taxon but it is not conspecific and ‘sp.’ (species) was used when the taxon showed no similarity with any known species after the literature review.

### 2.3.4 Data analysis

Only taxa with relative abundance ≥1% were included in the statistical analyses, resulting in 105 diatom taxa. Diatom abundances were transformed using Hellinger's transformation, which is suited to large abundance datasets with lots of low counts and zeros (Legendre and Gallagher 2001).

From the initial dataset composed of 42 samples, only 39 were used for the analysis of running waters, i.e. those streams with water velocity records in at least one of the sampling campaigns; the three samples of site 10 were omitted since no water velocity was recorded in this spring-fed creek at any of the three sampling campaigns, with 10 cm/s being the detection limit of the water velocity meter. All the environmental variables, except for temperature, pH and water velocity were transformed using $\log_{10}$ (x+1) because they had skew distributions. Distribution tests were run in STATISTICA 8.0.

Multivariate analyses were performed to explore gradients in diatom composition and its relation to environmental factors. Detrended Correspondence Analysis (DCA) was used to estimate gradient lengths. The first four axes showed lengths of 5.7, 3, 2.3 and 2.2, suggesting a strong unimodal response, meaning that a method based on unimodal models like Canonical Correspondence Analysis (CCA) would be appropriate for subsequent ordination. CCA was run to identify variation in species composition and abundance that can be determined by environmental variables. Since not all the environmental variables influence diatom distributions independently, CCA with forward selection and unrestricted Montecarlo permutation tests was used (999 permutations, $p<0.05$). All ordinations were done using CANOCO 4.5 for Windows (ter Braak and Šmilauer 2002), with downweighting of rare species in all cases.

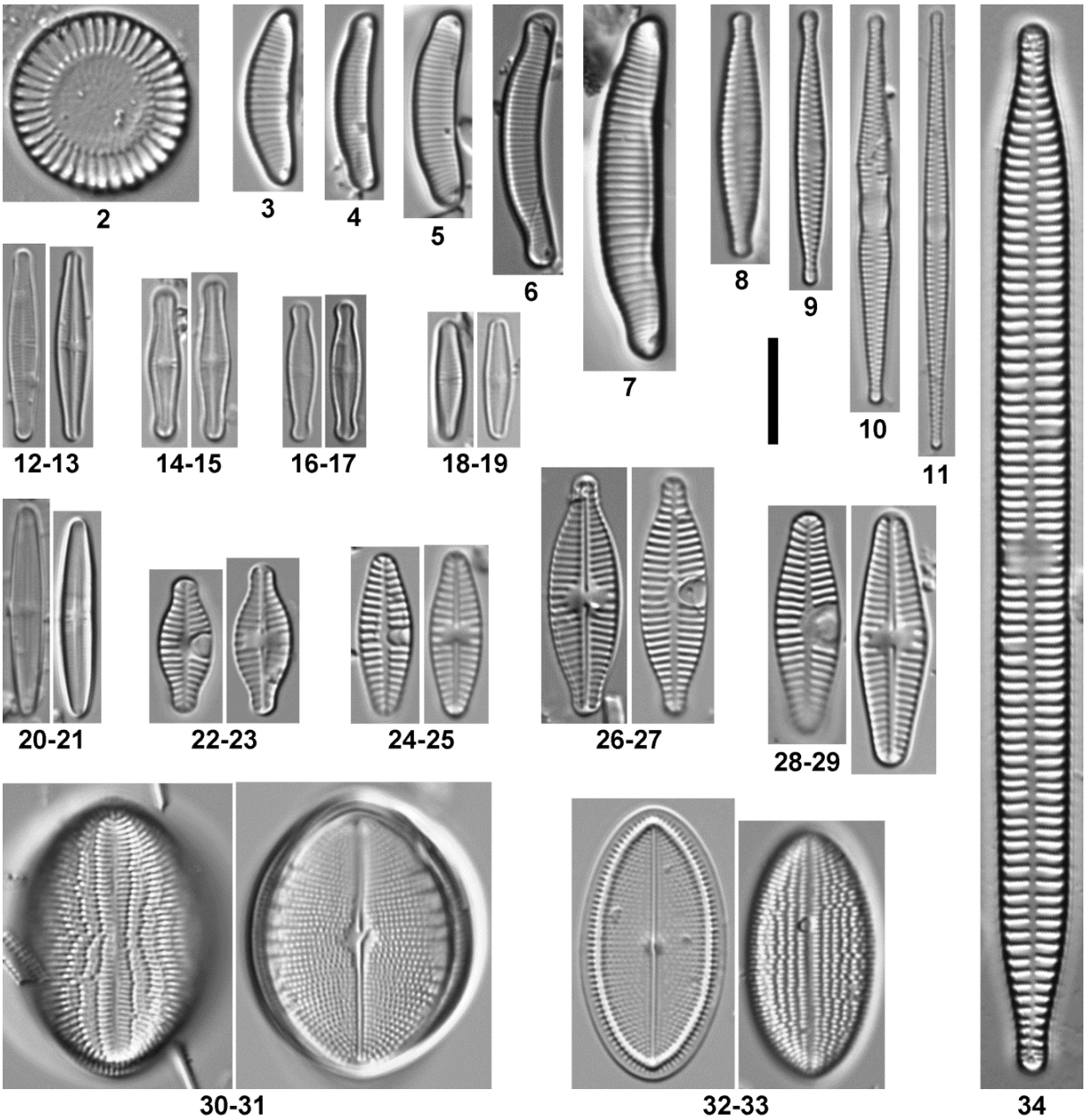
The indicator value method (IndVal) (Dufrêne and Legendre 1997) was used to identify the most characteristic species of the groups visualized after the CCA. This method combines the specificity (relative abundance) and fidelity (relative frequency) of a species to a given group. The indicator value of a species is given in percentage, reaching its maximum when all the individuals of a species are present at all the sites of a single group. Species with high indicator values >50% are considered to be good indicators; species with values between 25–50% might be regarded as detector species of change, therefore detector species can be in more than one group (Tornés et al. 2007, Carmona-Jiménez et al. 2016). IndVal calculations were run in PC–ORD 4 (McCune and Mefford 1999) with untransformed abundance data. The statistical significance of the IndVal was tested with a randomization Montecarlo test (10,000 permutations, $p<0.05$). The Shannon-Wiener diversity index and Pielou evenness index were calculated as in Peet (1974) for the groups visualized after the CCA.

## 2.4 Results

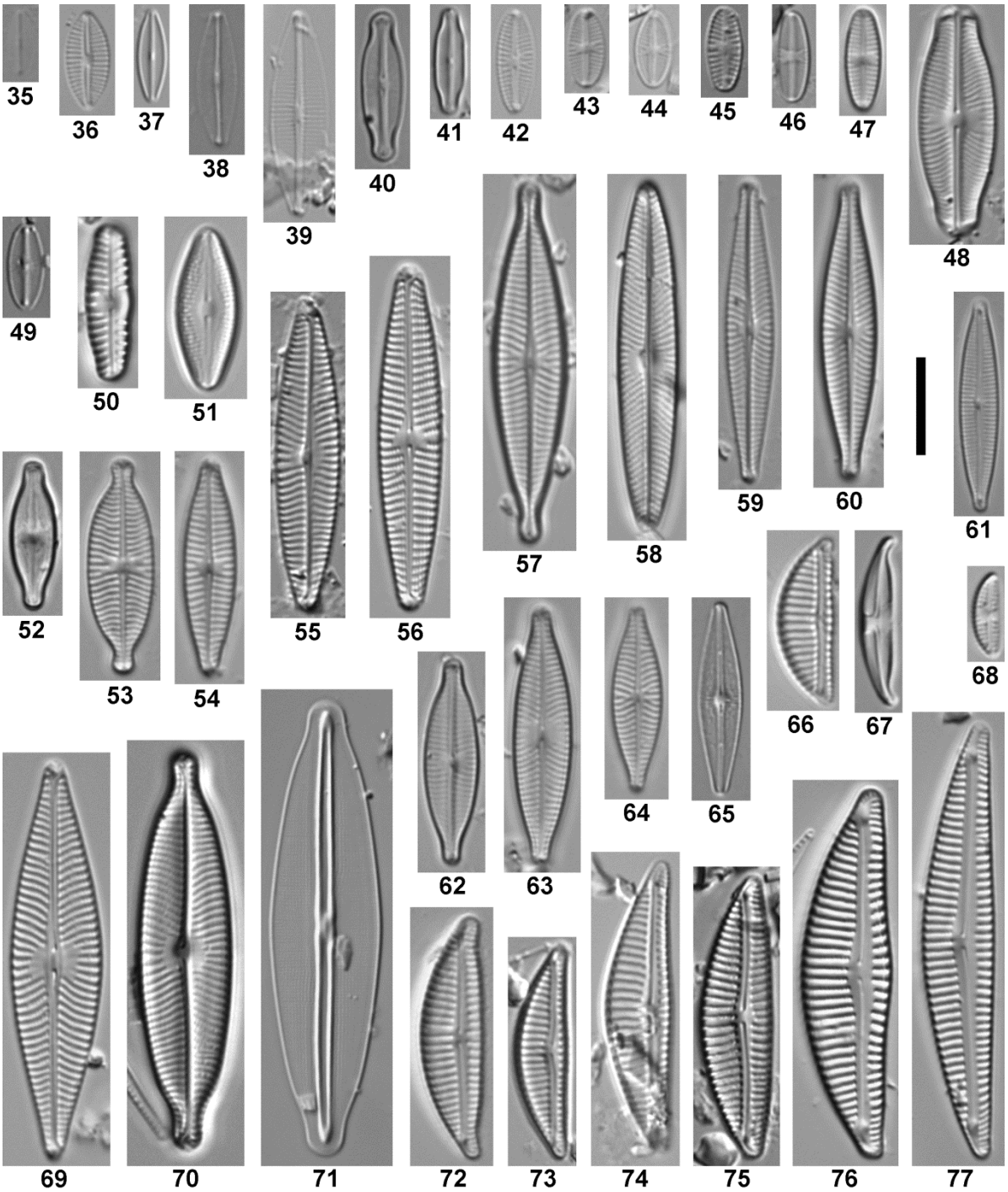
### 2.4.1 Species composition and taxonomy.

A total of 196 taxa (species and varieties) were found while performing the counts to determine relative abundances. Seventy-eight additional taxa were observed by scanning the whole slides looking for rare taxa, bringing the total diversity to 274 taxa belonging to 48 genera (supplementary material 1 – Appendix 1). Sixty-three taxa are new records for the Lerma–Chapala Basin. The most common taxa (relative abundances ≥1% in at least one sample), illustrated here (Figs 2–117), were included in subsequent statistical analyses.

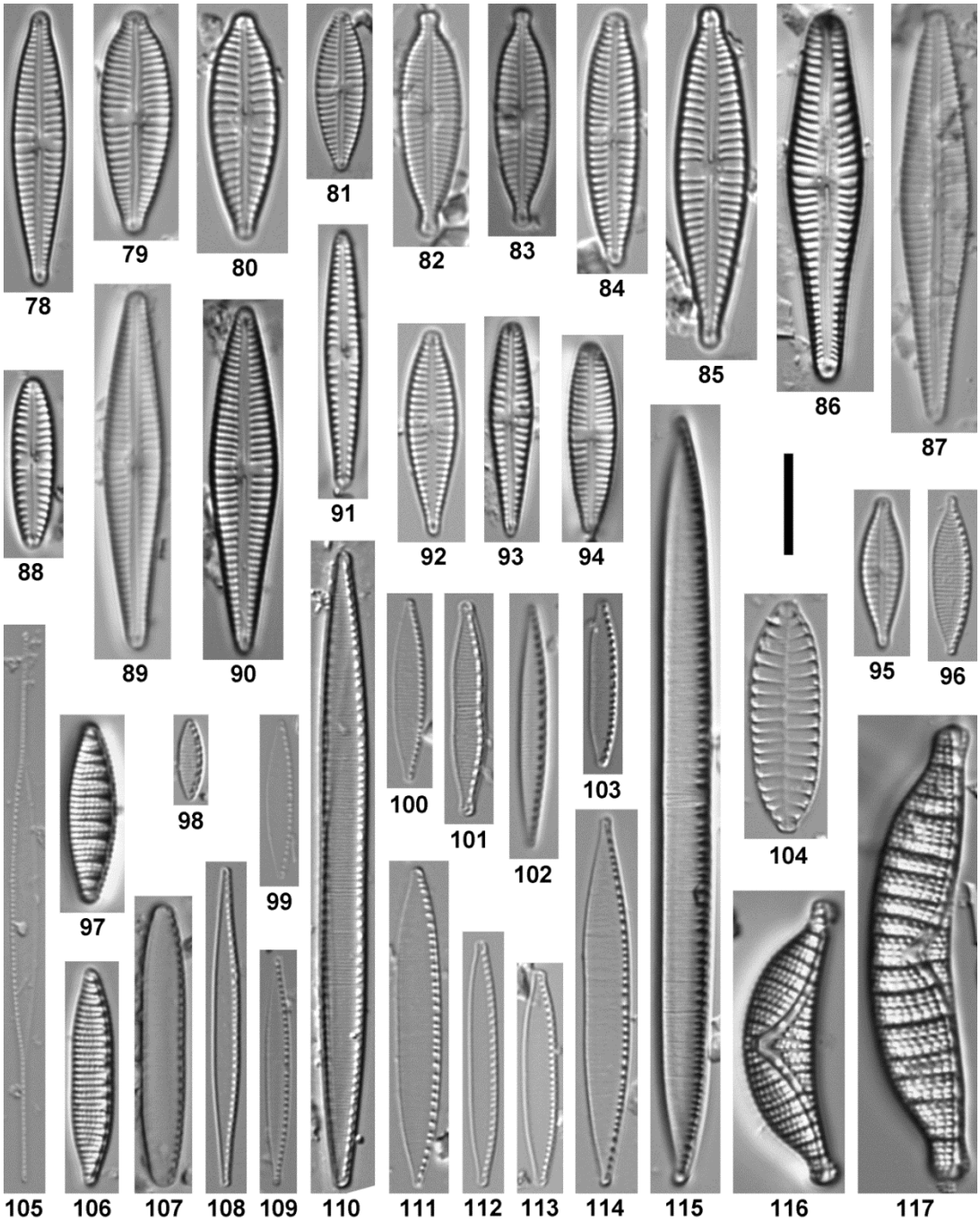
A high specific taxa richness was found among the genera *Nitzschia* (35 taxa), *Gomphonema* (26 taxa), *Pinnularia* (21 taxa), *Navicula* (19 taxa), *Sellaphora* (18 taxa) and *Eunotia* (16 taxa). About a third of the diversity found, 94 taxa, did not fit completely into already described species. Most of the taxa were found in relatively low abundances while further scanning the slides under the LM after the enumeration of 500 valves; when scanning samples under the SEM, some of those rare unidentified taxa were found but in several cases not. When the taxa were found under the SEM, not enough valves were observed for reliable identification. This is why only two new species from those 94 unidentified taxa are here described as new, one belonging to the genus *Brachysira* and the other to *Sellaphora*. Furthermore, one *Eolimna* species is transferred to *Sellaphora*, this species sharing the same morphology of areolae as the *Sellaphora* species here described as new.




Figs 2–34. Overview of the most abundant taxa (≥ 1% relative abundance in at least one sample). (2) *Cyclotella meneghiniana*; (3) *Eunotia* cf. *meridiana*; (4) *Eunotia* sp. 1; (5) *Eunotia* sp. 3; (6) *Eunotia* sp. 2; (7) *Eunotia minor*; (8) *Fragilaria pectinalis*; (9) *Fragilaria austriaca*; (10) *Fragilaria bidens*; (11) *Fragilaria tenera*; (12–13) *Achnanthidium* sp. 5; (14–15) *Achnanthidium* aff. *catenatum*; (16–17) *Achnanthidium* sp. 1; (18–19) *Achnanthidium minutissimum*; (20–21) *Achnanthidium* sp. 4; (22–23) *Planothidium rostratum*; (24–25) *Planothidium victori*; (26–27) *Planothidium incuriatum*; (28–29) *Planothidium cryptolanceolatum*; (30–31) *Cocconeis pediculus*; (32–33) *Cocconeis* sp. 2; (34) *Ulnaria ulna*. Scale bar 10 µm.




Figs 35–77. Overview of the most abundant taxa (≥ 1% relative abundance in at least one sample). (35) *Fistulifera saprophila*; (36) *Craticula subminuscula*; (37) *Craticula* sp. 2; (38) *Craticula molestiformis*; (39) *Craticula* cf. *pumilio*; (40) *Sellaphora cosmopolitana*; (41) *Sellaphora* sp. 3; (42) *Eolimna* sp. 1; (43) *Sellaphora nigri*; (44) *Sellaphora madida*; (45) *Sellaphora queretana*; (46) *Sellaphora atomoides*; (47) *Sellaphora saugerresii*; (48) *Sellaphora pupula*; (49) *Mayamaea permitis*; (50) *Reimeria sinuata*; (51) *Diadesmis confervacea*; (52) *Nupela wellneri*; (53) *Geissleria decussis*; (54) *Navicula veneta*; (55) *Navicula erifuga*; (56) *Navicula libonensis*; (57) *Navicula capitatoradiata*; (58) *Navicula symmetrica*; (59) *Navicula notha*; (60) *Navicula* cf. *cryptocephala*; (61) *Encyonopsis* cf. *thienemannii*; (62) *Navicula gregaria*; (63) *Navicula cryptocephala*; (64) *Navicula reichardtiana*; (65) *Brachysira altepetlensis*; (66) *Encyonema minutum*; (67) *Halamphora montana*; (68) *Amphora pediculus*; (69) *Navicula trivialis*; (70) *Navicula rostellata*; (71) *Frustulia crassinervia*; (72) *Encyonema brevicapitatum*; (73) *Encyonema minutiforme*; (74) *Encyonema* cf. *minutiforme*; (75) *Encyonema* cf. *hebridiforme*; (76) *Encyonema jemtlandicum*; (77) *Encyonema pergracile*. Scale bar 10 µm.




Figs 78–117. Overview of the most abundant taxa (≥ 1% relative abundance in at least one sample). (78) *Gomphonema exilissimum*; (79) *Gomphonema parvuliforme*; (80) *Gomphonema* cf. *parvuliforme*; (81) *Gomphonema parvulum*; (82) *Gomphonema lagenula*; (83) *Gomphonema* cf. *lagenula*; (84) *Gomphonema* aff. *sarcophagus*; (85) *Gomphonema* aff. *mariovense*; (86) *Gomphonema subclavatum*; (87) *Gomphonema stonei*; (88) *Gomphonema pumilum*; (89) *Gomphonema graciledictum*; (90) *Gomphonema naviculoides*; (91) *Gomphonema minusculum*; (92) *Gomphonema* sp. 4 ; (93) *Gomphonema* sp. 2; (94) *Gomphonema innocens*; (95) *Gomphonema* aff. *parvulius*; (96) *Nitzschia desertorum*; (97) *Nitzschia semirobusta*; (98) *Nitzschia inconspicua*; (99) *Nitzschia* sp. 1; (100) *Nitzschia supralitorea*; (101) *Nitzschia* cf. *hantzschiana*; (102) *Nitzschia fonticola*; (103) *Nitzschia perminuta*; (104) *Surirella angusta*; (105) *Nitzschia acicularis*; (106) *Nitzschia amphibia*; (107) *Nitzschia communis*; (108) *Nitzschia gracilis*; (109) *Nitzschia paleacea*; (110) *Nitzschia intermedia*; (111) *Nitzschia palea*; (112) *Nitzschia palea* var. *tenuirostris*; (113) *Nitzschia palea* var. *debilis*; (114) *Nitzschia balcanica*; (115) *Nitzschia linearis*; (116) *Epithemia sorex*; (117) *Epithemia adnata*. Scale bar 10 µm.

*Brachysira altepetlensis* D. Mora, R. Jahn et N. Abarca sp. nov. (Figs 118–132)

Holotype: B 40 0042006; Figure 121 represents the holotype.

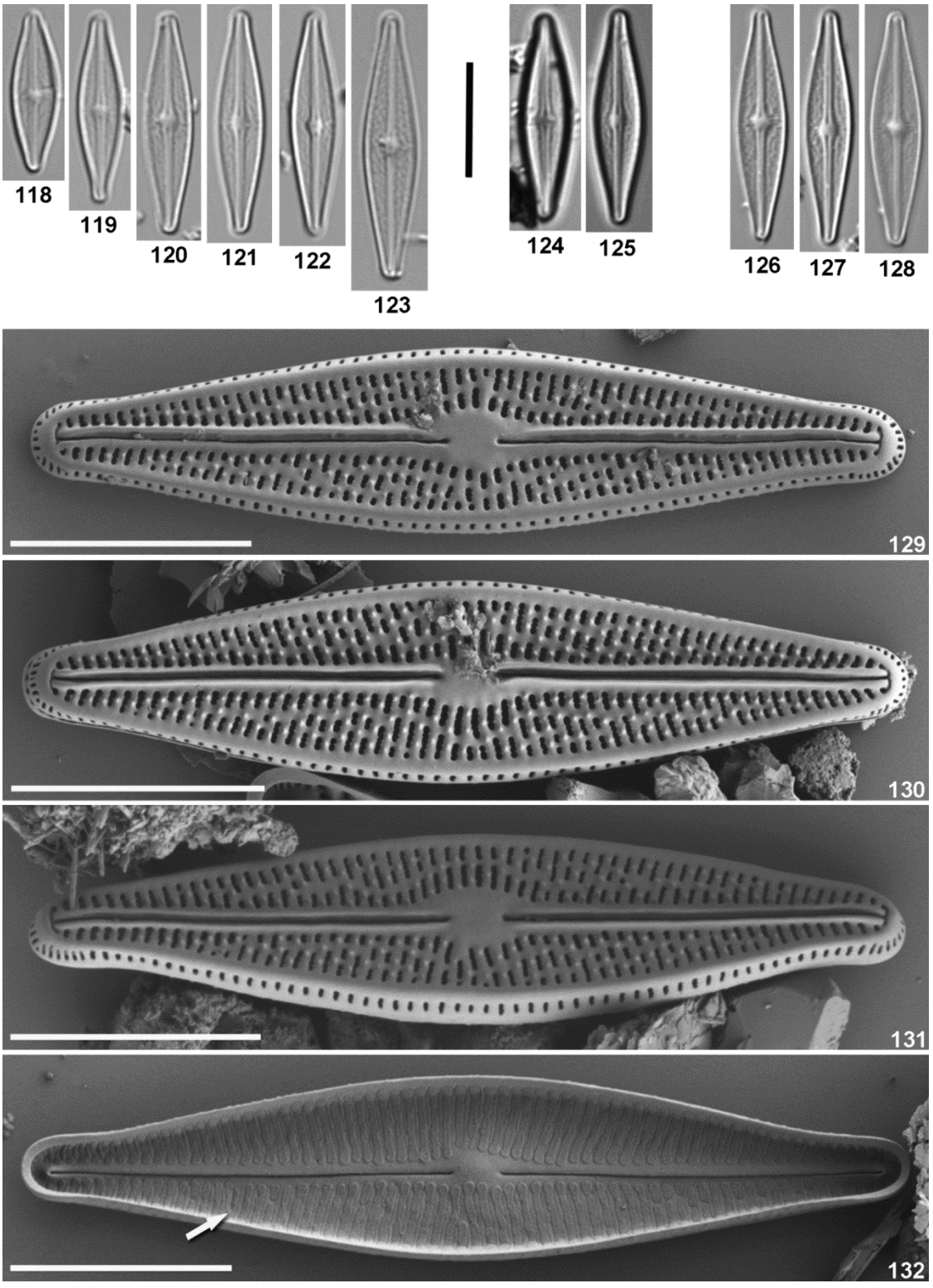
Isotypes: B 40 0042007 (SEM stub), QMEX DIAT0001 (Slide).

Cleaned unmounted material is available under the numbers B 40 0042008 and QMEX DIAT0002.

Type locality: Paredones stream, on the outskirts of Paredones village, Dolores Hidalgo, Guanajuato, Mexico (21°11'20.60"N; 101°06'53.40"W; 2089 m a.s.l). Collected by Demetrio Mora on 07.09.2014.

Registration: http://phycobank.org/100101

Description: the valves are lanceolate to linear–lanceolate with rostrate apices. The axial area is narrow–linear throughout the valve and the central area round to elliptical (Figs 118–128). Length: 12.6–23.1 µm, width: 3.2–4.5 µm, length/width ratio: 3.2–5.4; striae in 10 µm: 34–37. The raphe is filiform, slightly sinuous, bordered by a thickened longitudinal siliceous rib on both sides (Figs 129–131). The proximal raphe ends are straight, while the distal raphe endings are T-shaped (Figs 129–131). Internally, the proximal raphe endings are slightly bent to the same side of the valve and distally end in helictoglossa (Fig. 132). The striae are uniseriate and radiate throughout; composed of 2–3 transapically elongated areolae except close to the apices where only one elongated areola is present (Figs 129–131). Striae in the valve mantle are composed of single elongated areola (Fig. 131). In some valves the Voigt discontinuity can be seen (Fig. 132). Internally the areolae are occluded by hymens (Fig. 132). The virgae have irregularly spaced papillae (Figs 129–131).




Figs 118–132. *Brachysira altepetlensis* D. Mora, R. Jahn et N. Abarca sp. nov. LM (118–128) and SEM (129–132).118–123 type material, from Paredones stream, Guanajuato, Mexico, collected on 07.09.2014. 121 designated as holotype. 124–125 collected from type location but on 06.10.2013. 126–128 collected from type location but on 09.02.2014. 129–132 from type material: 129–130 external view of entire valves; 131 external view of an entire valve showing elongated areolae in the valve mantle; 132 internal view of entire valve, showing occlusion of the areolae by hymens. The arrow points at Voigt discontinuity. Scale bars 10 µm (118–128); 5 µm (129–132).

Differential diagnosis: *Brachysira procera* Lange-Bertalot et Gerd Moser is the species which most closely resembles *B. altepetlensis* in valve outline but is larger (25–60 µm), wider at valve center (4.5–6 µm) and has less striae in 10 µm (27–30) (Lange-Bertalot and Moser 1994). The valve outline of *Brachysira neglectissima* Lange-Bertalot also resembles that of *B. altepetlensis* but the valves of *B. neglectissima* are wider (4.3–4.5 µm), have more striae (36–40), the areolae are arranged in a way that they give the appearance of waves and each single areola is comparatively not as elongated as in *B. altepetlensis* (Lange-Bertalot and Moser 1994). *Brachysira guarrerai* Vouilloud, Sala et Núñez-Avellaneda is also similar in valve outline but the valves are wider (5.5–7 µm), have less striae (26–32) and lack papillae in the interstriae (Vouilloud et al. 2014).

The valve dimensions as well as the striae density of the new species fall within the range of the *Brachysira neoexilis* Lange-Bertalot species complex, but the type population of *B. neoexilis* has clear capitate apices and the larger specimens have a very slightly triundulate valve margins (Lange-Bertalot and Moser 1994). All the other populations from *B. neoexilis* species complex depicted in the original description (Lange-Bertalot and Moser 1994) have subcapitate to capitate apices, not matching at all the outline of *B. altepetlensis*. The specimens depicted in Rumrich et al. (2000), identified as *B. neoexilis* (Pl. 89: figs 18–20), closely resemble *B. altepetlensis* in valve outline but they clearly differ from specimens depicted in the type description of *B. neoexilis* (Lange-Bertalot and Moser 1994). The specimens of *Brachysira* found by Abarca-Mejía (2010) in a spring also in the Lerma–Chapala Basin, closely resemble *B. altepetlensis* in LM, but her identification was based on Rumrich et al. (2000), which led her to identify those valves as *B. neoexilis*.

Etymology: this new *Brachysira* species takes the name from the word “āltepētl” which means “water mountain” in Náhuatl language, that is how the surrounding mountains were used to be named by native people 500 years ago, at the time Spaniards first came to the region.

Distribution: apart from the type locality, this species was also found in four streams sampled for this study, namely Peña Colorada (site 4), San Martín (site 5), La Laborcilla 1 (site 7) and La Laborcilla 2 (site 8), all of these sites were characterized by low specific conductivity (≤ 100 µS/cm) and pH values going from acidic to slightly alcaline (5.1–7.9). But *B. altepetlensis* only reached high relative abundances (>10%) in acidic waters (pH= 5.1–5.8) with low specific conductivity (42–53 µS/cm).

*Sellaphora queretana* D. Mora, N. Abarca et J. Carmona sp. nov. (Figs 133–144)

Holotype: B 40 0042009; Figure 137 represents the holotype.

Isotypes: B 40 0042010 (SEM stub), QMEX DIAT0003 (Slide),.

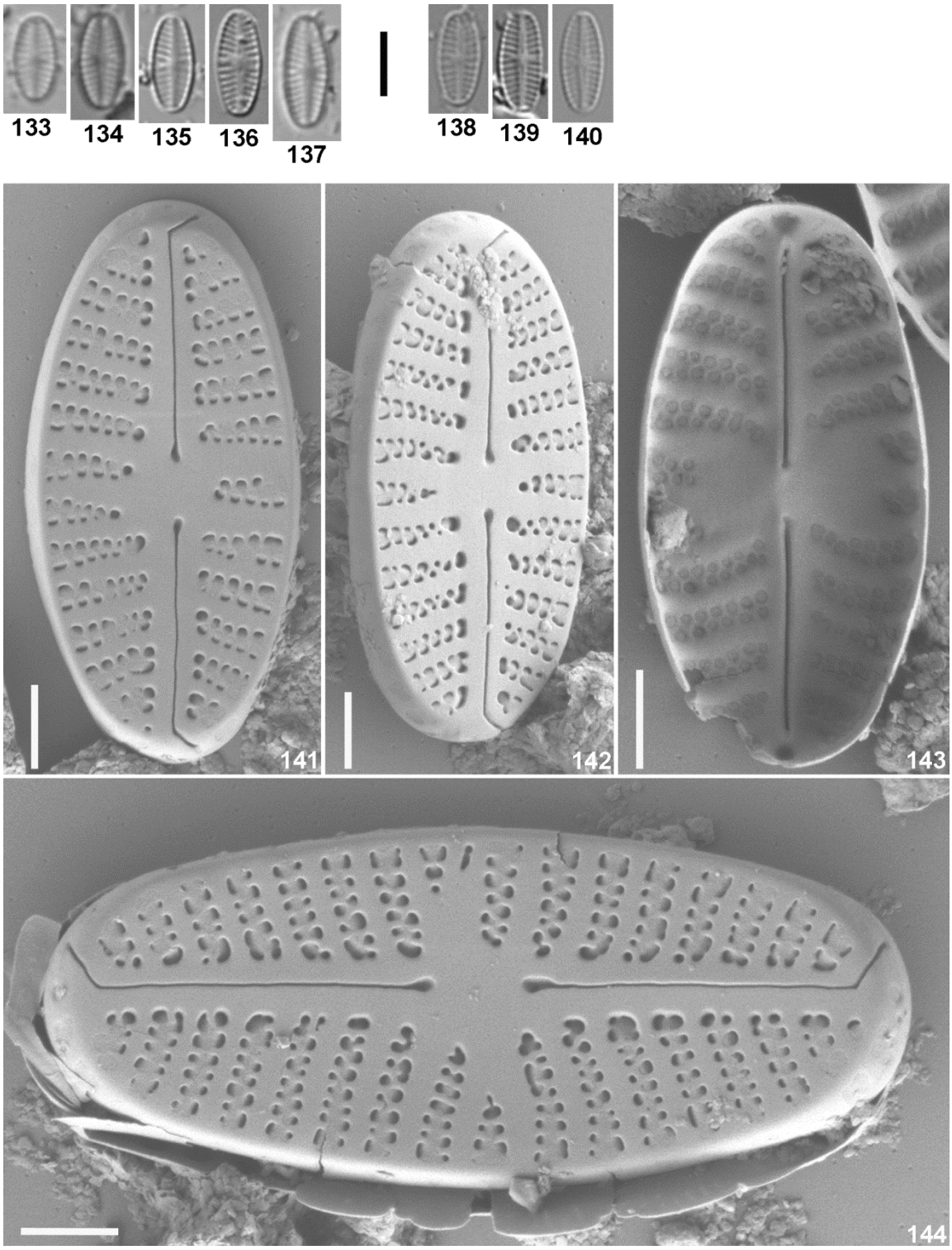
Cleaned unmounted material is available under the numbers B 40 0042011 and QMEX DIAT0004.

Type locality: stream Los Ailes 1, close to the town San Pedro, Huimilpan, Querétaro, Mexico (20°19'58.72"N; 100°15'17.09"W; 2358 m a.s.l). Collected by Demetrio Mora on 18.09.2013.

Registration: http://phycobank.org/100102

Description: the valves are linear–elliptical with broadly rounded apices (Figs 133–140). The axial area is narrow–linear throughout most of the valve, slightly widening close to the central area. The central area is asymmetrical due to irregular shortenings of the striae bordering it (Figs 141, 142 and 144). Length: 5.6–8.4 µm, width: 2.8–3.9 µm, length/width ratio: 1.9–2.4; striae in 10 µm: 19–22. The raphe is filiform with enlarged proximal raphe endings and slightly deflected to the same side of the valve; the distal raphe endings are strongly bent to the same side of the valve and extended onto the mantle (Figs 141, 142 and 144); the deflection of both proximal and distal raphe endings in external valve face is in the same direction (Figs 141, 142 and 144). Internally, the proximal raphe endings are straight and distally the raphe ends in helictoglossa (Fig. 143). The striae are biseriate and radiate throughout, however becoming uniseriate near the central area (Figs 141, 142 and 144). The areolae are lunate in form and are internally occluded by a hymen (Fig. 143). The hymenes are close to the foramina (seen on external view) (Figs 141, 142 and 144).

Differential diagnosis: there are no known taxa with the same combination of valve outline and areola type. The outline of *S. queretana* resembles that of *Sellaphora chistiakovae* (Kulikovskiy et Lange–Bertalot) C.E. Wetzel, Ector, Van de Vijver, Compère et D.G. Mann; the linear–elliptical forms of *Sellaphora crassulexigua* (E. Reichardt) C.E. Wetzel et Ector; and that of *Sellaphora nigri* (De Notaris) C.E. Wetzel et Ector. But *S. chistiakovae* has uniseriate to irregularly biseriate striae (Kulikovskiy et al. 2010); *S. crassulexigua* and *S. nigri* have uniseriate striae (Wetzel et al. 2015). Taxa with similar striae, with hymenes close to the foramina, include *Sellaphora labernardierei* Beauger, C.E.Wetzel et Ector, *Sellaphora rhombelliptica* (Gerd Moser, Lange–Bert. et Metzeltin) C.E. Wetzel et Ector, *Sellaphora rhombica* (Gerd Moser, Lange–Bert. et Metzeltin) D. Mora, N. Abarca et R. Jahn comb. nov. (see new combination below) and *Sellaphora thioense* (Gerd Moser, Lange–Bert. et Metzeltin) C.E. Wetzel, Ector, Van de Vijver, Compère et D.G. Mann. But the valves of *S. labernardieri* are linear to linear–elliptical, slightly inflated at the center and have consistently more striae 10 µm (20–28, mainly 24–25) (Beauger et al. 2016). *Sellaphora rhombelliptica* has more striae (25), which are uniseriate and the valves are rhomboelliptic (Moser et al. 1998). *Sellaphora rhombica* has similar number of striae (17–21) but the valve outline is rhombic to rhombic–lanceolate (Moser et al. 1998). *Sellaphora thioense* has slender elliptical valves (2.5–2.8) with higher striae density (27–28) (Moser et al. 1998).




Figs 133–144. *Sellaphora queretana* D. Mora, N. Abarca et J. Carmona sp. nov.
LM (133–140) and SEM (141–144). 133–137 type material, from stream Los Ailes 1, Querétaro, Mexico, collected on 18.09.2013. 137 designated as holotype. 138–140 population from stream Laguna de Servín 2, collected on 29.09.2013. 141–144 from type material: 141, 142 and 144 external views of entire valves; 143 internal view of an entire valve. Scale bars 5 µm (133–140); 1 µm (141–144).

Etymology: this new *Sellaphora* species takes its name from the demonym of the Mexican state Querétaro, from where it was collected.

Distribution: so far only known from the type locality (sampling site 11 in this study) and from stream Laguna de Servín 2 (site 13) located 4 km away from the type location, in acidic waters (pH 5.9–6.2) with low conductivity (77–88 µS/cm).

Based on morphological similarities with other small *Sellaphora* species, *Eolimna rhombica* Gerd Moser, Lange–Bertalot et Metzeltin is transferred to *Sellaphora*:

*Sellaphora rhombica* (Gerd Moser, Lange-Bertalot et Metzeltin) D. Mora, N. Abarca et R. Jahn comb. nov.

Basionym: *Eolimna rhombica* Gerd Moser, Lange–Bertalot et Metzeltin 1998, Bibliotheca Diatomologica, vol. 38, p. 156, pl. 23, figs 11–20.

Registration: http://phycobank.org/100103

### 2.4.2 Community analysis

The physical and chemical composition of the water from the sampling sites, as well as QBR values are enlisted in Table 2. From the original dataset of 14 environmental variables used in the DCA, total dissolved solids and total alkalinity were highly correlated with specific conductivity and therefore removed from the analysis. Dissolved oxygen and dissolved oxygen saturation percentage were also highly correlated, the latter being removed from further analysis. Dissolved inorganic nitrogen was also removed because it correlated strongly with nitrates. CCA with forward selection and unrestricted Monte Carlo permutations tests (999 permutations, $p<0.05$) identified temperature ($F=1.60$, $p=0.028$), pH ($F= 2.53$, $p=0.0010$), specific conductivity ($F= 5.07$, $p=0.0010$), soluble reactive phosphorous ($F=1.68$, $p=0.0060$) and the Riparian Forest Quality Index ($F=2.47$, $p=0.0010$) as the variables that significantly explained variation in the diatom data. The first two CCA axes accounted for 66.5 % of the cumulative variance of the species – environmental relation, both axes being significant ($p=0.0010$). The first CCA axis was strongly correlated with specific conductivity (inter–set correlation $r= 0.93$) and pH ($r= 0.80$). The second CCA axis was negatively correlated with QBR ($r= -0.61$) and positively correlated with temperature ($r= 0.44$).

On the CCA biplot three groups of samples were visualized (Fig. 145). The first group, situated at the bottom left part of the plot is composed of sites with the most acidic waters and lowest specific conductivity on average. The average number of species for this group was 16 (Table 3). This group was characterized by *Achnanthidium* sp. 1, the only taxon with a high indicator value (IndVal >50%).

Other indicator taxa (IndVal 20–50%) for this group were *Achnanthidium* aff. *catenatum* (J.Bílý et Marvan) Lange–Bertalot, *Brachysira altepetlensis*, *Eunotia* sp. 3, *Fragilaria austriaca* (Grunow) Lange-Bertalot, *Frustulia crassinervia* (Brébisson) Lange–Bertalot et Krammer and *Gomphonema exilissimum* (Grunow) Lange-Bertalot et E. Reichardt (Table 4).

The second group, found on the upper middle side of the plot contains samples with circumneutral waters, low in specific conductivity and the highest mean temperature and soluble reactive phosphorous concentrations. The mean number of species was 17 (Table 3). These sites were characterized by *Craticula molestiformis* (Hustedt) Mayama, *Encyonema minutum* (Hilse) D.G. Mann, *Mayamaea permitis* (Hustedt) Bruder et Medlin and *Nitzschia palea* var. *tenuirostris* Grunow, all these taxa with high and significant IndVals (>50%) (Table 4).

Samples from the third group correspond to well mineralized waters with the highest pH values on average, and also the lowest nitrogen concentrations. The sites in this group scored the higher values for the QBR on average. The mean species richness was 17 (Table 3). This group was characterized by *Cocconeis* sp. 2, *Navicula reichardtiana* Lange–Bertalot , *Nitzschia inconspicua* Grunow, *Planothidium victori* Novis, Braidwood et Kilroy, *Reimeria sinuata* (W. Gregory) Kociolek et Stoermer and *Sellaphora atomoides* (Grunow) C.E. Wetzel et Van de Vijver.

The three sampling campaigns of eight sites are within the same groups of the CCA plot (Fig. 145), pointing out to stability of the diatom communities: samples from sites 6, 8, 12 and 13 are within group 1; sites 4 and 9 within group 2; and sites 2 and 3 in group 3.

In contrast, in 5 sites there were changes of the samples among the three groups. For site 7, one sample from the rainy season is together with the sample from the dry season in group 2, whereas the other rainy season sample is in group 1. The three samples of sites 11 and 14 are one in each of the three different groups observed in the CCA plot (Fig. 145). Only in sites 1 and 5, both rainy season samples are together within the same group, whereas the samples of the dry season are located in a different group.

Table 2: Physical and chemical composition of the water from the sampling sites in the Lerma–Chapala Basin. Samples were taken in September/October 2013 for sampling campaign I, in February 2014 for the campaign II and in September 2014 for campaign III. T= temperature in °C; Cond= specific conductivity corrected at 25°C (µS/cm); TDS= total dissolved solids as particles per million (ppm); TA= total alkalinity mg/L of $CaCO_3$; *v*= water velocity (cm/s); DO= dissolved oxygen (mg/L); DOS= dissolved oxygen saturation percentage; SRP = soluble reactive phosphorous (mg/L); $NO_{2^-}$–N= nitrite nitrogen (mg/L); $NO_{3^-}$–N= nitrate nitrogen (mg/L); $NH_4^+$–N= ammonium nitrogen (mg/L); DIN= dissolved inorganic nitrogen (mg/L); QBR= Riparian Forest Quality Index.

| Sampling campaign | Site | T | pH | Cond | TDS | TA | *v* | DO | DOS | SRP | $NO_{2^-}$–N | $NO_{3^-}$–N | $NH_4^+$–N | DIN | QBR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I Rainy season | 1 | 14.5 | 6.7 | 114 | 45 | 30 | 29 | 6.5 | 84 | 1.09 | 0.005 | 0.010 | 0.005 | 0.02 | 85 |
| | 2 | 16.1 | 7.5 | 417 | 173 | 91 | 33 | 8.2 | 107 | 0.92 | 0.005 | 0.010 | 0.050 | 0.06 | 75 |
| | 3 | 17.8 | 7.7 | 422 | 182 | 93 | 39 | 7.2 | 98 | 0.59 | 0.004 | 0.015 | 0.025 | 0.04 | 75 |
| | 4 | 26.3 | 7.4 | 59 | 30 | 12 | 11 | 6.5 | 103 | 0.59 | 0.004 | 0.010 | 0.020 | 0.03 | 55 |
| | 5 | 25.8 | 7.1 | 100 | 51 | 13 | 24 | 6.8 | 105 | 0.57 | 0.003 | 0.010 | 0.005 | 0.02 | 70 |
| | 6 | 20.0 | 5.8 | 48 | 21 | 9 | 15 | 6.7 | 94 | 0.49 | 0.002 | 0.010 | 0.000 | 0.01 | 75 |
| | 7 | 23.5 | 6.1 | 84 | 41 | 23 | 32 | 7.2 | 108 | 0.68 | 0.003 | 0.020 | 0.000 | 0.02 | 50 |
| | 8 | 25.4 | 6.3 | 70 | 35 | 18 | 38 | 5.9 | 91 | 0.50 | 0.004 | 0.020 | 0.010 | 0.03 | 75 |
| | 9 | 23.2 | 7.2 | 134 | 65 | 38 | 22 | 9.7 | 146 | 0.67 | 0.009 | 1.200 | 0.055 | 1.26 | 30 |
| | 10 | 20.7 | 7.6 | 777 | 357 | 369 | 0 | 16.2 | 233 | 0.55 | 0.176 | 8.800 | 0.140 | 9.12 | 35 |
| | 11 | 15.9 | 6.2 | 88 | 36 | 20 | 24 | 7.1 | 96 | 0.30 | 0.010 | 1.250 | 0.065 | 1.33 | 75 |
| | 12 | 16.3 | 5.8 | 58 | 24 | 10 | 9 | 7.2 | 98 | 0.36 | 0.005 | 0.750 | 0.000 | 0.76 | 60 |
| | 13 | 18.5 | 5.9 | 77 | 34 | 9 | 37 | 7.0 | 99 | 0.84 | 0.015 | 0.050 | 0.060 | 0.12 | 70 |
| | 14 | 16.4 | 6.5 | 96 | 40 | 26 | 38 | 7.3 | 99 | 0.83 | 0.018 | 0.140 | 0.105 | 0.26 | 65 |

| Sampling campaign | Site | T | pH | Cond | TDS | TA | *v* | DO | DOS | SRP | $NO_{2-}$–N | $NO_{3-}$–N | $NH_4^+$–N | DIN | QBR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| II Dry season | 1 | 13.8 | 7.5 | 432 | 170 | 152 | 18 | 8.3 | 105 | 0.24 | 0.015 | 0.025 | 0.010 | 0.05 | 85 |
| | 2 | 17.4 | 7.5 | 878 | 375 | 168 | 24 | 8.5 | 115 | 0.30 | 0.015 | 0.030 | 0.000 | 0.04 | 75 |
| | 3 | 18.5 | 7.7 | 857 | 376 | 168 | 13 | 8.7 | 119 | 0.23 | 0.016 | 0.025 | 0.000 | 0.04 | 75 |
| | 4 | 18.5 | 7.2 | 61 | 27 | 14 | 25 | 8.3 | 114 | 0.25 | 0.015 | 0.020 | 0.000 | 0.04 | 65 |
| | 5 | 20.5 | 6.8 | 79 | 36 | 16 | 19 | 9.3 | 131 | 0.28 | 0.016 | 0.020 | 0.015 | 0.05 | 70 |
| | 6 | 17.4 | 5.8 | 42 | 18 | 7 | 16 | 7.9 | 106 | 0.29 | 0.014 | 0.020 | 0.010 | 0.04 | 65 |
| | 7 | 25.6 | 7.4 | 71 | 36 | 21 | 14 | 9.2 | 144 | 0.26 | 0.016 | 0.030 | 0.035 | 0.08 | 50 |
| | 8 | 22.7 | 5.5 | 53 | 25 | 14 | 17 | 8.3 | 123 | 0.28 | 0.017 | 0.030 | 0.000 | 0.05 | 65 |
| | 9 | 14.7 | 6.1 | 283 | 113 | 83 | 0 | 4.2 | 54 | 0.32 | 0.016 | 0.030 | 0.000 | 0.05 | 30 |
| | 10 | 18.8 | 7.4 | 969 | 427 | 461 | 0 | 9.7 | 134 | 0.83 | 0.018 | 0.030 | 0.015 | 0.06 | 35 |
| | 11 | 9.9 | 6.4 | 279 | 99 | 91 | 0 | 6.1 | 71 | 0.28 | 0.015 | 0.025 | 0.000 | 0.04 | 75 |
| | 12 | 13.8 | 5.8 | 94 | 37 | 17 | 0 | 6.2 | 81 | 0.26 | 0.016 | 0.030 | 0.025 | 0.07 | 60 |
| | 13 | 12.0 | 6.3 | 129 | 48 | 22 | 0 | 6.4 | 80 | 0.26 | 0.017 | 0.030 | 0.015 | 0.06 | 60 |
| | 14 | 19.5 | 6.9 | 172 | 77 | 65 | 23 | 6.9 | 99 | 0.24 | 0.014 | 0.020 | 0.010 | 0.04 | 75 |

| Sampling campaign | Site | T | pH | Cond | TDS | TA | *v* | DO | DOS | SRP | $NO_{2-}$–N | $NO_{3-}$–N | $NH_4^+$–N | DIN | QBR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| III Rainy season | 1 | 16.5 | 7.7 | 125 | 53 | 41 | 32 | 5.3 | 71 | 0.38 | 0.007 | 0.015 | 0.040 | 0.06 | 75 |
| | 2 | 16.0 | 7.2 | 306 | 127 | 66 | 43 | 6.1 | 80 | 0.29 | 0.025 | 0.040 | 0.145 | 0.21 | 65 |
| | 3 | 18.0 | 7.7 | 313 | 136 | 72 | 68 | 6.1 | 82 | 0.29 | 0.017 | 0.025 | 0.115 | 0.16 | 75 |
| | 4 | 23.9 | 6.8 | 40 | 20 | 10 | 31 | 5.2 | 79 | 0.29 | 0.006 | 0.010 | 0.010 | 0.03 | 65 |
| | 5 | 26.1 | 7.9 | 65 | 33 | 19 | 27 | 5.5 | 86 | 0.31 | 0.005 | 0.010 | 0.005 | 0.02 | 70 |
| | 6 | 17.1 | 5.1 | 42 | 18 | 4 | 80 | 5.9 | 79 | 0.36 | 0.006 | 0.010 | 0.010 | 0.03 | 75 |
| | 7 | 19.9 | 5.3 | 55 | 25 | 16 | 62 | 5.2 | 73 | 0.27 | 0.011 | 0.020 | 0.055 | 0.09 | 50 |
| | 8 | 22.1 | 5.5 | 48 | 22 | 12 | 51 | 5.1 | 75 | 0.13 | 0.008 | 0.020 | 0.030 | 0.06 | 65 |
| | 9 | 24.2 | 6.8 | 138 | 68 | 56 | 36 | 5.1 | 77 | 0.50 | 0.005 | 0.010 | 0.015 | 0.03 | 30 |
| | 10 | 20.8 | 7.1 | 850 | 391 | 430 | 0 | 4.7 | 67 | 0.47 | 0.051 | 0.165 | 0.040 | 0.26 | 35 |
| | 11 | 15.2 | 6.8 | 91 | 37 | 34 | 18 | 5.0 | 66 | 0.71 | 0.010 | 0.030 | 0.140 | 0.18 | 65 |
| | 12 | 15.7 | 5.4 | 54 | 22 | 8 | 32 | 5.6 | 75 | 0.43 | 0.005 | 0.005 | 0.020 | 0.03 | 60 |
| | 13 | 15.8 | 5.9 | 78 | 32 | 15 | 45 | 5.9 | 80 | 0.55 | 0.010 | 0.030 | 0.015 | 0.05 | 70 |
| | 14 | 17.6 | 6.5 | 99 | 42 | 32 | 50 | 5.3 | 73 | 0.46 | 0.008 | 0.010 | 0.020 | 0.04 | 65 |

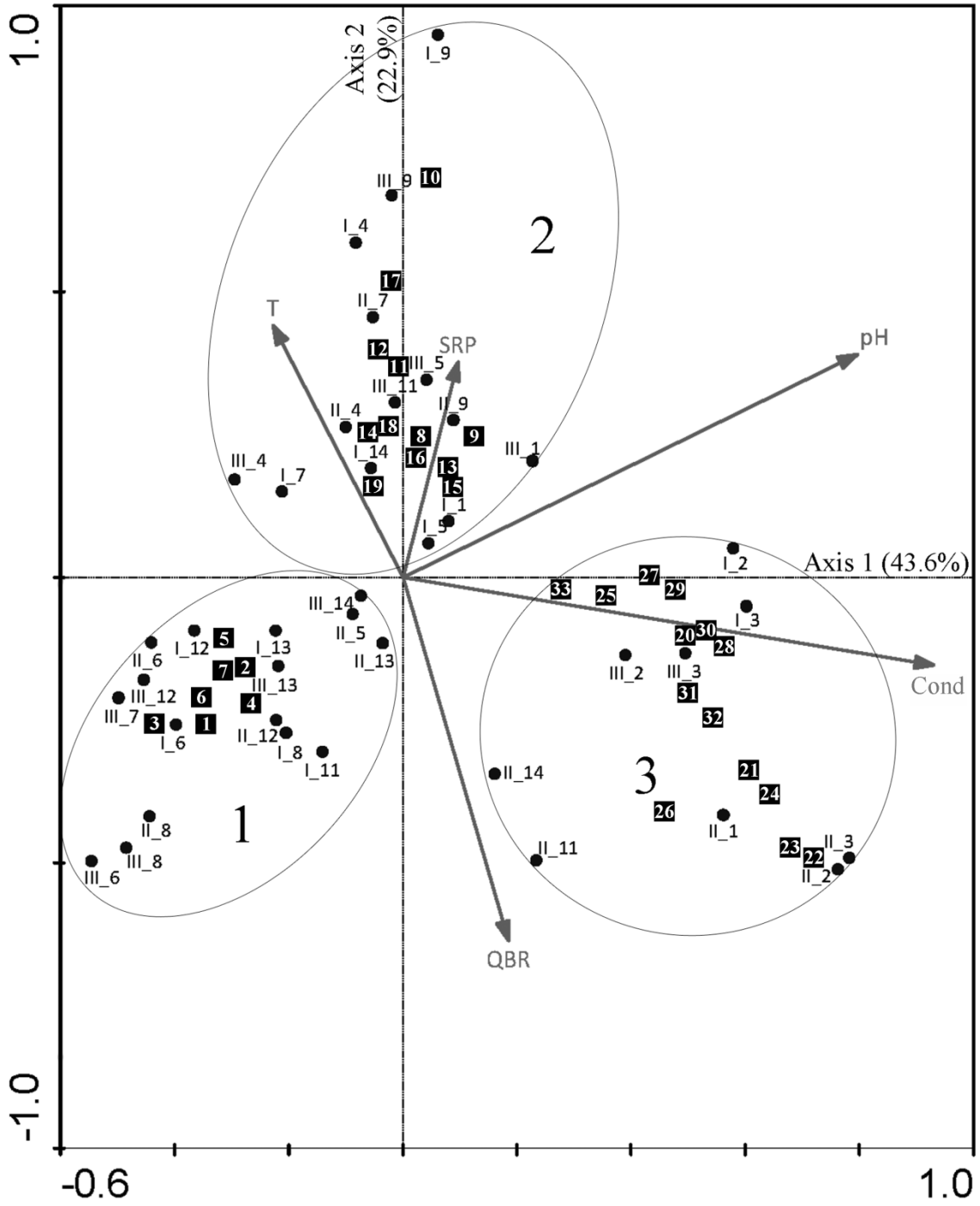



Fig. 145. Canonical Correspondence Analysis (CCA) ordination plot. Distribution of sampling sites based on diatom abundance data in relation to statistically significant environmental variables. Three groups of samples are depicted within ovals. For visualization purposes, only species with significant IndVals ($p$< 0.05) are included in the plot. Black squares correspond to species; numbers within the black squares refer to taxa names in Table 4. Sampling sites are codified as follows: a Roman numeral indicating the sampling campaign (I, II and III), followed by an underscore symbol and an Arabic numeral indicating the sampling site (sites 1 to 14). For abbreviations and units of the physical and chemical parameters refer to Table 2.

Table 3. Diversity indices and physical and chemical composition of the three groups visualized after the CCA. The mean value and standard deviation is provided for each variable. *S*= species richness; *H'*= Shannon-Wiener diversity index; *J'* = Pielou evenness index. For abbreviations and units of the physical and chemical variables refer to Table 2.

| | **Group 1** | **Group 2** | **Group 3** |
|---|---|---|---|
| *S* | 16±5 | 21±6 | 17±4 |
| *H'* | 2.43±0.33 | 2.75±0.40 | 2.53±0.30 |
| *J'* | 0.61±0.12 | 0.63±0.16 | 0.56±0.17 |
| T | 18.2 ± 3.5 | 21 ± 4.8 | 16 ± 2.9 |
| pH | 5.9 ± 0.5 | 7 ± 0.5 | 7.3 ± 0.5 |
| Cond | 70 ± 24 | 104 ± 59 | 453 ± 249 |
| TDS | 30 ± 9 | 47 ± 23 | 191 ± 110 |
| TA | 14 ± 7 | 30 ± 20 | 107 ± 43 |
| *v* | 31 ± 23 | 24 ± 11 | 29 ± 20 |
| DO | 6.6 ± 1.2 | 6.6 ± 1.7 | 7.3 ± 1.1 |
| DOS | 91 ± 18 | 95 ± 27 | 97 ± 16 |
| SRP | 0.38 ± 0.17 | 0.5 ± 0.25 | 0.37 ± 0.23 |
| $NO_{2-}$–N | 0.010 ± 0.005 | 0.008 ± 0.005 | 0.014 ± 0.006 |
| $NO_{3-}$–N | 0.14 ± 0.35 | 0.11 ± 0.32 | 0.02 ± 0.01 |
| $NH_4^+$–N | 0.022 ± 0.021 | 0.031 ± 0.043 | 0.039 ± 0.054 |
| DIN | 0.18 ± 0.35 | 0.15 ± 0.33 | 0.08 ± 0.06 |
| QBR | 66 ± 7 | 58 ± 18 | 75 ± 5 |

Table 4. Indicator taxa from the three groups visualized after the CCA. The indicator value of the taxa is accompanied by their relative abundance (RA) and relative frequency (RF) values. Significant IndVals ($p$< 0.05) are indicated in bold.

| | Taxa | Group 1 | | | Group 2 | | | Group 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **RA** | **RF** | **IndVal** | **RA** | **RF** | **IndVal** | **RA** | **RF** | **IndVal** |
| 1 | *Achnanthidium* sp. 1 | 99 | 69 | **68** | 0 | 7 | 2 | 1 | 11 | 0 |
| 2 | *Achnanthidium* aff. *catenatum* | 77 | 63 | **48** | 22 | 29 | 6 | 1 | 11 | 0 |
| 3 | *Brachysira altepetlensis* | 96 | 44 | **42** | 4 | 36 | 1 | 0 | 0 | 0 |
| 4 | *Eunotia* sp. 3 | 99 | 31 | **31** | 1 | 7 | 0 | 0 | 0 | 0 |
| 5 | *Fragilaria austriaca* | 71 | 56 | **40** | 29 | 7 | 2 | 0 | 0 | 0 |
| 6 | *Frustulia crassinervia* | 100 | 31 | **31** | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | *Gomphonema exilissimum* | 64 | 75 | **47** | 31 | 43 | 13 | 5 | 22 | 1 |
| 8 | *Craticula molestiformis* | 16 | 31 | 5 | 74 | 79 | **58** | 10 | 33 | 3 |
| 9 | *Craticula subminuscula* | 3 | 6 | 0 | 84 | 57 | **48** | 13 | 56 | 7 |
| 10 | *Cyclotella meneghiniana* | 0 | 0 | 0 | 100 | 21 | **21** | 0 | 0 | 0 |
| 11 | *Encyonema minutum* | 11 | 13 | 1 | 85 | 64 | **54** | 4 | 11 | 0 |
| 12 | *Eolimna* sp. 1 | 12 | 13 | 2 | 88 | 36 | **31** | 0 | 0 | 0 |
| 13 | *Fistulifera saprophila* | 4 | 6 | 0 | 79 | 57 | **45** | 17 | 22 | 4 |
| 14 | *Gomphonema* aff. *sarcophagus* | 3 | 13 | 0 | 96 | 43 | **41** | 1 | 11 | 0 |
| 15 | *Mayamaea permitis* | 9 | 31 | 3 | 69 | 86 | **59** | 22 | 67 | 15 |
| 16 | *Navicula rostellata* | 2 | 6 | 0 | 87 | 57 | **50** | 11 | 11 | 1 |
| 17 | *Nitzschia gracilis* | 0 | 0 | 0 | 100 | 29 | **29** | 0 | 0 | 0 |
| 18 | *Nitzschia palea* var. *debilis* | 8 | 25 | 2 | 91 | 50 | **45** | 1 | 11 | 0 |
| 19 | *Nitzschia palea* var. *tenuirostris* | 9 | 38 | 4 | 91 | 64 | **58** | 0 | 0 | 0 |
| 20 | *Amphora pediculus* | 0 | 0 | 0 | 3 | 7 | 0 | 97 | 44 | **43** |
| 21 | *Cocconeis* sp. 2 | 0 | 0 | 0 | 2 | 14 | 0 | 98 | 67 | **66** |
| 22 | *Cocconeis pediculus* | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 22 | **22** |
| 23 | *Epithemia adnata* | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 33 | **33** |
| 24 | *Epithemia sorex* | 0 | 0 | 0 | 4 | 7 | 0 | 96 | 44 | **43** |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 25 | *Gomphonema pumilum* | 0 | 6 | 0 | 28 | 21 | 6 | 72 | 67 | **48** |
| 26 | *Gomphonema minusculum* | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 33 | **33** |
| 27 | *Halamphora montana* | 0 | 0 | 0 | 16 | 14 | 2 | 84 | 56 | **46** |
| 28 | *Navicula reichardtiana* | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 56 | **56** |
| 29 | *Navicula gregaria* | 0 | 0 | 0 | 13 | 14 | 2 | 87 | 56 | **48** |
| 30 | *Nitzschia inconspicua* | 1 | 13 | 0 | 0 | 0 | 0 | 99 | 56 | **55** |
| 31 | *Planothidium victori* | 0 | 6 | 0 | 12 | 29 | 3 | 88 | 78 | **69** |
| 32 | *Reimeria sinuata* | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 67 | **67** |
| 33 | *Sellaphora atomoides* | 10 | 31 | 3 | 20 | 29 | 6 | 70 | 78 | **54** |

## 2.5 Discussion

### 2.5.1 Species composition and taxonomy

The species richness found, 274 taxa, was relatively high compared to previous studies on the basin: 209 taxa were found by Abarca-Mejía (2010) from 59 samples analyzed from three substrates; 178 taxa by Segura-García (2012) from 66 epilithon samples analyzed; 173 taxa by Mora et al. (2015) from 12 epilithon samples; and 70 taxa by Segura-García (2016) from 16 epilithon samples. This kind of comparison is difficult to make since it depends on the number of samples analyzed, the timing of the samplings, the physical and chemical composition of the waters, the number of substrates sampled and the taxonomic effort with which the diatom valves were analyzed (Morales et al. 2001, Veselá and Johansen 2009). Nevertheless, our results on taxa diversity are higher than the four previous studies conducted in the basin.

The resulting high diversity found in our study can be explained by the detail at which samples were analyzed under both LM and SEM, which resulted in the separation of several morphodemes instead of lumping them into species complexes. The fact that a third of the flora, 94 morphodemes, could not be assigned to described species is not surprising due to the nature of the samples, coming from within the tropics, for which no extensive identification floras have been produced yet, compared to northern temperate regions. Furthermore, it is encouraging to have such a big number of unidentified morphodemes, because they could be helpful in the quest of unravelling if the

freshwater diatom floras of Mexico have certain biogeographical affinities, as it would be expected due to the fact that the country lies within the so called Mexican Transition Zone, a complex area in which Neotropical and Nearctic biotic elements converge (Huidobro et al. 2006). This task could be facilitated by coupling detailed morphological examinations with molecular tools (Trobajo et al. 2009, Abarca et al. 2014, Zimmermann et al. 2014).

In most of the freshwater diatom floras generated for Mexico, there seems to be a high intrinsic cosmopolitism, with a large proportion of taxa from north temperate waters. Nowadays it seems unlikely to find large amounts of shared species with north temperate regions due to mounting evidence that even microorganisms like diatoms have biogeography (Kociolek and Spaulding 2000, Vanormelingen et al. 2008, Abarca et al. 2014). This raises the question of identification literature and the detail with which samples are analyzed, such as force–fitting identifications to north temperate taxa and lumping into broad species complexes due to limited high resolution microscopy tools.

On the other hand, finding a large proportion of cosmopolitan taxa should not be that surprising since isolated areas such the Andes have shown to have as much as 42% cosmopolitan taxa, but also a considerable proportion of newly described taxa (9.5%) plus seemingly endemic regionals (Lange-Bertalot 2007). So far these 9.5% of newly described taxa have not been the case for the flora from the Lerma–Chapala Basin or even Central Mexico, for which no species from streams have been described as new in the last 25 years. Within the basin, the number of unidentified taxa, potentially containing undescribed species varies from 6% in Segura-García (2012), to 19% in Abarca-Mejía (2010) and 22% in Mora et al. (2015) but as those authors pointed out, further examinations on some of those taxa are needed to determine if they really should be described as new species.

Another hypothesis that could explain the high species richness found in our study is the heterogeneity of environmental conditions of the study areas: a) the sampling campaigns were done in both rainy and dry seasons; b) varied geomorphologies of the streams from headwaters to the midlands and also from the plains, resulting in different riparian communities, reflected in the QBR index values obtained; c) streams ranging from perennial to temporary; d) heterogeneity of physical and chemical composition of the water. Environmental heterogeneity of habitats has been proposed in other studies as a determinant of species richness and distribution (Petrov and Nevrova, 2014).

An additional indicator of the heterogeneity of the studied sites is the fact that no single taxon was found in all samples, which contrast with previous findings on the Lerma–Chapala Basin, where the following taxa were found in all sites and seasons *Craticula subminuscula*, *Gomphonema parvulum*,

*Navicula veneta*, *Nitzschia amphibia*, *N. capitellata*, *N. palea* and *Sellaphora pupula* (Segura-García 2012, Mora et al. 2015).

When looking at the macroalgae of the studied streams, it is worth mentioning that sampling sites 11–14 host red algae like *Batrachospermum gelatinosum* (Linnaeus) De Candolle, *Paralemanea mexicana* (Kützing) Vis et Sheath and *Sirodotia suecica* Kylin, species typically found in headwater mountain streams of temperate regions (Bojorge–García et al. 2010). On the other hand, in sites 1–10 species rather associated to warmer waters were found, such as *Cladophora mexicana* P. Crouan et H. Crouan. This is another indicator of the heterogeneity of the sampling sites.

### 2.5.2 Diatom communities

The different diatom compositions found in the Lerma–Chapala Basin were mainly driven by specific conductivity and pH. Temperature, soluble reactive phosphorous and the Riparian Forest Quality Index were statistically significant but when analyzing the mean values and their standard deviations, the border between each group was not distinct.

For both specific conductivity and pH, the lowest values were recorded in the streams located in the headwaters, which is logical since water there has not gone deep into the geological matrix and therefore is not well mineralized. On the other hand, the higher values for both specific conductivity and pH were recorded on the midland and plains, where the streams received more contributions of well mineralized waters, for example from springs. There is no better example of this than what was recorded at sampling site 10, where pH values were high and specific conductivity values were the highest recorded for this study. This phenomenon is shown by Mahlknecht et al. (2004) in an aquifer recharge model for the same area where sampling sites 1–10 from our study are located. In the model, rain water normally has a pH of 5 but as water goes through the geological matrix it can reach pH values of up to 9 through several mineral dissolution processes and cation exchange, before it appears again at the surface i.e. springs.

No clear seasonal effect (rainy and dry seasons) was observed on the three groups of sampling sites observed after the CCA because in every group there are samples from both rainy seasons together with the dry season. Even though there were seasonal variations in physical and chemical factors such as specific conductivity, pH and water velocity, the community composition (species richness and abundance) apparently did not respond to those seasonal fluctuations (Rothfritz et al. 1997, Bojorge-García et al. 2014). This is well exemplified by the fact that the three samples of eight out of 13 sites included in the CCA remained within the same group during the three sampling periods, showing an overall stability of the diatom communities. This stability can be attributed to the fact

that seasonal changes, e.g. in water velocity, discharge and chemical variables do not have long term effects so communities revert to their pre-disturbance state after the disappearance of the perturbation (Connell and Sousa 1983, Soininen and Eloranta 2004). On the other hand, perturbations such as mine tailings spills can have long lasting effects on diatom communities due to heavy metal pollution (Sabater 2000). The time it takes for communities to revert to a pre-disturbance state will largely depend on life span, reproduction and recolonization rates of the organisms as well as on the magnitude of the perturbation (Townsend et al. 1997, Soininen and Eloranta 2004). In order to relate seasonal changes in the community structure to fluctuations in environmental conditions, the timing and scale at which samplings should be made has to be proportional to the life span of the organism in question and cover a complete turnover of all individuals or longer (Soininen and Eloranta 2004). Since diatoms have short life cycles, high reproduction rates and recolonization rates that are within weeks (Round 1991, Licursi and Gómez 2009, Lowe 2011), it should be necessary to conduct intensive samplings to demonstrate dependency of changes in community structure due to fluctuations in environmental factors. This could be a reason why we observed an overall stability of the diatom communities. On the other hand, there were changes in the samples from 5 sampling sites, which can be attributed to the timing, since at the time of sampling the community composition was representing the changes due to seasonal fluctuations and not in an overall stable state after reverting from a perturbation (e.g. major flood, drought).

Regarding the characteristic species of the three groups visualized from the CCA, there are several similarities with previous reports on the ecological preferences of these taxa. Some species were found in all three groups but with varying relative abundances, so only those with the largest abundances were taken as the representative for a group.

For group 1, species from genera such as *Brachysira, Eunotia* and *Frustulia* are well regarded as characteristic from acidic and electrolyte poor waters (van Dam et al. 1994; Wolfe and Kling 2001; Hofmann et al. 2013; Vouilloud et al. 2014), which fits well to the chemical composition of the waters from the sites of this group. *Fragilaria austriaca*, *Frustulia crassinervia* and *Gomphonema exilissimum* are also regarded as indicators of low nutrients (van Dam et al. 1994). It is interesting to notice the presence of three taxa with uncertain identity, namely *Achnanthidium* aff. *catenatum*, *Achnanthidium* sp. 1 and *Eunotia* sp. 3, characteristic taxa of this group, which hints at the possibility to regard them as characteristic of acidic, and electrolyte and nutrient poor waters. But before their taxonomic position is confirmed, no comparisons about ecological preferences can be made.

The representative species from group 2 were taxa well regarded as indicators of circumneutral and eutrophic waters with varying degrees of perturbation such as *Craticula molestiformis*, *Mayamaea*

*permitis* and *N. palea* var. *tenuirostris* (van Dam et al. 1994; Besse-Lototskaya et al. 2011; Hofmann et al. 2013). Other representatives of the beforehand conditions include *Craticula subminuscula* (Manguin) C.E. Wetzel et Ector, *Cyclotella meneghiniana* Kützing, *Fistulifera saprophila* (Lange-Bertalot et Bonik) Lange-Bertalot and *Navicula rostellata* Kützing (van Dam et al. 1994; Besse-Lototskaya et al. 2011; Hofmann et al. 2013). The exception for group 2 is *Encyonema minutum*, normally reported from oligo–mesotrophic waters, but the precise ecological preference of this taxon is difficult to tell since it has been long confounded with *Encyonema silesiacum* (Bleisch) D.G. Mann (Hofmann et al. 2013). In the sampling sites belonging to this group, the highest average phosphorous concentrations were recorded. Regarding the degree of perturbation, the QBR values for these sites scored the lowest values on average, which were related to human perturbation on the riparian forest. Some of these sites are in fact close to diffuse pollution sources such as cattle grazing and agriculture.

Regarding group 3, its characteristic species also confirm the meso-eutrophic, mineralized and alkaliphilous nature of its waters, with taxa such a *Cocconeis* sp. 2 (*C. placentula* Ehrenberg *sensu lato* based only on LM observations), *Navicula reichardtiana*, *Nitzschia inconspicua*, *Planothidium victori* (formerly within *Planothidium frequentissimum* (Lange-Bertalot) Lange-Bertalot *sensu lato*), *Reimeria sinuata* and *Sellaphora atomoides* (former *Eolimna tantula* (Hustedt) Lange-Bertalot) (van Dam et al. 1994; Lange-Bertalot 2001). Other taxa characteristic of this conditions include *Amphora pediculus* (Kützing) Grunow, *Epithemia adnata* (Kützing) Brébisson, *Epithemia sorex* Kützing, *Gomphonema pumilum* (Grunow) E. Reichardt et Lange-Bertalot, *Halamphora montana* (Krasske) Levkov and *Navicula gregaria* Donkin (van Dam et al. 1994; Lange-Bertalot 2001). When looking at the average dissolved inorganic nitrogen from the group, the lowest of all three groups, it is hard to explain that it is based on the seasonal inputs from the surrounding environment. But when looking at the algae present on the water, it is worth mentioning that on all of the sites from this group *Nostoc* spp. was found, in some cases blooming. The presence of these nitrogen-fixing cyanobacteria is regarded as an indicator of poor nitrogen concentrations since these algae can thrive under this condition by actively fixating atmospheric nitrogen (Grimm and Petrone 1997).

## 2.6 Conclusion

This work contributed to increase the knowledge of the diatom flora from the Lerma–Chapala Basin, Central Mexico, providing a diversity baseline and evidence of its distinctiveness from the floras of other areas in Mexico, with a large proportion of unidentified taxa to be described as new. The studied diatom communities are subjected to moderate environmental disturbance, representing a transition between warm and cold waters, with ionic composition, temperature and the quality of the riparian forest being the main factors defining the community composition observed. The next

approach to investigate the diatom diversity of the region would be by means of environmental DNA metabarcoding in combination with the development of a taxonomic reference database, in order to highlight the complementary aspect of classical taxonomy and eDNA metabarcoding, i.e. the importance of the reciprocal illumination (Visco et al. 2015; Zimmermann et al. 2015).

## 2.7 Acknowledgements

The work of DM was funded by the Mexican Government through doctoral grants from CONACYT, CONCYTEQ and DGRI–SEP. Samples were taken under permit CONAPESCA PPF/DGOPA–149/14. The Deutsche Forschungsgemeinschaft is thanked for Grant INST 130/839–1 FUGG concerning SEM funding. We gratefully acknowledge Carlos E. Wetzel who provided advice on the taxonomy of small *Sellaphora* species. Loren Bahls kindly adviced on the identification of some *Stauroneis* species. Wolf–Henning Kusber is thanked for his advice on nomenclature. Monika Lüchow and Kim Govers kindly assisted at the SEM. We thank Verónica Aguilar Zamora for creating the map. Martin Jagodzinski kindly advised on figure plate preparation. Mahinda Martínez is thanked for her assistance with vascular plant identifications.

## 2.8 References

Abarca-Mejía NC 2010 Diatom community analysis and quality assessment of the polluted tropical Lerma River (Mexico). Dissertation, Freie Universität Berlin, Germany.

Abarca N, Jahn R, Zimmermann J, Enke N (2014) Does the cosmopolitan diatom *Gomphonema parvulum* (Kützing) Kützing have a biogeography? PLoS ONE, 9, e86885. https://doi.org/10.1371/journal.pone.0086885.

Allan JD, Castillo MM (2007) Stream ecology: structure and functioning of running waters. Second edition. Springer, Dordrecht, 1–436.

Aparicio J (2001) Hydrology of the Lerma–Chapala watershed. In: Hansen AM, van Afferden M (Eds) The Lerma–Chapala watershed: evaluation and management. Kluwer Academic/PlenumPublishers, New York, 3–30.

APHA (American Public Health Association) (2005) Standard Methods for Examination of Water and Wastewater. 21st Ed. Port City Press, Washington D.C. 1–1200.

Arriaga L., Espinoza JM, Aguilar C, Martínez E, Gómez L, Loa E (2000) Regiones terrestres prioritarias de México. Comisión Nacional para el Conocimiento y uso de la Biodiversidad. México D.F. http://www.conabio.gob.mx/conocimiento/regionalizacion/doctos/rtp_099.pdf

Beauger A, Wetzel CE, Voldoire O, Garreau A, Ector L (2016) *Sellaphora labernardierei* (Sellaphoraceae, Bacillariophyta), a new epilithic species from French spring and four new combinations within the genus *Sellaphora*. Phytotaxa 260(3): 235–246. https://doi.org/10.11646/phytotaxa.260.3.3.

Bellinger EG, Sigee DC (2010) Freshwater algae and their use as bioindicators. John Wiley & Sons, Chichester, 1–271. https://doi.org/10.1002/9780470689554.

Besse-Lototskaya A, Verdonschot PFM, Coste M, van de Vijver B (2011) Evaluation of European diatom trophic indices. Ecological Indicators 11: 456–467. https://doi.org/10.1016/j.ecolind.2010.06.017.

Bojorge-García MG, Cantoral-Uriza EA (2007) Estructura comunitaria de diatomeas asociadas a talos de *Prasiola mexicana* (Chlorophyta) en el río Magdalena, D.F. Hidrobiológica 17(1): 11–24. http://hidrobiologica.izt.uam.mx/index.php/revHidro/article/view/948

Bojorge-García M, Carmona J, Beltrán Y, Cartajena M (2010) Temporal and spatial distribution of macroalgal communities of mountain streams in Valle de Bravo Basin, central Mexico. Hydrobiologia 641: 159–169. https://doi.org/10.1007/s10750-009-0074-5.

Bojorge-García M, Carmona J, Ramírez R (2014) Species richness and diversity of benthic diatom communities in tropical mountain streams of Mexico. Inland Waters 4: 279–292. https://doi.org/10.5268/IW-4.3.568.

Calderón de Rzedowski G, Rzedowski J (2001) Flora fanerogámica del Valle de México. Instituto de Ecología A.C., Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Pátzcuaro, 1–1406.

Cantoral-Uriza EA, Carmona-Jiménez J, Montejano G (1997) Diatoms of calcareous tropical springs in the central region of Mexico. Cryptogamie Algologie 18(1): 19–46.

Carmona-Jiménez J, Ramírez–Rodríguez R, Bojorge-García MG, González–Hidalgo B, Cantoral-Uriza EA (2016) Estudio del valor indicador de las comunidades de algas bentónicas: una propuesta de evalución y aplicación en el Río Magdalena, Ciudad de México. Revista Internacional de Contaminación Ambiental 32(2): 139–152. https://doi.org/10.20937/RICA.2016.32.02.01.

Carranza-González E (1995) Salicaceae. Flora del Bajío y de Regiones Adyacentes. Instituto de Ecología A.C., Pátzcuaro, Fascicle 37, 1–21. http://www1.inecol.edu.mx/publicaciones/resumeness/FLOBA/Flora%2037.pdf

Carranza-González E, Madrigal–Sánchez X (1995) Betulaceae. Flora del Bajío y de Regiones Adyacentes. Instituto de Ecología A.C., Pátzcuaro, Fascicle 39, 1–22. http://www1.inecol.edu.mx/publicaciones/resumeness/FLOBA/Flora%2039.pdf

CEPFa Biodiversity Hotspots, Mesoamerica. http://www.cepf.net/resources/hotspots/North-and-Central-America/Pages/Mesoamerica.aspx [accessed 14.06.2017]

CEPFb Biodiversity Hotspots, Madrean Pine-Oak Woodlands. http://www.cepf.net/resources/hotspots/North-and-Central-America/Pages/Madrean-Pine-Oak-Woodlands.aspx [accessed 14.06.2017]

CNA (Comisión Nacional del Agua, México) (2017a) Estación 00011008 Cañada de González. Normales Climatológicas, estado de Guanajuato. Servicio Meteorológico Nacional. http://smn1.conagua.gob.mx/index.php?option=com_content&view=article&id=181&tmpl=component [accessed 23.01.2017]

CNA (Comisión Nacional del Agua, México) (2017b) Estación 00011093 San Miguel de Allende. Normales Climatológicas, estado de Guanajuato. Servicio Meteorológico Nacional. http://smn1.conagua.gob.mx/index.php?option=com_content&view=article&id=181&tmpl=component [accessed 23.01.2017]

CNA (Comisión Nacional del Agua, México) (2017c) Estación 00022029 Huimilpan. Normales Climatológicas, estado de Querétaro. Servicio Meteorológico Nacional. http://smn1.conagua.gob.mx/index.php?option=com_content&view=article&id=177&tmpl=component [accessed 23.01.2017]

Connell JH, Sousa WP (1983) On the evidence needed to judge ecological stability or persistence. The American Naturalist 121(6): 789–824.

Cotler H, Mazari M, de Anda J (2006) Atlas de la cuenca Lerma–Chapala, construyendo una visión conjunta. Instituto Nacional de Ecología, Secretaría de Medio Ambiente y Recursos Naturales/Instituto de Ecología, Universidad Nacional Autónoma de México. México D.F. 1–196.

Dufrêne M, Legendre P (1997) Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs 67(3): 345–366. https://doi.org/10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2.

Giller PS, Malmqvist B (1998) The biology of streams and rivers. Oxford University Press, 1–296.

Grimm NB, Petrone KC (1997) Nitrogen fixation in a desert stream ecosystem. Biogeochemistry 37(1): 33–61. https://doi.org/10.1023/A:1005798410819.

HACH (2003) Water Analysis Handbook. 4th Ed. Hawking JH, Smith LM, LeBusque K, Davey C (Eds). Hach Co., Loveland, Colorado. 1–1260.

Hofmann G, Lange-Bertalot H, Werum M (2013) Diatomeen im Süßwasser–Benthos von Mitteleuropa. Bestimmungsflora Kieselalgen für die ökologische Praxis. Über 700 der häufigsten Arten und ihre Ökologie. Koeltz Scientific Books, Königstein, 1–908.

Huidobro L, Morrone JJ, Villalobos JL, Álvarez F (2006) Distributional patterns of freshwater taxa (fishes, crustaceans and plants) from the Mexican Transition Zone. Journal of Biogeography 33(4): 731–741. https://doi.org/10.1111/j.1365-2699.2005.01400.x.

Kelly MG (1998) Use of community–based indices to monitor eutrophication in European rivers. Environmental Conservation 25(1): 22–29. https://doi.org/10.1017/S037689299800006X.

Kociolek JP, Spaulding SA (2000) Freshwater diatom biogeography. Nova Hedwigia 71: 223–241.

Kulikovskiy MS, Lange-Bertalot H, Witkowski A (2010) Diatom flora of Polistovo–Lowatski Sphagnum tract (European Russia). I. *Eolimna chistiakovae* sp. nov. and further transfers to the genus *Eolimna* Lange-Bertalot & Schiller. Diatom Research 25(1): 77–85. https://doi.org/10.1080/0269249X.2010.9705830.

Lange-Bertalot H (2001) Diatoms of the European inland waters and comparable habitats. *Navicula* sensu stricto. 10 genera separated from *Navicula* sensu lato. *Frustulia*. Diatoms of Europe 2. A.R.G. Gantner Verlag K.G., Ruggell, 1–526.

Lange-Bertalot H (2007) Kosmopoliten versus Endemiten: Biogeographische Implikationen steigender Taxazahlen. In: Kusber W–H, Jahn R (Eds) Proceedings of the 1st Central European Diatom Meeting, Berlin. Botanic Garden and Botanical Musem Berlin–Dahlem, Freie Universität Berlin, 107–112. https://doi.org/10.3372/cediatom.122.

Lange-Bertalot H, Moser G (1994) *Brachysira*. Monographie der Gattung. Wichtige Indikator–Species für das Gewässer Monitoring und *Naviculadicta* nov. gen. Ein Lösungsvorschlag zu dem Problem *Navicula* sensu lato ohne *Navicula* sensu stricto. Bibliotheca Diatomologica 29. J. Cramer, Berlin–Stuttgart, 1–212.

Legendre P, Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. Oecologia 129(2): 271–280. https://doi.org/10.1007/s004420100716.

Licursi M, Gómez N (2009) Effects of dredging on benthic diatom assemblages in a lowland stream. Journal of Environmental Management 90(2): 973–982. https://doi.org/10.1016/j.jenvman.2008.03.004.

Lowe RL (2011) The importance of scale in understanding the natural history of diatom communities. In: Seckbach J, Kociolek JP (Eds) The Diatom World. Springer, Dordrecht, 293–311. https://doi.org/10.1007/978-94-007-1327-7.

Mahlknecht J, Steinich B, Navarro de León I (2004) Groundwater chemistry and mass transfers in the Independence aquifer, central Mexico, by using multivariate statistics and mass–balance models. Environmental Geology 45: 781–795. https://doi.org/10.1007/s00254-003-0938-3.

McCune B, Mefford MJ (1999) PC–ORD: multivariate analysis of ecological data. Version 4 for Windows User's Guide, MjM software design.

Mora D, Carmona J, Cantoral-Uriza EA (2015) Diatomeas epilíticas de la cuenca alta del río Laja, Guanajuato, México. Revista Mexicana de Biodiversidad 86(4): 1024–1040. https://doi.org/10.1016/j.rmb.2015.09.004.

Morales EA, Siver PA, Trainor FR (2001) Identification of diatoms (Bacillariophyceae) during ecological assessments: comparison between light microscopy and scanning electron microscopy techniques. Proceedings of the Academy of Natural Sciences of Philadelphia 151: 95–103. https://doi.org/10.1635/0097-3157(2001)151[0095:IODBDE]2.0.CO;2.

Moser G, Lange-Bertalot H, Metzeltin D (1998) Insel der Endemiten. Geobotanisches Phänomen Neukaledonien. Bibliotheca Diatomologica 38. J. Cramer, Berlin –Stuttgart. 1–464.

Munné A, Prat N, Solá C, Bonada N, Rieradevall M (2003) A simple field method for assessing the ecological quality of riparian habitat in rivers and streams: QBR index. Aquatic Conservation: Marine and Freshwater Ecosystems 13: 147–163. https://doi.org/10.1002/aqc.529.

Peet, R. K. (1974). The measurement of species diversity. Annual review of ecology and systematics 5(1): 285–307. https://doi.org/10.1146/annurev.es.05.110174.001441.

Petrov AN, Nevrova EL (2014) Numerical analysis of the structure of benthic diatom assemblages in replicate samples (Crimea, the Black Sea). Nova Hedwigia 143: 245–253. https://doi.org/10.1127/1438-9134/2014/012.

Potapova MG, Charles DF (2002) Benthic diatoms in USA rivers: distribution along spatial and environmental gradients. Journal of Biogeography 29: 167–187. https://doi.org/10.1046/j.1365-2699.2002.00668.x.

Potapova MG, Charles DF (2007) Diatom metrics for monitoring eutrophication in rivers of the United States. Ecological Indicators 7: 48–70. https://doi.org/10.1016/j.ecolind.2005.10.001.

Ramírez-Vázquez M, Beltrán-Magos Y, Bojorge-García M, Carmona-Jiménez J, Cantoral-Uriza EA, Valadez–Cruz F (2001) Flora algal del río La Magdalena, Distrito Federal, México. Boletín de la Sociedad Botánica de México 68: 45–67. http://www.redalyc.org/pdf/577/57706805.pdf

Ramírez-Vázquez M, Cantoral-Uriza EA (2003) Flora algal de ríos templados en la zona occidental de la cuenca del Valle de México. Anales del Instituto de Biología, Universidad Nacional Autónoma de México, Serie Botánica 74(2): 143–194. http://www.redalyc.org/articulo.oa?id=40074202

Rothfritz H, Jüttner I, Suren AM, Ormerod SJ (1997) Epiphytic and epilithic diatom communities along environmental gradients in the Nepalese Himalaya: implications for the assessment of biodiversity and water quality. Archiv für Hydrobiologie 138(4): 465–482.

Round FE (1991) Diatoms in river water-monitoring studies. Journal of Applied Phycology 3(2): 129–145. https://doi.org/10.1007/BF00003695.

Rumrich U, Lange-Bertalot H, Rumrich M (2000) Diatomeen der Anden. Von Venezuela bis Patagonien/Feuerland. Und zwei weitere Beiträge. Iconographia Diatomologica 9. A.R.G. Gantner Verlag K.G., Ruggell. 1–673.

Rzedowski J, Calderón de Rzedowski G (2004) Oleaceae. Flora del Bajío y de Regiones Adyacentes. Instituto de Ecología A.C., Pátzcuaro, Fascicle 124, 1–37. http://www1.inecol.edu.mx/publicaciones/resumeness/FLOBA/Oleaceae124.pdf

Sabater S (2000) Diatom communities as indicators of environmental stress in the Guadiamar River, SW. Spain, following a major mine tailings spill. Journal of Applied Phycology 12(2): 113–124. https://doi.org/10.1023/A:1008197411815.

Segura-García V, Israde–Alcantara I, Maidana NI (2010) The genus *Navicula sensu stricto* in the upper Lerma Basin, Mexico. I. Diatom Research 25(2): 367–383. https://doi.org/10.1080/0269249X.2010.9705857.

Segura-García V, Cantoral-Uriza EA, Israde I, Maidana N (2012) Epilithic diatoms (Bacillariophyceae) as indicators of water quality in the Upper Lerma River, Mexico. Hidrobiológica 22(1): 16–27. http://hidrobiologica.izt.uam.mx/index.php/revHidro/article/view/676

Segura-García V, Almanza–Álvarez JS, Ponce–Saavedra J (2016) Diversidad en comunidades de diatomeas epilíticas con relación a los parámetros fisicoquímicos en la cabecera del río Zinapécuaro, México. Hidrobiológica 26(2): 187–202. http://hidrobiologica.izt.uam.mx/index.php/revHidro/article/view/495

Smol JP, Stoermer EF (2010) The diatoms: applications for the environmental and earth sciences. Second edition. Cambridge University Press, Cambridge, 1–667.

Soininen J, Eloranta P (2004) Seasonal persistence and stability of diatom communities in rivers: are there habitat specific differences? European Journal of Phycology 39(2): 153–160. https://doi.org/10.1080/0967026042000201858.

Stevenson RJ, Pan Y, van Dam H (2010) Assessing environmental conditions in rivers and streams with diatoms. In: Smol JP, Stoermer EF (Eds) The diatoms: applications for the environmental and earth sciences. Cambridge University Press, Cambridge, 57–85.

Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH (2006) Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16(4): 1267–1276. https://doi.org/10.1890/1051-0761(2006)016[1267:SEFTEC]2.0.CO;2.

Tavera R, Elster J, Marvan P (1994) Diatoms from Papaloapan basin communities, Mexico. Algological Studies 74: 35–65.

ter Braak CJF, Šmilauer P (2002) CANOCO reference manual and CanoDraw for Windows User's guide: Software for Canonical Community Ordination (Version 4.5). Microcomputer Power, Ithaca, NY, 1–500.

Tornés E, Cambra J, Gomà J, Leira M, Ortiz R, Sabater S (2007) Indicator taxa of benthic diatom communities: a case study in Mediterranean streams. Annales de Limnologie–International Journal of Limnology 43(1): 1–11. https://doi.org/10.1051/limn:2007023.

Townsend C, Doledec S, Scarsbrook M. (1997) Species traits in relation to temporal and spatial heterogeneity in streams: a test of habitat templet theory. Freshwater Biology 37(2): 367–387. https://doi.org/10.1046/j.1365-2427.1997.00166.x.

Trobajo R, Clavero E, Chepurnov VA, Sabbe K, Mann DG, Ishihara S, Cox E (2009). Morphological, genetic and mating diversity within thewidespread bioindicator Nitzschia palea (Bacillariophyceae). Phycologia 48(6), 443–459. https://doi.org/10.2216/08-69.1.

Valadez-Cruz F, Carmona-Jiménez J, Cantoral-Uriza EA (1996) Algas de ambientes lóticos en el estado de Morelos, México. Anales del Institutode Biología, Universidad Nacional Autónoma de México, Serie Botánica 67: 227–282. http://www.revistas.unam.mx/index.php/bot/article/view/1892

van Dam H, Mertens A, Sinkeldam J (1994) A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. Netherlands Journal of Aquatic Ecology 28: 117f133. https://doi.org/10.1007/BF02334251.

Vanormelingen P, Verleyen E, Vyverman W (2008) The diversity and distribution of diatoms: from cosmopolitanism to narrow endemism. Biodiversity and Conservation 17: 393–405. https://doi.org/10.1007/s10531-007-9257-4.

Vázquez G, Aké-Castillo JA, Favila ME (2011) Algal assemblages and their relationship with water quality in tropical Mexican streams with different land uses. Hydrobiologia 667: 173–189. https://doi.org/10.1007/s10750-011-0633-4.

Veselá J, Johansen JR (2009) The diatom flora of ephemeral headwater streams in the Elbsandsteingebirge region of the Czech Republic. Diatom Research 24(2): 443–477. https://doi.org/10.1080/0269249X.2009.9705813.

Visco JA, Apothéloz-Perret-Gentil L, Cordonier A, Esling P, Pillet L, Pawlowski J (2015) Environmental monitoring: inferring the diatom index from next–generation sequencing data. Environmental Science & Technology 49(13): 7597–7605. https://doi.org/10.1021/es506158m.

Vouilloud AA, Sala SE, Núñez-Avellaneda M, Montoya-Moreno Y, Duque SR (2014) *Brachysira* (Naviculales, Bacillariophyceae) in lowland waters from Colombia. Diatom Research 29(2): 147–163. https://doi.org/10.1080/0269249X.2013.866909.

Wester P, Scott CA, Burton M (2005) River basin closure and institutional change in Mexico's Lerma–Chapala Basin. In: Irrigation and River Basin management: options for governance and institutions. CABI Publishing, Wallingford, 125–144.

Wetzel CE, Ector L, van de Vijver B, Compère P, Mann DG (2015) Morphology, typification and critical analysis of some ecologically important small naviculoid species (Bacillariophyta). Fottea 15(2): 203–234. https://doi.org/10.5507/fot.2015.020.

Wolfe AP, Kling HJ (2001) A consideration of some North American soft–water *Brachysira* taxa and description of *B. arctoborealis* sp. nov. In: Jahn R, Kociolek JP, Witkowski A, Compère P. Lange-Bertalot–Festschrift. Studies on Diatoms, dedicated to Dr. Dr. h.c. Horst Lange-Bertalot on the occasion of his 65th birthday. A.R.G. Gantner Verlag K.G., Ruggell, 243–264.

Zamudio S, Rzedowski J, Carranza E, Calderón de Rzedowski G (1992) La vegetación del estado de Querétaro. Instituto de Ecología A.C., Centro Regional del Bajío, Consejo de Ciencia y Tecnología del Estado de Querétaro, 1–92 pp.

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. https://doi.org/10.1371/journal.pone.0108793.

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. https://doi.org/10.1111/1755-0998.12336.

# 3 Diatoms from the Lerma-Chapala River Basin, Central Mexico: an identification guide

Nélida Abarca[1]*, Demetrio Mora[1]*, Isabel Israde Alcántara[2] and Regine Jahn[1]

[1]Botanischer Garten und Botanisches Museum Berlin–Dahlem, Freie Universität Berlin, Königin–Luise–Straße 6–8, 14195 Berlin, Germany

[2]Estratigrafía y Paleoambientes, Departamento de Geología y Mineralogía, Instituto de Ciencias de la Tierra, Edificio U. Universidad Michoacana de San Nicolás de Hidalgo. Ciudad Universitaria, Col. Felicitas del Río, 58060 México

* Shared first authorship

The information presented here summarizes a diatom identification monograph, in preparation for *Englera*, an international, monographic series published by the Botanic Garden and Botanical Museum Berlin (BGBM), Freie Universität Berlin.

## 3.1 Abstract

This study focuses on the diatom flora of the Lerma-Chapala River Basin, Central Mexico. The basin is one of the most important in Mexico and a crucial water source for over 15 million people. Rapid economic and social changes in the basin have exacerbated the already scarce natural water supply. Periphyton samples, mainly epilithon, were collected from 31 sites seasonally (rainy and dry season), with sampling campaigns performed in 2003-2005 and in 2013-2014. The taxonomic composition consists of a total of 307 infrageneric taxa in 62 genera. The genera with the highest specific taxa richness were found among *Nitzschia* (38), *Gomphonema* (30), *Navicula* (24) and *Pinnularia* (22). Ten species are described as new, belonging to *Cocconeis*, *Craticula*, *Gomphonema* and *Sellaphora*. For each taxon, the following information is provided: scientific name with author(s), synonyms, morphological description, frequency and distribution in the basin and worldwide distribution. All the taxa are illustrated with light microscope images and most of the taxa are also illustrated with scanning electron microscope images. This study is an important contribution to the knowledge of

the diatom diversity of Mexico, representing an identification guide as well as a baseline for environmental monitoring studies.

## 3.2 Introduction

The Lerma-Chapala Basin, located in Central Mexico, is one of the most important basins of Mexico and a crucial water source for over 15 million people. The main source of the river is located 3000 m above sea level (asl), and after 750 km, discharges into the biggest water body of the country, the Lake Chapala, located at 1510 m asl. The Lerma-Chapala Basin supports the biggest industrial zone in Mexico with about 9,200 industries, and an eighth of the total land area is used for agricultural activities. Rapid economic and social changes in the basin have exacerbated the already scarce natural water supply. Water quality is also decreasing due to heavy pressures of economic development with most of the waters of the basin contaminated with the discharge of effluents from human settlements, agriculture and industry (Cotler et al. 2006; Mestre R 1997; Sedeño-Díaz and López-López 2007).

The evaluation and monitoring of water quality in the basin, mainly for the Lerma River, for which long time records of physical and chemical measurements are available, indicate a cumulative impact of point and non-point pollution sources dependent of land uses in the basin. Even though there is a slow progress in quality after improvements achieved since 1992, the water quality of the Lerma River still ranks among the most degraded in Mexico (Sedeño-Díaz and López-López 2007).

Physical and chemical assessments of water quality have been used for decades and provide a punctual record of the time the measurements were conducted. But in order to provide a complete assessment, not only of water quality but of ecological integrity, there is an increasingly use of biological indicators worldwide. Biological indicators have the advantage over physical and chemical indicators that they record historical ecological conditions, e.g. in lakes and streams from weeks to millennia (Hering et al. 2006; Iliopoulou-Georgudaki et al. 2003; Mccormick and Cairns 1994; Stevenson et al. 2010; Williamson et al. 2008). Diatoms are among the most widely used biological indicators, with some quality assessments already conducted in the Lerma-Chapala Basin (Abarca-Mejía 2010; Segura-Garcia et al. 2012).

Even though the advantages of the use of diatoms as biological indicators are well known (Dell'Uomo and Torrisi 2011; Dixit et al. 1992; Hering et al. 2006; Kelly et al. 2009; Licursi and Gomez 2009; Lowe 2011), the main impediment of their use concerns identification. Taxonomic identification is a crucial step in the use of diatoms as biological indicators, because closely related species and even cryptic

species have different environmental optima and tolerances (Morales et al. 2001; Potapova and Charles 2007; Poulickova et al. 2008; Van Dam et al. 1994). But taxonomic identification of diatoms requires in-depth knowledge of the diatom diversity of the area to be analyzed and awareness of the morphological plasticity a species displays under varying environmental factors and during its life cycle.

Since little is known about the diatom flora of Mexico, the objective of this study is to present an identification baseline of diatoms from the Lerma-Chapala Basin, with a commented list illustrating the morphological variability with both light and scanning electron microscopical images. The sampling includes sites at the Lerma River, the main watercourse of the basin, but also tributaries with varying degrees of anthropogenic impacts.

## 3.3 Material and Methods

### 3.3.1 Study area

The Lerma-Chapala Basin is located in the central region of Mexico, between the two largest and economically most important cities of the country, Mexico City and Guadalajara. Its geographical location is defined between the northern latitudes 19° 3' and 2° 34' and the western longitudes 99° 16' and 103° 31' (Figure 1). The maximal altitude of this basin is located on the volcano Nevado of Toluca at 4,600 m asl; whereas its minimal altitude is located in the surroundings of Lake Chapala at 1,500 m asl. The Lerma-Chapala Basin has a catchment area of 53,591.3 $km^2$ which represents about 3 % of Mexico's territory. The River Lerma with a length of 750 km is the longest river of Mexico, and the main watercourse of the basin (Cotler et al. 2006).

The region studied includes a great portion of the Trans-Mexican Volcanic Belt, constituting a dividing belt between the plains of the semi-arid north and the humid mountains in the southern part. It is divided in two main hydro climatic zones (humid and dry climates). These originate in the distribution of air masses generated by the presence and heights of the mountain systems surrounding the river basin. The humid semi-cold temperatures are located in the high part of the river basin, the humid temperatures and barren dry temperatures occur in the average river basin, and the sub-humid semi-warm temperatures are in the lower part of the river basin (Priego et al. 2004). The precipitation of the Lerma-Chapala Basin is differentiated by dry and rainy seasons. The rainy season begins in May, with maximum rain in July, and ends in October. The other six months are the dry season and correspond to low water. There is a direct relation between precipitation and air temperature, so that the months of maximum precipitation are also those of maximum temperature

(INE 2003). The average precipitation is 711.5 mm (CONAGUA 2006); the annual average temperature is 18.3 °C, with monthly variations of ± 3 °C.

### 3.3.2 Sampling

Samples were taken in two annual sampling campaigns. Sampling sites 1-17 were collected in November-December (end of rainy season) and April-June (dry season) of 2003, 2004 and 2005. Samples 18-31 were collected in September-October (rainy season) of 2013 and 2014, and in February (dry season) of 2014. For details of the physical and chemical composition of the waters, refer to Abarca-Mejía (2010) and Mora et al. (2017).

For the microscopical analysis of the diatom samples, sampling, cleaning, identification and counting procedures followed standard procedures (Kelly et al. 1998; Mora et al. 2017).

Benthic diatom samples were collected mainly from cobbles (epilithon), but when cobbles were not available samples were taken from macrophytes (epiphyton) and sediments (epipsammon). The substrates with visible algal growths were scrapped using disposable toothbrushes. Substrates located at 10-30 cm depth and in varying water flow conditions were sampled, in order obtain a representative cross section of the local diatom flora. The material was preserved with 4% formaldehyde (samples from 2003-2005) and with 70% alcohol (samples from 2013-2014) for transportation to the laboratory at the BGBM Berlin-Dahlem, Freie Universität Berlin.

### 3.3.3 Diatom analysis

Diatom frustules were cleaned in the laboratory by boiling (80 °C) in $H_2O_2$ for several hours. The peroxide was removed by repeated centrifugation and washing events using distilled water. Cleaned subsamples were dispersed on cover glasses, dried, and embedded in the diatom resin Naphrax. Counts and identification of diatoms valves were carried out using a Leitz DIALUX 20 microscope under the 100X objective for the 2003-2005 samples. Samples from 2013-2014 were counted and identified under a Zeiss Axioscope-microscope, using the 100X objective. All samples were photographed under a Zeiss Axioscope-microscope with DIC equipped with an AXIOAM MRc camera.

For scanning electron microscopy (SEM) analysis, aliquots of cleaned sample material for scanning electron microscopy observations were mounted on stubs, sputter-coated with gold-palladium. Observations of the 2003-2005 samples were performed under a Philips 515 scanning electron microscope (SEM), whereas samples of 2013-2014 were observed under a Hitachi FE 8010 SEM.

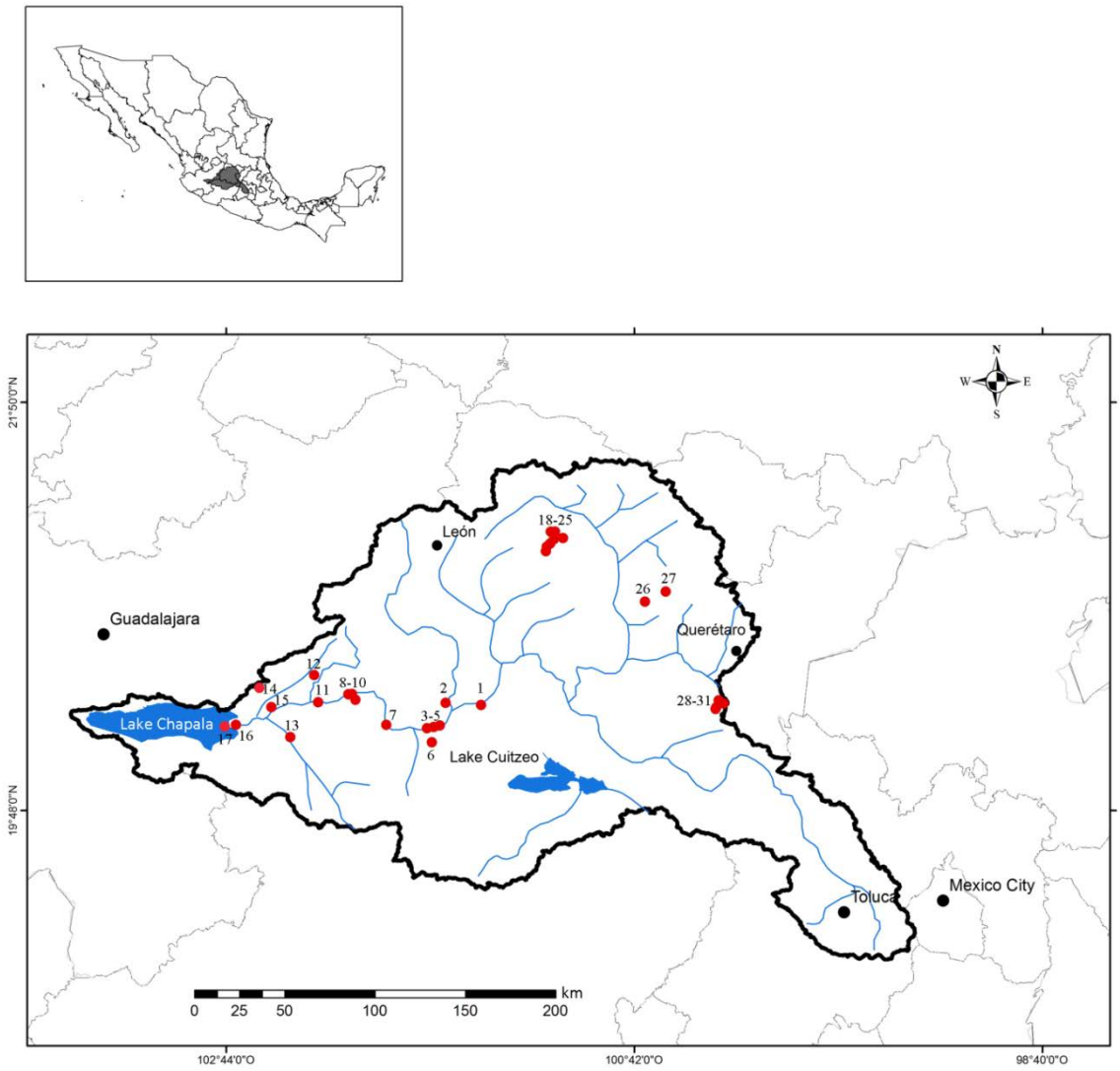



Figure 1. Location of the 31 sampling sites in the Lerma-Chapala Basin, Central Mexico. The numbers next to the red dots refer to the name of the sampling site in Table 1.

Table 1. Sampling sites in the Lerma-Chapala Basin, including site number and name, type of water body, width of the stream/river, geographical coordinates and elevation.

| Site | | Water body | Width (m) | Latitude (N) | Longitude (W) | Elevation (m) |
|---|---|---|---|---|---|---|
| 1 | Paso de Cobos | River | 79-90 | 20°19’28” | 101°27’40” | 1709 |
| 2 | Río Turbio | Stream | 15 | 20°20’11” | 101°38’20” | 1700 |
| 3 | Puerta de Agua Caliente | River | 79-90 | 20°13’24” | 101°40’02” | 1722 |
| 4 | El Mármol | River | 3.5-20 | 20°12’35” | 101°43’52” | 1700 |
| 5 | El Mármol | Spring | 1 | 20°12’56” | 101°41’53” | 1700 |
| 6 | Río Duero | Stream | 10-36 | 20°08’22” | 101°42’24” | 1531 |
| 7 | Hornitos | River | 25 | 20°13’34” | 101°56’02” | 1680 |
| 8 | Ojo de Agua | Spring | 40 | 20°21’04” | 102°05’13” | 1718 |
| 9 | San Juan del Fuerte | River | 67 | 20°22’43” | 102°07’23” | 1700 |
| 10 | Charapuato-Casas Blancas | River | 58-70 | 20°22’48” | 102°06’14” | 1626 |
| 11 | La Concepción | River | 35 | 20°20’17” | 102°16’23” | 1545 |
| 12 | Mesa el Salero | Spring | 1 | 20°28’27” | 102°16’93” | 1700 |
| 13 | Río Angulo | Stream | 10 | 20°09’57” | 102°24’41” | 1700 |
| 14 | Barranca del Aguacate | Spring | 1 | 20°26’58” | 102°34’93” | 1750 |
| 15 | Zalamea | River | 35 | 20°18’51” | 102°30’22” | 1504 |

| Site | | Water body | Width (m) | Latitude (N) | Longitude (W) | Elevation (m) |
|---|---|---|---|---|---|---|
| 16 | Maltaraña | River | n.a. | 20°13’36” | 102°40’57” | 1524 |
| 17 | Chapala-La Zapotera | Lake | Chapala Lake | 20°12’28” | 102°46’39” | 1531 |
| 18 | La Mesa | Stream | 1-4 | 21°05’29” | 101°08’19” | 2215 |
| 19 | Calvillo | Stream | 2-7 | 21°06’50” | 101°08’04” | 2138 |
| 20 | Ojo de Agua de Calvillo | Stream | 1-6 | 21°07’42” | 101°07’05” | 2102 |
| 21 | Peña Colorada | Stream | 1-5 | 21°09’04” | 101°05’59” | 2110 |
| 22 | San Martín | Stream | 2-7 | 21°09’25” | 101°03’11” | 2017 |
| 23 | Paredones | Stream | 2-5 | 21°11’21” | 101°06’53” | 2089 |
| 24 | La Laborcilla 1 | Stream | 1-7 | 21°11’05” | 101°06’15” | 2076 |
| 25 | La Laborcilla 2 | Stream | 2-13 | 21°11’20” | 101°05’38” | 2065 |
| 26 | El Membrillo | Stream | 2-4 | 20°50’21” | 100°38’43” | 2114 |
| 27 | Guanajuatito | Spring-fed creek | 1 | 20°53’24” | 100°32’31” | 2120 |
| 28 | Los Ailes 1 | Stream | 5 | 20°19’59” | 100°15’17” | 2358 |
| 29 | Laguna de Servín 1 | Stream | 2-4 | 20°18’18” | 100°17’38” | 2409 |
| 30 | Laguna de Servín 2 | Stream | 1-2 | 20°18’45” | 100°17’26” | 2409 |
| 31 | Los Ailes 2 | Stream | 1-7 | 20°20’50” | 100°16’46” | 2317 |

## 3.4 Results

A total of 307 infrageneric taxa in 63 genera were found (Table 2). The genera with the highest specific taxa richness were found among *Nitzschia* (38), *Gomphonema* (30), *Navicula* (24) and *Pinnularia* (22). Ten taxa are described as new, with the genera *Cocconeis*, *Craticula* and *Sellaphora* with one new species described each, and seven new species of *Gomphonema*.

The taxa in Table 2 are listed alphabetically. The morphological variability of all the taxa are illustrated in LM plates and in most cases SEM images are also provided. For each taxon, the following information is provided:

- Taxon name: scientific name with author name(s).
- Synonyms: taxonomic synonyms are given only in the case of recent nomenclatural changes or when used for particular reasons.

  ≡ homotypic synonym (objective synonym): both names share the same type specimen.

  = heterotypic synonym (subjective synonym): names with different type specimen.

  – concept synonym according to a given reference (e.g. misapplied name).
- Identification: one or two references are indicated that contain illustrations that best match the specimens of the Lerma-Chapala Basin.
- Morphology: own measurements of length, width and number of striae and areolae (when resolvable). The number of measured valves is given as (n=x).
- Frequency and distribution in the basin: the localities where the taxon was observed are given; the total valves observed in all sampling sites are given in parenthesis. The occurrence of each taxon is classified as follows:

  - Very rare: 3 valves or less were found in total.

  - Rare: less than 2% of relative abundance in at least one sample, but at least 4 valves found.

  - Uncommon: 2-5% relative abundance in at least one sample.

  - Common (c): ≥5-20% relative abundance in at least one sample.

  - Very common (cc): ≥20% relative abundance in at least one sample.
- Worldwide distribution: geographic distribution according to the reports in the literature.

- Additional observations: a short description is given if there are any morphological differences with the descriptions presented in the literature consulted.

Table 2. Diatom taxa list of the Lerma-Chapala Basin, Central Mexico. * Indicates taxa that are being described as new.

***Achnanthes inflata*** (Kützing) Grunow
***Achnanthidium* aff. *catenatum*** (J.Bílý & Marvan) Lange–Bertalot
***Achnanthidium exiguum*** (Grunow) Czarnecki
***Achnanthidium exile*** (Kützing) Round & Bukhtiyarova
***Achnanthidium minutissimum*** (Kützing) Czarnecki
***Achnanthidium* sp. 1**
***Achnanthidium* sp. 2**
***Achnanthidium* sp. 3**
***Amphipleura chiapasensis*** Metzeltin & Lange-Bertalot
***Amphora copulata*** (Kützing) Schoeman & R.E.M. Archibald
***Amphora pediculus*** (Kützing) Grunow
***Anomoeoneis sphaerophora*** Pfitzer
***Aulacoseira distans*** (Ehrenberg) Simonsen
***Aulacoseira granulata*** (Ehrenberg) Simonsen
***Aulacoseira italica*** (Ehrenberg) Simonsen
***Bacillaria paxillifera*** (O.F. Müller) Hendey var. paxillifera
***Biremis circumtexta*** (F.Meister ex Husttedt) Witkowski & Lange-Bertalot
***Brachysira altepetlensis*** D. Mora, R. Jahn & N. Abarca
***Brachysira brebissonii*** Ross
***Brachysira microcephala*** (Grunow) Compère
***Caloneis bacillum*** (Grunow) Cleve
***Caloneis clevei* var. *uruguayensis*** Frenguelli
***Caloneis schumanniana* var. *biconstricta*** (Grunow) Reichelt
***Caloneis silicula*** (Ehrenberg) Cleve
***Caloneis* cf. *silicula* var. *elliptica*** Frenguelli
***Caloneis* sp.**
***Cavinula scutelloides*** (W.Smith) Lange-Bertalot
***Chamaepinnularia submuscicola (Krasske)*** Lange-Bertalot
***cf. Chamaepinnularia* sp.**
***Cocconeis* cf. *neodiminuta*** Krammer
***Cocconeis pediculus*** Ehrenberg
***Cocconeis placentula*** Ehrenberg
***Cocconeis* cf. *placentula* var. *euglyta*** (Ehrenberg) Grunow
***Cocconeis placentula* var. *lineata*** (Ehrenberg) Van Heurck
**** Cocconeis placentula*** Ehrenberg **var. *placentula***
***Craticula accomoda*** (Hustedt) D.G. Mann
***Craticula acidoclinata*** Lange-Bertalot & Metzeltin

***Craticula ambigua*** (Ehrenberg) D.G. Mann
***Craticula buderi*** (Hustedt) Lange-Bertalot
***Craticula citrus*** (Krasske) E.Reichardt
***Craticula cuspidata*** (Kützing) D.G.Mann
***Craticula molestiformis*** (Hustedt) Lange-Bertalot
**** Craticula* cf. *pumilio*** Lange-Bertalot & U.Rumrich, nom. inval.
***Craticula subminuscula*** (Manguin) C.E. Wetzel & Ector
***Craticula submolesta*** (Hustedt) Lange-Bertalot
***Craticula* sp.**
***Cyclostephanos dubius*** (Hustedt) Round
***Cyclostephanos invisitatus*** (M.H.Hohn & Hellerman) Stoermer, E.C.Theriot & Håkansson
***Cyclotella atomus*** Hustedt
***Cyclotella meneghiniana*** Kützing
***Cymbella kolbei*** Hustedt
***Cymbella mexicana*** (Ehrenberg) Cleve
***Cymbella tropica*** Krammer & Metzeltin
***Cymbella tumida*** (Brébisson) Van Heurck
***Cymbella* sp.**
***Cymbopleura naviculiformis*** (Auerswald) Krammer
***Diadesmis confervacea*** Kützing
***Diploneis ovalis*** (Hilse) Cleve
***Diploneis subovalis*** Cleve
***Diploneis* sp.**
***Discostella pseudostelligera*** (Hustedt) Houk & Klee
***Discostella stelligera*** (Cleve & Grunow) Houk & Klee
***Encyonema neomesianum*** Krammer
***Encyonema silesiacum*** (Bleisch) D.G.Mann
***Encyonema triangulum*** (Ehrenberg) Kützing
***Encyonema brevicapitatum*** Krammer
***Encyonema* cf. *hebridiforme*** Krammer
***Encyonema jemtlandicum*** Krammer
***Encyonema jemtlandicum*** var. ***venezolanum*** Krammer
***Encyonema minutiforme*** Krammer
***Encyonema* cf. *minutiforme*** Krammer
***Encyonema minutum*** (Hilse) D.G. Mann
***Encyonema pergracile*** Krammer
***Encyonopsis microcephala*** (Grunow) Krammer
***Encyonopsis subminuta*** Krammer & E.Reichardt
***Encyonopsis* cf. *thienemannii*** (Hustedt) Krammer
***Encyonopsis* sp. 1**
***Epithemia adnata*** (Kützing) Brébisson
***Epithemia sorex*** Kützing
***Epithemia turgida*** (Ehrenberg) Kützing
***Eunotia bidens*** Ehrenberg

***Eunotia bilunaris*** (Ehrenberg) Schaarschmidt
***Eunotia* cf. *bigibba* var. *pumila*** Grunow
***Eunotia kruegeri*** Lange–Bertalot
***Eunotia major*** (W. Smith) Rabenhorst **var. *major***
***Eunotia* cf. *meridiana*** Metzeltin & Lange–Bertalot
***Eunotia metamonodon*** Lange–Bertalot
***Eunotia minor*** (Kützing) Ehrenberg
***Eunotia monodon*** Ehrenberg
***Eunotia mucophila***(Lange-Bertalot, Nörpel-Schempp & Alles) Lange-Bertalot
***Eunotia tridentula*** Ehrenberg
***Eunotia* sp. 1**
***Eunotia* sp. 2**
***Eunotia* sp. 3**
***Fallacia pygmaea*** (Kützing) Stickle & D.G.Mann
***Fistulifera saprophila*** (Lange-Bertalot & Bonik) Lange-Bertalot
***Fragilaria austriaca*** (Grunow) Lange-Bertalot
***Fragilaria bidens*** Heiberg
***Fragilaria crotonensis*** Kitton
***Fragilaria exigua*** Grunow
***Fragilaria goulardii*** (Brébisson) Lange-Bertalot
***Fragilaria pectinalis*** (O.F. Müller) Lyngbye
***Fragilaria pinnata*** Ehrenberg **var. *pinnata***
***Fragilaria rumpens*** (Kützing) Carlson
***Fragilaria tenera*** (W. Smith) Lange-Bertalot
***Fragilaria vaucheriae*** (Kützing) J. B. Petersen
***Frustulia crassinervia*** (Brébisson) Lange–Bertalot & Krammer
***Frustulia neomundana*** Lange–Bertalot & Rumrich
***Frustulia* cf. *spicula* ssp. *spicula*** Amossé
***Frustulia* cf. *undosa*** Melzeltin & Lange–Bertalot
***Frustulia vulgaris*** (Thwaites) De Toni
***Geissleria decussis*** (Østrup) Lange-Bertalot & Metzeltin
***Gomphonema acuminatum*** Ehrenberg **var. *acuminatum***
***Gomphonema affine*** Kützing **var. *affine***
**** Gomphonema* cf. *augur*** Ehrenberg
***Gomphonema brasiliense*** Grunow
***Gomphonema exilissimum*** (Grunow) Lange-Bertalot & E.Reichardt
***Gomphonema graciledictum*** E. Reichardt
***Gomphonema innocens*** E. Reichardt
***Gomphonema insigne*** W.Gregory
***Gomphonema kobayashiae*** Metzeltin & Lange-Bertalot
***Gomphonema lagenula*** Kützing
***Gomphonema laticollum*** E.Reichardt
***Gomphonema* aff. *mariovense*** Levkov & Tofilovska
***Gomphonema mexicanum*** Grunow

***Gomphonema minusculum*** Krasske
***Gomphonema naviculoides*** W. Smith
***Gomphonema parvuliforme*** Levkov, Mitic–Kopanja & E.Reichardt
***Gomphonema parvulum Kützing***
***Gomphonema pseudoaugur*** Lange-Bertalot
***Gomphonema pumilum* var. *rigidum*** E.Reichardt & Lange-Bertalot
**** Gomphonema* cf. *salae*** Lange-Bertalot. & E.Reichardt
***Gomphonema* aff. *sarcophagus*** W. Gregory
***Gomphonema stonei*** E.Reichardt
***Gomphonema subclavatum*** (Grunow) Grunow
**** Gomphonema* sp. 1 cf. *apicatum*** Ehrenberg
**** Gomphonema* sp. 2 cf. *gracile*** Ehrenberg
**** Gomphonema* sp. 3**
**** Gomphonema* sp. 4**
**** Gomphonema* sp. 5**
***Gomphonema* sp. 6**
***Gomphonema* sp. 7**
***Gomphosphenia lingulatiformis*** (Lange-Bertalot & E.Reichardt) Lange-Bertalot
***Gomphosphenia tenerrima*** (Hustedt) E.Reichardt
***Gyrosigma kuetzingii*** (Grunow) Cleve
***Gyrosigma obtusatum*** (Sullivant & Wormley) Boyer
***Halamphora montana*** (Krasske) Levkov
***Halamphora* cf. *pseudomontana*** (Cholnoky) Levkov
***Halamphora veneta*** (Kützing) Levkov
***Hantzschia abruptirostrata*** Lange-Bertalot & Metzeltin
***Hantzschia abundans*** Lange-Bertalot
***Hantzschia amphioxys*** (Ehrenberg) Grunow
***Hantzschia uruguayensis*** Metzeltin, Lange-Bertalot & García-Rodríguez
***Hantzschia vivacior*** Lange-Bertalot
***Hantzschia* cf. *vivax*** (W.Smith) Peragallo
***Hantzschia* sp.**
***Hippodonta capitata*** (Ehrenberg) Lange-Bertalot, Metzeltin & Witkowski
***Humidophila contenta*** (Grunow) Lowe, Kociolek, Johansen, Van de Vijver, Lange-Bertalot & Kopalová
***Lemicola hungarica*** (Grunow) Round & Basson
***Luticola goeppertiana*** (Bleisch) D.G.Mann
***Luticola mutica*** (Kützing) D.G.Mann
***Luticola nivalis*** (Ehrenberg) D.G.Mann
***Luticola* cf. *peguana*** (Grunow) D.G.Mann
***Luticola ventricosa*** (Kützing) D.G.Mann
***Luticola* sp. 1**
***Luticola* sp. 2**
***Mayamaea* cf. *crassistriata*** Lange–Bertalot, Cavacini, Tagliaventi & Alfinito
***Mayamaea permitis*** (Hustedt) Bruder & Medlin

***Mayamaea* sp. 1**
***Melosira varians*** C.Agardh
***Navicula antonii*** Lange-Bertalot
***Navicula* cf. *arvensis*** Hustedt
***Navicula capitatoradiata*** H.Germain
***Navicula cryptocephala*** Kützing
***Navicula cryptotenella*** Lange-Bertalot
***Navicula erifuga*** Lange-Bertalot
***Navicula germainii*** J.H.Wallace
***Navicula gregaria*** Donkin
***Navicula kotschyi*** Grunow
***Navicula leptostriata*** E.G.Jørgensen
***Navicula radiosa*** Kützing
***Navicula recens*** (Lange-Bertalot) Lange-Bertalot
***Navicula riediana*** Lange-Bertalot & U.Rumrich
***Navicula rostellata*** Kützing
***Navicula symmetrica*** R.M.Patrick
***Navicula tripunctata*** (O.F.Müll.) Bory
***Navicula trivialis*** Lange-Bertalot
***Navicula* cf. *veneta*** Kützing
***Navicula* sp. 1**
***Navicula* sp. 2**
***Navicula* cf. *cryptocephala*** Kützing
***Navicula libonensis*** Schoeman
***Navicula notha*** Wallace
***Navicula reichardtiana*** Lange–Bertalot
***Neidium* cf. *productum*** (W.Smith) Cleve
***Nitzschia acicularis*** (Kützing) W.Smith
***Nitzschia amphibia*** Grunow
***Nitzschia amphibia* f. *frauenfeldii*** (Grunow) Lange-Bertalot
***Nitzschia brevissima*** Grunow
***Nitzschia capitellata*** Hustedt
***Nitzschia clausii*** Hantzsch
***Nitzschia communis*** Rabenhorst
***Nitzschia constricta*** (Kützing) Ralfs
***Nitzschia desertorum*** Hustedt
***Nitzschia dissipata*** (Kützing) Grunow
***Nitzschia dissipata* var. *media*** (Hantzsch) Grunow
***Nitzschia filiformis*** (W.Smith) Van Heurck
***Nitzschia fonticola*** (Grunow) Grunow
***Nitzschia frustulum*** (Kützing) Grunow
***Nitzschia gracilis*** Hantzsch
***Nitzschia inconspicua*** Grunow
***Nitzschia intermedia*** Hantzsch ex Cleve & Grunow

***Nitzschia lanceolata*** W.Smith
***Nitzschia levidensis* var*. victoriae*** (Grunow) Cholnoky
***Nitzschia linearis*** (C.Agardh) W.Smith **var. *linearis***
***Nitzschia lorenziana*** Grunow
***Nitzschia microcephala*** Grunow
***Nitzschia palea*** (Kützing) W.Smith
***Nitzschia palea* var*. debilis*** (Kützing) Grunow
***Nitzschia paleacea*** (Grunow) Grunow
***Nitzschia rautenbachiae*** Cholnoky
***Nitzschia reversa*** W.Smith
***Nitzschia semirobusta*** Lange-Bertalot
***Nitzschia sigma*** (Kützing) W.Smith
***Nitzschia* cf. *simplex*** Hustedt
***Nitzschia supralitorea*** Lange-Bertalot
***Nitzschia umbonata*** (Ehrenberg) Lange-Bertalot
***Nitzschia palea* var. *tenuirostris*** Grunow
***Nitzschia perminuta*** (Grunow) Peragallo
***Nitzschia tubicola*** Grunow
***Nitzschia* sp. 1**
***Nitzschia* sp. 2**
***Nitzschia* sp. 3**
***Nupela praecipua*** (E. Reichardt) E. Reichardt
***Nupela wellneri*** (Lange–Bertalot) Lange–Bertalot
***Opephora olsenii*** M.Møller
***Pinnularia acrosphaeria*** (Brébisson) Rabenhorst
***Pinnularia anglica*** morphodeme 1 Krammer
***Pinnularia anglica*** morphodeme 2 Krammer
***Pinnularia borealis* var. scalaris** (Ehrenberg) Rabenhorst
***Pinnularia* cf. *altiplanensis*** Lange-Bertalot, Krammer & Rumrich
***Pinnularia brebissonii*** (Kützing) Rabenhorst
***Pinnularia* cf. *brebissonii* var. *acuta*** Cleve–Euler
***Pinnularia divergens*** W.Smith
***Pinnularia divergens* var. *media*** Krammer
***Pinnularia gibba*** (Ehrenberg) Ehrenberg
***Pinnularia mayeri*** Krammer
***Pinnularia microstauron*** (Ehrenberg) Cleve
***Pinnularia saprophila*** Lange–Bertalot, H. Kobayasi & Krammer
***Pinnularia* cf. *subcapitata* var. *elongata*** Krammer
***Pinnularia subgibba*** Krammer
***Pinnularia viridiformis*** Krammer
***Pinnularia* sp. 1**
***Pinnularia* sp. 2**
***Pinnularia* sp. 3**
***Pinnularia* sp. 4**

***Pinnularia* sp. 5**
***Pinnularia* sp. 6**
***Placoneis* cf. *constans*** (Hustedt) E.J. Cox
***Placoneis undulata*** (Østrup) Lange–Bertalot
***Placoneis* sp.**
***Planothidium frequentissimum*** (Lange-Bertalot) Lange-Bertalot
***Planothidium lanceolatum*** (Brébisson ex Kützing) Lange-Bertalot
***Planothidium incuriatum*** C.E. Wetzel, Van de Vijver & Ector
***Planothidium rostratum*** (Østrup) Lange–Bertalot
***Pleurosira laevis*** (Ehrenberg) Compère
***Pseudostaurosira brevistriata*** (Grunow) D.M.Williams & Round
***Reimeria sinuata*** (W.Gregory) Kociolek & Stoermer
***Rhoicosphenia abbreviata*** (C.Agardh) Lange-Bertalot
***Rhopalodia acuminata*** Krammer
***Rhopalodia gibba*** (Ehrenberg) O.Müller
***Rhopalodia operculata*** (C.Agardh) Håkansson
***Rhopalodia rupestris*** (W.Smith) Krammer
***Sellaphora atomoides*** (Grunow) C.E. Wetzel & Van de Vijver
***Sellaphora bacilloides*** Hustedt
***Sellaphora cosmopolitana*** (Lange-Bertalot) C.E. Wetzel & Ector
***Sellaphora elorantana*** (Lange-Bertalot) C.E. Wetzel
***Sellaphora laevissima*** (Kützing) D.G.Mann
***Sellaphora madida*** (Kociolek) C.E. Wetzel
***Sellaphora nigri*** (De Notaris) C.E. Wetzel & Ector
***Sellaphora pupula*** (Kützing) Mereschkowski
***Sellaphora queretana*** D. Mora, N. Abarca & J. Carmona
***Sellaphora* cf. *saugerresii*** (Desmazières) C.E. Wetzel & D.G. Mann
***Sellaphora stauroneioides*** Lange-Bertalot
***Sellaphora wallacei*** (Reimer) Potapova & Ponader
***Sellaphora* sp. 1**
***Sellaphora* sp. 2 cf. *pupula*** (Kützing) Mereschk
**** Sellaphora* sp. 3**
***Stauroneis anceps*** Ehrenberg
***Staurosira construens*** Ehrenberg var. ***construens***
***Staurosira longirostris*** (Frenguelli) Metzeltin & García-Rodríguez, nom. inval
***Stenopterobia delicatissima*** (Lewis) Van Heurck
***Stephanodiscus medius*** Håkansson
***Stephanodiscus* sp.**
***Surirella angusta*** Kützing
***Surirella apiculata* var. *panduriformis*** Frenguelli
***Surirella brebissonii*** Krammer & Lange-Bertalot
***Surirella minuta* var. *peduliformis*** Frenguelli
***Surirella ovalis*** Brébisson
***Surirella splendida*** (Ehrenberg) Kützing

***Surirella* sp.**
***Thalassiosira* cf. *fauri*** (Gasse) Hasle
***Trybionella calida*** (Grunow) D.G.Mann
***Tryblionella hungarica*** (Grunow) D.G.Mann
***Ulnaria acus*** (Kützing) M. Aboal
***Ulnaria lanceolata*** (Kützing) Compère
***Ulnaria ulna*** (Nitzsch) compère

Here, an example of a taxon description is provided, accompanied by a plate containing LM and SEM images (Fig. 2).

***Cymbopleura naviculiformis*** (Auerswald) Krammer (Fig. 2)

≡ *Cymbella naviculiformis* Auerswald ex Heiberg

*Identification*: Krammer (2003) pl. 83: 9-11; Hofmann et al. (2013), pl. 83: 20–23.

*Morphology*: valve length: 31-38.3 µm; width: 8.5-10.4 µm; length/width ratio: 3.4-3.9; ventral striae at valve center in 10 µm: 13-16; dorsal striae at valve center in 10 µm: 10-13; ventral areolae at valve center in 10 µm: 30-34; dorsal areolae at valve center in 10 µm: 28-36 (n= 15).

*Frequency and distribution in the basin*: rare in each of the locations it was found (Paredones, La Laborcilla 1, La Laborcilla 2, Laguna de Servín 1 and Laguna de Servín 2) (33 total valves observed).

*Worldwide distribution*: widely distributed in temperate to subarctic zones, but also at higher elevations in the tropics (Krammer 2003).

*Additional observations*: in external view, the areolae are apically elongated except at the apices where they are transapically elongated. Krammer (2003) reports areolae slightly elongated transapically, whereas Van de Vijver et al. (2011) report oval areolae. The breadth of the central area from the Lerma-Chapala specimens is narrower than in Krammer (2003) and Bahls (2012).

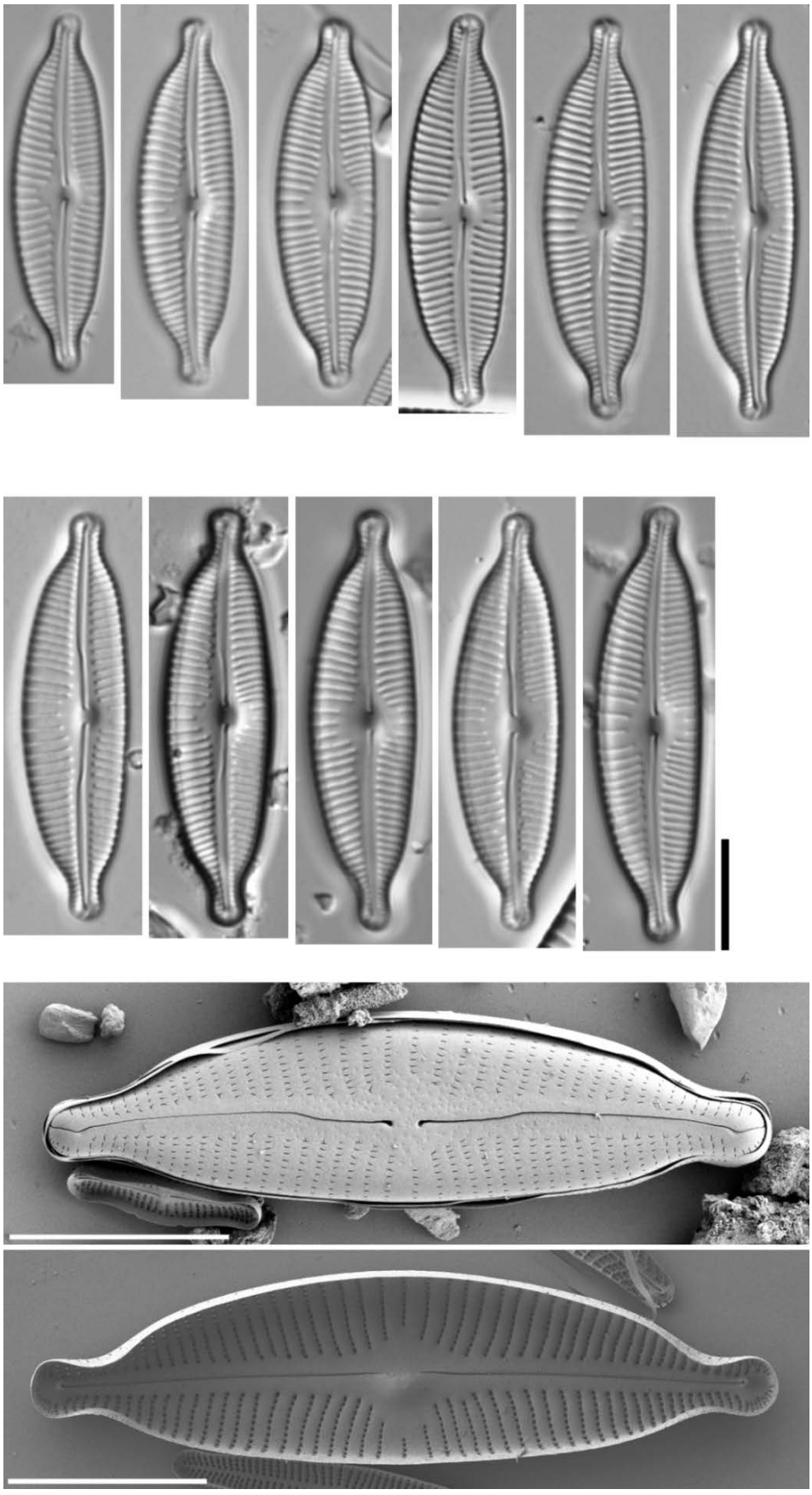


Figure 2. Morphological diversity of *Cymbopleura naviculiformis* (Auerswald) Krammer illustrated in LM and SEM. Scale bars 10 µm.

## 3.5 Conclusion

The taxonomic richness presented here, 307 infrageneric taxa, sets an important identification baseline for the diatoms of the Lerma-Chapala Basin and adjacent regions, with ten species described as new. Even though several studies have already been conducted in the Lerma-Chapala Basin (Abarca-Mejía 2010; Mora et al. 2015; Mora et al. 2017; Segura-García 2011; Segura-García et al. 2010), our study is the most comprehensive for river diatoms so far conducted in the basin and in Mexico, in terms of taxa number and illustrating the morphological variability of each taxa in both LM and SEM. Previous studies report less taxa and illustrate each taxon with one or a couple of images only (mostly LM), but not the broad variability a taxon might display, which is important for correct diatom identification. A further important point of our study are the broad spectra of environmental conditions that were sampled, from the polluted Lerma River and its large tributaries, to the non-polluted streams from the north of the basin as well as some springs, setting the baseline for the flora of different environmental conditions, important for comparisons in water quality monitoring studies. It is hoped that this study will serve as an important contribution to the aquatic biodiversity of Mexico, an identification guide as well as a baseline for monitoring studies.

## 3.6 Acknowledgements

The work of DM was funded by the Mexican Government through doctoral grants from CONACYT, CONCYTEQ and DGRI-SEP. Samples were taken under permit CONAPESCA PPF/DGOPA-149/14. The Deutsche Forschungsgemeinschaft is thanked for Grant INST 130/839-1 FUGG concerning SEM funding. Monika Lüchow and Kim Govers kindly assisted at the SEM. Verónica Aguilar Zamora is thanked for creating the map.

## 3.7 References

Abarca-Mejía N (2010) Diatom community analysis and quality assessment of the polluted tropical Lerma River (Mexico). Dissertation, Freie Universität Berlin, Germany.

Bahls L (2012).*Cymbopleura naviculiformis*. In: Diatoms of the United States. http://westerndiatoms.colorado.edu/taxa/species/cymbopleura_naviculiformis [accessed 31.10.2017]

CONAGUA (2006) Estadísticas del Agua en México 2006. Sistema Nacional de Información sobre cantidad, calidad, usos y conservación del agua 4ª edición, México.

Cotler H, Mazari M, De Anda J (2006) Atlas de la cuenca Lerma-Chapala, construyendo una visión conjunta. INE-UNAM, 191p.

Dell'Uomo A, Torrisi M (2011) Rheophile Apennine diatoms and their use as bioindicators of water quality. The Diatom World. Springer, 401-417.

Dixit SS, Smol JP, Kingston JC, Charles DF (1992) Diatoms - Powerful Indicators of Environmental-Change. Environmental science & technology 26(1): 22-33. https://doi.org/10.1021/es00025a002.

Hering D, Johnson RK, Kramm S, Schmutz S, Szoszkiewicz K, Verdonschot PFM (2006) Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. Freshwater Biology 51(9): 1757-1785. https://doi.org/10.1111/j.1365-2427.2006.01610.x.

Hofmann G, Lange-Bertalot H, Werum M (2013) Diatomeen im Süßwasser–Benthos von Mitteleuropa. Bestimmungsflora Kieselalgen für die ökologische Praxis. Über 700 der häufigsten Arten und ihre Ökologie. Koeltz Scientific Books, Königstein, 908 pp.

Iliopoulou-Georgudaki J, Kantzaris V, Katharios P, Kaspiris P, Georgiadis T, Montesantou B (2003) An application of different bioindicators for assessing water quality: a case study in the rivers Alfeios and Pineios (Peloponnisos, Greece). Ecological Indicators 2(4): 345-360. https://doi.org/10.1016/S1470-160X(03)00004-9.

INE (2003) Paisajes hidrológicos y balance hídrico de la cuenca Lerma Chapala. Escala 1: 250 000.

Kelly M, Bennett C, Coste M, Delgado C, Delmas F, Denys L, Ector L, Fauville C, Ferreol M, Golub M, Jarlman A, Kahlert M, Lucey J, Ni Chathain B, Pardo I, Pfister P, Picinska-Faltynowicz J, Rosebery J, Schranz C, Schaumburg J, van Dam H, Vilbaste S (2009) A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. Hydrobiologia 621(1): 169-182. https://doi.org/10.1007/s10750-008-9641-4.

Kelly MG, Cazaubon A, Coring E, Dell'Uomo A, Ector L, Goldsmith B, Guasch H, Hürlimann J, Jarlman A, Kawecka B, Kwandrans J, Laugaste R, Lindstrøm E-A, Leitao M, Marvan P, Padisák J, Pipp E, Prygiel J, Rott E, Sabater S, van Dam H, Vizinet J (1998) Recommendations for the routine sampling of diatoms for water quality assessments in Europe. Journal of Applied Phycology 10(2): 215. https://doi.org/10.1023/a:1008033201227.

Krammer K (2003) Diatoms of Europe: Diatoms of the European inland waters and comparable habitats. *Cymbopleura*, *Delicata*, *Navicymbula*, *Gomphocymbellopsis*, *Afrocymbella*. A.R.G. Gantner Verlag K.G., Ruggell, 530 pp.

Licursi M, Gomez N (2009) Effects of dredging on benthic diatom assemblages in a lowland stream. J Environ Manage 90(2): 973-982. https://doi.org/10.1016/j.jenvman.2008.03.004.

Lowe RL (2011) The importance of scale in understanding the natural history of diatom communities. The Diatom World. Springer, 293-311.

Mccormick PV, Cairns J (1994) Algae as Indicators of Environmental-Change. Journal of Applied Phycology 6(5-6): 509-526. https://doi.org/10.1007/Bf02182405.

Mestre R JE (1997) Integrated approach to river basin management: Lerma-Chapala case study—attributions and experiences in water management in Mexico. Water International 22(3): 140-152. https://doi.org/10.1080/02508069708686693.

Mora D, Carmona J, Cantoral-Uriza EA (2015) Diatomeas epilíticas de la cuenca alta del río Laja, Guanajuato, México. Revista Mexicana de Biodiversidad 86(4): 1024-1040. https://doi.org/10.1016/j.rmb.2015.09.004.

Mora D, Carmona J, Jahn R, Zimmermann J, Abarca N (2017) Epilithic diatom communities of selected streams from the Lerma-Chapala Basin, Central Mexico, with the description of two new species. PhytoKeys 88(39-69. https://doi.org/10.3897/phytokeys.88.14612.

Morales EA, Siver PA, Trainor FR (2001) Identification of diatoms (Bacillariophyceae) during ecological assessments: Comparison between Light Microscopy and Scanning Electron Microscopy techniques. Proceedings of the Academy of Natural Sciences of Philadelphia 151(1): 95-103. https://doi.org/10.1635/0097-3157(2001)151[0095:Iodbde]2.0.Co;2.

Potapova M, Charles DF (2007) Diatom metrics for monitoring eutrophication in rivers of the United States. Ecological Indicators 7(1): 48-70. https://doi.org/10.1016/j.ecolind.2005.10.001.

Poulickova A, Spackov J, Kelly MG, Duchoslav M, Mann DG (2008) Ecological variation within Sellaphora species complexes (Bacillariophyceae): specialists or generalists? Hydrobiologia 614(1): 373-386. https://doi.org/10.1007/s10750-008-9521-y.

Priego ÁG, Morales H, Enríquez C (2004) Paisajes físico-geográficos de la cuenca Lerma-Chapala. Gaceta ecológica 71): 11-22.

Sedeño-Díaz JE, López-López E (2007) Water Quality in the Río Lerma, Mexico: An Overview of the Last Quarter of the Twentieth Century. Water Resources Management 21(10): 1797-1812. https://doi.org/10.1007/s11269-006-9128-x.

Segura-García V (2011) Taxonomía y ecología de las diatomeas epilíticas de la Cuenca del Alto Lerma, México. Dissertation, Universidad Nacional Autónoma de México, Mexico.

Segura-Garcia V, Cantoral-Uriza EA, Israde I, Maidana N (2012) Epilithic diatoms (Bacillariophyceae) as indicators of water quality in the Upper Lerma River, Mexico. Hidrobiológica 22(1): 16-27.

Segura-García V, Israde-Alcántara I, Maidana NI (2010) The genus Navicula sensu stricto in the upper Lerma Basin, Mexico. I. Diatom Research 25(2): 367-383. https://doi.org/10.1080/0269249X.2010.9705857.

Stevenson R, Pan Y, van Dam H (2010) Assessing environmental conditions in rivers and streams with diatoms. The Diatoms: Applications for the Environmental and Earth Sciences, 2nd ed Cambridge University Press, Cambridge.

Van Dam H, Mertens A, Sinkeldam J (1994) A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. Aquatic Ecology 28(1): 117-133. https://doi.org/10.1007/BF02334251.

Van de Vijver B, Jüttner I, Gurung S, Sharma C, Sharma S, Haan M, Cox E (2011) The genus *Cymbopleura* (Cymbellales, Bacillariophyta) from high altitude freshwater habitats, Everest National Park, Nepal, with the description of two new species. Fottea 11(2): 245-269. https://doi.org/10.5507/fot.2011.025.

Williamson CE, Dodds W, Kratz TK, Palmer MA (2008) Lakes and streams as sentinels of environmental change in terrestrial and atmospheric processes. Frontiers in Ecology and the Environment 6(5): 247-254. https://doi.org/10.1890/070140.

# 4 Morphology and metabarcoding! A test with stream diatoms from Mexico highlights the complementarity of methods

Demetrio Mora[1], Nélida Abarca[1], Sebastian Proft[1], José H. Grau[2-3], Neela Enke[1], Javier Carmona[4], Oliver Skibbe[1], Regine Jahn[1] and Jonas Zimmermann[1]

[1]Botanischer Garten und Botanisches Museum Berlin–Dahlem, Freie Universität Berlin, Königin-Luise-Straße 6–8, 14195 Berlin, Germany

[2]Museum für Naturkunde Berlin, Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Invalidenstraße 43, 10115, Berlin, Germany

[3]BEGENDIV - Berlin Center for Genomics in Biodiversity Research, Berlin, Germany

[4]Facultad de Ciencias, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, Coyoacán, 04510 México, D.F., México

*submitted to Freshwater Science*

## 4.1 Abstract

Diatoms are widely used as biological indicators, with diatom-based indices developed to monitor environmental change, e.g. biotic integrity and water quality. Thus correct species identification is a crucial step in abstracting sound and meaningful results from those indices. As a test case, epilithic diatoms of streams from Central Mexico were studied by morphology and metabarcoding in order to compare how these two identification methods perform. In parallel, a regional taxonomic reference library was assembled based on clonal cultures, resulting in 188 strains belonging to 70 species in 24 genera. The morphological analysis of environmental samples resulted in the identification of 205 taxa in 43 genera, while the metabarcoding approach resulted in the identification of 266 infrageneric taxa belonging to 35 genera. The taxonomic assignment of the taxa inferred from metabarcoding led to the identification of 94 infrageneric taxa being confidently assigned. One quarter of the taxonomic assignations from High-Throughput Sequencing (HTS) data were due to our taxonomic reference library (24 %, 23 out of the 94 assigned taxa). The comparison of relative abundances of valves and sequence reads showed big disparities between both methods. The

prospect of using HTS data as a source of barcodes is supported by our results, since we were able to recover barcodes for *Iconella delicatissima* and *Navicula notha*, with *N. notha* being the second most abundant taxon retrieved from HTS data across all samples. Our results led us to conclude that the combination of morphological and molecular methods increases the detection and identification of diatoms.

**Keywords:** DNA barcoding, eDNA metabarcoding, epilithic diatoms, High-Throughput Sequencing (HTS), taxonomic reference libraries, V4 18S rRNA gene.

## 4.2 Introduction

Though freshwater only comprises a small fraction of the Earth's water and surface, it supports a high share of the global biodiversity and provides invaluable ecosystem services to humanity. Therefore, its conservation, assessment, monitoring and management are of global concern. Despite its importance, freshwater ecosystems and their related services experience even bigger threats than terrestrial ecosystems (Dodds et al. 2013; Dudgeon et al. 2006; Postel et al. 1996). In order to counteract degradation, national and international legislation have been adopted for the conservation, assessment, monitoring and management of freshwater ecosystems, including the Water Framework Directive (WFD, Directive 2000/60/EC) and the Marine Strategy Framework Directive (MSFD, Directive 2008/56/EC) of the European Union, the Water Protection Ordinance (WPO, Swiss Federal Council 1998) of Switzerland and the Clean Water Act (CWA, https://www.epa.gov/cwa-404/) of the Environmental Protection Agency in the USA.

Concerning the assessments and monitoring of freshwater ecosystems, in the aforementioned legislations the use of not only physical and chemical indicators has been established as it has been done for decades, but also the incorporation of biological indicators. Biological indicators have the advantage over physical and chemical indicators in that they record the historical ecological conditions, e.g. in lakes and streams from weeks to millennia (Mccormick and Cairns 1994; Stevenson et al. 2010; Williamson et al. 2008). Diatoms are among the most commonly used biological indicators of water quality because of their high diversity, widespread distribution, short life cycles, species level optima and tolerances to water quality variables and pollutants, and their relative ease of sampling, processing and storage (Dell'Uomo and Torrisi 2011; Dixit et al. 1992; Hering et al. 2006; Kelly et al. 2009; Licursi and Gomez 2009; Lowe 2011), with several local and regional monitoring programs and networks using diatom-based indices to monitor biotic integrity and water quality (Berthon et al. 2011; Kelly et al. 2009; Kelly 1998; Potapova and Charles 2007; Potapova and Charles 2002; Prygiel 2002; Smucker and Vis 2011).

Though the advantages of the use of diatoms as biological indicators are acknowledged, there are also some issues regarded as disadvantageous, such as identification. Taxonomic identification is a crucial step in recording sound and meaningful results from those indices since closely related species and even cryptic species have different environmental optima and tolerances (Morales et al. 2001; Potapova and Hamilton 2007; Poulickova et al. 2008; Van Dam et al. 1994). But taxonomic identification of diatoms is time consuming and requires in-depth knowledge of the diatom diversity of the area to be analyzed and awareness of the morphological plasticity a species displays under varying environmental factors and during its life cycle.

DNA barcoding (Hebert et al. 2003) and environmental DNA metabarcoding (Taberlet et al. 2012b) have been proposed as an alternative to morphology-based identifications of organisms, including diatoms, through DNA "barcode" sequences. A barcode consist of a molecular marker that can be easily sequenced in one Sanger run, unambiguously identifying a taxon independent of its life cycle (Hebert et al. 2003; Moritz and Cicero 2004; Zimmermann et al. 2011). In diatoms, the main sources for barcodes are clonal cultures, coming from single cell isolations. This approach has the advantage of correlating the barcode sequence to the material for morphological examination (Stachura-Suchoples et al. 2015; Zimmermann et al. 2014), but diatom culturing is highly time-consuming. Single cell PCR amplifications have been proposed as an alternative to culturing for obtaining barcodes, which is a reasonable alternative for taxa that exhibit recalcitrance to laboratory culturing conditions (Chen et al. 2013; Hamilton et al. 2015; Lang and Kaczmarska 2011), but corroboration of taxon identity is difficult because the valves of the single cell isolated are normally destroyed in the DNA extraction process. It also has the disadvantage of missing DNA availability for further amplification of different markers if, for example, a multi-marker phylogeny study is intended. Most recently, the use of High-Throughput Sequencing (HTS) data as a source of barcodes has been proposed to overcome the incompleteness of reference databases, setting up criteria to ensure that the proposed barcodes truly correspond to microscopy observations (Rimet et al. 2018).

Environmental DNA (eDNA) is the DNA than can be extracted from bulk samples of soil, water or air, which contain a "soup of biodiversity" (Taberlet et al. 2012a; Yu et al. 2012).

eDNA metabarcoding (Taberlet et al. 2012b) relies on the amplification of specific DNA barcode regions of the extracted DNA using universal primers. The standard metabarcoding approach consists of several steps that involve processing of environmental samples (water, soil, sediment) to obtain DNA sequences of organisms present in those samples. These steps include: (1) the isolation of eDNA, (2) the PCR amplification of the DNA barcode targeting the biotic community to be analyzed, followed by (3) HTS of obtained amplicons, (4) filtering of sequence data to remove sequencing errors and dereplication of identical sequences to obtain Individual Sequence Units (ISU), (5) the

clustering of the ISUs, (6) the clustering of Molecular Operational Taxonomic Units (MOTUs) and (7) the assignation of MOTUs to morphotaxa. The taxa list can then be used to calculate biotic indices.

Under the current rapid development of HTS technologies and metabarcoding methods for diatom identification, morphological methods for species identification seem to lag behind DNA methods in terms of timely and efficient analyses of large sets of samples (Apotheloz-Perret-Gentil et al. 2017; Hajibabaei et al. 2011; Kermarrec et al. 2013b; Vasselon et al. 2017b; Visco et al. 2015; Zimmermann et al. 2015). There are studies that are already calibrating indices under a taxonomy-free approach (Apotheloz-Perret-Gentil et al. 2017; Visco et al. 2015).

In order to compare the performance of morphological and metabarcoding approaches regarding identification and quantification of abundances of diatoms, the objectives of this study are: 1) to evaluate diatom richness from environmental samples by morphology and metabarcoding; 2) to create a regional specific morphological and molecular taxonomic reference library to aid in the assignment of the HTS data; 3) to compare taxon abundances obtained from morphology and HTS; and 4) to test the suitability of HTS data to retrieve barcode sequences.

## 4.3 Methods

### 4.3.1 Study area

The Lerma-Chapala Basin is located in Central Mexico, covering an area of 53,590 $km^2$. It lies within two biodiversity hotspots, namely Mesoamerica and the Madrean Pine–Oak Woodlands (Cotler et al. 2006; Myers et al. 2000; Sloan et al. 2014). It is geologically and climatically heterogeneous, and has well defined rainy (June to October) and dry seasons (November to May). This basin is one of the most important centers in the country for agriculture and industry, and has a population of more than 15 million inhabitants, but the basin is also one of the most environmentally degraded basins in the country (Aparicio 2001; Cotler et al. 2006; Wester et al. 2005).

### 4.3.2 Sampling

Epilithon samples were collected in nine streams from the Lerma–Chapala Basin, Central Mexico (Fig. 1, Table 1). Each sampling site was sampled twice, once in February 2014 and once in September 2014. For more details of the study area and of the physical and chemical composition of the waters, refer to Mora et al. (2017).

Each epilithon sample was collected from five cobbles across a transversal section of the stream, brushing with a disposable toothbrush ten square centimeters of epilithic growth from each of the five cobbles to make a composite sample, suspended in a total volume of 60 ml. The sample was homogenized and divided into three subsamples of 20 ml each: a) deep frozen (-24°C) for HTS; b) for the establishment of clone cultures to build a regional morphological and molecular taxonomic reference library; c) fixed in 70% alcohol for morphological analyses.

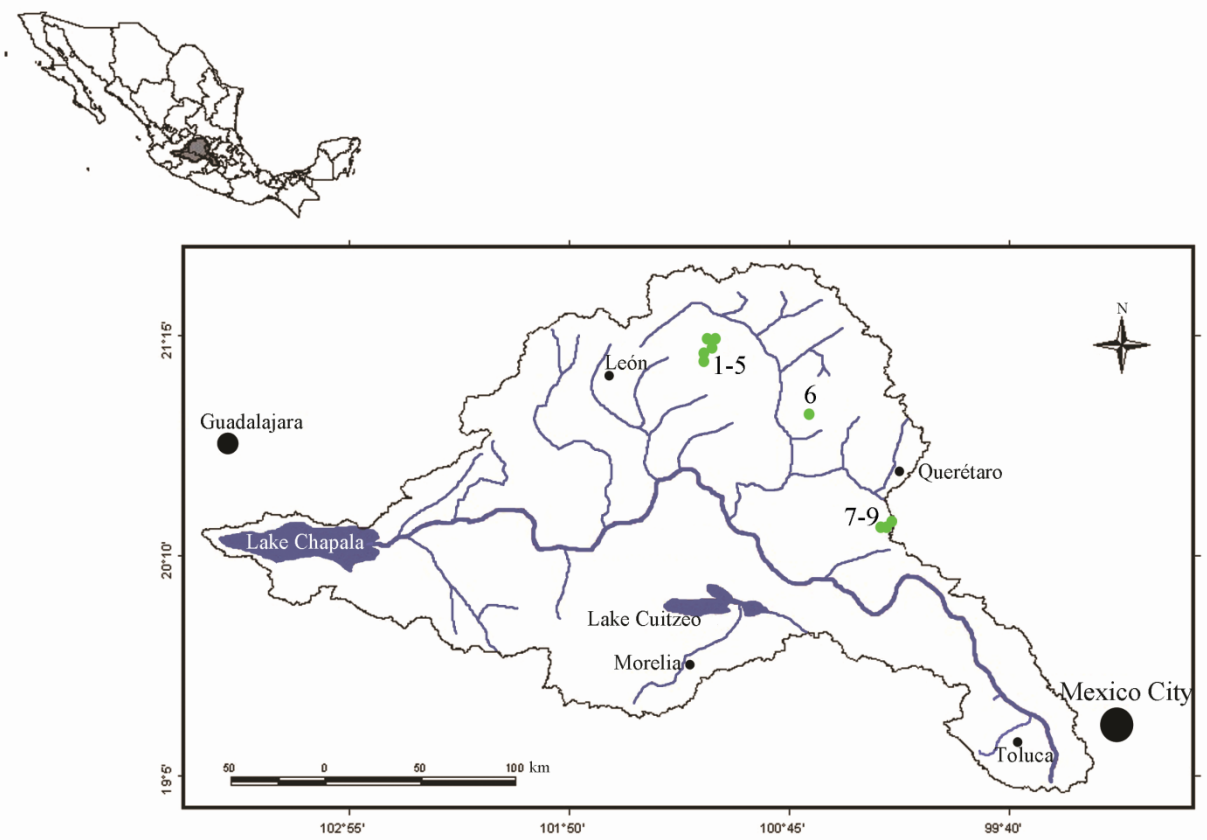



Figure 1. Location of the nine sampling sites collected within the Lerma-Chapala Basin, in Central Mexico, indicated by green dots. The numbers next to the green dots refer to the name of the sampling site in Table 1.

### 4.3.3 Morphological analysis from environmental samples

Sample and slide preparation, observations at the light microscope (LM) and scanning electron microscope (SEM), taxa identification and counting to determine abundance were performed as in Mora et al. (2017).

Table 1. Streams sampled in the Lerma-Chapala Basin, Central Mexico, including site number, name, geographical coordinates, elevation and sample numbers collected at each stream.

| Site | | Latitude (N) | Longitude (W) | Elevation (m a.s.l) | Samples |
|---|---|---|---|---|---|
| 1 | La Mesa | 21° 05' 28.69" | 101° 08' 18.98" | 2215 | 1, 10 |
| 2 | Calvillo | 21° 06' 50.40" | 101° 08' 04.10" | 2138 | 2, 11 |
| 3 | Peña Colorada | 21° 09' 03.84" | 101° 05' 58.96" | 2110 | 3, 12 |
| 4 | Paredones | 21° 11' 20.60" | 101° 06' 53.40" | 2089 | 4, 13 |
| 5 | La Laborcilla 1 | 21° 11' 04.70" | 101° 06' 14.60" | 2076 | 5, 14 |
| 6 | El Membrillo | 20° 50' 21.22" | 100° 38' 43.46" | 2114 | 6, 15 |
| 7 | Los Ailes 1 | 20° 19' 58.72" | 100° 15' 17.09" | 2358 | 7, 16 |
| 8 | Laguna de Servín 1 | 20° 18' 18.10" | 100° 17' 38.10" | 2409 | 8, 17 |
| 9 | Laguna de Servín 2 | 20° 18' 45.20" | 100° 17' 25.60" | 2409 | 9, 18 |

### 4.3.4 Taxonomic Reference Library of the Lerma-Chapala River Basin

#### 4.3.4.1 Isolation, cultivation and harvesting of clonal cultures

Single-cell isolations were performed from aliquots of environmental samples (subsamples b) using micro-capillary glass pipettes, mainly under the light microscope but also under a stereo light microscope. Samples were diluted to decrease diatom density as well as other algae and flagellates. Series of isolation and re-isolation on microscope slides were performed in order to ensure only single-cell isolations. The isolates were then transferred to 5 cm diameter Petri dish containing the culture medium AlgaGrow (Plagron Weert, Netherlands). The culture medium was used at the final concentration recommended by the manufacturer but also at a half and a quarter of that concentration. The cultures were grown in Memmert® Growth Chambers at 17-20 °C and a 12h day/night photoperiod. After a successful clonal culture had been established, the culture was divided into three subsamples: A) for DNA extraction; B) reserve and C) for morphological analysis. Each of the subsamples was transferred to an independent Petri dish, maintained in the growth chamber for one to three weeks and then harvested.

#### 4.3.4.2 Molecular analysis of clonal cultures

The cultured material from subsamples A was transferred to 15 mL plastic centrifuge tubes, centrifuged at 2000 rpm for 10 minutes, the supernatant was removed and the pellet was

transferred to 1.5 mL tubes. DNA was isolated using NucleoSpin® Plant II Mini Kit (Macherey and Nagel, Düren, Germany) following the product instructions. DNA concentrations were checked using gel electrophoresis (1.5% agarose gel) and Nanodrop (PeqLab Biotechnology LLC; Erlangen, Germany). DNA samples were stored at −20°C for future use. The V4 region of the 18S locus was amplified using the primers and PCR regime from Zimmermann et al. (2011). PCR products were visualized in 1.5% agarose gel and cleaned with MSB Spin PCRapace® (Invitek LLC, Berlin, Germany) following the product instructions. DNA concentrations were measured using Nanodrop® (PeqLab Biotechnology) and samples were normalized to a total DNA content >100 ng µL−1 for sequencing. M13 tails were used as sequencing primers following Zimmermann et al. (2014) and Ivanova et al. (2007). Sanger sequencing conducted by Starseq (GENterprise LLC; Mainz, Germany). The sequences were edited in PhyDE (Müller et al. 2005).

As with subsamples A, reserve subsamples (B) were centrifuged and transferred to 1.5 mL tubes. The tubes are stored in the deep-freezer (-24°C) as reserve material in case further DNA extractions are needed.

#### 4.3.4.3 Morphological analysis of clonal cultures.

The cultivated material from subsamples C was transferred to 15 mL plastic centrifuge tubes and filled with 35% hydrogen peroxide in order to oxidize the organic material. Two days later, the peroxide remnants were removed by rinsing four times with distilled water, with one day between every rinse. With the cleaned samples, one permanent slide per sample was made using the high refraction index mounting medium Naphrax®. The slides were observed and the diatoms photographed under the light microscope (LM), using a Zeiss Axio Imager.M2 with an implemented AxioCam HRc (Zeiss, Oberkochen, Germany). Aliquots of cleaned sample material for scanning electron microscopy observations were air dried, mounted on stubs and observed under a Hitachi FE 8010 (Hitachi, Tokyo, Japan) scanning electron microscope (SEM) operated at 1.0 kV. Taxon identification was conducted using the same identification references from the Supplementary Material as by Mora et al. (2017). This reference library was established not only from the 18 samples analyzed in this study, but from all the samples studied by Mora et al. (2017). Accession numbers pending.

### 4.3.5 HTS from environmental samples

Samples were defrosted and transferred to 15 mL plastic centrifuge tubes, centrifuged at 5,000 rpm for 5 minutes, the supernatant was removed and the pellet was transferred to 1.5 mL tubes. DNA was isolated using NucleoSpin® Plant II Mini Kit (Macherey and Nagel, Düren, Germany) following the

product instructions. DNA concentrations were quantified using a Qubit 2.0 fluorometer (Invitrogen, Carlsbad, California, USA) and adjusted to a concentration of 20 ng/µL for PCR. The amplification of the hypervariable V4 locus was done using nextera primers DIV4for: 5'-GCGGTAATTCCAGCTCCAATAG-3' and DIV4rev3: 5'-CTCTGACAATGGAATACGAATA-3' from Zimmermann et al. (2011) with the modification of Visco et al. (2015) to fit for sequencing at Illumina MiSeq. PCR amplifications were performed in duplicate per sample, each in a total volume of 25 µL: 0.5 µL dNTP mix (25 mM each dNTP), 0.25 µL BSA (10 mg/mL), 0.25 µL DMSO, 1 µL of each forward and reverse primers (10 pm/µL), 0.4 µL of Herculase II Fusion DNA Polymerase (Agilent Technologies Inc., Santa Clara, California, USA), 5 µL Herculase II reaction buffer, 1 µL of template DNA (20 ng/µL) and 15.6 µL of HPLC grade water. The PCR regime included an initial denaturation at 94°C (2 minutes), then 35 cycles consisting of denaturation at 94°C (45 seconds), annealing at 52°C (45 seconds), elongation at 72°C (1 minute) and a final elongation at 72°C (10 minutes). PCR products were visualized by electrophoresis on 1% agarose gels. Duplicates of PCR products were pooled into a final volume of 50 µL. Aliquots of 25 µL of the amplicons were purified using HighPrep PCR paramagnetic beads (Magbio Genomics, Gaithersburg, Maryland, USA). A second PCR (indexing PCR) run was conducted, in order to ligate a unique combination of tags to the 5´ end of the primer. Indexing PCR reactions of 25 µL were conducted as follows: 0.25 dNTP mix, 1 µL DMSO, 0.625 µL of each primer, 0.25 µL of Herculase, 5 µL Herculase II reaction buffer, 10 µL of template DNA and 7.25 µL of HPLC grade water. The indexing PCR regime started with denaturation at 94°C (2 minutes), then 8 cycles consisting of denaturation at 95°C (20 seconds), annealing at 52°C (30 seconds), elongation at 72°C (30 seconds) and a final elongation at 72°C (3 minutes). Products were purified using HighPrep PCR paramagnetic beads and quantified using Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, Carlsbad, California, USA). Library preparation was performed using MiSeq Reagent Kit V3 (Illumina, San Diego, California, USA) following manufacturer instructions.

Bioinformatics analysis were performed using MetBaN (Proft et al. 2017), a bioinformatic pipeline which implements a modular and flexible phylogenetic based species delimitation approach by streamlining metabarcoding and phylogenetic software packages. Within MetBaN only the first modules were used, most of them implemented from the OBITOOLS package (Boyer et al. 2016). Thus the samples consisting of paired-end reads were initially merged using "illuminapairedend". Only merged reads longer than 150 bp and with complete primer sequences were retained. Subsequently primers were removed with the "ngsfilter" module function, and identical sequences were then merged in order to prevent redundant classification. Sequences that appeared only once in all the samples were filtered out, in addition to chimeric and low quality sequences. Finally, all filtered sequences from the 18 samples were pooled and clustered into Molecular Operational Taxonomic Units (MOTUs) at 6 bp difference identity threshold (98%). Subsequently taxonomic

assignation was performed by matching to the EMBL nucleotide sequence database (Kanz et al. 2005), MOTUs without hits were retained as unclassified.

From the taxonomic assignments obtained from the comparison against the EMBL database, datasets of sequences were created in order to refine the identifications based on a phylogenetic-based coalescent model approach (PCMA) after Zimmermann et al. (2015), in order to identify taxon boundaries from variation in branching rates of a tree (Monaghan et al. 2009). The datasets were constructed for: 1) Achnanthidiaceae; 2) Bacillariaceae; 3) Centrales; 4) Cocconeidaceae; 5) Cymbellales; 6) Eunotiaceae; 7) Fragilariophycidae; 8) Mastogloiales; 9) Naviculales, excluding *Caloneis*, *Mayamaea*, Pinnulariaceae and Stauroneidaceae; 10) *Caloneis*, *Mayamaea*, Pinnulariaceae and Stauroneidaceae; 11) Surirellales-Rhopalodiales; and 12) Bacillariophyta. For each dataset, the respective sequences produced in this study, from the taxonomic reference library, as well as the BGBM Diatom Sequence Reference Database (unpublished) were added, as well as annotated diatom sequences from the NCBI nucleotide database. The datasets were aligned using the software MEGA (Tamura et al. 2013) and the implemented MUSCLE (Edgar 2004) alignment algorithm following the recommendations from Zimmermann et al. (2015). Alignments were visualized and manually improved in PhyDE (Müller et al. 2005).

The phylogenetic analyses were conducted by Maximum Likelihood as implemented in RAxML (Stamatakis 2006; 2014; Stamatakis et al. 2008) using the CIPRES platform (Miller et al. 2010). The model of sequence evolution used was the general time reversible (GTR) with gamma distribution (Γ) and a proportion of invariable sites (I) (Tavaré 1986), with 1000 replicates for the bootstrap analysis. The taxonomic assignment of the MOTUs to infrageneric level was done as in Zimmermann et al. (2015), with well-supported clades (≥60) considered independent taxa, and when a morphological correlation was possible according to the recorded diatom composition from environmental samples and clonal cultures. A 95% identity threshold was set for genus assignation.

#### 4.3.5.1 Comparison of abundance

The five taxa with the highest abundance from morphology as well as HTS were compared to their respective molecular or morphological counterparts with bar graphs. In order to make pair-wise comparisons, the abundances from microscopy counts as well as read abundances were transformed to relative abundance. In order to assess improvements in taxa quantification, the abundance correction factors (CF) for HTS data inferred from cell biovolume proposed by Vasselon et al. (submitted) we applied, even though there are taxonomical differences but similarities in size and therefore in biovolume. For *Achnanthidium* sp. 1+5 and *Achnanthidium* aff. *catenatum*, the CF for *Achnanthidium minutissimum* was applied, since these taxa have similar size. The CF for *Cocconeis*

*placentula* was used for *Cocconeis* sp. 2 and the CF for *Navicula cryptotenella* was applied to *Navicula notha*.

## 4.4 Results

### 4.4.1 Taxonomic Reference Library of the Lerma-Chapala River Basin

A total of 188 clonal cultures were established, resulting in the identification of 70 taxa in 24 genera. When taking into account the cultures isolated from the 18 samples analyzed in this study through microscopy and HTS only, from 111 strains 45 taxa were identified belonging to 21 genera (Supplementary material 1 – Appendix 2).

From the 188 sequences produced for the 18S V4 locus, 100 sequences are novel. The nine sequences generated from cultures of *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot are the first records of this genus for the 18S locus in the INSDC databases (DDBJ, EMBL-EBI and NCBI). Apart from *Simonsenia* Lange-Bertalot, there are other diatom genera underrepresented in the INSDC databases, with less than 10 entries for the 18S locus, listed below. The two sequences generated in our study for *Brachysira altepetlensis* D. Mora, R. Jahn et N. Abarca are added to the single sequence for this genus in INSDC databases. Regarding *Diademis* Kützing, there are 3 sequences of the genus in INSDC, our study contributing four new sequences generated for *Diadesmis confervacea* Kützing. There is another sequence listed as belonging to *Diadesmis* in INSDC, i.e. *Diadesmis gallica* W.Smith, but this taxon has been recently transferred to *Humidophila* (Lange-Bertalot & Werum) R.L.Lowe, Kociolek, J.R.Johansen, Van de Vijver, Lange-Bertalot et Kopalová (Lowe et al. 2017). Regarding *Diploneis* (Ehrenberg) Cleve, the two sequences generated here for *Diploneis* sp. add to the existing five sequences of this genus in INSDC. The sequence generated for *Nupela wellneri* (Lange-Bertalot) Lange-Bertalot adds to other two sequences for this genus in INSDC databases. Finally, the two sequences generated here for *Tryblionella* W. Smith, one for *Tryblionella calida* (Grunow) D.G. Mann and the other for *Tryblionella hungarica* (Grunow) D.G. Mann, add to five existing sequences for the genus in INSDC.

### 4.4.2 Morphological analysis from environmental samples

A total richness of 148 taxa (species and varieties) in 38 genera was found while performing the counts to determine abundance (Table 2). The most abundant taxa (relative abundance ≥ 2%) across all samples were, in decreasing order, *Achnanthidium* sp. 5, *Achnanthidium* sp. 1, *Gomphonema parvulum* (Kützing) Kützing, *Cocconeis* sp. 2, *Achnanthidium* aff. *catenatum* (J.Bílý et Marvan) Lange-Bertalot, *Mayamaea permitis* (Hustedt) Bruder et Medlin, *Fragilaria austriaca* (Grunow) Lange-

Bertalot, *Planothidium victori* Novis, Braidwood et Kilroy, *Gomphonema lagenula* Kützing, *Craticula subminuscula* (Manguin) C.E. Wetzel et Ector, *Reimeria sinuata* (W. Gregory) Kociolek et Stoermer and *Planothidium cryptolanceolatum* R. Jahn et N. Abarca (Supplementary material 2 – Appendix 2). Fifty-seven additional taxa were observed by scanning the whole slides looking for rare taxa after the counts, elevating the richness to 205 taxa in 43 genera. The additional five genera were *Achnanthes* Bory and *Neidium* Pfitzer only observed in LM; *Thalassiosira* Cleve was only observed under SEM; *Cymbella* Agardh and *Iconella* Jurilj were found both under LM and SEM (Table 2).

### 4.4.3 Morphological diversity from environmental samples and clonal cultures

As stated above, a total richness of 205 taxa was observed in the 18 environmental samples analyzed from diatoms counts under LM and after a thorough revision of slides and stubs under LM and SEM respectively. From the same samples, 45 taxa were identified from clonal cultures, but only 33 taxa from those 45 taxa were found in the observations of environmental samples. This result increases the total richness to 217 taxa in 44 genera (Figure 2A, Table 2, Suppplementary material 2). *Simonsenia* was the only genus which was isolated and cultured but neither observed by LM nor by SEM in environmental samples.

### 4.4.4 HTS from environmental samples

The Illumina MiSeq sequencing run generated 2,738,628 reads from the 18 libraries sequenced. After singleton and chimera deletion, 1,156,360 quality reads were retained, clustering the reads into MOTUs (6 base-pair similarity threshold or 98%) with at least two reads. A total of 43,703 (3.8 %) of those reads correspond to diatom sequences according to the BLASTn (Altschul et al. 1990) conducted against the EMBL nucleotide database. Only diatom sequences were further analyzed. A total of 2181 MOTUs were obtained from those 43,703 reads.

The taxonomic assignment following the phylogenetic-based coalescence model approach (PCMA) resulted in 350 units from here on called taxa (infrageneric). In order to further remove potential sequencing noise, taxa made up solely of one doubleton or one tripleton, and without any correlation to morphology or to a reference sequence were removed, reducing the number of taxa to 331 in 35 genera. There were 65 taxa that could not be assigned to already described genera according to the threshold of 95% identity, further reducing the richness to 266 taxa in 35 genera. From those 266 taxa, 94 were assigned a specific epithet because a morphological correlation was possible, or a correlation to a sequence from our own reference databases or correlation to sequences downloaded from NCBI in well-supported (≥ 60) clades.

The most abundant taxa (sequence relative abundance ≥ 2%) across all samples were, in decreasing order, *Gomphonema parvulum* sensu lato, *Navicula notha* J.H. Wallace, *Cocconeis* sp. 2, *Nitzschia palea* (Kützing) W. Smith, *Ulnaria* cf. *ulna* (Nitzsch) Compère, *Nitzschia* cf. *linearis* (Agardh) W. Smith and *Cocconeis* sp. 8 (Supplementary material 2 – Appendix 2).

### 4.4.5 Diatom composition inferred from morphology and HTS

In most cases, HTS recovered more taxa than the morphological approach: 217 taxa by morphology and 266 by HTS (only one third assigned to morphologically identified species). At the genus level, 45 genera were recovered by morphology whereas 35 genera were identified by HTS. From the genera identified by morphology, 14 were not recovered by HTS. On the other hand, five genera were recovered by HTS but not by morphology, both methods detecting 49 genera in total (Figure 2B, Table 2). The combination of the total morphological richness of 217 taxa with the 94 infrageneric taxa assigned from HTS data, resulted in 250 taxa. From those taxa, 62 were detected by both methods. The here presented regional reference library allowed the identification of 23 taxa that otherwise would have been left with assignation to the genus level only (Figure 2C).

Table 2. Richness of infrageneric taxa detected by microscopy and HTS in 18 samples from streams of the Lerma-Chapala Basin, Central Mexico. M1 = obtained from LM counts ; M2 = after LM counts and/or SEM; TRL = taxonomic reference library for the Lerma-Chapala Basin; M3 = combined categories M1, M2 and TRL); HTS = High-Throughput Sequencing.

| **Genus** | **M1** | **M2** | **TRL** | **M3** | **HTS** |
|---|---|---|---|---|---|
| *Achnanthes* | - | 1 | 1 | 1 | - |
| *Achnanthidium* | 8 | 10 | 3 | 10 | 14 |
| *Amphora* | 1 | 1 | - | 1 | 2 |
| *Anomoeoneis* | - | - | - | - | 1 |
| *Brachysira* | 3 | 5 | 1 | 5 | 1 |
| *Caloneis* | 4 | 5 | 1 | 5 | 1 |
| *Chamaepinnularia* | 2 | 2 | - | 2 | - |
| *Cocconeis* | 2 | 3 | - | 3 | 10 |
| *Craticula* | 4 | 5 | - | 5 | 8 |
| *Cyclostephanos* | 1 | 1 | - | 1 | - |
| *Cyclotella* | 2 | 2 | - | 2 | 3 |

| Genus | M1 | M2 | TRL | M3 | HTS |
|---|---|---|---|---|---|
| *Cymbella* | - | 1 | - | 1 | 4 |
| *Cymbopleura* | 1 | 1 | - | 1 | 2 |
| *Diadesmis* | 1 | 1 | 1 | 1 | 1 |
| *Diatoma* | - | - | - | - | 3 |
| *Encyonema* | 8 | 9 | 1 | 9 | 9 |
| *Encyonopsis* | 2 | 3 | - | 3 | - |
| *Eolimna* | 3 | 3 | - | 3 | - |
| *Epithemia* | 3 | 3 | - | 3 | 3 |
| *Eunotia* | 4 | 12 | 1 | 12 | 7 |
| *Fistulifera* | 1 | 1 | 1 | 1 | 8 |
| *Fragilaria* | 3 | 5 | 2 | 5 | 17 |
| *Frustulia* | 2 | 4 | - | 4 | - |
| *Geissleria* | 1 | 1 | - | 1 | 1 |
| *Gomphonema* | 15 | 20 | 5 | 22 | 22 |
| *Halamphora* | 3 | 3 | - | 3 | - |
| *Humidophila* | 1 | 1 | - | 1 | - |
| *Iconella* | - | 1 | - | 1 | 1 |
| *Luticola* | 3 | 5 | - | 5 | - |
| *Mayamaea* | 3 | 3 | 2 | 4 | 7 |
| *Melosira* | - | - | - | - | 1 |
| *Navicula* | 14 | 15 | 2 | 16 | 42 |
| *Navigiolum* | 1 | 1 | - | 1 | - |
| *Neidium* | - | 3 | - | 3 | - |
| *Nitzschia* | 23 | 29 | 5 | 30 | 48 |
| *Nupela* | 1 | 2 | 1 | 2 | - |
| *Pinnularia* | 6 | 13 | 3 | 14 | 7 |
| *Planothidium* | 4 | 4 | 2 | 4 | 7 |
| *Pseudofallacia* | 1 | 1 | - | 1 | - |

| **Genus** | **M1** | **M2** | **TRL** | **M3** | **HTS** |
|---|---|---|---|---|---|
| *Reimeria* | 1 | 1 | - | 1 | 4 |
| *Rhopalodia* | 1 | 1 | - | 1 | 2 |
| *Sellaphora* | 11 | 14 | 8 | 19 | 7 |
| *Simonsenia* | - | - | 1 | 1 | - |
| *Stauroneis* | 1 | 3 | 1 | 3 | 1 |
| *Stephanodiscus* | - | - | - | - | 1 |
| *Surirella** | 1 | 3 | 1 | 3 | 4 |
| *Thalassiosira* | - | 1 | - | 1 | 1 |
| *Tryblionella* | - | - | - | - | 2 |
| *Ulnaria* | 2 | 2 | 2 | 2 | 14 |
| Σ | 148 | 205 | 45 | 217 | 266 |

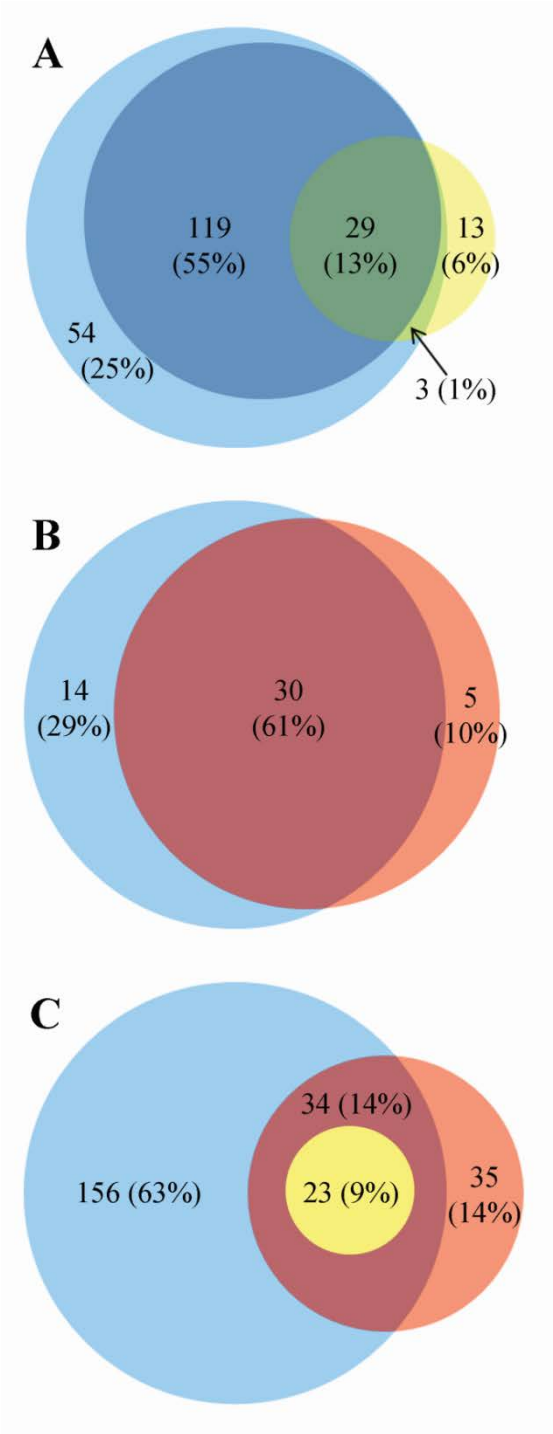



Figure 2. Venn diagrams comparing the proportion each identification method contributed to the total taxa richness recorded. A) Morphological richness across all samples and clonal cultures: taxa identified by counting 500 valves per sample under the LM (dark blue); taxa identified after additional scanning of the slides after the counts as well as taxa observed during SEM examinations (blue); taxa identified from clonal cultures isolated from the 18 samples focus of this study (yellow). B) Genera identified by morphology (blue) and metabarcoding (orange). C) Taxa identified by morphology and metabarcoding (only assigned taxa shown): taxa identified by morphology (blue); metabarcoding (red); taxa retrieved from metabarcoding that was assigned with the Lerma-Chapala taxonomic reference library (yellow).

### 4.4.6 Comparison of relative abundances

*Gomphonema parvulum*, *Navicula notha*, *Cocconeis sp. 2*, *Nitzschia palea* and *Ulnaria ulna* recorded the highest relative abundance of reads. In morphology, *Achnanthidium* sp. 5, *Achnanthidium* sp. 1, *Gomphonema parvulum*, *Cocconeis* sp. 2 and *Achnanthidium* cf. *catenatum* scored the highest relative abundance of valves. Since the barcoding marker employed did not discriminate among closely related taxa, the relative abundances of *Achnnanthidium* sp. 1 *Achnnanthidium* sp. 5 obtained by morphology were pooled. The same was done for *Gomphonema parvulum* since the barcoding marker did not differentiate between *Gomphonema exilissimum*, *G. parvulum* and *G. lagenula*. *Nitzschia palea* was also treated in broad sense, sensu lato. In all the resulting seven comparative graphs, there were disparities among the abundances for both methods (Fig. 3). After application of the correction factors to HTS abundance data, there were mixed results. On the one hand, there were reductions in the difference between morphology and HTS by 94% in *Navicula notha*, 95% in *Nitzschia palea*, 83% in *Ulnaria* cf. *ulna*, 7% in *Achnanthidium* aff. *catenatum* and 4% in *Achnanthidium* sp. 1+5 (Fig. 4). On the other hand, there were increases in the difference in abundance for two taxa, by as much as 191% in *Gomphonema parvulum* and 425% in *Cocconeis* sp. 2 (Fig. 4).

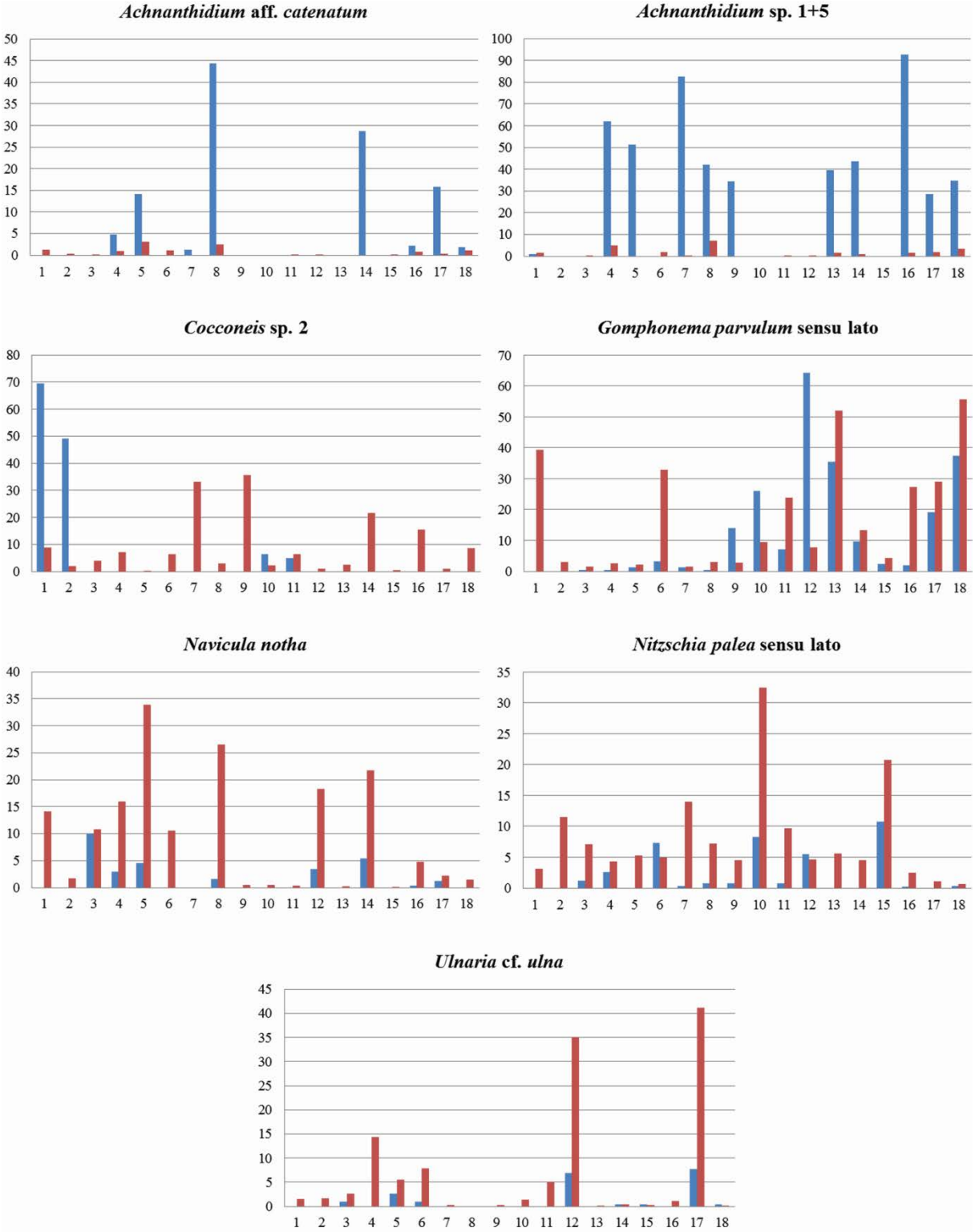



Figure 3. Relative abundance of the most abundant taxa obtained from morphology (blue) and metabarcoding (red) across samples. The vertical axis in each graph shows the relative abundance in percentage, whereas the horizontal axis shows sample number.

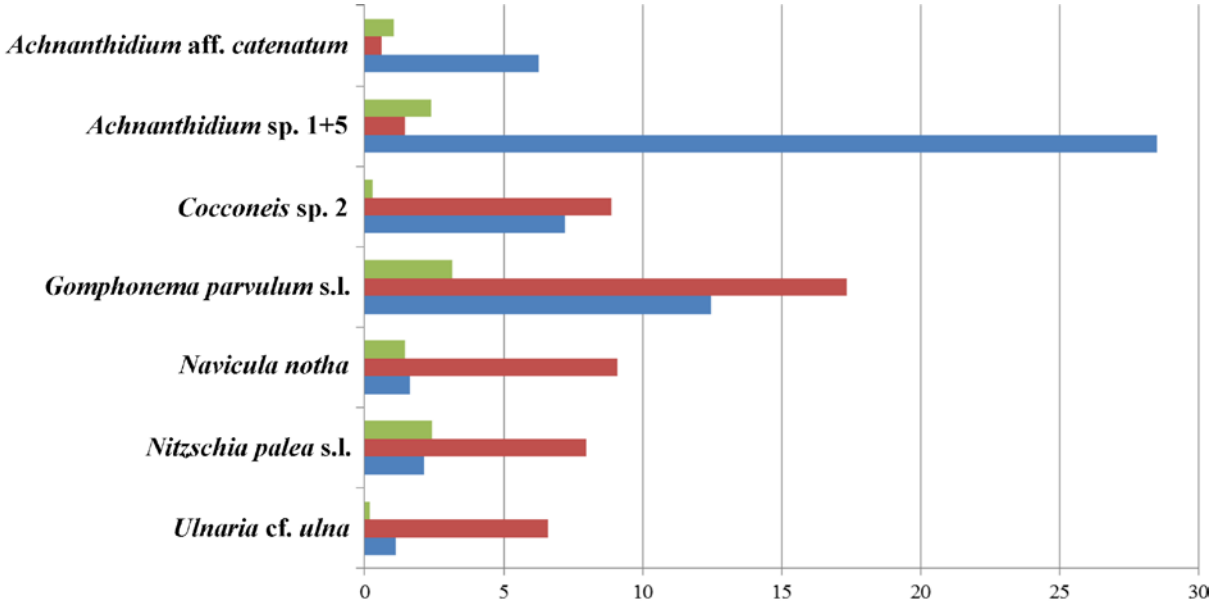



Figure 4. Cumulative relative abundance of the most abundant taxa obtained from morphology (blue), metabarcoding (red) and metabarcoding after application of correction factors (green).

### 4.4.7 HTS from environmental samples as a source of barcodes

After in-depth examination of HTS data obtained in our study, we are proposing 18SV4 barcode sequences for two taxa following some of the criteria proposed by Rimet et al. (2018) such as: 1) being among the most abundant sequences in the sample; and 2) phylogenetic neighbours' belonging to the same neighbour taxa expected from morphological observations; 3) identification of this taxon in microscopy examinations; 4) since 18S rDNA gene is a non-coding region, our sequences did not meet the criteria of Rimet et al. (2018) of neither indels nor insertions in the proposed sequences for their barcoding marker *rbc*L, which is a coding region. We therefore relaxed these criteria to a maximum of 1 indel/insertion after aligning with closely related taxa.

*Iconella delicatissima* (F.W.Lewis) Ruck et Nakov (Fig. 5 A and B)

≡ Stenopterobia delicatissima (Lewis) Van Heurck

The genus *Iconella* has been recently resurrected to accommodate *Stenopterobia* and the "robustoid" members of *Surirella* and *Campylodiscus* (Jahn et al. 2017b; Ruck et al. 2016a; Ruck et al. 2016b). Sample 18 was the only sample among our 18 samples to have sequences of *Iconella*. The morphological examination of this sample affirmed the morphological presence of *Iconella delicatissima*. *Surirella angusta* was the only other member of the Surirellales found in this sample,

but this species belongs to *Surirella* sensu stricto (Ruck et al. 2016a). A total of 52 sequence reads were obtained for this taxon from a total of 6873 reads obtained from sample 18 (Laguna de Servín 2), making it the 13$^{th}$ most abundant taxon in this sample. After blasting this sequence on NCBI, the three most similar sequences belonged to the genus *Stenopterobia*, with 98% similarity to the two most similar sequences (*Stenopterobia pumila* and *S. curvula*) and 95% of similarity to the third most similar sequence (*Stenopterobia* sp. 50). Other sequences with 95% similarity include a sequence of *Surirella* sp. and two sequences of *Campylodiscus levanderi*. After aligning our sequence with sequences of the before mentioned taxa, our sequence had only one insertion regarding the three available sequences available from *Stenopterobia* and neither indels nor insertions with respect to *Surirella* cf. *tenuissima*, which clustered together in the phylogeny Ruck et al. (2016b) as part of the resurrected genus *Iconella*.

*Navicula notha* Wallace (Fig. 5 C-G)

Sequences of this taxon were found in 17 out of the 18 samples analyzed here by HTS. The read abundance of this taxon across all samples was the second highest only after *Gomphonema parvulum* sensu lato. The LM and SEM examinations confirmed the morphological presence of *Navicula notha*. We selected samples 4, 5 and 14 as source of barcodes, since in those samples this taxon was the only *Navicula* representative found during the valve counts and in those samples it reached high relative abundances, 2.9%, 4.5% and 5.4% respectively. The relative abundance of sequences in those samples was as high as 16%, 34% and 22%. After blasting the sequences from this taxon on NCBI, the closest sequence corresponded to *Navicula cryptotenelloides*, with a four base-pair difference (99%). Other sequences with high similarity (98%) to our sequence include *Navicula cryptotenella*, *N. reinhardtii*, *Navicula* sp. AT-201Gel01 and *Hippodonta capitata*. We discard the possibility of our here proposed barcode sequence to be a sequence from *Navicula cryptotenelloides* or *N. reinhardtii* because those taxa were not observed morphologically in our samples. We also discard the possibility of our sequence being from *Navicula cryptotenella* because even though this taxon was observed by microscopy after the counts, it should have been detected in several sites and with relatively high abundances to support the appearance of such a large number of sequence reads as the ones observed. We rule out our sequence being *Hippodonta capitata* as well, since it was not observed by microscopy.

Our sequence aligns close to other *Navicula* sequences and neither indels nor deletions were observed after the alignment.

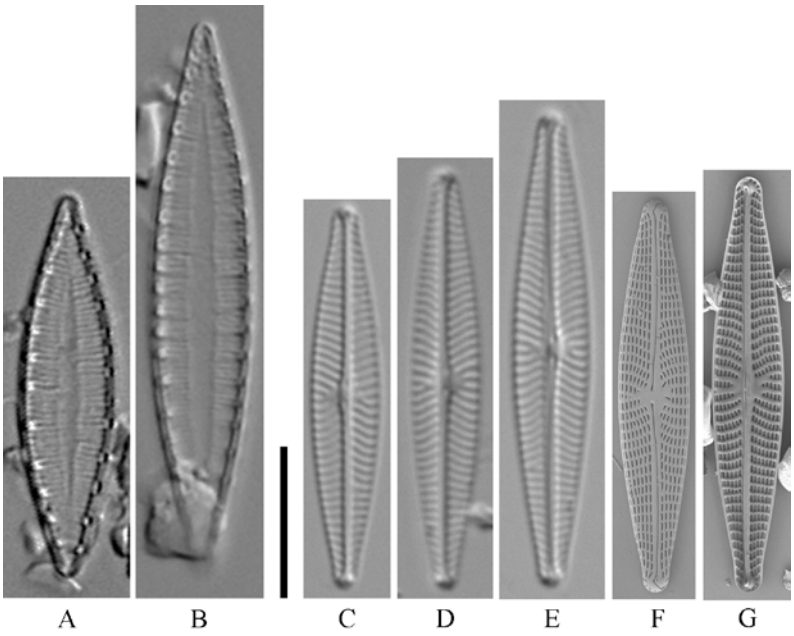



Figure 5. Taxa for which barcodes were retrieved from HTS data. A- B *Iconella delicatissima*, LM, from Laguna de Servín 1, sample 18. C-G *Navicula notha*, LM and SEM, from La Laborcilla 1, sample 5. Scale bar 10 µm.

## 4.5 Discussion

### 4.5.1 Taxonomic Reference Library of the Lerma-Chapala River Basin

The regional taxonomic reference library presented here is the first morphological and molecular characterization of stream diatoms from Mexico. The only published molecular characterization of epicontinental diatoms from Mexico are: three strains from *Gomphonema parvulum*, *Nitzschia* cf. *semirobusta* and *Pinnularia divergens* in the barcoding study from Zimmermann et al. (2011); two strains of *Gomphonema lagenula* in the phylogeographic study of Abarca et al. (2014); and the eight strains of *Planothidium lanceolatum* and *Planothidium victori* in the phylogenetic study of Jahn et al. (2017a). There are other entries in INSDC databases referring to diatoms from epicontinental locations of Mexico, but all of them correspond to uncultured material, for which no morphological correlation can be made since neither valves nor DNA for further studies are available. These facts highlight the importance of having vouchered material to allow traceability of the data, and the availability of reserve material for both morphological and molecular studies for other diatom studies, i.e. ongoing phylogenetic studies using strains from this study include the genera *Achnanthidium*, *Caloneis*, *Cocconeis*, *Diadesmis*, *Encyonema*, *Gomphonema*, *Pinnularia* and *Surirella* (pers. comm.)

From the 188 strains established here, 99 strains correspond to already described species, 48 strains were named as closely related (cf. *confer*) to already described species, nine as similar (aff. *affinis*) to other species, and 32 strains were only named to the genus level after a thorough morphological and bibliographical examination. It is highly possible that those unnamed strains should be described as new to science, as well as those where the abbreviations cf. and aff. were used. It is no surprise to have such a large fraction (47%) of unidentified taxa because our samples come from within the tropics, from Central Mexico, where no comprehensive identification monographs have been published yet. A similar proportion of unidentified taxa has been found for polar diatoms (58%) coming from a small sampling of 26 strains (Stachura-Suchoples et al. 2015). Even in thoroughly investigated regions like Berlin regarding its diatom flora 10% of the identified species by Zimmermann et al. (2014) were new to science.

### 4.5.2 Diatom composition detected by microscopy and HTS

The composition found after LM and SEM examinations from environmental samples is the same as that found by Mora et al. (2017) since in this study a subset of samples of that study was analyzed. The only addition to the flora reported by Mora et al. (2017) is *Thalassiosira weissflogii*, which was only found in the material prepared for SEM.

The overall microscopy (LM and SEM) observations from 18 environmental samples led to the identification of 205 taxa in 43 genera. When only taking into account the counts of 500 valves at the LM to determine abundance, the richness falls to 148 taxa (72%). Our results indicate that basing diversity evaluations solely on valve counts of the traditional fixed number of 400 or 500 valves considerably decreases the diversity, in our case by 28 %. Similar results were found by Von Falkenhayn (2008), with her findings indicating that the number of valves which need to be counted in order to catch the whole diversity in a sample depends on the sample, with some samples needing 100 or 300 valve counts to get 100% of the diversity, but most samples requiring up to 600 valve counts to capture the whole diversity. The results from Von Falkenhayn (2008) showed that the commonly used 400-500 valve counts in diatom studies detects between 70% to 100% of the diversity dependent of the sample, mostly around 80-90%. The commonly used 400 valve counts, particularly used in monitoring studies, are adequate for monitoring, since indices are based on the most abundant taxa (Bate and Newall 1998). Rare species, normally those with less than 1 or 2% or relative abundance are discarded in the calculation of most indices. But in diversity assessments, our results and those from Von Falkenhayn (2008) indicate that 500 valve counts display a fraction of the diversity (72% in our study). So it is highly recommended to do further scanning of the slides, e.g. using a fixed amount of time as in Apotheloz-Perret-Gentil et al. (2017) if the same scanning effort is

intended to make all samples comparable in terms of scanning effort, otherwise it is recommended to scan at least one or two whole slides after the counts as it was done in this study.

The taxonomic composition after the PCMA from HTS results showed a larger diversity than that resulted from morphological analyses, 331 taxa (266 assigned at genus level plus 65 not assigned to a genus) from HTS versus 217 taxa identified morphologically. When only taking into account the assigned taxa (94), two thirds correspond to taxa found by microscopy in this study, whereas for one third there were sequence similarity that allowed assignation but no morphological matches were observed in the microscopy examination. But most of the taxa not found here by microscopy have been found in other locations of the basin (Abarca-Mejía 2010; Mora et al. 2015; Segura-Garcia et al. 2012). Even though it still represents a fraction of the diatom diversity of the region, the reference library presented herein represents a milestone for stream diatoms of Mexico. In our study, it allowed the taxonomic assignation of 23 taxa recovered from HTS and PCMA that otherwise would have been left unassigned.

Although it can be argued that HTS can lead to an inflation/overestimation of diversity since it normally retrieves more MOTUs than morphologically identified taxa (Vasselon et al. 2017b; Zimmermann et al. 2015), eDNA metabarcoding can be regarded as a good initial biodiversity screening, showing the gaps in the hitherto given floras or reference libraries leading to further refined diversity assessments. The main reason cited in the literature for incongruence of taxa lists obtained from morphology and HTS analyses is the incompleteness and lack of accuracy of reference databases that impedes correct taxonomic assignment of environmental sequences. Taxa absent from databases could never be identified in environmental sequences, while sequences with wrong taxonomy in databases will generate wrong identifications (Kermarrec et al. 2014; Lejzerowicz et al. 2015; Visco et al. 2015; Zimmermann et al. 2014).

### 4.5.3 Richness overestimation

Even after a thorough morphological examination of samples by microscopy, the taxa list retrieved by HTS (266 genus-assigned plus 65 unassigned to genus) was larger than the 217 morphology-based list. Generally, the number of MOTUs generated by eDNA sequencing considerably deviates from the number of taxa observed morphologically in the same environmental samples (Cowart et al. 2015; Groendahl et al. 2017; Pawlowski et al. 2014). There are several biological, environmental and technical factors that contribute to this over- or underestimation of taxonomic richness in metabarcoding data.

The most important biological factor that influences richness overestimation is the natural intraspecific and intragenomic variability of the barcoding marker. This is particularly problematic when a single traditionally recognized species or bioindicator taxon comprises a variety of different genotypes (Balint et al. 2016; Brown et al. 2015). Sequences corresponding to different genotypes within the same taxon may cluster into different MOTUs, and thus artificially inflate taxonomic richness. High intraspecific genetic variations are well documented in practically all bioindicator groups such as aquatic insects (Alp et al. 2012; Elbrecht et al. 2014; Previšić et al. 2009; Sweeney et al. 2011) and diatoms (Rimet et al. 2014; Rynearson and Armbrust 2000; Trobajo et al. 2009). Moreover, taxa show high intragenomic polymorphism, such as nematodes (Bik et al. 2013), foraminifera (Weber and Pawlowski 2014) and prokaryotes (Sun et al. 2013), additionally contributing to the increase of MOTUs number.

Another factor that can lead to richness overestimation is the presence of so-called “ghost” MOTUs, corresponding to the taxa represented by extracellular DNA only. Free DNA molecules can be preserved for a long time in aquatic ecosystems, either bound to the sediment (Mao et al. 2014; Turner et al. 2015; Torti et al. 2015) or transported in water over large distances (Deiner and Altermatt 2014).

MOTU richness can also be artificially inflated through technical errors at different steps of sample processing. Most of these errors are generated during PCR amplification and amplicon sequencing. One of the most commonly cited causes of richness overestimation are the formation of chimeric sequences during PCR amplification (Balint et al. 2016; Fonseca et al. 2012), yet the use of a proofreading polymerase in amplification, as done in our study, reducing the overestimation of MOTUs by 15% (Oliver et al. 2015). Tag switching can be another source of chimera when combining amplicons from different samples (Balint et al. 2016; Carlsen et al. 2012; Esling et al. 2015; Schnell et al. 2015) further increasing MOTUs richness. Technical errors can also be generated during the sequencing step (Meacham et al. 2011; Schirmer et al. 2015). On the other hand, DNA extraction methods, as reported in diatoms, do not affect MOTUs richness (Vasselon et al. 2017a).

The MOTU delimitation approach is another factor that can affect richness estimations and interpretations. The most common threshold for MOTUs delimitation relies on algorithms normally clustering sequences at similarity thresholds of 97-99% but also 95% (Kermarrec et al. 2014; Vasselon et al. 2017b). But MOTUs not necessarily correspond to species, failing to straightforwardly identify meaningful ecologically or phylogenetically units (Balint et al. 2016; Ryberg 2015). To ameliorate this, in some studies MOTUs are further analyzed by phylogenetic-based approaches to assign them to specific taxa (Monaghan et al. 2009; Visco et al. 2015; Zimmermann et al. 2015).

Microscopy-based assessments can also lead to richness overestimation because for identification diatoms are oxidized, with diversity estimations relying on the valves of dead cells. As with free DNA molecules, diatom valves can also be transported in lentic and lotic environments, leading to overestimate the real community composition (Potapova and Charles 2005; Sawai 2001). Gillett et al. (2011) estimated that the percentage of live diatom cells in rivers to be very variable, ranging from 2% to 98%. Gillett et al. (2009) also determined that diatom assemblages generated from alive or dead cells showing overall similarity, both correlating well with physical stream conditions. Even though the use of live diatoms in diversity assessments provides ecological reliability, the conventional method of identification of dead cells provides taxonomic confidence at species level, which is not achievable from live cell identification because the micro-morphological features of the diatom valves are not visible in live cells to the necessary extent (Gillett et al. 2009).

Another source of error in richness estimations relates to the sampling techniques, with the common use of toothbrushes when sampling benthos potentially leading to cross-contaminations. This was avoided due to single use of disposable toothbrushes in the samplings (Kelly and Zgrundo 2013).

### 4.5.4 Concealed diversity revealed by clonal culturing

The strains which were established from the 18 samples by clonal culturing and analyzed in this study, led to the identification of 45 infrageneric taxa, with 12 taxa in 7 genera not observed after a meticulous examination by LM and SEM. At the genus level, *Simonsenia* was isolated but not detected by microscopy. *Diploneis* is a further example of a genus that was isolated and cultured in the reference library presented here, but not reported by Mora et al. (2017) after thorough examination by LM and SEM. This leads to the question how much of the diversity in a sample remains concealed even after exhaustive microscopical examination. The 12 extra taxa, 6% of the entire diversity of 218 taxa, were only detected by culturing. This can be attributed to the selectivity that the culture media and culturing conditions (i.e. light, day/night cycle and temperature) exert on the cells that had abundances low enough to pass undetected after thorough microscopy examinations. The culturing conditions would make them reach abundances high enough to be picked by the person doing the isolations. It has been reported that culture media at very low concentrations, even at three orders of magnitude lower, can lead to the identification of a diversity concealed in morphological examinations and by standard culturing techniques in cyanobacteria from a hot spring (Ferris et al. 1996) and in marine bacteria (Connon and Giovannoni 2002). In our study, the culture media was not only used at the recommended concentration by the manufacturer, but also at half and a quarter of the recommended concentration. This could explain the concealed diversity of diatoms which is only revealed by clonal culturing.

### 4.5.5 Discrepancies in abundance data

Discrepancies in abundances retrieved by microscopy counts and sequence reads are an issue that has been debated in metabarcoding, particularly when it comes to the use of sequence numbers as a proxy in biological monitoring (Apotheloz-Perret-Gentil et al. 2017; Elbrecht and Leese 2015; Groendahl et al. 2017; Pawlowski et al. 2014; Vasselon et al. 2017b; Visco et al. 2015). The barcoding marker and its ability to discriminate among closely related species, as well as primer specificity, are of major importance and can hinder the otherwise straightforward use of sequence reads to determine species abundances (Elbrecht and Leese 2015; Elbrecht et al. 2017b). This is evident in our results, e.g. *Gomphonema* and *Nitzschia*, because our barcoding marker does not have the discriminatory power to differentiate among closely related species. As for *Gomphonema parvulum*, the existence of cryptic diversity has been demonstrated (Kermarrec et al. 2013a), with only a multi-marker phylogeny coupled with detailed micromorphology being able to disentangle this once considered cosmopolitan species (Abarca et al. 2014). For *Nitzschia palea*, both morphology and HTS retrieved three taxa each, but this taxon was also treated in a broad sense because it was not possible to correlate the nominate variety and the varieties *debilis* and *tenuirostris* obtained from the morphological analysis, to the three taxa retrieved from HTS. Morphological, genetic and mating studies of *Nitzschia palea* concluded this taxon is made up of three or more species, with molecular and mating experiments not separating taxa into the traditional varieties recognized morphologically (Trobajo et al. 2009; Trobajo et al. 2010).

The discrepancies in abundances observed in *Nitzschia palea* and *Gomphonema parvulum* have also consequences in terms of bioindication inference if using abundance data from morphology or HTS, since the varieties *palea* and *debili*s are regarded as indicators of different environmental conditions, with *Nitzschia palea* regarded as an indicator of eutrophic waters and tolerant to heavy metal pollution whereas *Nitzschia palea* var. *debilis* is an indicator of oligotrophic environments (Potapova and Charles 2007; Sabater 2000; Van Dam et al. 1994). Similarly, *Gomphonema parvulum* is regarded as indicator of eutrophic waters, tolerant to organic and heavy metal pollution while *Gomphonema exilissimum* as indicator of oligotrophic waters (Kelly and Whitton 1995; Sabater 2000; Van Dam et al. 1994).

Cell size is another determinant factor in the abundance disparity observed, as shown in our results for *Achnanthidium* and *Ulnaria*, which represent opposites in cell size. On one hand, both species of *Achnanthidium* from our example represent 28.5% of the total diatom abundance but make up only 1.5% of the total read abundance. On the other hand, *Ulnaria* cf. *ulna* represents 1.1% of the total valve abundance, whereas it makes up 6.6% of the total read abundance. A correlation between cell

size and biovolume and gene copies of the SSU rDNA has been suggested (Godhe et al. 2008; Zhu et al. 2005) and could explain why *Achnanthidium* was underrepresented and *Ulnaria* overrepresented in read abundances compared to the abundances obtained by morphology.

The findings of Vasselon et al. (submitted) indicate that correcting factors based on cell biovolume for diatoms considerably reduce the disparity between morphological and molecular abundance data in as much as 45%, both in mock communities and environmental samples. After the application of CFs proposed by Vasselon et al. (submitted), there were improvements in the difference between abundances obtained from morphology and HTS, with reductions in differences going from 4% to 97%, pointing out to the usefulness of CFs to improve taxa quantification. But there were also increases as large as 425% for *Cocconeis* sp. 2, indicating further improvements are needed to make comparable abundance data retrieved by morphology and metabarcoding.

During sample processing the taxonomic composition is mainly altered at the PCR step by differential primer efficiency, specificity and template competition. Studies comparing molecular and morphological taxonomic inventories in bulk samples found primer bias as primary source of variation and a common factor resulting in false negatives in metabarcoding data (Elbrecht and Leese 2015; Elbrecht et al. 2017a; Elbrecht et al. 2017b). Among the technical factors, the PCR is considered as the main source of quantitative biases. The final amount of sequences assigned to a given species is highly dependent on the number of amplicons generated during PCR reaction but primer efficiency differs between species (Elbrecht and Leese 2015; Kermarrec et al. 2013b; Piñol et al. 2015). Primer biases might also be responsible for preferential amplification of selected taxa that leads to a common situation when most of sequence reads belong to few species that are easily amplified compared to the others, with 80% of sequence reads originate from one taxon in Dowle et al. (2016). In our study no such an extreme was observed, but the 10 most abundant taxa accounted for 62% of sequence reads. The difference between highly abundant and rare taxa in molecular assessments can easily span several orders of magnitudes, impeding the correct quantitative analysis. Moreover, PCR primer efficiency likely differs between samples in response to the sampled community, resulting in incomparable results of molecular biodiversity and abundance assessments.

### 4.5.6 HTS as a source of barcodes

There are several challenges and limitations to establish complete barcode libraries, which currently rely mostly on clonal cultures. We can mention among others the time consuming process of single cell isolation, culturing, recalcitrance of some species and maintenance of the cultures, without mentioning that culturing can often be unsuccessful (Mann and Chepurnov 2004; Rimet et al. 2018). As an alternative, HTS has been proposed as a source of barcodes (Rimet et al. 2018). As it was

demonstrated in our study with *Iconella delicatissima* and *Navicula notha*, HTS can be used as a source of barcodes if data is carefully analyzed.

Retrieving *Iconella delicatissima* was easy since it was the only representative of the genus in the sample where it was found. In this sample there was another representative of the Surirellales, *Surirella angusta*, which might have hindered our findings, but for *Surirella angusta* we have reference sequences obtained via culturing and Sanger sequencing so it was impossible to misidentify it for *Iconella delicatissima*.

Retrieving *Navicula notha* was more complicated due to the high abundance of *Navicula* taxa in our samples, represented not only by sequences derived from HTS but also from microscopy observations. Sequences of this taxon were retrieved from 3 (4, 5 and 14) out of our 18 samples, since in those samples *Navicula notha* was the only representative of *Navicula* found after microscopy observations with relative abundances as high as 5.4% in sample 14.

## 4.6 Conclusion

Our study demonstrates that the combination of morphological (LM and SEM) and molecular (Sanger and HTS) methods applied to environmental samples, combined with a regional taxonomic reference library, increases the detection and identification of diatom species, highlighting the complementary aspects of classical taxonomy and eDNA metabarcoding, i.e. the importance of their reciprocal illumination. Even in the advent of big advances and successes in the development and standardization of molecular tools for biodiversity assessments and monitoring (Apotheloz-Perret-Gentil et al. 2017; Elbrecht et al. 2017a), the role of morphology in species discovery and detection remains central along with genomic advancements.

## 4.7 Acknowledgements

The work of DM was funded by the Mexican Government through doctoral grants from CONACYT, CONCYTEQ and DGRI-SEP. Samples were taken under permit CONAPESCA PPF/DGOPA-149/14. The Deutsche Forschungsgemeinschaft is thanked for Grant INST 130/839-1 FUGG concerning SEM funding. Maja Grubisic, Kirsten Richter, Susan Mbedi and Verena Deutschmeyer are gratefully acknowledged for their help with HTS laboratory procedures. Jana Bansemer is thanked for helping with diatom cultivation and retrieval of molecular data. Monika Lüchow, Kim Govers and Juliane Bettig kindly assisted at the SEM. Verónica Aguilar Zamora is thanked for creating the map.

## 4.8 References


Abarca-Mejía N (2010) Diatom community analysis and quality assessment of the polluted tropical Lerma River (Mexico). Freie Universität Berlin.

Abarca N, Jahn R, Zimmermann J, Enke N (2014) Does the cosmopolitan diatom *Gomphonema parvulum* (Kutzing) Kutzing have a biogeography? PLoS One 9(1): e86885. https://doi.org/10.1371/journal.pone.0086885.

Alp M, Keller I, Westram AM, Robinson CT (2012) How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. Freshwater Biology 57(5): 969-981. https://doi.org/10.1111/j.1365-2427.2012.02758.x.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3): 403-410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Aparicio J (2001) Hydrology of the Lerma-Chapala watershed. The Lerma-Chapala Watershed. Springer, 3-30.

Apotheloz-Perret-Gentil L, Cordonier A, Straub F, Iseli J, Esling P, Pawlowski J (2017) Taxonomy-free molecular diatom index for high-throughput eDNA biomonitoring. Mol Ecol Resour. https://doi.org/10.1111/1755-0998.12668.

Balint M, Bahram M, Eren AM, Faust K, Fuhrman JA, Lindahl B, O'Hara RB, Opik M, Sogin ML, Unterseher M, Tedersoo L (2016) Millions of reads, thousands of taxa: microbial community structure and associations analyzed via marker genes. FEMS Microbiol Rev 40(5): 686-700. https://doi.org/10.1093/femsre/fuw017.

Bate N, Newall P (1998) Techniques for the use of diatoms in water quality assessment: how many valves. In: Proceedings of the 15th international diatom symposium. 153-160 pp.

Berthon V, Bouchez A, Rimet F (2011) Using diatom life-forms and ecological guilds to assess organic pollution and trophic level in rivers: a case study of rivers in south-eastern France. Hydrobiologia 673(1): 259-271. https://doi.org/10.1007/s10750-011-0786-1.

Bik HM, Fournier D, Sung W, Bergeron RD, Thomas WK (2013) Intra-genomic variation in the ribosomal repeats of nematodes. PLoS One 8(10): e78230. https://doi.org/10.1371/journal.pone.0078230.

Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E (2016) obitools: a unix-inspired software package for DNA metabarcoding. Mol Ecol Resour 16(1): 176-182. https://doi.org/10.1111/1755-0998.12428.

Brown EA, Chain FJ, Crease TJ, MacIsaac HJ, Cristescu ME (2015) Divergence thresholds and divergent biodiversity estimates: can metabarcoding reliably describe zooplankton communities? Ecol Evol 5(11): 2234-2251. https://doi.org/10.1002/ece3.1485.

Carlsen T, Aas AB, Lindner D, Vrålstad T, Schumacher T, Kauserud H (2012) Don't make a mista (g) ke: is tag switching an overlooked source of error in amplicon pyrosequencing studies? Fungal Ecology 5(6): 747-749. https://doi.org/10.1016/j.funeco.2012.06.003.

Chen XG, Zhang J, Huang Y, Hou YP (2013) Diatom taxa identification based on single-cell isolation and rDNA sequencing. Forensic Science International: Genetics Supplement Series 4(1): e308-e309. https://doi.org/10.1016/j.fsigss.2013.10.157.

Connon SA, Giovannoni SJ (2002) High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. Appl Environ Microbiol 68(8): 3878-3885. https://doi.org/10.1128/AEM.68.8.3878-3885.2002

Cotler H, Mazari M, De Anda J (2006) Atlas de la cuenca Lerma-Chapala, construyendo una visión conjunta. INE-UNAM, 191p.

Cowart DA, Pinheiro M, Mouchel O, Maguer M, Grall J, Miné J, Arnaud-Haond S (2015) Metabarcoding is powerful yet still blind: a comparative analysis of morphological and molecular surveys of seagrass communities. PLoS One 10(2): e0117562. https://doi.org/10.1371/journal.pone.0117562.

Deiner K, Altermatt F (2014) Transport distance of invertebrate environmental DNA in a natural river. PLoS One 9(2): e88786. https://doi.org/10.1371/journal.pone.0088786.

Dell'Uomo A, Torrisi M (2011) Rheophile Apennine diatoms and their use as bioindicators of water quality. The Diatom World. Springer, 401-417.

Dixit SS, Smol JP, Kingston JC, Charles DF (1992) Diatoms - Powerful Indicators of Environmental-Change. Environmental science & technology 26(1): 22-33. https://doi.org/10.1021/es00025a002.

Dodds WK, Perkin JS, Gerken JE (2013) Human impact on freshwater ecosystem services: a global perspective. Environ Sci Technol 47(16): 9061-9068. https://doi.org/10.1021/es4021052.

Dowle EJ, Pochon X, J CB, Shearer K, Wood SA (2016) Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: a case study using freshwater macroinvertebrates. Mol Ecol Resour 16(5): 1240-1254. https://doi.org/10.1111/1755-0998.12488.

Dudgeon D, Arthington AH, Gessner MO, Kawabata Z, Knowler DJ, Leveque C, Naiman RJ, Prieur-Richard AH, Soto D, Stiassny ML, Sullivan CA (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. Biol Rev Camb Philos Soc 81(2): 163-182. https://doi.org/10.1017/S1464793105006950.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5): 1792-1797. https://doi.org/10.1093/nar/gkh340.

Elbrecht V, Feld CK, Gies M, Hering D, Sondermann M, Tollrian R, Leese F (2014) Genetic diversity and dispersal potential of the stonefly Dinocras cephalotes in a central European low mountain range. Freshwater Science 33(1): 181-192. https://doi.org/10.1086/674536.

Elbrecht V, Leese F (2015) Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass--Sequence Relationships with an Innovative Metabarcoding Protocol. PLoS One 10(7): e0130324. https://doi.org/10.1371/journal.pone.0130324.

Elbrecht V, Peinert B, Leese F (2017a) Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. Ecol Evol 7(17): 6918-6926. https://doi.org/10.1002/ece3.3192.

Elbrecht V, Vamos EE, Meissner K, Aroviita J, Leese F, Yu D (2017b) Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. Methods in Ecology and Evolution 8(10): 1265-1275. https://doi.org/10.1111/2041-210x.12789.

Esling P, Lejzerowicz F, Pawlowski J (2015) Accurate multiplexing and filtering for high-throughput amplicon-sequencing. Nucleic acids research 43(5): 2513-2524. https://doi.org/10.1093/nar/gkv107.

Ferris MJ, Ruff-Roberts AL, Kopczynski ED, Bateson MM, Ward DM (1996) Enrichment culture and microscopy conceal diverse thermophilic Synechococcus populations in a single hot spring microbial mat habitat. Appl Environ Microbiol 62(3): 1045-1050.

Fonseca V, Nichols B, Lallias D, Quince C, Carvalho G, Power D, Creer S (2012) Sample richness and genetic diversity as drivers of chimera formation in nSSU metagenetic analyses. Nucleic acids research 40(9): e66-e66. https://doi.org/10.1093/nar/gks002.

Gillett N, Pan Y, Parker C (2009) Should only live diatoms be used in the bioassessment of small mountain streams? Hydrobiologia 620(1): 135-147. https://doi.org/10.1007/s10750-008-9624-5.

Gillett ND, Pan YD, Manoylov KM, Stevenson RJ (2011) The role of live diatoms in bioassessment: a large-scale study of Western US streams. Hydrobiologia 665(1): 79-92. https://doi.org/10.1007/s10750-011-0606-7.

Godhe A, Asplund ME, Harnstrom K, Saravanan V, Tyagi A, Karunasagar I (2008) Quantification of diatom and dinoflagellate biomasses in coastal marine seawater samples by real-time PCR. Appl Environ Microbiol 74(23): 7174-7182. https://doi.org/10.1128/AEM.01298-08.

Groendahl S, Kahlert M, Fink P (2017) The best of both worlds: A combined approach for analyzing microalgal diversity via metabarcoding and morphology-based methods. PLoS One 12(2): e0172808. https://doi.org/10.1371/journal.pone.0172808.

Hajibabaei M, Shokralla S, Zhou X, Singer GA, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS One 6(4): e17497. https://doi.org/10.1371/journal.pone.0017497.

Hamilton PB, Lefebvre KE, Bull RD (2015) Single cell PCR amplification of diatoms using fresh and preserved samples. Front Microbiol 6(1084. https://doi.org/10.3389/fmicb.2015.01084.

Hebert PD, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc Biol Sci 270(1512): 313-321. https://doi.org/10.1098/rspb.2002.2218.

Hering D, Johnson RK, Kramm S, Schmutz S, Szoszkiewicz K, Verdonschot PFM (2006) Assessment of European streams with diatoms, macrophytes, macroinvertebrates and fish: a comparative metric-based analysis of organism response to stress. Freshwater Biology 51(9): 1757-1785. https://doi.org/10.1111/j.1365-2427.2006.01610.x.

Ivanova NV, Zemlak TS, Hanner RH, Hebert PDN (2007) Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes 7(4): 544-548. https://doi.org/10.1111/j.1471-8286.2007.01748.x.

Jahn R, Abarca N, Gemeinholzer B, Mora D, Skibbe O, Kulikovskiy M, Gusev E, Kusber WH, Zimmermann J (2017a) *Planothidium lanceolatum* and *Planothidium frequentissimum* reinvestigated with molecular methods and morphology: four new species and the taxonomic importance of the sinus and cavum. Diatom Research 32(1): 75-107. https://doi.org/10.1080/0269249x.2017.1312548.

Jahn R, Kusber WH, Cocquyt C (2017b) Differentiating *Iconella* from *Surirella* (Bacillariophyceae): typifying four Ehrenberg names and a preliminary checklist of the African taxa. PhytoKeys 82): 73-112. https://doi.org/10.3897/phytokeys.82.13542.

Kanz C, Aldebert P, Althorpe N, Baker W, Baldwin A, Bates K, Browne P, van den Broek A, Castro M, Cochrane G, Duggan K, Eberhardt R, Faruque N, Gamble J, Diez FG, Harte N, Kulikova T, Lin Q, Lombard V, Lopez R, Mancuso R, McHale M, Nardone F, Silventoinen V, Sobhany S, Stoehr P, Tuli MA, Tzouvara K, Vaughan R, Wu D, Zhu W, Apweiler R (2005) The EMBL Nucleotide Sequence Database. Nucleic Acids Res 33(Database issue): D29-33. https://doi.org/10.1093/nar/gki098.

Kelly M, Bennett C, Coste M, Delgado C, Delmas F, Denys L, Ector L, Fauville C, Ferreol M, Golub M, Jarlman A, Kahlert M, Lucey J, Ni Chathain B, Pardo I, Pfister P, Picinska-Faltynowicz J, Rosebery J, Schranz C, Schaumburg J, van Dam H, Vilbaste S (2009) A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water

Framework Directive: results of an intercalibration exercise. Hydrobiologia 621(1): 169-182. https://doi.org/10.1007/s10750-008-9641-4.

Kelly M, Whitton B (1995) The trophic diatom index: a new index for monitoring eutrophication in rivers. Journal of Applied Phycology 7(4): 433-444.

Kelly MG (1998) Use of community-based indices to monitor eutrophication in European rivers. Environmental Conservation 25(1): 22-29. https://doi.org/10.1017/S037689299800006x.

Kelly MG, Zgrundo A (2013) Potential for cross-contamination of benthic diatom samples when using toothbrushes. Diatom Research 28(4): 359-363. https://doi.org/10.1080/0269249x.2013.806959.

Kermarrec L, Bouchez A, Rimet F, Humbert JF (2013a) First evidence of the existence of semi-cryptic species and of a phylogeographic structure in the *Gomphonema parvulum* (Kützing) Kützing complex (Bacillariophyta). Protist 164(5): 686-705. https://doi.org/10.1016/j.protis.2013.07.005.

Kermarrec L, Franc A, Rimet F, Chaumeil P, Frigerio JM, Humbert JF, Bouchez A (2014) A next-generation sequencing approach to river biomonitoring using benthic diatoms. Freshwater Science 33(1): 349-363. https://doi.org/10.1086/675079.

Kermarrec L, Franc A, Rimet F, Chaumeil P, Humbert JF, Bouchez A (2013b) Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: a test for freshwater diatoms. Mol Ecol Resour 13(4): 607-619. https://doi.org/10.1111/1755-0998.12105.

Lang I, Kaczmarska I (2011) A protocol for a single-cell PCR of diatoms from fixed samples: method validation using *Ditylum brightwellii* (T. West) Grunow. Diatom Research 26(1): 43-49. https://doi.org/10.1080/0269249x.2011.573703.

Lejzerowicz F, Esling P, Pillet L, Wilding TA, Black KD, Pawlowski J (2015) High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. Sci Rep 5(13932. https://doi.org/10.1038/srep13932.

Licursi M, Gomez N (2009) Effects of dredging on benthic diatom assemblages in a lowland stream. J Environ Manage 90(2): 973-982. https://doi.org/10.1016/j.jenvman.2008.03.004.

Lowe R, Kociolek JP, You QM, Wang QX, Stepanek J (2017) Diversity of the diatom genus Humidophila in karst areas of Guizhou, China. Phytotaxa 305(4): 269-284. https://doi.org/10.11646/phytotaxa.305.4.3.

Lowe RL (2011) The importance of scale in understanding the natural history of diatom communities. The Diatom World. Springer, 293-311.

Mann DG, Chepurnov VA (2004) What have the Romans ever done for us? The past and future contribution of culture studies to diatom systematics. Nova Hedwigia 79(1): 237-291. https://doi.org/10.1127/0029-5035/2004/0079-0237.

Mccormick PV, Cairns J (1994) Algae as Indicators of Environmental-Change. Journal of Applied Phycology 6(5-6): 509-526. https://doi.org/10.1007/Bf02182405.

Meacham F, Boffelli D, Dhahbi J, Martin DI, Singer M, Pachter L (2011) Identification and correction of systematic error in high-throughput sequence data. BMC bioinformatics 12(1): 451. https://doi.org/10.1186/1471-2105-12-451.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Gateway Computing Environments Workshop (GCE), 2010. Ieee, 1-8 pp.

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJ, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst Biol 58(3): 298-311. https://doi.org/10.1093/sysbio/syp027.

Mora D, Carmona J, Cantoral-Uriza EA (2015) Diatomeas epilíticas de la cuenca alta del río Laja, Guanajuato, México. Revista Mexicana de Biodiversidad 86(4): 1024-1040. https://doi.org/10.1016/j.rmb.2015.09.004.

Mora D, Carmona J, Jahn R, Zimmermann J, Abarca N (2017) Epilithic diatom communities of selected streams from the Lerma-Chapala Basin, Central Mexico, with the description of two new species. PhytoKeys 88(39-69. https://doi.org/10.3897/phytokeys.88.14612.

Morales EA, Siver PA, Trainor FR (2001) Identification of diatoms (Bacillariophyceae) during ecological assessments: Comparison between Light Microscopy and Scanning Electron Microscopy techniques. Proceedings of the Academy of Natural Sciences of Philadelphia 151(1): 95-103. https://doi.org/10.1635/0097-3157(2001)151[0095:Iodbde]2.0.Co;2.

Moritz C, Cicero C (2004) DNA barcoding: promise and pitfalls. PLoS Biol 2(10): e354. https://doi.org/10.1371/journal.pbio.0020354.

Müller J, Müller K, Neinhuis C, Quandt D (2005) PhyDE-Phylogenetic Data Editor. Program distributed by the authors http:\\www phyde de.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature 403(6772): 853-858. https://doi.org/10.1038/35002501.

Oliver AK, Brown SP, Callaham MA, Jumpponen A (2015) Polymerase matters: non-proofreading enzymes inflate fungal community richness estimates by up to 15 %. Fungal Ecology 15(86-89. https://doi.org/10.1016/j.funeco.2015.03.003.

Pawlowski J, Esling P, Lejzerowicz F, Cedhagen T, Wilding TA (2014) Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. Mol Ecol Resour 14(6): 1129-1140. https://doi.org/10.1111/1755-0998.12261.

Piñol J, Mir G, Gomez-Polo P, Agustí N (2015) Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. Molecular Ecology Resources 15(4): 819-830. https://doi.org/10.1111/1755-0998.12355.

Postel SL, Daily GC, Ehrlich PR (1996) Human appropriation of renewable fresh water. Science: 785-788.

Potapova M, Charles DF (2005) Choice of substrate in algae-based water-quality assessment. Journal of the North American Benthological Society 24(2): 415-427. https://doi.org/10.1899/03-111.1.

Potapova M, Charles DF (2007) Diatom metrics for monitoring eutrophication in rivers of the United States. Ecological Indicators 7(1): 48-70. https://doi.org/10.1016/j.ecolind.2005.10.001.

Potapova M, Hamilton PB (2007) Morphological and ecological variation within the Achnanthidium minutissimum (Bacillariophyceae) species complex. Journal of Phycology 43(3): 561-575. https://doi.org/10.1111/j.1529-8817.2007.00332.x.

Potapova MG, Charles DF (2002) Benthic diatoms in USA rivers: distributions along spatial and environmental gradients. Journal of biogeography 29(2): 167-187. https://doi.org/10.1046/j.1365-2699.2002.00668.x.

Poulickova A, Spackov J, Kelly MG, Duchoslav M, Mann DG (2008) Ecological variation within Sellaphora species complexes (Bacillariophyceae): specialists or generalists? Hydrobiologia 614(1): 373-386. https://doi.org/10.1007/s10750-008-9521-y.

Previšić A, Walton C, Kučinić M, Mitrikeski PT, Kerovec M (2009) Pleistocene divergence of Dinaric Drusus endemics (Trichoptera, Limnephilidae) in multiple microrefugia within the Balkan Peninsula. Molecular ecology 18(4): 634-647. https://doi.org/10.1111/j.1365-294X.2008.04046.x.

Proft S, Grau J, Caswara C, Mazzoni C, Mora D, Zimmermann J (2017) MetBaN: Automated pipeline for metabarcoding data using taxonomical/phylogenetical classification of organisms. https://doi.org/10.5281/zenodo.1035193.

Prygiel J (2002) Management of the diatom monitoring networks in France. Journal of Applied Phycology 14(1): 19-26. https://doi.org/10.1023/A:1015268225410.

Rimet F, Abarca N, Bouchez A, Kusber WH, Jahn R, Kahlert M, Keck F, Kelly MG, Mann DG, Piuz A, Trobajo R, Tapolczai K, Vasselon V, Zimmermann J (2018) The potential of high throughput sequencing (HTS) of natural samples as a source of primary taxonomic information for reference libraries of diatom barcodes. Fottea 18(1): in press. https://doi.org/10.5507/fot.2017.013.

Rimet F, Trobajo R, Mann DG, Kermarrec L, Franc A, Domaizon I, Bouchez A (2014) When is sampling omplete? The effects of geographical range and marker choice on perceived diversity in

*Nitzschia palea* (Bacillariophyta). Protist 165(3): 245-259. https://doi.org/10.1016/j.protis.2014.03.005.

Ruck E, Nakov T, Alverson A, Theriot E (2016a) Nomenclatural transfers associated with the phylogenetic reclassification of the Surirellales and Rhopalodiales. Notulae algarum 10:(17 August 2016).

Ruck EC, Nakov T, Alverson AJ, Theriot EC (2016b) Phylogeny, ecology, morphological evolution, and reclassification of the diatom orders Surirellales and Rhopalodiales. Mol Phylogenet Evol 103(155-171. https://doi.org/10.1016/j.ympev.2016.07.023.

Ryberg M (2015) Molecular operational taxonomic units as approximations of species in the light of evolutionary models and empirical data from Fungi. Molecular ecology 24(23): 5770-5777. https://doi.org/10.1111/mec.13444.

Rynearson TA, Armbrust E (2000) DNA fingerprinting reveals extensive genetic diversity in a field population of the centric diatom Ditylum brightwellii. Limnology and Oceanography 45(6): 1329-1340. https://doi.org/10.4319/lo.2000.45.6.1329.

Sabater S (2000) Diatom communities as indicators of environmental stress in the Guadiamar River, SW. Spain, following a major mine tailings spill. Journal of Applied Phycology 12(2): 113-124. https://doi.org/10.1023/A:1008197411815.

Sawai Y (2001) Distribution of living and dead diatoms in tidal wetlands of northern Japan: relations to taphonomy. Palaeogeography, Palaeoclimatology, Palaeoecology 173(3): 125-141. https://doi.org/10.1016/S0031-0182(01)00313-3.

Schirmer M, Ijaz UZ, D'Amore R, Hall N, Sloan WT, Quince C (2015) Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. Nucleic acids research 43(6): e37-e37. https://doi.org/10.1093/nar/gku1341.

Schnell IB, Bohmann K, Gilbert MTP (2015) Tag jumps illuminated–reducing sequence-to-sample misidentifications in metabarcoding studies. Molecular Ecology Resources 15(6): 1289-1303. https://doi.org/10.1111/1755-0998.12402.

Segura-Garcia V, Cantoral-Uriza EA, Israde I, Maidana N (2012) Epilithic diatoms (Bacillariophyceae) as indicators of water quality in the Upper Lerma River, Mexico. Hidrobiológica 22(1): 16-27.

Sloan S, Jenkins CN, Joppa LN, Gaveau DLA, Laurance WF (2014) Remaining natural vegetation in the global biodiversity hotspots. Biological Conservation 177(12-24. https://doi.org/10.1016/j.biocon.2014.05.027.

Smucker NJ, Vis ML (2011) Diatom biomonitoring of streams: Reliability of reference sites and the response of metrics to environmental variations across temporal scales. Ecological Indicators 11(6): 1647-1657. https://doi.org/10.1016/j.ecolind.2011.04.011.

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. https://doi.org/10.1007/s00300-015-1683-2.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22(21): 2688-2690. https://doi.org/10.1093/bioinformatics/btl446.

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9): 1312-1313. https://doi.org/10.1093/bioinformatics/btu033.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol 57(5): 758-771. https://doi.org/10.1080/10635150802429642.

Stevenson R, Pan Y, van Dam H (2010) Assessing environmental conditions in rivers and streams with diatoms. The Diatoms: Applications for the Environmental and Earth Sciences, 2nd ed Cambridge University Press, Cambridge.

Sun D-L, Jiang X, Wu QL, Zhou N-Y (2013) Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. Applied and Environmental Microbiology 79(19): 5962-5969. https://doi.org/10.1128/AEM.01282-13.

Sweeney BW, Battle JM, Jackson JK, Dapkey T (2011) Can DNA barcodes of stream macroinvertebrates improve descriptions of community structure and water quality? Journal of the North American Benthological Society 30(1): 195-216. https://doi.org/10.1899/10-016.1.

Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012a) Environmental DNA. Mol Ecol 21(8): 1789-1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012b) Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol 21(8): 2045-2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30(12): 2725-2729. https://doi.org/10.1093/molbev/mst197.

Tavaré S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on mathematics in the life sciences 17(2): 57-86.

Torti A, Lever MA, Jørgensen BB (2015) Origin, dynamics, and implications of extracellular DNA pools in marine sediments. Marine genomics 24(185-196. https://doi.org/10.1016/j.margen.2015.08.007.

Trobajo R, Clavero E, Chepurnov VA, Sabbe K, Mann DG, Ishihara S, Cox EJ (2009) Morphological, genetic and mating diversity within the widespread bioindicator *Nitzschia palea* (Bacillariophyceae). Phycologia 48(6): 443-459. https://doi.org/10.2216/08-69.1.

Trobajo R, Mann DG, Clavero E, Evans KM, Vanormelingen P, McGregor RC (2010) The use of partial cox1, rbcL and LSU rDNA sequences for phylogenetics and species identification within the *Nitzschia palea* species complex (Bacillariophyceae). European Journal of Phycology 45(4): 413-425. https://doi.org/10.1080/09670262.2010.498586.

Turner CR, Uy KL, Everhart RC (2015) Fish environmental DNA is more concentrated in aquatic sediments than surface water. Biological Conservation 183(93-102. https://doi.org/10.1016/j.biocon.2014.11.017.

Van Dam H, Mertens A, Sinkeldam J (1994) A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. Aquatic Ecology 28(1): 117-133. https://doi.org/10.1007/BF02334251.

Vasselon V, Bouchez A, Rimet F, Jacquet S, Trobajo R, Corniquel M, Tapolczai K, Domaizon I (submitted) A correction factor inferred from cell biovolume improves quantification in diatom metabarcoding. Molecular Ecology Resources.

Vasselon V, Domaizon I, Rimet F, Kahlert M, Bouchez A (2017a) Application of high-throughput sequencing (HTS) metabarcoding to diatom biomonitoring: Do DNA extraction methods matter? Freshwater Science 36(1): 162-177. https://doi.org/10.1086/690649.

Vasselon V, Rimet F, Tapolczai K, Bouchez A (2017b) Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). Ecological Indicators 82(1-12. https://doi.org/10.1016/j.ecolind.2017.06.024.

Visco JA, Apotheloz-Perret-Gentil L, Cordonier A, Esling P, Pillet L, Pawlowski J (2015) Environmental Monitoring: Inferring the Diatom Index from Next-Generation Sequencing Data. Environ Sci Technol 49(13): 7597-7605. https://doi.org/10.1021/es506158m.

Von Falkenhayn L (2008) An assessment of the use of Bacillariophyceae as biological monitors of heavy metal pollution in Australian tropical streams. The University of Adelaide.

Weber AA-T, Pawlowski J (2014) Wide occurrence of SSU rDNA intragenomic polymorphism in foraminifera and its implications for molecular species identification. Protist 165(5): 645-661. https://doi.org/10.1016/j.protis.2014.07.006.

Wester P, Scott CA, Burton M (2005) River basin closure and institutional change in Mexico's Lerma-Chapala Basin. Irrigation and River Basin management: options for governance and institutions: 125-144.

Williamson CE, Dodds W, Kratz TK, Palmer MA (2008) Lakes and streams as sentinels of environmental change in terrestrial and atmospheric processes. Frontiers in Ecology and the Environment 6(5): 247-254. https://doi.org/10.1890/070140.

Yu DW, Ji YQ, Emerson BC, Wang XY, Ye CX, Yang CY, Ding ZL (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods in Ecology and Evolution 3(4): 613-623. https://doi.org/10.1111/j.2041-210X.2012.00198.x.

Zhu F, Massana R, Not F, Marie D, Vaulot D (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. FEMS Microbiol Ecol 52(1): 79-92. https://doi.org/10.1016/j.femsec.2004.10.006.

Zimmermann J, Abarca N, Enke N, Skibbe O, Kusber WH, Jahn R (2014) Taxonomic reference libraries for environmental barcoding: a best practice example from diatom research. PLoS One 9(9): e108793. https://doi.org/10.1371/journal.pone.0108793.

Zimmermann J, Glöckner G, Jahn R, Enke N, Gemeinholzer B (2015) Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Mol Ecol Resour 15(3): 526-542. https://doi.org/10.1111/1755-0998.12336.

Zimmermann J, Jahn R, Gemeinholzer B (2011) Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. Organisms Diversity & Evolution 11(3): 173-192. https://doi.org/10.1007/s13127-011-0050-6.

# 5 *Planothidium lanceolatum* and *Planothidium frequentissimum* reinvestigated with molecular methods and morphology: four new species and the taxonomic importance of the sinus and cavum

Jahn R.[1], Abarca N.[1],Gemeinholzer B.[2], Mora D.[1], Skibbe O.[1], Kulikovskiy M.[3], Gusev E.[3], Kusber W.-H.[1], Zimmermann J.[1]

[1] Botanischer Garten und Botanisches Museum Berlin-Dahlem, Freie Universität Berlin, Königin-Luise-Str. 6-8, 14195 Berlin, Germany.

[2]Justus-Liebig-Universität Giessen, AG Systematic Botany, Heinrich-Buff-Ring 38, 35392 Giessen, Germany.

[3] I.D. Papanin Institute for Biology of Inland Waters Russian Academy of Sciences, 152742, Borok, Yaroslavl, Russia.

Diatom Research 32(1): 75-107. https://doi.org/10.1080/0269249x.2017.1312548

## 5.1 Abstract

Molecular data (*rbc*L and 18S) of 36 *Planothidium* strains were analyzed. 27 strains were also studied morphologically by LM and SEM: six strains from Berlin, two strains from the Faroe Islands, four strains from Lake Baikal, seven strains from Korea, and eight strains from Mexico. The findings were compared to INSDC data of strains from New Zealand, Germany, France and the USA. The molecular and morphological data differentiated eight species, and the molecular trees underlined the clear differentiation between the two clades, with taxa possessing an asymmetrical central area on the sternum valve which is devoid of striae, either in the form of a sinus or of a cavum clustering together, versus the one taxon lacking a sinus or a cavum. In addition to *Planothidium lanceolatum* and *P*. cf. *subantarcticum*, the clade with a sinus contains two new species, *P. cryptolanceolatum* and *P. taeansa*, whereas the other clade with a cavum contains, in addition to *P. frequentissimum* and the recently described *P. victori* and *P. caputium,* and the new species *P. naradoense*. The species without a sinus or cavum is also described as new, *P. suncheonmanense*. With respect to their distribution, *P. victori* is the most common, with 11 strains from the studied continents, Europe (Berlin), Asia (Lake Baikal), Americas (Mexico, USA), Australia/Oceania (New Zealand), whereas *P. frequentissimum,* represented by five strains, was restricted to Germany, France and New Zealand. A

different geographical pattern seems to apply to the *P. lanceolatum* clade, with four strains occurring in Germany, the Faroe Islands, USA and Lake Baikal. *Planothidium cryptolanceolatum* (eight strains) only occurred in Korea, Mexico, and USA. One strain from the Faroe Islands, morphologically very similar to *P. subantarcticum* and recently described as new from the Subantarctic, was also identified.

**Key Words:** Bacillariophyceae, *Planothidium*, molecular data, morphology, sinus or cavum, new species

## 5.2 Introduction

The monoraphid diatom, *Planothidium lanceolatum* (Brébisson ex Kützing) Lange-Bertalot, was first described (as *Achnanthidium lanceolatum* Kützing) in 1846 from material of Brébisson (Kützing 1846). Grunow recombined it as *Achnanthes lanceolata* (Kützing) Grunow (Cleve and Grunow 1880) and it was known under this name for over a century as occurring throughout the world. It took almost 1.5 centuries and the use of scanning electron microscopy to discern that the typical horse-shoe shape (in German: hufeisenförmiger Fleck) on one side of the rapheless or sternum valve can be differentiated into specimens with a single or a double horse-shoe shaped hoofmark, named a sinus or a cavum respectively (Moss and Carter 1982), or with a rimmed depression and a hood Spaulding et al. (2008). Krammer & Lange-Bertalot (1991) and Lange-Bertalot (1993) did not completely agree with this distinction since they saw transitional stages in the outline of the valves. Nevertheless they differentiated *A. lanceolata* ssp. *lanceolata* with a sinus from *A. lanceolata* ssp. *frequentissima* Lange-Bertalot with a cavum. Later they transferred both taxa into the newly established genus *Planothidium* and raised *frequentissimum* to species rank as *P. frequentissimum* (Lange-Bertalot) Lange-Bertalot (Bukhtiyarova and Round 1996; Krammer and Lange-Bertalot 2004; Lange-Bertalot 1999). Morales (2006) reiterated the importance of the characteristic cavum in the central region of the sternum valve and speculated about its function. A recent paper has suggested that it may function as a lens for sunlight (Bukhtiyarova and Lyakh 2014). Morales (2006) proposed four groups of *Planothidium* of which two have either a sinus or a cavum, and also questioned the phylogenetic implications of this feature for the genus as a whole. In addition to valve outline, multiseriate striae on both raphe and sternum valves (Morales 2006), as well as unilateral terminal raphe fissure deflection (Spaulding et al. 2008) have been considered salient features of the genus.

In diatom preparations from mixed environmental samples it is very difficult to differentiate these features and evaluate their variability and consistency since monoraphid diatom species are heterovalvar having two different valve types. But unialgal cultures provide sufficient numbers of

both valve types (sternum and raphe) of the same species to study the variability of features such as form of sinus and cavum, raphe endings, number of areola rows in a multiseriate stria, as well as valve outline with LM and SEM. Thus, clonal cultures not only provide DNA for molecular investigation but also show some of the phenotypic plasticity of a single genotype.

Since *P. lanceolatum* is the generitype of *Planothidium*, it is important to understand its taxonomy, both morphologically and molecularly. Recently, this taxon was lectotypified by Van de Vijver et al. (2013) and its parallel taxon, *P. frequentissimum*, is in the process of being typified (Wetzel pers. comm.).

Although some of the molecular data are already published (Zimmermann et al. 2014), for this study we investigated six clones of *P. lanceolatum*, *P. frequentissimum* and *P. caputium* J. Zimmermann & R. Jahn from Berlin waters in more detail (molecular and morphological). They were compared to morphologically similar clones from the Faroe Islands (Denmark), Lake Baikal (Russia), Korea and Mexico from our own collection, and to molecularly related clonal culture data from INSDC (The International Nucleotide Sequence Database Collaboration) of taxa from New Zealand (Novis et al. 2012), Germany (Brinkmann et al. 2015), France (Keck et al. 2016), and the USA (Medlin and Kaczmarska 2004), T. Nakov et al. unpub.).

The aim of our study was to elucidate both the taxonomy and the systematics of the *P. lanceolatum/P. frequentissimum* taxon complex, utilizing both morphological and molecular data for phylogenetic analysis. The results of both methods are used for reciprocal illumination to discover the concealed diversity in this group, where morphological characters alone are not distinctive and where the addition of molecular findings is expected to yield significantly enhanced taxonomic resolution. In particular, we were interested to see if the distinction between a sinus versus a cavum is a stable taxonomic character, and whether there are similar biogeographic patterns for these taxa to those discussed for *Gomphonema parvulum* s.l. (Abarca et al. 2014).

## 5.3 Material & Methods

### 5.3.1 *Field collection and culturing* (for detailed data see Table 1)

Freshwater samples were collected from Denmark, Germany, Korea, Mexico and Russia between 2004 - 2014. 36 strain data are included in the present study; 27 strains were established by the authors. The sequence data for the other nine strains were downloaded from The International

Nucleotide Sequence Database Collaboration (INSDC). All sequences downloaded from INSDC were BLASTed against the INSDC database to test for taxonomic consistency.

Clonal strains were established by micropipetting single cells under an inverted light microscope. All strains cultivated and harvested in Berlin were treated according to Romero and Jahn (2013). The four Lake Baikal strains (labelled B) were cultivated in WC liquid medium (Guillard and Lorenzen 1972). Non-axenic unialgal cultures were maintained at 10$^{o}$C in a growth chamber with a 12:12 h light/dark photoperiod.

### 5.3.2 Documentation and vouchering

For all newly established strains the frustule preparation and morphological documentation were executed following Zimmermann et al. (2014). LM and SEM images were either taken with a Zeiss AxioImager.M2 and Hitachi FE SEM 8010 in Berlin, Germany, or a Zeiss Axiovert and a JSM-6510LV FE SEM in Borok, Russia. Vouchers for strains starting with D, plus the strains Ko0408 and Ko8A0610-1, are deposited in B (Herbarium Berolinense); vouchers for strains starting with B (Lake Baikal, Russia) are deposited in IBIW (Herbarium of I.D. Papanin Institute for Biology of Inland Waters, Russian Academy of Science; collection of M. Kulikovskiy). DNA samples are stored in the Berlin DNA Bank (Gemeinholzer et al. 2009+); data are available through AlgaTerra (Jahn and Kusber 2005+).

### 5.3.3 DNA extraction, sequencing and alignment

Cultured material was transferred to 1.5 ml tubes. DNA was isolated using NucleoSpin ® Plant II Mini Kit (Macherey and Nagel, Düren, Germany) or Qiagen® Dneasy Plant Mini Kit (Qiagen Inc.; Valencia, CA) following the respective product instructions. DNA fragment size and concentrations were measured via gel electrophoresis (1.5% agarose gel) and Nanodrop® (PeqLab Biotechnology LLC; Erlangen, Germany) respectively. DNA samples were stored at –20 °C for future use and finally deposited in the Berlin collection of the DNA bank network (Gemeinholzer et al. 2009+). PCR for *rbc*L was conducted following Abarca et al. (2014). The 18S SSU rRNA gene locus was amplified in two overlapping parts using two different primer pairs Algen F (CTG GTT GAT CCT GCC AGT AG, start of 18S) and Algen iR (TTC GAT CCC CTA ACT TTC GTT, position 1150) as well as Algen iF (TTG TCA GAG GTG AAA TTC TTG GA, position 1088) and D1800R (GCT TGA TCC TTC TGC AGG T, end of 18S; Brinkmann et al. 2015) following the PCR regime in Zimmermann et al. (2011). PCR products were visualized in a 1.5% agarose gel and cleaned with MSB Spin PCRapace® (Invitek LLC; Berlin, Germany) following standard procedures. DNA concentrations were measured using Nanodrop® (PeqLab Biotechnology) and samples were normalized to a total DNA content >100 ng/µl for sequencing.

Table 1. List of Materials: Strain Number, Voucher Code, Taxon Name, Sampling Data, INSDC Accession Numbers for 18S, [18S V4], *rbc*L, and Morphometric Data of the Studied Strains

| Strain | Voucher | Taxonname | Origin, date and collector | DNA Bank no. | INSDC no. 18S [V4] *r*RNA | INSDC no. *rbc*L | Length µm [mean] | Width µm [mean] | Striae/10 µm [mean] | n (SV / RV) |
|---|---|---|---|---|---|---|---|---|---|---|
| B086-3 | IBIW s.n. | *Planothidium victori* Novis, Braidwood & Kilroy = *Planothidium caputium* J. Zimmermann & R. Jahn | Russia; Lake Baikal Bay at 8 km from Enkhaluk village, 52.450694° N, 106.886917° E, 456m asl, 2011-07-14, leg. & isol. M. Kulikovskiy. | DB 26836 | KY650777 | KY650806 | 10.0-10.7 | 4.0-4.2 | 15 | 20 (SV) |
| | | | | | | | 10.0-10.7 | 4.0-4.2 | 14-15 | 20(RV) |
| B141 | IBIW s.n. | *Planothidium victori* = *Planothidium caputium* | Russia; Lake Baikal, near Enkhaluk village, 52.483278° N, 106.960028° E, 2011-07-15, leg. & isol. M. Kulikovskiy. | DB 26837 | KY650778 | KY650807 | | | | |
| B144 | IBIW s.n. | *Planothidium victori* = *Planothidium caputium* | Russia; Lake Baikal, near Enkhaluk village, 52.483278° N, 106.960028° E, 2011-07-15, leg. & isol. M. Kulikovskiy. | DB 26838 | KY650779 | KY650808 | 9.3-12.0 | 4.0-4.7 | 16-17 | 14 (SV) |
| | | | | | | | 9.3-11.3 | 4.0-4.7 | 15-16 | 12(RV) |
| B146 | IBIW s.n. | *Planothidium lanceolatum* (Bréb. ex Kütz.) Lange-Bert. | Russia; Lake Baikal, near Enkhaluk village, 52.483278° N, 106.960028° E, 2011-07-15, leg. & isol. M. Kulikovskiy. | DB 26839 | KY650780 | KY650809 | 18.7-24.6 | 7.3-8.0 | 15 | 21 (SV) |
| | | | | | | | 20.7-24.0 | 7.3-8.0 | 15-16 | 14(RV) |
| D06_014 | B 40 0040871 | *Planothidium victori* = *Planothidium caputium* | Germany; Berlin, Creek Wuhle, 52.520778° N, 13.577806° E, 2004 April, leg. & isol. O. Skibbe. | DB 8687 | KY650780 [KM084876]* | KM084938* | 19.6-23.5 [21.7] | 5.5-6.6 | 13 | 20 |
| | *Epitype | | | | | | | [5.9] | [13] | |
| D06_047 | B 40 0040874 | *Planothidium lanceolatum* | Germany; Berlin, Creek Wuhle, 52.520778° N, 13.577806° E, 2004 April, leg. & isol. O. Skibbe. | DB 8689 | KY650782 [KM084879]* | KM084943* | 10.1-12.5 | 4.6-5.1 | 13-14 | 20 |
| D06_113 | B 40 0040875 | *Planothidium victori* = *Planothidium caputium* | Germany, Berlin, Creek Wuhle, 52.520778° N, 13.577806° E, 2004 April, leg. & isol. O. Skibbe. | DB 8688 | KY650783 [KM084893]* | KM084959* | 12.8-21.5 | 5.4-7.3 | 12-14.5 | 23 |
| | | | | | | | [20.1] | [6,0] | [12.8] | |
| D06_117b | B 40 0040237 | *Planothidium frequentissimum* (Lange-Bert.) Lange-Bert. | Germany; Berlin, Creek Wuhle, 52.520778° N, 13.577806° E, 2004 April, leg. & isol. O. Skibbe. | DB 9678 | KY650784 | KY650813 | 15.6-16.4 [15.9] | 4.9-5.2 | 13.0-14.5 [13.6] | 14 |
| | | | | | | | | [5.0] | | |
| D06_138 | B 40 0040872 | *Planothidium frequentissimum* | Germany; Berlin, Creek Wuhle, 52.520778° N, 13.577806° E, 2004 April, leg. & isol. O. Skibbe. | DB 8685 | KY650785 [KM084895]* | KM084961* | 14.6-16.7 | 4.8-5.4 | 13-14 | 23 |
| | | | | | | | [15,6] | [5,2] | [13,3] | |
| D06_139 | B 40 0040873 | *Planothidium* | Germany; Berlin, Creek Wuhle, | DB 8686 | KY650786 | KM084962* | 14.7-15.7 | 4.9-5.5 | 13-14 | 20 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | *frequentissimum* | 52.520778° N, 13.577806° E, 2004 April, leg. & isol. O. Skibbe. | | [KM084896]* | | [15.5] | [5.2] | [13.7] | |
| D16_002 | B 40 0040797 | *Planothidium lanceolatum* | Denmark; Faroe Islands, Viöoy, Waterfall below church, 62.360333° N, 6.542833° W, 9m asl, 2004-08-04, leg. & isol. J. Bansemer. | DB 9679 | KY650792 | KY650821 | 13.1-37.4 [33.6] | 6.5-9.9 [9.2] | 12.05.2014 [13.3] | 27 (3 small) |
| D17_002 | B 40 0040798 | *Planothidium* cf. *subantarcticum* Van de Vijver & C.E.Wetzel | Denmark, Faroe Islands, Lambi, Eysturoy, isle creek, 62.139707° N, 6.725514° W, 80m asl, 2004-08-02, leg. & isol. J. Bansemer. | DB 9680 | KY650793 | KY650822 | 30.6-32.0 [31.4] | 8.2-8.6 [8.4] | 12.0-12.5 [12.1] | 15 |
| D21_002 | B 40 0040799 Holotype** | *Planothidium cryptolanceolatum* R. Jahn & N. Abarca sp. nov. | Korea; Cholla Namdo, creek at TaeAnSa, 35.131286° N, 127.388056° E, 299m asl, 2004-10-13, leg. R. Jahn & B.M. Suh, isol. O. Skibbe. | DB 9681 | KY650794 | KY650823 | 15.3-20.7 [18.1] | 4.8-5.9 [5.3] | 13.5-14.5 [13.9] | 20 |
| D23_024 | B 40 0040800 Holotype** | *Planothidium naradoense* R. Jahn & J. Zimmermann sp. nov. | Korea; Cholla Namdo, NaeNarodo Island, Spring, 34.533056 N, 127.463672° E, 114m asl, 2004-10-14, leg. R. Jahn & B.M. Suh, isol. O. Skibbe. | DB 9682 | KY650795 | KY650824 | 15.0-16.0 [15.4] | 4.7-5.2 [4.9] | 13.5-14.5 [14.2] | 20 |
| D26_002 | B 40 0040801 Holotype** | *Planothidium taeansa* R. Jahn & N. Abarca sp. nov. | Korea; Cholla Namdo, creek at TaeAnSa, 35.131286° N, 127.388056° E, 299m asl, 2004-10-13, leg. R. Jahn & B.M. Suh, isol. O. Skibbe. | DB 9683 | KY650796 | KY650825 | 18.5-19.4 [19.1] | 6.7-7.5 [7.0] | 12.5-13.0 [12.9] | 22 |
| D26_014 | B 40 0040802 | *Planothidium cryptolanceolatum* sp. nov. | Korea; Cholla Namdo, creek at TaeAnSa, 35.131286° N, 127.388056° E, 299m asl, 2004-10-13, leg. R. Jahn & B.M. Suh, isol. O. Skibbe. | DB 9684 | KY650797 | KY650826 | 23.7-27.3 [25.4] | 6.5-7.4 [6.9] | 12.5-13.5 [13.1] | 21 |
| D26_017 | B 40 0040803 | *Planothidium cryptolanceolatum* sp. nov. | Korea; Cholla Namdo, creek at TaeAnSa, 35.131286° N, 127.388056° E, 299m asl, 2004-10-13, leg. R. Jahn & B.M. Suh, isol. O. Skibbe. | DB 9685 | KY650798 | KY650827 | 22.0-24.4 [23.2] | 6.4-7.3 [6.7] | 12.05.2014 [12.9] | 15 |
| D31_010 | B 40 0040804 | *Planothidium cryptolanceolatum* sp. nov. | Mexico; spring Barranca del Aguacate, 20.449444°N, 102.592500°W, 2004-12-23, leg. N. Abarca, isol. O. Skibbe. | DB 9686 | KY650799 | KY650828 | 17.2-20.2 [18.6] | 5.0-6.5 [5.6] | Dez 14 [12.8] | 22 |
| D31_019 | B 40 0040805 | *Planothidium cryptolanceolatum* sp. nov. | Mexico; spring Barranca del Aguacate, 20.449444°N, 102.592500W, 2004-12-23, leg. N. Abarca, isol. O. Skibbe. | DB 9687 | KY650800 | KY650829 | 23.1.-25.0 [24.1] | 6.4-7.1 [6.7] | 12.0-13.5 [12.6] | 23 |
| D31_043 | B 40 0040806 | *Planothidium victori* = *Planothidium caputium* | Mexico; spring Barranca del Aguacate, 20.449444°N, 102.592500W, 2004-12-23, leg. N. Abarca, isol. O. Skibbe. | DB 9688 | KY650801 | KY650830 | 22.0-23.6 [22.8] | 5.2-6.3 [5.7] | 11.5-14.0 [12.4] | 23 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| D100_015 | B 40 0041750 | *Planothidium victori* = *Planothidium caputium* | Mexico; stream La Mesa, 21.091303 N, 101.138606° W, 2215 m asl, 2014-09-06, leg. D. Mora & isol. J. Bansemer. | DB 9689 | KY650787 | KY650816 | 6.3-21.7 [16.6] | 4.0-6.0 [5.1] | Dez 14 [12.4] | 22 |
| D101_022 | B 40 0041506 | *Planothidium victori* = *Planothidium caputium* | Mexico; stream Calvillo, 21.114000° N, 101.134472° W, 2138 m asl, 2014-09-06., leg. & isol. D. Mora. | DB 9690 | KY650788 | KY650817 | 5.1-7.3 [6.2] | 3.6-4.2 [3.9] | Dez 16 [14.8] | 21 |
| D108_021 | B 40 0041752 | *Planothidium cryptolanceolatum* sp. nov. | Mexico; stream El Membrillo, 20.839228° N, 100.645406° W, 2114 m asl, 2014-09-07, leg. D. Mora & isol. J. Bansemer. | DB 9691 | KY650789 | KY650818 | 15.1-17.3 [16.5] | 4.9-6.3 [5.6] | Nov 13 [12.1] | 22 |
| D109_018 | B 40 0041502 | *Planothidium victori* = *Planothidium caputium* | Mexico; Spring-fed Creek Guanajuatito, 20.889994° N, 100.541867° W, 2120 m asl, 2014-09-07, leg. D. Mora & isol. J. Bansemer. | DB 9692 | KY650790 | KY650819 | 7.9-23.1 [18.6] | 4.0-6.2 [5.4] | 11.0-14.0 [12.1] | 25 |
| D109_020 | B 40 0041503 | *Planothidium victori* = *Planothidium caputium* | Mexico; Spring-fed Creek Guanajuatito, 20.889994° N, 100.541867° W, 2120m asl, 2014-09-07, leg. D. Mora & isol. J. Bansemer. | DB 9693 | KY650791 | KY650820 | 9.0-24.3 [20.9] | 4.7-6.6 [5.7] | 11.5-14.5 [12.2] | 21 |
| Ko0408 | B 40 0041403 Holotype** | *Planothidium suncheonmanense* R.Jahn & J. Zimmermann sp. nov. | Korea; CholaNamdo, Bay SunCheonMan, shore, 34.874400° N, 127.500083° E, 1m asl, 2008-11-08, leg. B.M. Suh, isol. O. Skibbe. | DB 9694 | KY650802 | KY650831 | 5.5-10.7 [7.0] | 3.2-4.6 [3.7] | 15.0-17.0 [14.9] | 15 |
| Ko8A0610-1 | B 40 0041404 | *Planothidium cryptolanceolatum* sp. nov. | Korea; Kuwolsan, waterfall DanPyokPo, 38.489001° N, 125.299810° E, uncertainty 2000 m, 900 m asl, 2012-05-10, leg. R. Jahn & B.M. Suh, isol. O. Skibbe. | DB 9695 | KY650803 | KY650832 | 20.3-21,8 [20.9] | 6.2-7.0 [6.6] | 11.5-13.0 [12.5] | 24 |
| **Strains from GenBank** | | | | | | | | | | |
| LCR-S-2-1-1 | Novis et al. 2012 Holotype** | *Planothidium frequentissimum* | New Zealand, Christchurch, Styx River, 2009 November. | | [KX889991] | JQ610173 | 12.2-13.6 | 4.1-4.7 | 17(-19) | |
| LCR-S-18-1-1 | Novis et al. 2012 | *Planothidium victori* | New Zealand, Christchurch, Canterbury, Styx River, lat. 43.46333° S, 172.603550° E, periphyton, leg. P.M. Novis & J. Braidwood, 2009-11-4. | | [JQ610164] | JQ610172 | 14.0-16.5 | 5.0-5.2 | 14-15(-16) | |
| RK12 | Brinkmann et al. 2015 | *Planothidium frequentissimum* | Germany; Franken (49.383333° N, 11.466667° 28’ E), Creek Deinschwanger Bach & Harz (51.750000° N, 10.83333° E), Creek | | KF417663 | | | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Westerhöfer Bach, tufa forming biofilm, 2005 & 2006. | | | | | | | |
| TF-2014 clone 05DB5_12 | Brinkmann et al. 2015 | *Planothidium* sp. | Germany; Franken (49.383333° N, 11.466667E), Creek Deinschwanger Bach & Harz 51.750000° N, 10.83333° E ), Creek Westerhöfer Bach, tufa forming biofilm, 2005 & 2006. | | KF417664 | | | | | |
| TCC615 | Keck et al. 2015 | *Planothidium frequentissimum* | France, Lentigny, Rivière Le Lourdon, 46.000477° N, 3.980930° E. | | KT072986 | KT072932 | | | | |
| strain L1249 | Medlin & Kaszmarska 2004 | *Planothidium lanceolatum* | USA. | | AJ535189 | | | | | |
| PF1 | Nakov et al. unpublished | *Planothidium frequentissimum* | USA, NY (New York), Pack Forest Lake. | | KJ658409 | 658392 | | | | |
| PL2 | Nakov et al. unpublished | *Planothidium lanceolatum* | USA, PA (Pennsylvania), Ridley Creek. | | KJ658410 | 658393 | | | | |
| PL3 | Nakov et al. unpublished | *Planothidium lanceolatum* | USA, MT (Montana), Big Creek. | | KJ658411 | 658389 | | | | |

** Holotype for Planothidium caputium first published in Zimmermann et al. 2014, epitype for Planothidium victori published here,*

*** Molecular data of the authentic strain (see Holotype)*

Sanger sequencing was conducted by Starseq® (GENterprise LLC; Mainz, Germany). To sequence the 18S SSU rRNA gene the same primers were used as for the amplification. The *rbc*L gene was sequenced accordingly to Abarca et al. (2014). The new sequences (Table 1) were edited with Phyde® (Müller et al. 2010) and aligned using the Muscle algorithm (Edgar 2010) as implemented in MEGA6 (Tamura et al. 2013) for both markers. The alignments also included nine sequences from International Nucleotide Sequence Database Collaboration (INSDC) (Table 1). The alignment was subsequently improved manually in Phyde® (Müller et al. 2010). The newly generated sequences are deposited in the INSDC (accession numbers: KY650777-KY650835).

### 5.3.4 Morphological criteria

Besides valve outline and morphometric measurements of each clone (length, width, number of striae in 10 µm (Table 1)) the strains were investigated under SEM (unsputtered) : internal and external sternum valve views, as well as internal and external raphe valve views. Special attention was laid on the presence of a sinus or cavum, or neither, its form and size, the number of areola rows and their size on both the sternum and raphe valves, and merged and/or offset areolae on the valve mantle. These data are summarized in Table 2, and the matrix developed the morphological phylogenetic analysis is shown in Table 3.

### 5.3.5 Phylogenetic analyses

Four different data sets (18S, *rbc*L, 18S+*rbc*L concatenated and binary morphological matrix) were used for the phylogenetic analyses, each data set being analyzed using two different approaches: Maximum Likelihood (ML) as implemented in RAxML (Stamatakis 2006; Stamatakis 2014; Stamatakis et al. 2008) using the CIPRES platform (Miller et al. 2010), as well as Maximum Parsimony (MP) (calculated with MEGA6; (Tamura et al. 2013). The data sets for both markers (18S, *rbc*L) were first run independently. Afterwards a concatenated data set was generated using Mesquite (Maddison and Maddison 2016) and analyzed under the same rule, considering the methods of Souffreau et al. (2011); Mesquite (Maddison and Maddison 2016) was also used to explore character distribution in the trees based on combined data.

For the ML and MP analyses of the molecular data sets, the optimal model of sequence evolution that best fits the sequence data was calculated under the hierarchical likelihood ratio test (hLRT) and the Akaike information criterion (AIC) using model test 3.7 (Posada and Crandall 1998).The best fitting model was GTR + G + I (Tavaré 1986). The ML analysis was conducted using RAxML 8.2.8 (Stamatakis 2006; Stamatakis 2014; Stamatakis et al. 2008), ML search option (GTR + G + I) and 10,000 bootstrap replicates (model GTRCAT as implemented in RAxML for the rapid bootstrap

algorithm). For the MP analysis MEGA6 (Tamura et al. 2013) was used with heuristic search settings and 10,000 bootstrap replications. The strict consensus MP trees and the best ML tree (found by RAxML) were compared to each other.

The ML analysis of the morphological matrix was executed using RAxML 8.2.8 (Stamatakis 2006; Stamatakis 2014; Stamatakis et al. 2008) using the binary GAMMA model. Trees were drawn using FigTree v1.4.2 (Rambaut 2008) and Adobe Illustrator (Adobe Systems, San Jose, CA). Genetic distances for 18S and *rbc*L were calculated using MEGA6 (Tamura et al. 2013) and the implemented p-distance option.

**Table 2. Summarized morphological data of all studied strains**

| | Length [µm] | Width [µm] | Striae /10µm | Valve & Apices outline | Cavum | SV: number of areolae rows | SV: size of areolae inner row |
|---|---|---|---|---|---|---|---|
| *Planothidium lanceolatum* | 10.1-37.4 | 4.6-9.9 | 12.5-16.0 | Lanceolate to elliptical-lanceolate & broadly rounded apices | sinus only | up to 3 | smaller |
| *Planothidium* cf. *subantarcticum* | 30.6-32.0 | 8.2-8.6 | 12.0-12.5 | Lanceolate & weakly drawn out and rounded apices | sinus only | up to 3 | smaller |
| *Planothidium taeansa* | 18.5-19.4 | 6.7-7.5 | 12.5-13.0 | Elliptic-lanceolate & rostrate, drawn-out apices | sinus only | mostly 2 | n.a. |
| *Planothidium cryptolanceolatum* | 15.1-27.3 | 4.8-7.4 | 11.0-14.5 | Asymetric; elliptical-lanceolate & slightly rostrate and rounded apices | sinus only | up to 3 | smaller |
| *Planothidium frequentissimum** | 14.0-16.7 | 4.1-5.5 | 13.0-15.0 | Asymetric; lanceolate & slightly rostrate and rounded apices | roundish & tight opening | up to 4 | same |
| *Planothidium naradoense* | 15.0-16.0 | 4.7-5.2 | 13.5-14.5 | Asymetric; lanceolate & slightly rostrate-rounded apices. | parallel sides & tight opening | up to 4 | same |
| *Planothidium victori* = P. caputium* | 5.1-24.6 | 3.6-7.3 | 11.5-17 | Linear-lanceolate elliptical-lanceolate and elliptical & slightly rostrate-rounded apices | V-form & wide opening | up to 4 | same |
| *Planothidium suncheonmanense* | 5.5-10.7 | 3.2-4.6 | 15.0-17.0 | Asymetric; elliptic-lanceolate to elliptical & slightly drawn-out and rounded apices | none | up to 4 | same |

* including morphometric data of Novis et al (2012)

Table 3: Morphological Character Matrix of own Studied Strains.

| | Cavum | | | | Sinus | | | Areolae sternum valve | | | | | Outline |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Strain** | cavum | roundish, tight opening | v-form, wide opening | parallel sides | sinus | round depression | Irregular halt-moon depression | up to 3 rows of areolae | up to 4 rows of areolae | inner row smaller | same size | offset/merged areolae at mantle | asymmetric |
| B086_3 *Planothidium victori = P.* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| B141 *Planothidium victori = P. caputium* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| B144 *Planothidium victori = P. caputium* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| B146 *Planothidium lanceolatum* | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| D016_002 *Planothidium lanceolatum* | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| D017_002 *Planothidium* cf. | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| D021_002 *Planothidium* | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| D023_024 *Planothidium naradoense* | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| D026_002 *Planothidium taeansa* | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| D026_014 *Planothidium* | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| D026_017 *Planothidium* | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| D031_010 *Planothidium* | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| D031_019 *Planothidium* | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| D031_043 *Planothidium victori = P.* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| D06_014 *Planothidium victoriv* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| D06_047 *Planothidium lanceolatum* | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| D06_113 *Planothidium victori = P.* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| D06_117b *Planothidium* | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| D06_138 *Planothidium frequentissimum* | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| D06_139 *Planothidium frequentissimum* | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| D100_015 *Planothidium victori = P.* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| D101_022 *Planothidium victori = P.* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| D108_021 *Planothidium* | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |
| D109_018 *Planothidium victori = P.* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| D109_020 *Planothidium victori = P.* | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| Ko0408 *Planothidium* | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| Ko8A0610_1 *Planothidium* | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 |

## 5.4 Results

### 5.4.1 Molecular Data

#### 5.4.1.1 INSDC Data

The comparison of our data with the available INSDC data showed that the 18S data for KF417663 and KF417664 (Brinkmann et al. 2015) places KF417663 (the strain named RK12 *P. frequentissimum*) within our *P. frequentissimum* clade, and KF417664 (sequence generated via cloning 05DB5_12 *P.* sp.) within our *P. lanceolata* clade. But since the data for 18SV4 are incomplete and no *rbc*L data are available, these strains cannot be assigned to any subclade so we have excluded them from further analyses. The 18S data of AJ535189, strain L1249 *P. lanceolatum* (see (Medlin and Kaczmarska 2004)), are complete and this strain clusters within the *P. lanceolatum* clade. We included the data in our 18S tree. The *rbc*L and 18S data of the *Planothidium_frequentissimum*_isolate_TCC615 (Keck et al. 2016) are complete and cluster with our *P. frequentissimum* strains. The *rbc*L data of Novis et al. (2012) is complete, but only V4 data are available for 18S. Nevertheless, the two taxa cluster within the *P. frequentissimum* clade (F); LCR-S-2-1-1 *P. frequentissimum* with our *P. frequentissimum* in subclade F1, and LCR-S-18-1-1 *P. victori* Novis, Braidwood & Kilroy (Novis et al. 2012) within subclade F2. The 18S and *rbc*L data of the three further strains, PF1_ *P. frequentissimum*, PL2 *P. lanceolatum*, PL3 *P. lanceolatum* (T. Nakov et al., unpubl.) are complete and cluster in the two main clades in our trees; PL3 clusters within subclade L2.

#### 5.4.1.2 Phylogenetic analyses of concatenated and individual rbcL- & 18S-trees.

The results of the molecular phylogenetic analyses are shown in Fig. 1. The strict consensus tree of the ML analysis includes the concatenated dataset of both 18S and *rbc*L alignments, bootstrap values (>50) of ML Likelihood Bootstraps (LB) and MP Parsimony Bootstraps (PB). The individual results for the two genes are available for 18S (Supplementary Fig. 1 – Appendix 3) and *rbc*L (Supplementary Fig. 2 – Appendix 3), showing also strict consensus trees of ML analyses including bootstrap values (>50) of ML (LB) and MP (PB).

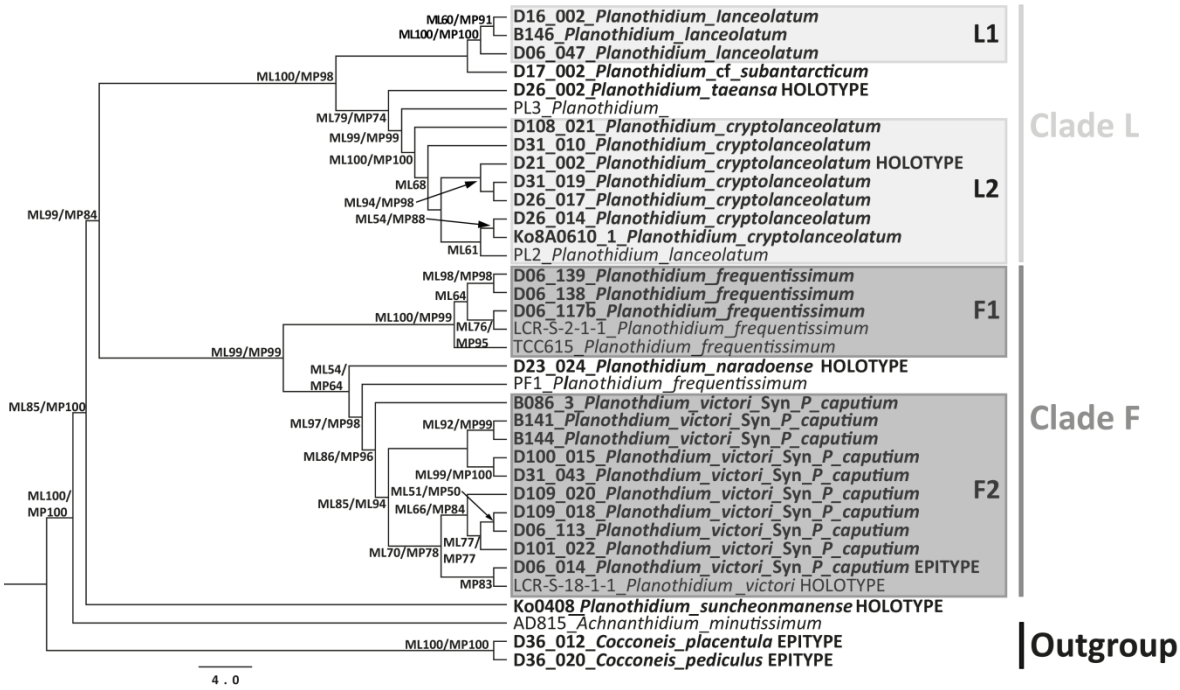



**Fig. 1.** Concatenated Strict Consensus Tree of the Combined Dataset of the Molecular Markers rbcL and 18S with the Results of Bootstrap Statistics (>50) for ML (LB) and MP (PB). Bold: strains cultured by the authors.

The three trees show no contradictory topologies for the two clades, **L** and **F,** or for their division into the subclades **L1**, **L2**, **F1** and **F2.** There are also no conflicts with respect to taxa excluded from the subclades (D23_024, D26_002, Ko0408, PF1 and PL3). Only D17_002, which is the sister taxon to the **L1** subclade in the concatenated, *rbc*L and the morphological trees (Fig. 2), clusters with D26_002 in the 18S tree, but without any statistical support. Slight variations in the tree topologies are only seen for relationships within the subclades, but without consistent statistical support.

Bootstrap values for the major split between strains with a sinus (Clade L, *P. lanceolatum* sensu lato and Clade F, *P. frequentissimum* sensu lato) (Fig. 1, Supplementary Figs 1, 2 – Appendix 3) are generally high ( ML (LB) 99, MP (PB) 84 in the concatenated tree; ML (LB) 73, MP (PB) 98 in the 18S tree; ML (LB) 77, MP (PB) 85 in the *rbc*L tree) Strain Ko0408, without a sinus or cavum, appears as sister to both clades with very good support (Concatenated ML (LB) 85; MP (PB) 100, 18S ML (LB) 78; MP (PB) 82 and *rbc*L ML (LB) 87; MP (PB) 90). The p- distances delimit Ko0408 with average differences of 5.5% for 18S and 6.6% for *rbc*L from these sister groups.

The *P. lanceolatum* subclade **L1** (ML [LB] 100; MP [PB] 100 in the concatenated tree) comprises four strains in 18S; one from Lake Baikal (B146), one from the Faroe Islands (D16_002), one from Berlin (D06_047), and one from the USA (L1249). *rbc*L data are not available for the last, therefore it is missing from the concatenated tree. With respect to the 18S tree, p-differences within this subclade are 0.0-0.28%, 0.0 - 0.37% for *rbc*L .

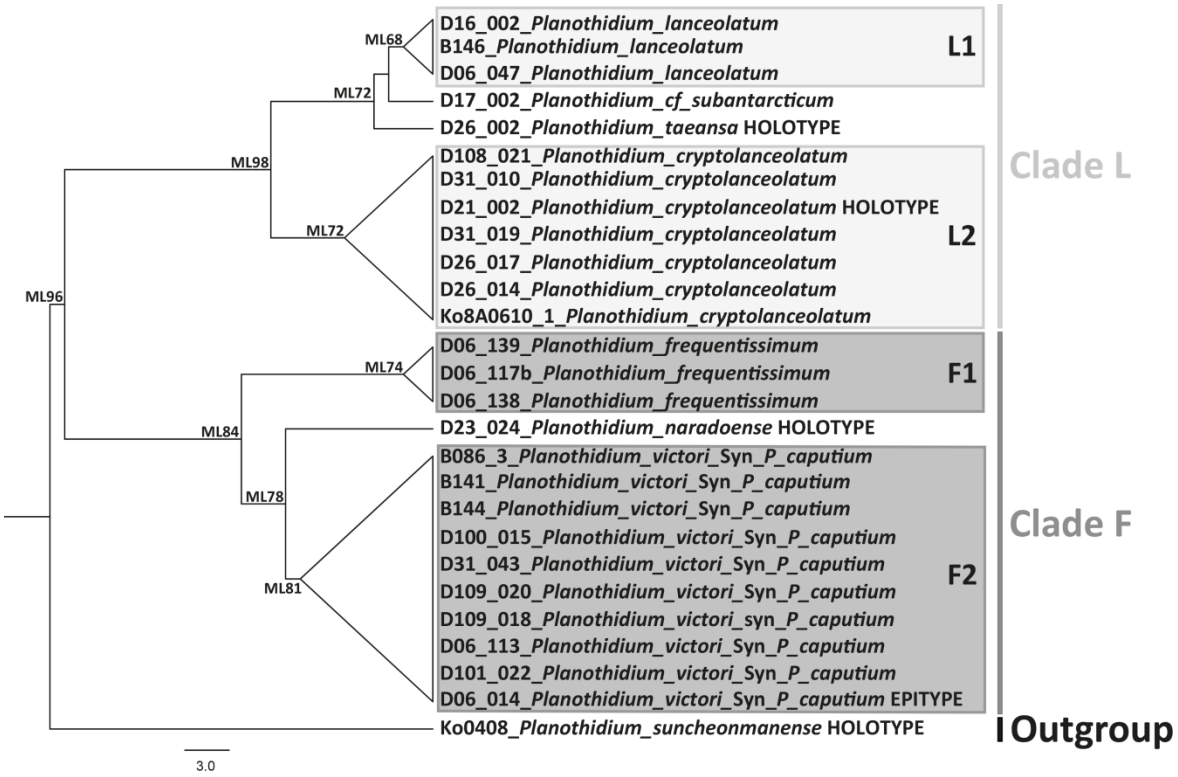



**Fig. 2.** Strict Consensus Tree for the ML analysis of the Morphological Character Matrix (Table 3) with the Results of Bootstrap Statistics (>50).

Strain D17_002 appears as sister to subclade **L1** (= *P. lanceolatum* s.s.) in the concatenated and *rbc*L trees, but sister to all other strains in this clade (L) in the 18S tree. Concerning p-distance data, the differences to subclade **L1** for 18S are at least 4.4% and 2.2% for *rbc*L, to subclade L2 they are at least 5% for 18S and 2.94% for *rbc*L.

Subclade **L2** (ML [LB] 100, MP [PB] 100 in the concatenated tree) contains eight strains from Korea, Mexico and USA; 0.0 – 1.1% differences for *rbc*L and 0.0 – 0.28% for 18S separate the four Korean strains (D21_002, D26_017, D26_014, Ko8A0610-1), the three Mexican strains (D31_010, D31_019) and the one USA strain (PL2_*Planothidium_lanceolatum*). The greatest within group variability (0.7 – 1.1%) is seen for *rbc*L for the Mexican strains.

In the concatenated and *rbc*L-trees, the two strains, D26_002 (Korea) and PL3_*Planothidium_lanceolatum* are each sister to subclade **L2;** D26_002 is sister to both subclades **L1** and **L2** in 18S. The strains differ by 0.37% for *rbc*L and 1.97% for 18S. They differ from subclade **L2** by at least 0.74% for *rbc*L and always 1.68% for 18S.

The *P. frequentissimum* subclade (**F1)** (ML [LB] 100, MP [PB] 99 in the concatenated tree) incorporates five strains: three from the same site and time from Berlin (D06_117b, D06_138, D06_139), one from New Zealand (LCR/Styx_2_1_1) and one from France (TCC615). There is 0-0.74% variation in the *rbc*L p-data, and 0-0.28% in the18S p-data.

The subclade **F2** (ML [LB] 97, MP [PB] 98 in the concatenated tree) includes 11 strains: two from Berlin (D06_014, D06_113), three from Lake Baikal (B086_3, B141, B144), five from Mexico (D31_043, D101_022, D109_018, D109_020, D100_015) and one from New Zealand (LCR_S_18_1_1). Their *rbc*L p-data vary between 0.0-1.1% and their 18S p-data between 0.0 - 0.56%. This subclade shows up to 0.37% molecular diversity within the *rbc*L p-data of the Mexican strains, and up to 1.1% between the Lake Baikal strains.

In all three trees, strain PF1_*Planothidium_frequentissimum* is sister to subclade **F2,** and D23_024 from Korea is the sister taxon to PF1, with 2.2% distance with *rbc*L. The *rbc*L-data show a distance of at least 2.6% between the Korean strain and subclade **F1**, and at least 1.4% with *s*ubclade **F2**. PF1_*Planothidium_frequentissimum* clusters within **F2**, with only 0.0-0.29% distance from the other strains..

### 5.4.2 Morphology & taxonomy

#### 5.4.2.1 Differences between the two main clades

Since no, or only limited (Novis et al. 2012), morphometric and SEM data are available for the INSDC strains, this part of the study focusses on our own strains. As in the gene trees (Fig, 1, Supplementary Figs 1, 2 – Appendix 3) the morphological tree (Fig. 2), has one main split (strongly supported by ML [LB] 96) between strains with a (Clade **L** = *P. lanceolatum* sensu lato) and those with a cavum (Clade **F** = *P. frequentissimum* sensu lato). The one strain without this feature is well separated from both clades. This supports the taxonomic value of differentiating on this morphological feature (sinus or cavum), which is easily seen in LM.

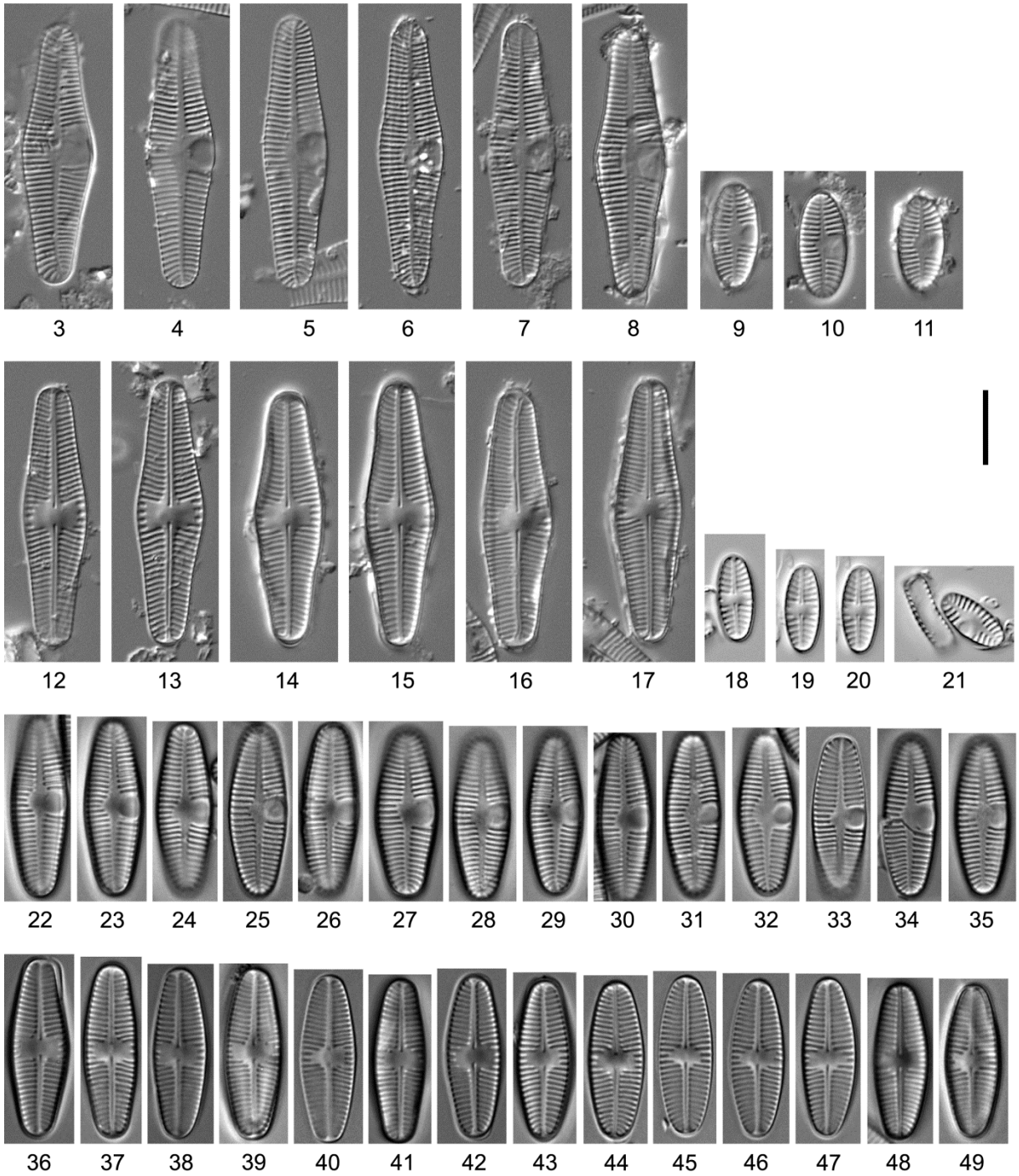
A further feature, seen only with SEM, is the number and size of the areolae in the multiseriate striae of the sternum (rapheless) valve. Taxa in the *P. lanceolatum* sensu lato clade (**L**) have multiseriate striae made up of two to three rows of areolae, with much smaller areolae in the middle row (of three). Taxa of the *P. frequentissimum* sensu lato clade (**F**) have multiseriate striae with three to four rows of similar sized areolae.

#### 5.4.2.2 The 'sinus' clade (Figs 3-191)

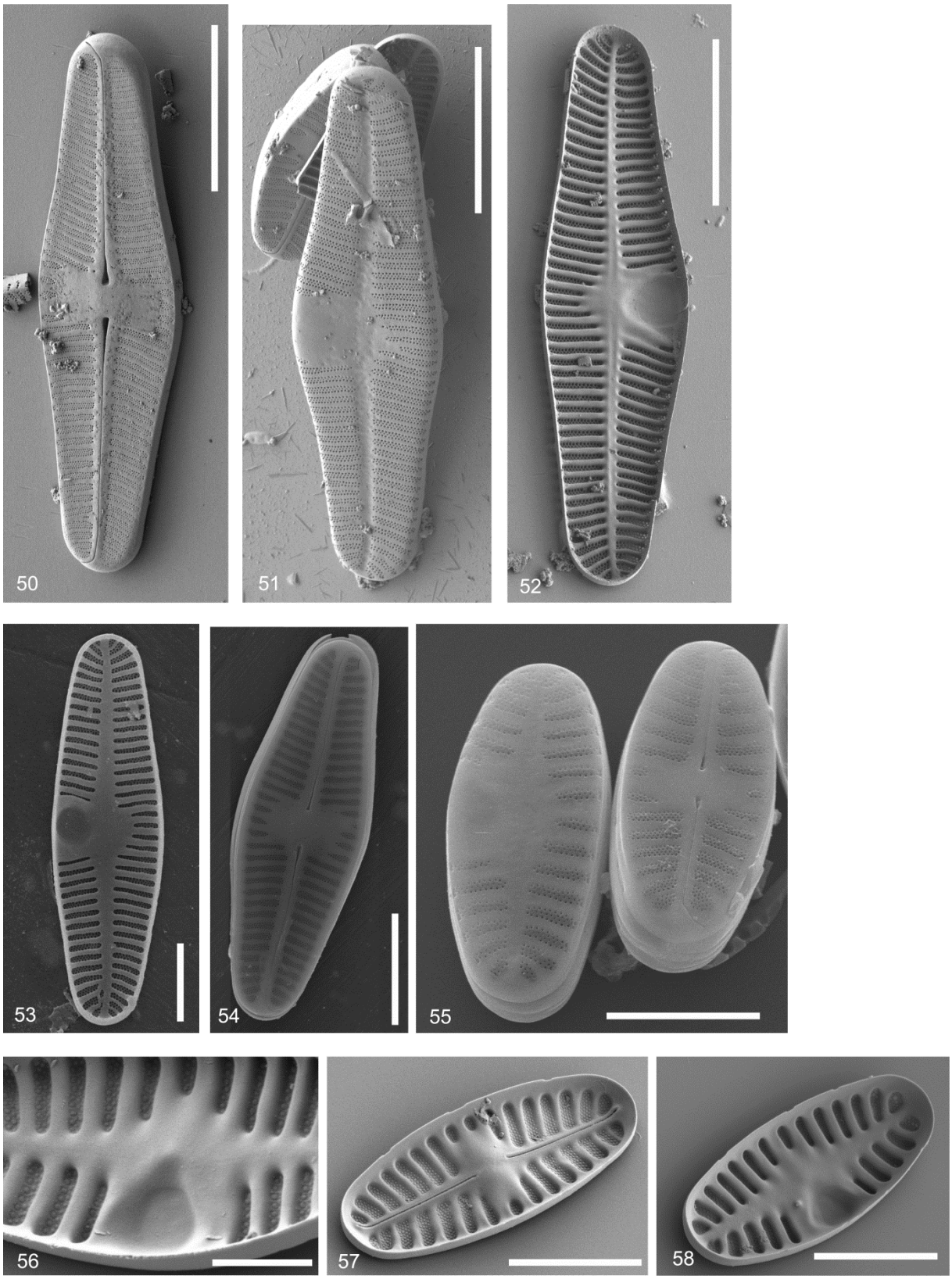
Valves of *P. lanceolatum* s.s. strains [subclade **L1**], especially from the Faroe Islands (D16_002) (Figs 3-17), are larger than those of the type population (Potapova 2010b; Van de Vijver et al. 2013). But since we also found small valves, from the other end of the cell cycle of this strain (Figs 9-11), it is clear that the large valves belong to this species. The small valves of this strain are morphologically very similar to those of strain D06_047 from Berlin (Figs 18-21), which has only small valves. The valves of strain B146 from Lake Baikal (Figs 22-49) are intermediate in size, but match morphologically with the other two strains. This shows that this species has a considerable size range across its complete cell cycle (Tables 1 and 2). Thus the size data for this species are herewith extended to 10.1-37.4 µm long, 4.6-9.9 µm wide, with 12-16 striae per 10 µm (see Van de Vijver et al, 2013, for a description of the type of *P. lanceolatum*).

Regarding the differentiating features between the subclades **L1** (Figs 3-58) and **L2** (Figs 59-156) the outline differences are very subtle, in that the latter is more slender and slightly asymmetric. The size and form of the sinus of the sternum valve seems to be a discriminating feature: *P. lanceolatum* (**L1;** Figs 52, 53, 56, 58) has a larger, roundish sinus, whereas *P. cryptolanceolatum* (L2; Figs 75, 149, 154) has a rather half-moon shaped sinus, with a deeper depression at the mantle. The number of shortened striae at the centre the raphe valve might also be a useful feature, with 2-5 striae for **L1** (*P. lanceolatum)* versus 1-3 for **L2**. In all these strains, the striae stop at the valve edge in the raphe valve (Figs 50, 55, 72, 147, 152). However, on the sternum valve they continue onto the mantle (Figs 51, 55, 74, 153). In **L2** (Figs 74, 153) and in D26_002 (Fig 190) the mantle areolae are offset and merge into a bigger oval pore or line; in *P. lanceolatum* (Figs 51, 55) and D17_002 (Fig. 166) the merged oval pores or lines are not offset at the valve mantle.

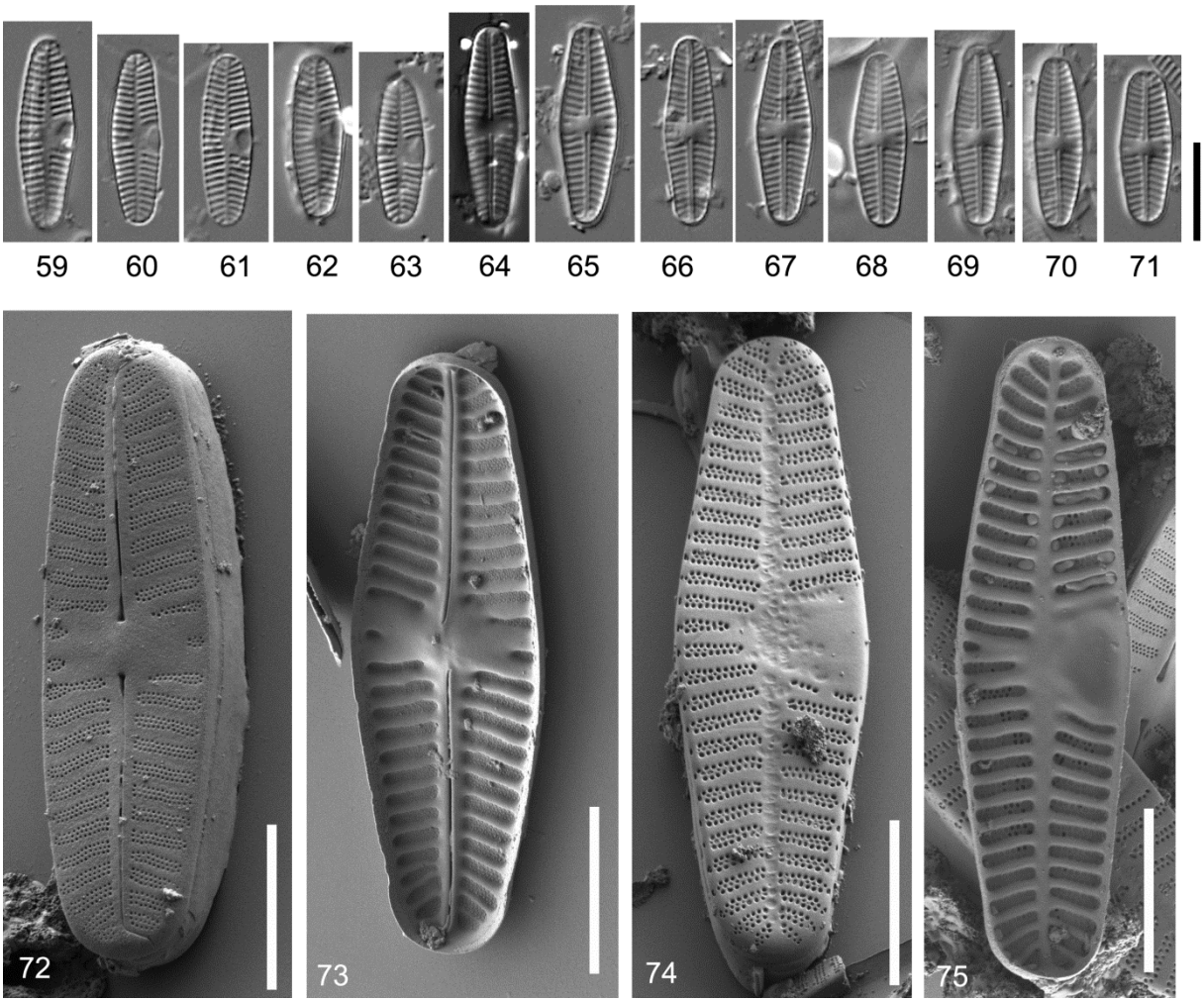
The two taxa which are represented by single strains only can be differentiated by their outline in LM from **L1** and **L2**, as well as from each other. D17_002 (Figs 156-171) has a rhombic outline, D26_002 (Figs 172-191) is broadly elliptic with rostrate apices. The morphological data for D17_002 match well to *P. subantarcticum* Van de Vijver & C. E. Wetzel (Van de Vijver, pers. comm.: sinus is deeper in *P. subantarcticum* but shallower in our species, apices in our species are protracted while in *P. subantarcticum* the apices are narrow gradually without being protracted). We have therefore decided to designate this strain as ***P.* cf. *subantarcticum***. Morphometric data of this strain can be found in Table 1. The areolae, which are internally covered by hymenes, have a diameter of 100–150 nm (Fig. 169). The stated differences in the central area of the raphe valve by (Van de Vijver et al. (2013) do not quite fit with our data since strain D16_002 (Figs 3-17) is of similar shape and size as our strain D17_002 (Figs 156-164).




**Figs 3-49.** *Planothidium lanceolatum*; LM. Figs 3-17: Strain D16_002, Faroe Islands. Figs 18-21: Strain D06_047, Berlin. Figs 22-49: Strain B146, Lake Baikal. Figs 3-11, 21-35: Sternum valves; Figs 12-20, 36-49: Raphe valves. Scale bar: 10 µm.




**Figs 50-58**. *Planothidium lanceolatum*; SEM. Figs 50-52: Strain D16_002, Faroe Islands. Figs 53-54: Strain B146, Lake Baikal. Figs 55-58: Strain D06_047, Berlin. Figs 51-53, 55 left, 56, 58: Sternum valves; Figs 50, 54, 55 right, 57: Raphe valves. Figs 50-51, 54-55: External valve view; Figs 52-53, 56-58: Internal valve view. Scale bars: Figs 50-52 = 10 µm; Figs 53-55, 57-58: 5 µm; Fig. 56: 2 µm.



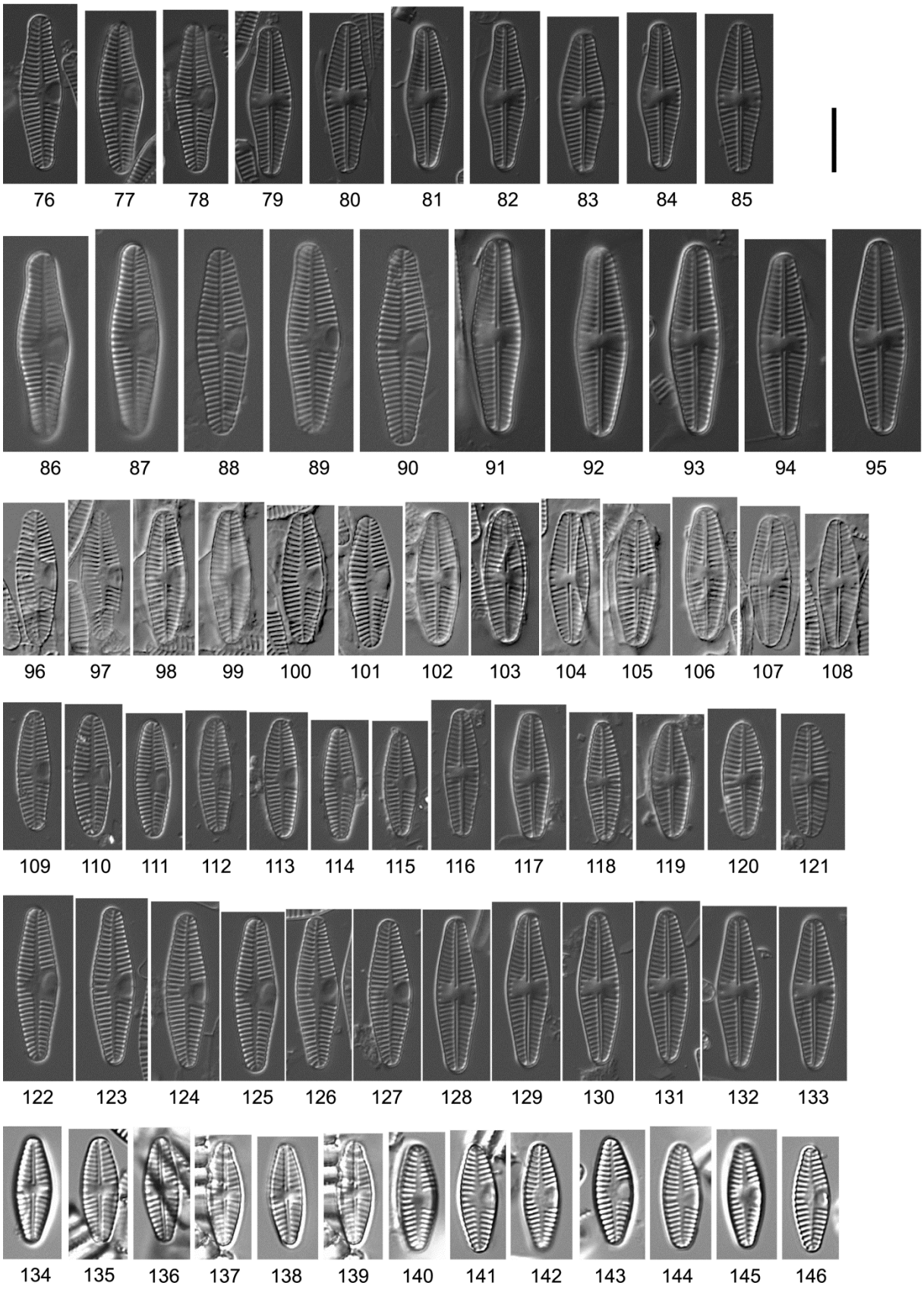

Figs 59-75: *Planothidium cryptolanceolatum*; Strain D21_002, Korea, Holotype. Figs 59-63, 74-75: Sternum valves; Figs 64-73: Raphe valves. Figs 72, 74: External valve view; Figs 73, 75: Internal valve view. Figs 59-71 LM; Figs 72-75 SEM. Scale bars: Figs 59-71 = 10 µm; Figs 72-75 = 5 µm.

Although the sternum valve of *P. amphibium* C.E. Wetzel, Ector & L. Pfister (Wetzel et al. 2014) is very similar having multiseriate striae with three rows of areolae, the central one with smaller areolae, and offset areolae on the mantle, its outline with capitate, protracted apices differs from our strains (Wetzel and Ector 2014).Thus, the morphological and molecular differences between the subclade **L2** (= *P. cryptolanceolatum* R. Jahn & N. Abarca) and D26_002 (= *P. taeansa* R. Jahn & N. Abarca) make it necessary to describe them as new (see section on nomenclatural and taxonomical consequences).

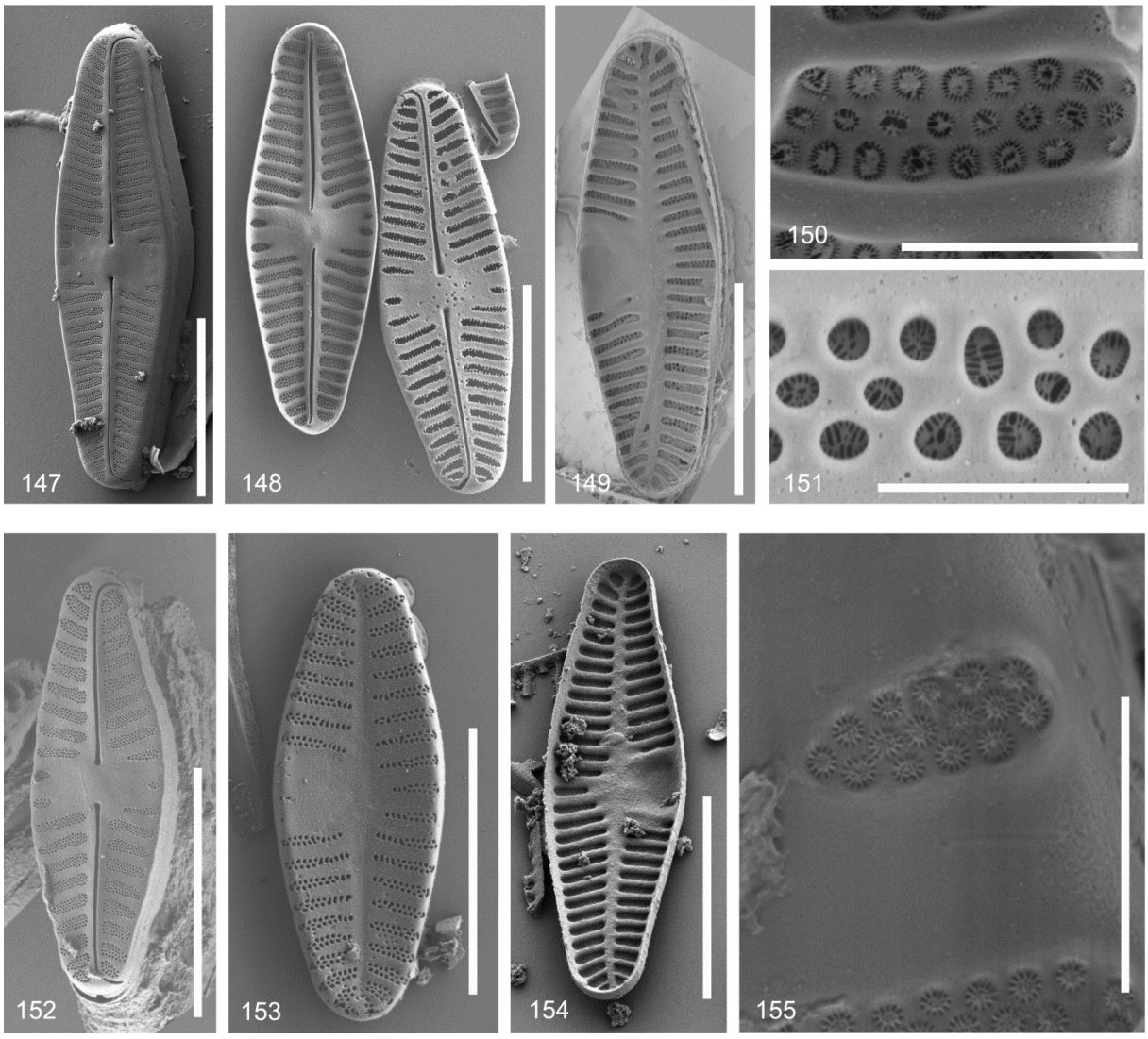
### 5.4.2.3 The 'cavum' clade (Figs 192- 375)

The *P. frequentissimum* sensu lato clade is also divided into two subclades, with *P. frequentissimum* [**F1**] represented by three strains from Berlin (Figs 192-239). The morphology matches the very broad concept of Hofmann et al. (2013), based on Krammer and Lange-Bertalot (2004), and includes the concepts of Straub (1990) as well as that of *P. frequentissimum* sensu Potapova (2010a). Morphometric data of our three strains are less variable: length 12.1-16.7 µm; width 4.1-5.5, 13-14.5 striae per 10 µm (Tables 1, 2). According to C.E. Wetzel (pers. comm.) our strains seem to be slightly slenderer than the type population, but since one of the decisive features, the form of the cavum opening as illustrated by SEM (Krammer and Lange-Bertalot 2004), fig. 45:18), matches well with our strains (Fig. 237), we have decided to defer formal naming until molecular data become available for the type population.

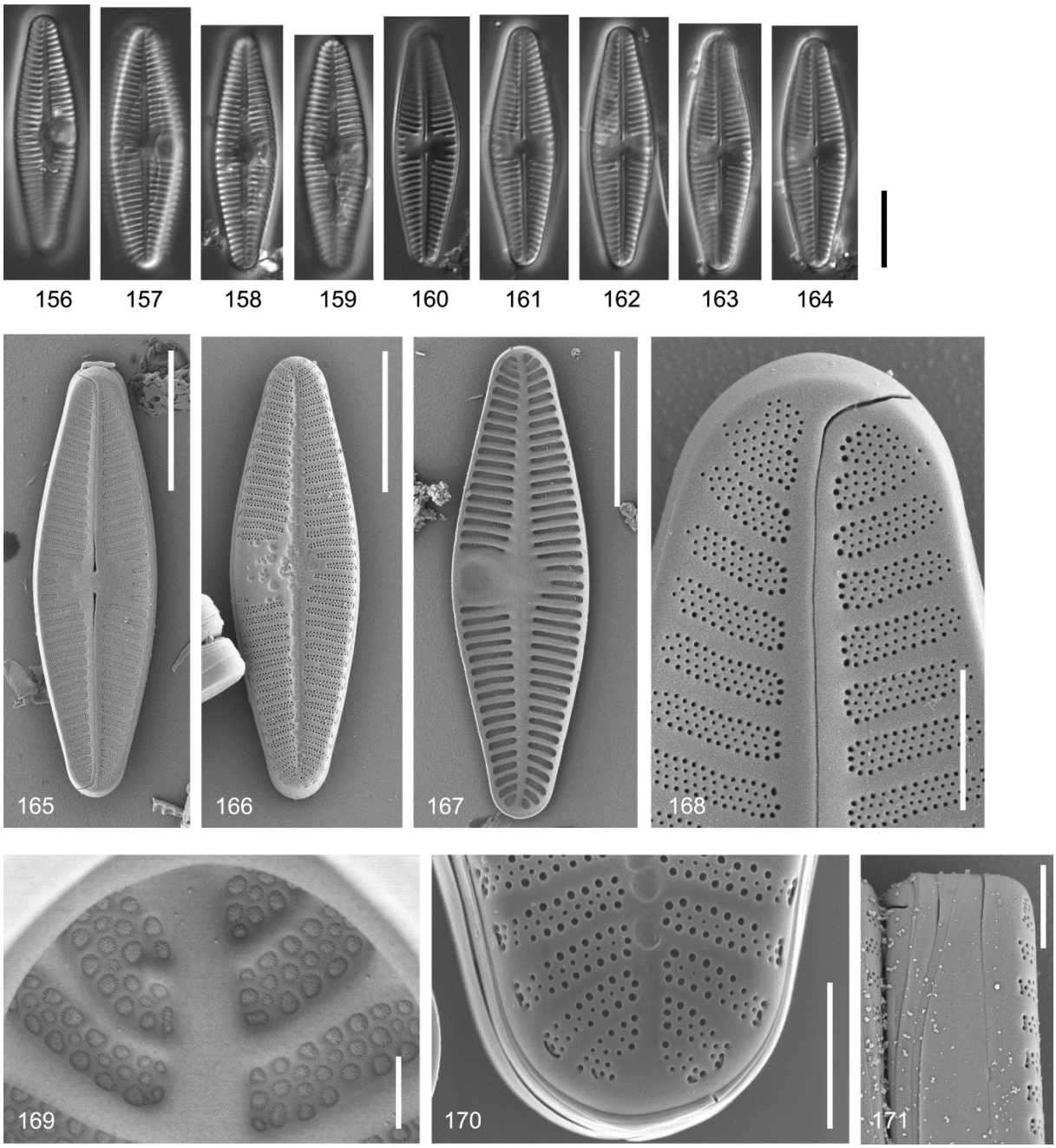
The species of clade **F2** have been described twice, as ***P. victori*** by Novis et al. (2012) from Styx river in New Zealand, and as *P. caputium* from Berlin by us (D06_014: Figs 240-250, D06_113; Figs 251-261) (Zimmermann et al. 2014). Although the morphometric data do not overlap, the 18SV4 and *rbc*L molecular data for *P. victori* and our holotype strain D06_014 *P. caputium* are identical. *Planothidium victori* takes priority over *P. caputium.* Since the orphological data of the holotype of *P. victori* are limited, we are using our strains (*P. caputium)* to characterise *P. victori* by an epitype designation (see below). The outline has been described as similar to *P. frequentissimum* but the valves are not as asymmetrical. The main differentiating feature is the form of the cavum and its opening: *P. frequentissimum* has a roundish cavum with a narrow opening (Fig. 237), seen as a curved line close to mantle (LM), whereas *P. victori* (= *P. caputium*) has a broader V-form with a wider cavum opening (Figs 342, 346), seen as an almost straight line further from the mantle (LM). The strains from Lake Baikal (B141, B144, B086-3; Figs 310-340) and the Mexican strains (D109_018, D101_022, D109_020, D31_043, D100_015; Figs 263-303 LM and Figs 341-350 SEM) are variable in size and outline. Some of the Berlin and Mexican strains look very similar in outline to typical *P. lanceolatum*; the more elliptical ones from Lake Baikal and the deformed specimens from New Zealand seem to occur not only in culture but also in the environment (Novis et al. 2012). These 11 strains (B086-3, B141, B144, D06_014-holotype of *P. caputium*, D06_113, D31_043, D100_015, D101_022, D109_018, D109_020, LCR_S_18_1_1 holotype of *P. victori*) extend the morphometric range of *P. victori* to 5.1-24.6 µm long, 3.6-7.3 µm wide, with 11.5-17 striae per 10 µm (Tables 1, 2). The Mexican strains D100_015 (Figs 272-277), D109_018 (Figs 278-284, 341-348) and D109_020 (Figs 294-303) show valves from the entire cell cycle.




**Figs 76-146.** *Planothidium cryptolanceolatum*; LM. Figs 76-85: Strain D26_017, Korea; Figs 86-95: Strain D26_014, Korea; Figs 96-108: Strain Ko8A0610_1, Korea. Figs 109-121: Strain D31_010, Mexico; Figs 122-133: Strain D31_019, Mexico; Figs 134-146: Strain D108_021, Mexico. Scale bar: 10 µm.




**Figs 147-155.** *Planothidium cryptolanceolatum*, SEM. Figs 147-148: Strain D26_014, Korea; Figs 150-151: Strain D26_017, Korea; Figs 152-153, 155: Strain D108_021, Mexico; Figs 154: Strain D31_019, Mexico. Figs 149-151, 153-154: Sternum valves; Figs 147-148, 152, 155: Raphe valves. Fig 150, 155: Internal valve view of one stria with three rows of areolae with hymenate occlusions. Fig. 151: External valve view of one stria with three rows of areolae with hymenate occlusions. Scale bars: Figs 147, 152-154 = 10 µm; Figs 150-151, 155 = 1 µm.

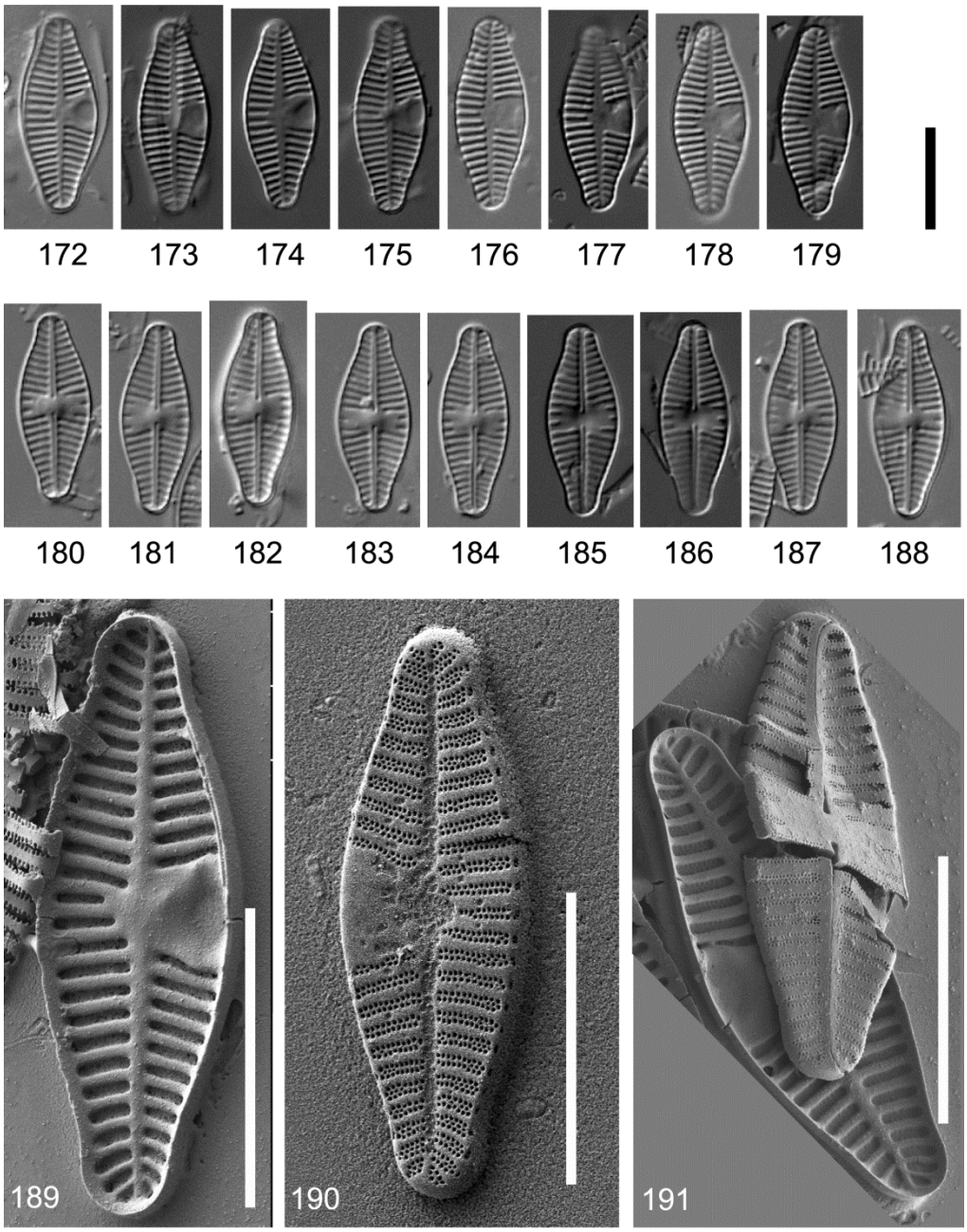



**Figs 156 -171.** *Planothidium* cf. *subantarcticum* (Strain D17_002, Faroe Islands); Figs 156-159, 166-167, 169-170: Sternum valves; Figs 160-165, 168: Raphe valves. Figs 165-166, 168, 170: external valve view; Figs 167, 169 internal valve view; Figs 171 girdle view. Figs 169, 170: note the smaller areolae size of the inner row of areolae within one stria. Fig 169: the hymenate areolae occlusions have a diameter of 100 – 150 nm. Figs 156-164 LM; Figs 165-171 SEM. Scale bars: Fig. 156-167 = 10 µm; Figs 168, 170,171 = 2 µm; Fig. 169 = 0.5 µm.

The outline of the isolated strain, D23_024 (*P. naradoense* R. Jahn & J. Zimmermann; Figs 351-375), is somewhat similar to *P. frequentissimum* but the cavum is oblong with parallel sides and a narrow

opening (Figs 367, 371) seen as a curved line close to mantle (LM). This species is described as new (below).

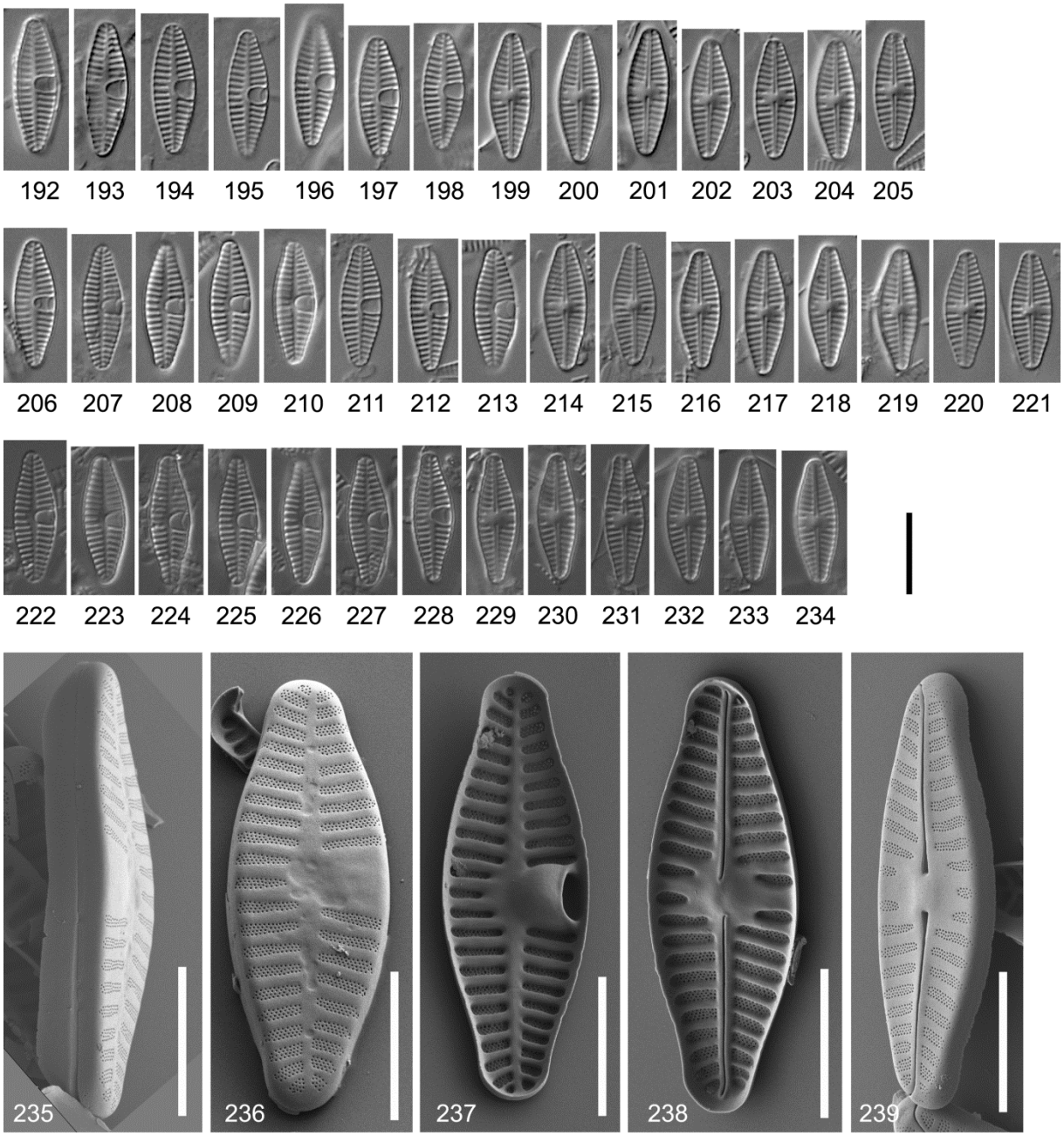
Comparing recently described *Planothidium* taxa with a cavum, for which the micromorphology of the cavum and the multiseriate striae have been illustrated, such as *P. comperei* C.E. Wetzel, N'Guessan & Tison-Rosebery and *P. piaficum* (J.R. Carter & Denny) C.E. Wetzel & Ector (N'Guessan et al. 2014), it is clear that both have similar multiseriate striae to the taxa in our study, but differ in their sternum valves. The inner areola row in *P. comperei* contains smaller areolae, which from our study are typical for the sinus clade. In *P. piaficum* the areola rows are raised above the virgae, very different from our taxa. In addition, both have a completely different outline, *P. comperei* is elliptic-lanceolate and *P. piaficum* is elliptic-lanceolate to oval, with protracted capitate apices. With respect to the cavum valve, both have a cavum with a small aperture; in *P. piaficum* the borders fuse with the neighboring virgae, a feature never observed in our taxa. Also *P. bagualensis* C.E. Wetzel & Ector (Wetzel and Ector 2014), *P. biporomum* (M.H. Hohn & Hellerman) Lange-Bertalot and *P. incuriatum* W.E. Wetzel, Van de Vijver & Ector (see Wetzel et al. (2013) have a wide open cavum with its borders attached to neighboring striae and 2-3 areola rows per stria on the sternum valve as well as offset areolae on the mantle. None of these species fit morphologically to our taxa.




**Figs 172-191.** *Planothidium taeansa;* strain D26_002, Holotype. Figs 172-179, 189-190, 191 lower valve: Sternum valves; Figs 180-188, 191 upper valve: Raphe valves. Figs 172-188: LM; Figs 180-191: SEM. Scale bars: Figs 172-188 = 10 µm; Figs 180-191 = 10 µm.

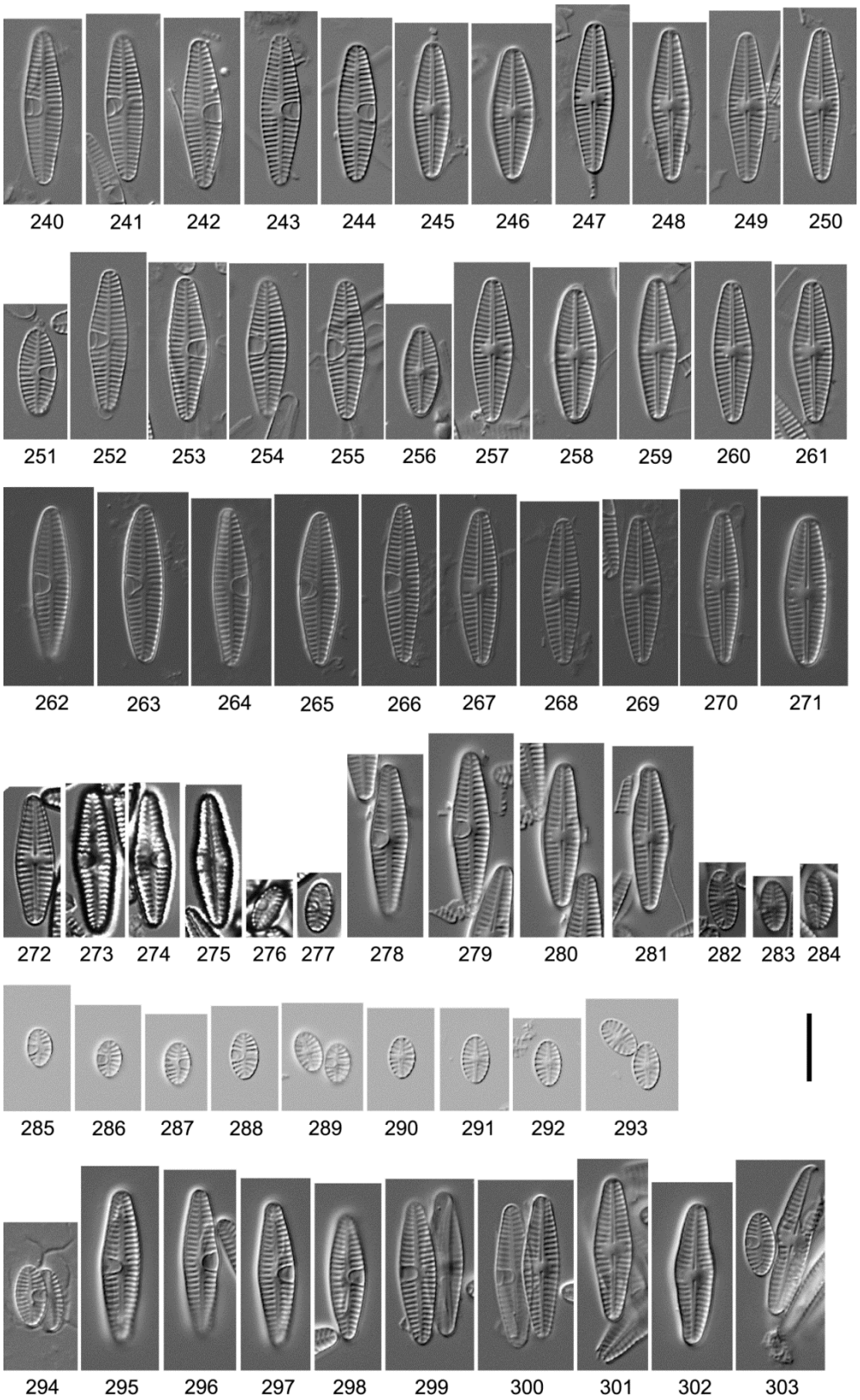
### 5.4.2.4 Strain without a Sinus or Cavum (Figs 376-385)

The strain Ko0408 which has neither a sinus nor a cavum is a morphological and molecular outlier. With respect to valve outline, terminal raphe endings (bent unilaterally), central raphe endings (externally drop like, internally bent very slightly to one side), as well as 3-4 areola rows per multiseriate stria, this strain is quite similar to those from clades L and (especially) F (Figs 376-385).

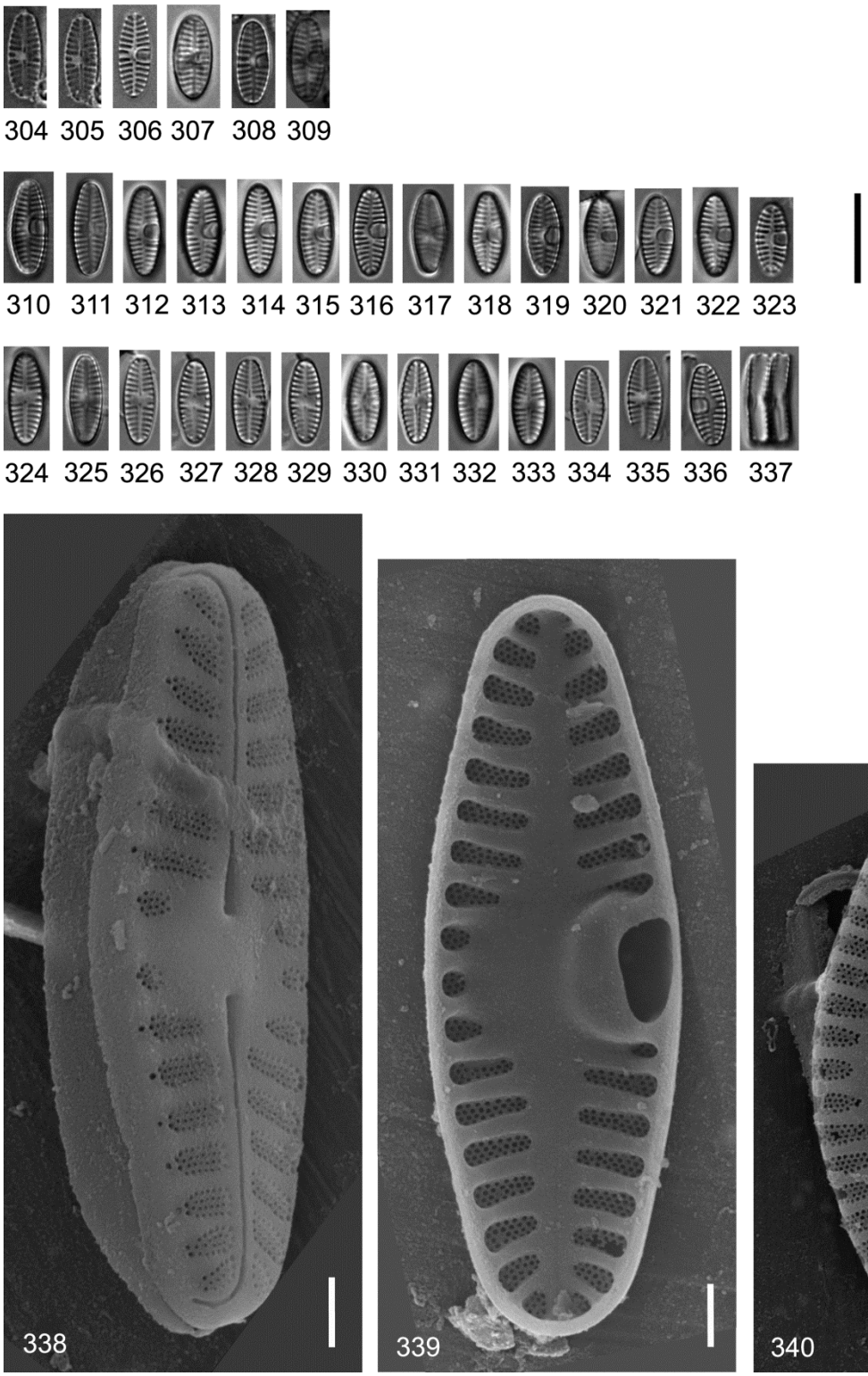



Figs 192-239: *Planothidium frequentissimum.* Figs 192-205: Strain D06_138, Berlin; Figs 206-221: Strain D06_139, Berlin; Figs 222-239: Strain D06_117b, Berlin. Figs 192-198, 206-213, 222-228, 235-237: Sternum valves; Figs 199-205, 214-221, 229-234, 238-239: Raphe valves. Figs 192-234: LM; Figs 235-239: SEM. Scale bars: Figs 192-234 = 10 µm; Figs 235-239 = 5 µm.

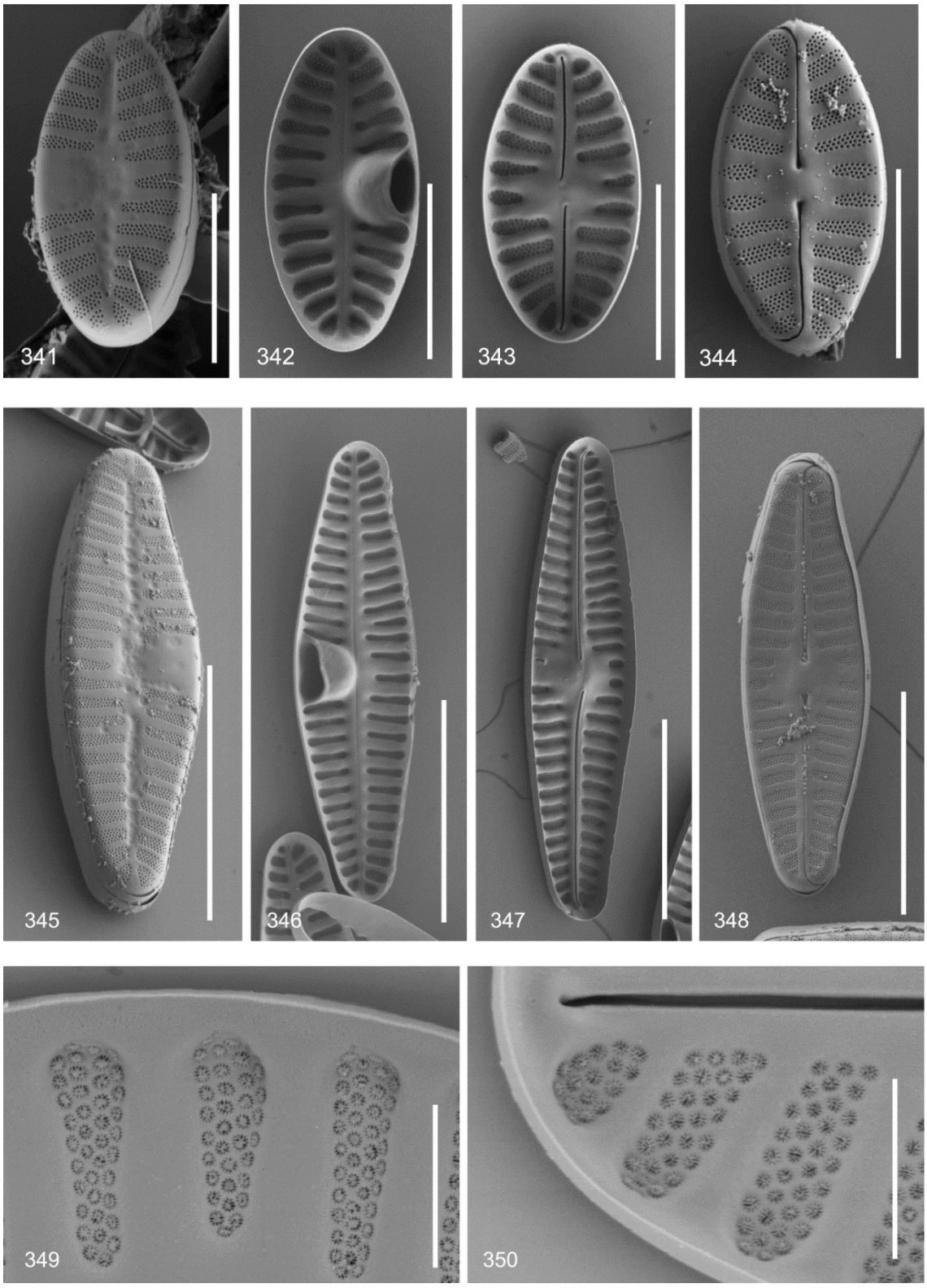
Bąk and Lange-Bertalot (2014) differentiate three small *Planothidium* taxa without sinus or cavum or other forms of a pore-free central area. Whereas *P. werumianum* Lange-Bertalot & Bąk and *P. rhombicum* Lange-Bertalot, Bąk & G. Hofmann have straight external raphe endings, *P. pumilum* Bąk & Lange-Bertalot has bent external raphe endings but the striae on the sternum valve are elevated between narrow virgae. *Planothidium engelbrechtii* (Cholnoky) Round et Bukhtiyarova (see Compère and Van de Vijver (2009)) has a similar outline but the external stria and virga patterns on the sternum valve are different.



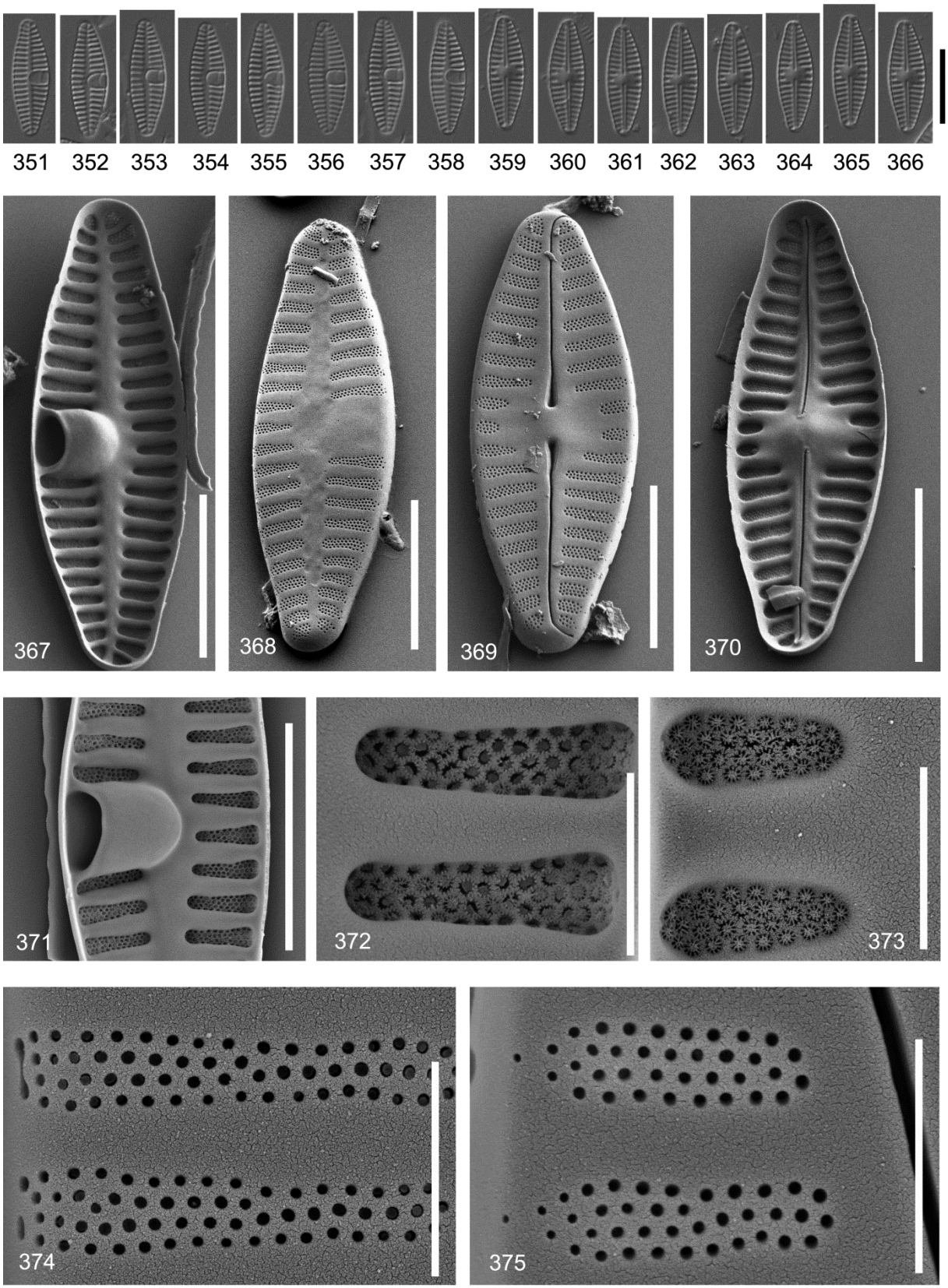

**Figs 240-303.** *Planothidium victori* (Syn.: *P. caputium*), LM. Figs 240-250: Strain D06_014, Berlin (epitype slide of *P. victori* = holotype slide of *P. caputium*); Figs 251-261: Strain D06_113, Berlin. Fig. 262-271: Strain D31_043, Mexico. Figs 272-277: Strain D100_015, Mexico; Figs 278-284: Strain D109_018, Mexico; Figs 285-293: Strain D101_022, Mexico; Figs. 294-303: Strain D109_020, Mexico. Scale bar = 10 µm.



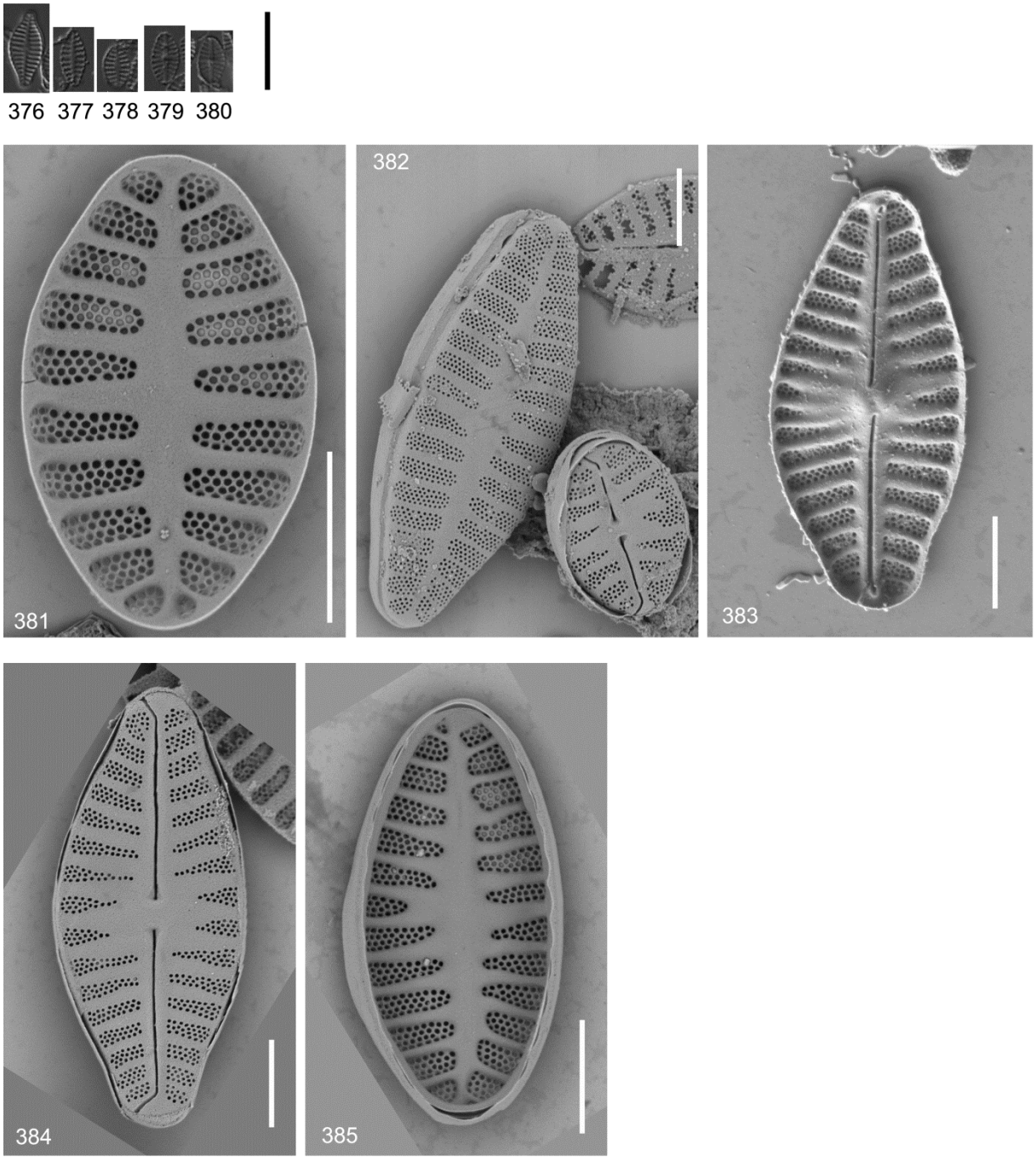

**Figs 304-340.** *Planothidium victori* (Syn.: *P. caputium*).Figs 304-309: Strain B086-3, Lake Baikal. Figs 310-340: Strain B144, Lake Baikal. Figs 338: Raphe valve; Figs 339-340: Sternum valves. Figs 304-337: LM; Figs 338-340: SEM. Scale bars: Figs 304-337 = 10 µm; Figs 338-340 = 1µm.




**Figs 341-350.** *Planothidium victori* (Syn.: *P. caputium*). Strain D109_018, Mexico, SEM. Figs 341-342, 345-346, 349: Sternum valves, Figs 343-344, 347-348, 350: Raphe valves; Figs 342-343, 346-347, 349-350: Internal view, Figs 341, 344, 345, 348: External valve view. Fig 349: Central area of sternum valve showing four rows of same size areolae in one stria. Fig 350: Apex area of raphe valve showing four rows of same size areolae in one stria. Scale bars: Figs 341-344 = 5 µm; 345-348= 10 µm; 349-350= 1 µm.




**Figs 351-375.** *Planothidium naradoense,* strain D23_024, Korea, Holotype. Figs 351 – 358, 367 – 368, 371-372, 374: Sternum valves; Figs 359-366, 369-370, 373, 375: Raphe valves. Figs 367, 370-373: Internal valve view; Figs 368, 369, 374, 375: External valve views. Fig 371: detail of cavum; Fig. 372-375: detail of two striae with four rows of same size areolae in one stria. Figs 351- 366: LM; Figs 367-375: SEM. Scale bars: Figs 351- 366 = 10 µm; 367-371 = 10 µm; 372 - 375 = 1 µm.




**Figs 376-385.** *Planothidium suncheonmanense.* Strain Ko0408, Korea, Holotype. Figs 375-380: LM; Figs 381-385: SEM. Figs 381, 381 left, 385: Sternum valves, Figs 382 right, 383, 384: Raphe valves; Figs 381, 383, 385: Internal valve view, Figs 382, 384: External valve view, Scale bars Figs 375-380 = 10 µm; Figs 381-385 = 2 µm.

*Planothidium lacustre* Álvares-Blanco, Cejudo-Figueiras & S. Blanco (Blanco et al. 2013) also has a similar outline but the sternum valve has only two rows of areolae, whereas strain Ko0408 has four

rows of areolae per multiseriate stria. *Planothidium galaicum* Álvares-Blanco & S. Blanco (Álvarez-Blanco and Blanco 2013) from a similar habitat, a marine coast, also has a similar outline but the external sternum valve is very different, with siliceous granules and unstructured striae. Areola rows could not be counted due to the poor quality of the SEM pictures. Morales (2006) added another species to this group of small *Planothidium* species: *P. daui* (Foged) Lange-Bertalot, *P. granum* (M.H. Hohn & Hellerman) Lange-Bertalot. Only *P. lemmermannii* (Hustedt) E. Morales lacks elevated striae between narrow virgae. There are also fewer areola rows within the multiseriate striae, three on the raphe valve and two on the sternum valve versus 3-4 on both valves in strain Ko0408 (Figs. 15: 6-10). We are therefore describing our taxon as a new species (*P. suncheonmanense* R. Jahn & J. Zimmermann, see below).

### 5.4.3 Nomenclatural and Taxonomic Consequences

*Planothidium cryptolanceolatum* R. Jahn & N. Abarca sp.nov. (Figs 59-75)

*Holotype*: B 40 0040799, strain D21_002, represented by Figure 59.

*Type locality*: Korea; CholIaNamdo, creek at TaeAnSa, collected 13 October 2004.

*Registration*: http://phycobank.org/ 100001

*Diagnosis*. Differs from *P. lanceolatum* in its more slender and slightly asymmetric outline; the sinus is less pronounced; i.e. it is shallower and rather irregularly half-moon shaped towards the mantle; the raphe valve has 1-3 shortened striae beside the central area versus 2-5 striae in *P. lanceolatum*; in the sternum valve the number of rows of areolae in the multiseriate striae are 2 (3) versus 3 (4) for *P. lanceolatum* and the striae continue onto the valve mantle, with merged 2-3 areolae offset from the valve striae. It differs from *P. amphibium* C.E. Wetzel, Ector & L. Pfister in its outline. The latter has capitate to subcapitate, as well as some valves with protracted apices (Wetzel et al. (2014).

Valve dimensions of type strain (n=20): length 15.3–20.7 µm, width 4.8–5.9 µm; 13.5-14.5 striae per 10µm. Dimensions based on all seven strains: length 15.1–27.3 µm, width 4.8–7.4 µm; 11.0-14.5 striae per 10µm (Tables 1, 2).

Valve outline is elliptical to ellipticallanceolate and slightly asymmetric; the apices are slightly rostrate and rounded. The axial area is narrow on both valves. The central area on the raphe valve forms an asymetrical rectangular fascia, with 1- 3 shortstriae on each side. The central area on the sternum valve has on one side a shallow half moon shaped sinus which is bounded at the other side

by (0)-1-3 shorter striae forming a somewhat small and roundish central area. Striae radiate, more strongly radiate towards the apices. Raphe branches are externally straight with drop-like proximal raphe endings; internally they are weakly deflected to opposite sides. Distal raphe fissures strongly unilaterally deflected externally, internally terminating in a very small helictoglossae.

The multiseriate striae are composed of several rows of small round areolae. There are generally two rows of areolae per stria on the sternum valve; if a third row is present, the areolae of the middle row are markedly smaller. On the raphe valve, the striae comprise 3-4 rows of same size areolae. In the raphe valve, The striae end at the valve edge on the raphe valve, but continue some distance onto the mantle of the sternum valve. The two or three areolae on the mantle are offset from the valve striae and forming a bigger oval or linear opening. Internally, each areola is covered by a hymenate occlusion.

*Planothidium taeansa* R. Jahn & N. Abarca sp.nov. (Figs 172-191)

*Holotype*: B 40 0040801, strain D26_002, represented by Figure 172.

*Type locality*: Korea; ChollaNamdo, creek at TaeAnSa, collected 13 October 2004.

*Registration*: http://phycobank.org/ 100002

*Diagnosis*: Differs from *P. lanceolatum,* (Van de Vijver et al. 2013) and *P. cryptolanceolatum* by its broad elliptic outline with rostrate and narrowly rounded apices; the outline resembles somewhat *P. rostrolanceolatum* Van de Vijver, Kopalová & Zidarova (Van de Vijver et al. 2013) and *P. dubium* (Grunow) Round & Bukhtiyarova but differs from them in the lesser number of areolae rows in the multiseriate striae (see Potapova (2011). It differs from *P. amphibium* C.E. Wetzel, Ector & L. Pfister in its different outline with capitate as well as protracted apices also differs in its breadth (6.7-7.5 µm: compare to 5-6 µm for *P. amphibium*) and striae density (12,5-13 striae per 10µm: compare to 13-16 striae per 10µm in *P. amphibium* (Wetzel et al. (2014).

Valve dimensions of type strain (n=22): length 18.5–19.4 µm, width 6.7–7.5 µm; 12.5-13.0 striae per 10µm.

Valve outline is elliptic-lanceolate with rostrate, drawn-out apices. The axial area is narrow on both valves. The central area on the raphe valve forms a roundish to rectangular fascia, with 2-3 more or less shorter irregularly spaced striae on each side. The central area on the sternum valve has on one side a slightly rimmed more or less half-moon shaped sinus which is bordered at the other side by 1-2 (3) slightly shortened striae to form a vague roundish area. The striae are slightly radiate, more radiate towards the apices. The raphe branches are externally straight with drop-like proximal raphe

endings; internally they are weakly deflected to opposite sides. Distal raphe fissures strongly unilaterally deflected externally, internally terminating in a small helictoglossae.

The multiseriate striae are composed of several rows of small round areolae. There are generally two rows of areolae per striae on the sternum valve; if a third row is present, the areolae of the middle row are markedly smaller. On the raphe valve, the striae comprise made up of 3 (4) rows of same size areolae. In the raphe valve, the striae end at the valve edge on the raphe valve, but continue some distance onto the mantle of the sternum valve. The two or three areolae on the mantle are offset from the valve striae and are merged into a bigger oval or linear opening. Internally, each areola is covered by a hymenate occlusion which have been destroyed in our specimen.

*Planothidium naradoense* R. Jahn & J. Zimmermann sp.nov. (Figs 351-375)

Holotype: B 40 0040800, strain D23_024, represented by Figure 352.

Type locality: Korea; ChollaNamdo, NaeNarado Island, Spring, Lat 34.533042° Lon 127.463672 ± 50 m, 114m asl, collected 14 October 2004.

Registration: http://phycobank.org/100003

Diagnosis: Differs by the form of the cavum which is oblong with parallel sides reaching slightly over the axial area versus a round cavum in *P. frequentissimum* and V-form in *P. caputium;* as well as in the tight hood opening (SEM) seen as a roundish line close to mantle (LM) versus a wider opening of the hood (SEM) seen as an almost straight line not close to the mantle (LM) as for *P. victori (= P. caputium)*. In the feature cavum it differs also from *P. comperei* and *P. piaficum* each having a small aperture; in *P. piaficum* its borders are joining the neighboring interstriae; *P. bagualensis*, *P. biporomum* and *P. incuriatum* show a wide open cavum with its borders attached to neighboring striae.

Valve dimensions (n=20): length 15.0–16.0 µm, width 4.7–5.2 µm; 13.5-14.5 striae per 10µm.

The valves are lanceolate to elliptic-lanceolate and slightly asymmetric, the apices are slightly rostrate and rounded. The axial area is narrow on both valves. The central area on the raphe valve forms a roundish fascia, with two shorter regular striae on each side. The central area on the sternum valve has on one side an oblong shaped cavum with parallel sides which reaches into and sometimes over the axial area and nearly hits the striae on the other side. The opening of the hood is quite tight seen as a roundish line close to the mantle (LM). Striae slightly radiate, more radiate towards the apices. Raphe branches are externally straight with drop-like proximal raphe endings;

internally they are weakly deflected to opposite sides. Distal raphe fissures strongly unilaterally deflected externally, internally terminating in a very small helictoglossae.

The multiseriate striae are composed of several rows of small round areolae. On both valves there are generally 3-4 rows of similar sized areolae. On both valves, the striae continue some distance onto the mantle. Internally, each areola is covered by hymenate occlusions (Figs 372, 373).

*Planothidium suncheonmanense* R. Jahn & J. Zimmermann sp.nov. (Figs 376-385)

*Holotype*: B 40 0041403, strain Ko0408, represented by Figure 376.

*Type locality*: Korea; ChollaNamdo, SunCheonMan, collected October 2008.

*Registration*: http://phycobank.org/ 100004

*Diagnosis*: Differs from *P. lanceolatum* and *P. frequentissimum* by the missing sinus and cavum. It also has no other marked feature in the center of the sternum valve such as distant striae as in *Planothidium minutissimum* (Krasske) Lange-Bertalot. It differs from taxa with a similar outline mainly in different areolae and costae patterns on the external sternum valve such as from *P. pumilum* Bąk & Lange-Bertalot (Bąk and Lange-Bertalot 2014) and from *Planothidium engelbrechtii* (Cholnoky) Round et Bukhtiyarova (see Compère and Van de Vijver (2009)) by their elevated costae; from *P. galaicum* Álvares-Blanco & S. Blanco (Álvarez-Blanco and Blanco 2013) from a similar habitat by its siliceous granules (striae are undiscernible), and from *Planothidium lacustre* Álvares-Blanco, Cejudo-Figueiras & S. Blanco (Blanco et al. 2013) by only 2 rows of areolae whereas *P. suncheonmanense* has 4 rows in the multiseriate stria.

Valve dimensions (n=15): 5.5-10.7 µm length, 3.2-4.6 µm width, 15-17 radial striae per 10µm.

The valve outline is elliptic-lanceolate to almost elliptical; in larger valves, the apices are slightly drawn-out and rounded. The axial area is narrow on both valves but slightly expanding at the central area. The central area on the raphe valve forms a slightly rectangular fascia, with 1- 2 slightly shorter and pointed striae on both sides. A central area on the sternum valve is none existent. Striae slightly radiate, more strongly radiate towards the apices. Raphe branches are externally straight with drop-like proximal raphe endings; internally they are weakly deflected to opposite sides. Distal raphe fissures strongly unilaterally deflected externally, internally terminating in a very small helictoglossae.

The multiseriate striae are composed of several rows of small round areolae. There are generally 4-5 rows of same sized areolae on the sternum valve. On the raphe valve, the striae comprise of up to 4

rows of same sized areolae. Except for the central area, the virgae are narrower than the striae. Internally, each areola is covered by a pore occlusion which has been destroyed in our specimens.

*Planothidium victori* Novis, Braidwood & Kilroy in Phytotaxa 64: 33. 2012.

*Holotype*: CHR618408; cleaned frustules made from culture LCR-S:18:1:1. = *Planothidium caputium* J. Zimmermann & R. Jahn in PlosOne 9: 16. 2014, syn. nov.

*Epitype* (hic designatus): B 40 0040871; cleaned frustules made from authentic strain D06_014 (Holotype of *Planothidium caputium*).

*Registration*: http://phycobank.org/ 100005

*Comments:* When *P. victori* was described on the basis of a teratological clone (Novis et al. 2012) with morphometrics that did not overlap with our measurements for *P. caputium* from Berlin (Zimmermann et al. 2014), we were not aware of the conspecificity of the species. According to additional molecular data for 18SV4, provided by P. Novis using the primers published in Zimmermann et al. (2011) (pers. comm. P. Novis, data accordingly also deposited in INSCD), and the *rbc*L data, we are now convinced that the taxa are conspecific, with *P. victori* having priority. Since we find it very difficult to identify *P. victori* from the images and morphometrics in its protologue, we think the best solution for diatomists is to interpret the teratological type specimen of *P. victori* with a new epitype showing the holotype of the now synonymized *P. caputium*.

## 5.5 Discussion

### 5.5.1 Taxonomy

The most obvious split in the molecular data is between the two clades of *Planothidium* strains, with a sinus or with a cavum. Until about 30 years ago this group of taxa was considered part of *A. lanceolata* (Cassie (1989); Germain (1981); Hustedt (1930); Patrick and Reimer (1966), including many varieties mainly based on their valve outlines. The discovery (Moss & Carter 1982) of the difference between a simple depression or hood for *A. lanceolata* (with sinus) and *A. rostrata* Østrup (with cavum) respectively, led to a paradigm shift. The implications began to be implemented (Krammer and Lange-Bertalot (1991) for *A. lanceolata* ssp. *lanceolata* and *A. lanceolata* ssp. *frequentissima* (with many varieties) now also based on this new feature, seen as a simple or double lined horseshoe in LM.

Straub (1985) differentiated six demes (Sippen) from the Swiss Alps within the *A. rostrata* (cavum)-group; one of these may fall within *P. victori*. Other workers have recently followed suit, clarifying type species as well as considering biogeographical occurrences (Spaulding et al. 2008; Van de Vijver et al. 2013).

In this study, we identified clear parallel clades, with either a sinus or a cavum, versus one strain lacking either feature, and postulate that this feature is of particular taxonomic significance. Further molecular studies on other *Planothidium* taxa will show if a differentiation between a sinus or cavum holds; for example if taxa with a cavum, such as *P. comperei* and *P. piaficum* (N'Guessan et al. (2014), *P. biporomum* and *P. incuriatum* (Wetzel et al. (2013), *P. bagualensis* (Wetzel and Ector 2014) cluster with *P. frequentissium* and *P. victori* (= *P. caputium*). We (Zimmermann et al. 2014) have shown that the form of the cavum opening in *P. caputium* seems to be an important feature when comparing this taxon to *P. frequentissimum* whose cavum is much more closed. We have never observed a cavum with its borders joining the neighboring striae, as seems typical for the above mentioned species *P. piaficum, P. biporomum, P. incuriatum, P. bagualensis*. The same can be postulated for the sinus clade: since *P.* cf. *subantarcticum* clusters within the sinus clade, will *P. amphibium*, *P. haynaldii* (Schaarschmidt) Lange-Bertalot and *P. capitatum* (O. Müller) Van de Vijver, Kopalová, C.E. Wetzel & Ector (Wetzel et al. 2014) also cluster within our sinus clade once molecular data become available? Except for the external sternum valve, the form of the cavum and their outline, all these taxa have similar micromorphologies of the multiseriate striae. The form of the sinus and its adjacent striae might also turn out to be a feature when enough data has been assembled.

Recently two new genera with a cavum have been published, but we are certain that this structure in *Gliwiszia* Kulikovskiy, Lange-Bertalot & Witkowski (Kulikovskiy et al. 2013) is not homologous with the cavum in *Planothidium* since it is present not only on the sternum valve but also on the raphe valve. Stria micromorphology in *Skabitschewskia* Kulikovskiy & Lange-Bertalot (Kulikovskiy et al. 2015a) and *Gliwiszia* is different from that in *Planothidium.* The former genera have uniseriate or biseriate striae with apparently differently structured hymenes.

Following the presence of a sinus or cavum, valve outline and morphometric data are considered the most important characters for differentiating taxa. Wetzel et al. (2013), (Wetzel and Ector 2014), Wetzel et al. (2014) and N'Guessan et al. (2014) have established a number of new species based on these features. But none of these new species fit our new taxa in outline. We think that these features are important but can be variable as can be seen with the extended morphological data of *P. lanceolatum* from the Faroe Islands, as well as *P. victori* (and its synonym *P. caputium*) from Mexico (Table 2). Van de Vijver et al. (2013) showed that within the multiseriate striae of *P. subantarcticum* the central row of areolae is much smaller. We have also found this feature in our strains, but this

seems to be common to all taxa and strains within the *P. lanceolatum* clade, but only for the sternum valve (*P. lanceolatum*, Fig. 51; *P. cryptolanceolatum,* Figs 74, 153; *P.* cf. *subantarcticum,* Figs 166, 169, 170; *P. taeansa,* Fig. 190). Van de Vijver et al. (2013) also described the presence of single pores on the valve mantle of type material of *P. lanceolatum.* We have also observed this in *P. lanceolatum* (Fig. 51) but not in *P.* cf. *subantarcticum* (Fig. 170). However, we think that the offset merged areolae on the valve mantle of the sternum valve, as seen in *P. cryptolanceolatum* (Figs 74, 153) and *P. taeansa* (Fig. 190), might be a good differentiating feature. Interestingly, *P. comperei* (N'Guessan et al. 2014) with a cavum also contains smaller areolae in the middle row, although from our study these are typical for the sinus clade.

Morales (2006) recognised four groups of *Planothidium* based on the characteristics of the central area of the sternum valve: 1. continuous striae with no central interruption (i.e. *P. daui*), 2. variable distant striae / clear space at the central region (i.e. *P. minutissimum*), 3. sinus or depression (i.e. *P. lanceolatum*), 4. cavum (i.e. *P. frequentissimum*). We basically agree with his groupings but we would like to split off taxa without terminal raphe fissures curved to the secondary side, following Spaulding et al. (2008) in their description of *Planothidium.* Taxa without curved terminal fissures would fall into *Platessa* Lange-Bertalot. In addition, or rather instead, we would recognise another group (or better clade) which also lacks a central interruption as for Morales' group 1, but does not have externally elevated striae above the sunken virgae in the sternum valve, i.e. our new species *P. suncheonmanense* (Figs 376-385). Based on the molecular data, this species constitutes a separate clade to the other two (*P. lanceolatum* and *P. frequentissimum* clades) as shown in Fig. 1. More molecular and better micro-morphological data are needed to understand how these groups cluster together, and to understand their phylogeny.

### 5.5.2 Distribution

As explained above, until about 30 years ago this group of taxa was considered part of a single species (as *A. lanceolata*) with a cosmopolitan distribution in all types of waters. This underlines the argument that coarse-grained taxonomy (Mann and Droop (1996) favours the ubiquitous dispersal hypothesis for small organisms (< 1 mm) in suitable habitats (Finlay et al. 2002), better known as "everything is everywhere, the environment selects" (Beijerinck 1913; De Wit and Bouvier 2006). Our combined morphological and molecular approach differentiated two main clades (**L** and **F),** four subclades (**L1**, **L2**, **F1** and **F2)** and three other taxa (D26_002, D23_024 and Ko0408) and showed that this group of 36 strains contains at least eight taxa from sites in eight countries on four continents (Germany, France & Denmark [Faroe Islands] in Europe, Russia [Lake Baikal] & Korea in Asia, Mexico & USA in North America, plus New Zealand), isolated from the benthos of different water types (two

springs, two waterfalls, two creeks, five streams and rivers and the shores of a huge lake). They show an interesting distribution pattern (see Table 4). The eighth species, *P. suncheonmanense*, serves as taxonomic comparison to the main group but was marine and will be excluded from further discussion.

The six strains from Germany (identified as *P. frequentissimum*, *P. victori* (Syn.:*P. caputium*) and *P. lanceolatum*) were isolated from one sample from the mesotrophic Wuhle stream, a tributary of the River Spree in Berlin. The strain from France ( as *P. lanceolatum*) was found in the River Le Lourdan (Keck et al. 2016). The two strains from the Faroe Islands, Denmark,(*P. lanceolatum* and *P.* cf. *subantarcticum*) were isolated from two oligotrophic samples (a stream and a waterfall) from different islands. The four strains from Lake Baikal (*P. victori* and *P. lanceolatum*) were taken from two shore sites. The six strains from Korea (*P. cryptolanceolatum*, *P. naradoense* and *P. taeansa*) were isolated from four different samples on two different occasions from three different eco-regions (one from a waterfall, one from a spring, and two from a stream). The nine Mexican strains were isolated from five samples on two different occasions (one from an oligotrophic spring, one from a eutrophic stream, three from two oligotrophic to mesotrophic streams, and one from a mesotrophic to eutrophic stream). One of the USA strains, L1249 from the Czarnecki Collection, identified as *P. lanceolatum*, was probably collected from the Des Moines River of Northern Central USA (Medlin and Kaczmarska 2004) and clusters with our *P. lanceolatum*. There is some habitat information for the other three strains, and the molecular data allow us to re-identify one of the strains. PL2_*P. lanceolatum* from Ridley Creek, PA (T. Nakov et al. unpub.) clusters within *P. cryptolanceolatum,* PL3_ *P. lanceolatum* from Big Creek, MT (T. Nakov et al. unpub.), clusters close to but separate from *P. taeansa,* and PF1_*P. frequentissimum* from Pack Forest Lake, NY (T. Nakov et al. unpub.) clusters close to but separate from *P. naradoense*. This shows that there is more genetic diversity in clades **L** and **F** than we determined.

*Planothidium amphibium* which is recorded as a dominant species in several samples from Oregon, USA (Wetzel et al. 2014) might be represented by the molecular data of PL3 from Montana, USA. The two strains from New Zealand (clustering within *P. frequentissimum* and *P. caputium*) were isolated from one sample from the polluted River Styx (Novis et al. 2012). If similar habitat is an important factor for the simultaneous occurrence of species, it should be noted that none of the studied habitats are really similar, except perhaps for the two springs in Mexico and Korea, and the two waterfalls in Korea and the Faroe Islands. However, neither of them host the same species and neither of them has the same ecological data. But it seems that the anthropogenically influenced water bodies in Berlin, Mexico, Lake Baikal, USA, and New Zealand were the most prone to host *P. victori*.

Table 4: Numbers and Distribution Patterns of the Studied *Planothidium* Strains Concerning their Identification and Countries of Occurrence.

| | **Germany** | **France** | **Faroe Islands** | **Baikal, Russia** | **Korea** | **Mexico** | **USA** | **NewZealand** |
|---|---|---|---|---|---|---|---|---|
| | **[6]** | **[1]** | **[2]** | **[4]** | **[7]** | **[8]** | **[4]** | **[2]** |
| *Planothidium lanceolatum* [4] | x | | x | x | | | x | |
| *Planothidium* cf. *subantarcticum* [1] | | | x | | | | | |
| *Planothidium taeansa* [1] | | | | | x | | | |
| *Planothidium cryptolanceolatum* [8] | | | | | xxxx | xxx | x | |
| *Planothidium frequentissimum* [5] | xxx | x | | | | | | x |
| *Planothidium naradoense* [1] | | | | | x | | | |
| *Planothidium victori = P. caputium* [11] | xx | | | xxx | | xxxxx | | x |
| *Planothidium* sp. [2] | | | | | | | xx | |
| *Planothidium suncheonmanense* [1] | | | | | x | | | |

*Planothidium frequentissimum* (five strains) was only found in Germany, France and New Zealand, and was not the most frequent *Planothidium* species. But *P. victori* (Syn.: *P. caputium*), its "look-alike", turned out to be most common, with 11 strains from five sites on all four studied continents. There is an overlap in two cases where both species are found in the same sample (in Berlin and New Zealand), both heavily anthropogenically disturbed. Neither taxon was found in Korea but a new species from the *P. frequentissimum* clade was described, as *P. naradoense.* A different geographical pattern seems to apply to the *P. lanceolatum* clade with four strains of *P. lanceolatum* found in Berlin, the Faroe Islands, Lake Baikal and the USA, whereas its "look-alike" *P. cryptolanceolatum* (seven strains) was restricted to Korea, Mexico and the USA. Except for the USA, there seems to be no geographical overlap between these two species. However, there is an overlap with another species found in the same sample from Korea, which is described as new, *P. taeansa*. The "look-alike" to the recently described *P. subantarcticum* from the Subantarctic was here identified as *P.* cf. *subantarcticum,* isolated from a sample from the Faroe Islands, in the Subarctic.

The eight Korean strains are the most diverse, with four species. The two purely Korean strains, one in each major clade are particularly interesting. The Korean Algal Flora (Joh 2012) lists *A. lanceolata* and its ssp. *frequentissima* (and *rostrata)* but none of the published pictures correspond to our concept of *P. lanceolatum* and *P. frequentissimum*, showing how force-fitting Central European Diatom concepts misrepresents an indigenous flora. In this study alone, four new species contained in Korean strains are described; three of them known only from Korea. The Korean Algal Flora (Joh 2012) illustrates a number of specimens of which only one (Joh 2012: fig. 44: J) shows some resemblance to our new species *P. cryptolanceolatum*; all the other illustrations seem to be of different species.

### 5.5.3 Biogeography

Well documented organisms from a variety of habitats throughout the world form the basis for biogeographical research. In the case of unicellular organisms such as diatoms this is nearly impossible to achieve, since sampling habitats and worldwide sites is difficult for specific single living microscopic cells, which need to be isolated and established as clonal cultures. Because cultures enable the variability of phenotypes to be linked to an individual genotype, they are fundamental for molecular and fine-grained morphological taxonomy. Therefore, our biogeographical conclusions can only be tentative.

Interpreting our findings biogeographically is difficult because there are different distribution patterns, some of which overlap. Just as for *Gomphonema narodoense* R. Jahn, N. Abarca, J. Zimmermann & Enke (Abarca et al. 2014), some species seem to be restricted to Korea (*P.*

*naradoense*, *P. taeansa*, *P. suncheonmanense*). And here the similarities end. *Planothidium lanceolatum* and *P. frequentissimum,* which used to be recorded from all over the world are not as common as thought and are restricted in our study to Europe (Germany and France), USA, and New Zealand. The one taxon restricted to Korea, Mexico and the USA is *P. cryptolanceolatum*, but this seems to reflect neither a similar geological habitat nor a disjunct distribution but probably reflects sampling gaps between these three countries. The central Mexican sites where it was found are too far inland to allow on distribution via ocean or wind currents.

Biogeographically, *P. victori* (*= P. caputium*) could be called a cosmopolitan taxon since it has been found in almost all our sampling places, even on the shores of Lake Baikal, but not in the Subarctic or Korea. As for plants there also seems to be the category of "ruderal species" in diatoms, i.e. taxa that colonize preferably anthropogenically enriched water bodies (see also Kulikovskiy et al. (2015b). But this clade is also turns the most diverse molecularly and morphologically. There are differences in the molecular data but they are not above 1.1 %. Nevertheless this highly variable clade might be in the process of speciation.

Most interesting is the finding of *P.* cf. *subantarcticum* in the subarctic. Although there are slight morphological differences, we hesitated to describe this species as new just because it occurs on the other side of the world. Molecular data of true *P. subantarcticum* are needed to prove whether we are dealing with look-alikes or cryptic and discrete species (see also (Cox 1995; Mann 1999).

With a polyphasic approach which incorporates valve variability offered by cultures, micro-morphology as discovered with SEM, and diverse molecular data we finally have tools to move from coarse grained diatom taxonomy to a more refined taxonomy which enables us to discover that diatoms have some biogeography (Cocquyt and Jahn 2007; Moser et al. 1998; Poulícková et al. 2010). However, the problem of under-sampling remains a deficit in this endeavor, since sampling and culturing of freshwater diatoms is currently rather random, only scratching the surface of diatom bio- or phylogeography. Nevertheless, this study offers a few more pieces to help unravel the puzzle.

## 5.6 Acknowledgements

We gratefully acknowledge Eileen Cox polishing our English, and Phil Novis, Bart Van de Vijver and Carlos Wetzel who provided comments and data on the taxonomy of *Planothidium* species which they had studied. The Mexican samples from 2014 were taken by DM under permit CONAPESCA PPF/DGOPA-149/14. Additionally, the Berlin Group acknowledges the help of Jana Bansemer for the retrieval of molecular data and of Monika Luchow at the SEM.

## 5.7 Funding

The work of MK and EG was partly supported by the Russian Science Foundation (14-14-00555). The work of DM was supported by the Mexican Government through CONACYT and CONCYTEQ. The Mexican samples from 2014 were taken by DM under permit CONAPESCA PPF/DGOPA-149/14. Additionally, the Berlin Group acknowledges the help of Jana Bansemer for the retrieval of molecular data and of Monika Lüchow at the SEM; it also acknowledges funding by the Deutsche Forschungsgemeinschaft Grant INST 130/839-1 FUGG for SEM, by DFG Grant JA 874/8-1 for Names Registration, and by BMBF Grant 01LI1501E for GBOL-2 Diatoms.

## 5.8 Supplemental data

Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/0269249X.2017.1312548.

## 5.9 References

Abarca N, Jahn R, Zimmermann J, Enke N (2014) Does the cosmopolitan diatom *Gomphonema parvulum* (Kützing) Kützing have a biogeography? PLoS ONE 9(1): e86885. https://doi.org/10.1371/journal.pone.0086885.

Álvarez-Blanco I, Blanco S (2013) *Planothidium galaicum* sp. nov. (Bacillariophyta, Achnanthidiaceae) , a new diatom species from Galician coast, Spain. Phytotaxa 151(1): 44-52. https://doi.org/10.11646/phytotaxa.151.1.4.

Bąk M, Lange-Bertalot H (2014) Four small-celled *Planothidium* species from Central Europe proposed as new to science. Oceanological and Hydrobiological Studies 43(4): 346-359. https://doi.org/10.2478/s13545-014-0152-9.

Beijerinck MW (1913) De infusies en de ontdekking der backteriën. Jaarboek van de Koninklijke Akademie voor Wetenschappen. Amsterdam, the Netherlands: Müller. (Reprinted in Verzamelde geschriften van M.W. Beijerinck, vijfde deel, pp. 119–140. Delft, 1921).

Blanco S, Álvarez-Blanco I, Cejudo-Figueiras C, Espejo JMR, Barrera CB, Bécares E, del Olmo FD, Artigas RC (2013) The diatom flora in temporary ponds of Doñana National Park (southwest Spain): five new taxa. Nordic Journal of Botany 31(4): 489-499. https://doi.org/10.1111/j.1756-1051.2013.01691.x.

Brinkmann N, Hodač L, Mohr KI, Hodačová A, Jahn R, Ramm J, Hallmann C, Arp G, Friedl T (2015) Cyanobacteria and Diatoms in Biofilms of Two Karstic Streams in Germany and Changes of

Their Communities Along Calcite Saturation Gradients. Geomicrobiology Journal 32(3-4): 255-274. https://doi.org/10.1080/01490451.2014.901438.

Bukhtiyarova LM, Lyakh AM (2014) Functional morphology of the horseshoe spot in the frustule of *Planothidium* species (Bacillariophyta). Ukrainian Botanical Journal 71(1): 223-227.

Bukhtiyarova LM, Round FE (1996) Revision of the genus *Achnanthes* sensu lato section *Marginulatae* Bukh. sect. nov. of *Achnanthidium* Kütz. Diatom Research 11(1): 1-30.

Cassie V (1989) A contribution to the study of New Zealand Diatoms. Bibliotheca Diatomologica 17(1-266.

Cleve PT, Grunow A (1880) Beiträge zur Kenntniss der arctischen Diatomeen Kongliga Svenska Vetenskaps-Akademiens Handlingar 17(1-121.

Cocquyt C, Jahn R (2007) Endemic taxa related to *Surirella ovalis* from Tanzania, East Africa: Typification of O. Müller's taxa and description of a new species. Cryptogamie Algologie 28(107-116.

Compère P, Van de Vijver B (2009) *Planothidium engelbrechtii* (Cholnoky) Round & Bukhtiyarova: Identity and lectotypification (Bacillariophyta). Fottea 9(2): 187-192. https://doi.org/10.5507/fot.2009.019.

Cox EJ (1995) Studies on the diatom genus *Navicula* Bory. VII. The identity and typification of *Navicula gregaria* Donkin, *N. cryptocephala* Kütz. and related taxa. Diatom Research 10(1): 91-111. https://doi.org/10.1080/0269249X.1995.9705330.

De Wit R, Bouvier T (2006) 'Everything is everywhere, but, the environment selects'; what did Baas Becking and Beijerinck really say? Environmental Microbiology 8(4): 755-758. https://doi.org/10.1111/j.1462-2920.2006.01017.x.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19): 2460-2461. https://doi.org/10.1093/bioinformatics/btq461.

Finlay BJ, Monaghan EB, Maberly SC (2002) Hypothesis: the rate and scale of dispersal of freshwater diatom species is a function of their global abundance. Protist 153(null): 261-273.

Gemeinholzer B, Droege G, Zetzsche H, Knebelsberger T, Raupach M, Borsch T, Klenk HP, Haszprunar G, Waegele JW (2009+).DNA Bank Network Webportal. In. http://www.dnabank-network.org

Germain H (1981) Flore des diatomées. Diatomophycées eaux douces et saumâtres du Massif Armoricain et des contrées voisines d'Europe occidentale. Société Nouvelle des Éditions Boubée, Paris, France.

Guillard RRL, Lorenzen CJ (1972) Yellow-green algae with chlorophyllide C. Journal of Phycology 8(1): 10-14. https://doi.org/10.1111/j.1529-8817.1972.tb03995.x.

Hofmann G, Werum M, Lange-Bertalot H (2013) Diatomeen im Süßwasser-Benthos von Mitteleuropa. Koeltz Scientific Books, Königstein, Germany.

Hustedt F (1930) Die Süsswasser-Flora Mitteleuropas. Bacillariophyta (Diatomeae) Verlag von Gustav Fischer, Jena, Germany.

Jahn R, Kusber WH (2005+) AlgaTerra Information System (online). Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin. http://www.algaterra.org. https://doi.org/http://www.algaterra.org.

Joh G (2012) Algal Flora of Korea, Vol.3, Nr. 7; Freshwater diatoms V. National Institute of Biological Resources.

Keck F, Rimet F, Franc A, Bouchez A (2016) Phylogenetic signal in diatom ecology: perspectives for aquatic ecosystems biomonitoring. Ecological Applications 26(3): 861-872. https://doi.org/10.1890/14-1966.

Krammer K, Lange-Bertalot H (1991) Bacillariophyceae 4. Teil Achnanthaceae. Kritische Ergänzungen zu *Navicula* (Lineolatae) und *Gomphonema*. – In: Ettl, H., Gärtner, G., Gerloff, J., Heynig, H. & Mollenhauer, D. (eds): Süsswasserflora von Mitteleuropa 2/4. Gustav Fischer Verlag, Stuttgart-Jena. 1-437 pp.

Krammer K, Lange-Bertalot H (2004) Bacillariophyceae 4. Teil Achnanthaceae. Kritische Ergänzungen zu *Navicula* s. str. und *Gomphonema*. – In: Ettl, H., Gärtner, G., Gerloff, J., Heynig, H. & Mollenhauer, D. (eds): Süsswasserflora von Mitteleuropa 2/4. – 468 pp., Spektrum Akademischer Verlag, Heidelberg, Berlin.

Kulikovskiy M, Lange-Bertalot H, Kuznetsova IV (2015a) Lake Baikal: Hotspot of endemic diatoms II. Iconographia Diatomologica 26. Koeltz Scientific Books, Oberreifenberg, Germany, 656 pp.

Kulikovskiy M, Lange-Bertalot H, Witkowski A (2013) *Gliwiczia* gen. nov. a new monoraphid diatom genus from Lake Baikal with a description of four species new for science. Phytotaxa 109(1): 1-16. https://doi.org/10.11646/phytotaxa.109.1.1.

Kulikovskiy M, Lange-Bertalot H, Witkowski A, Khursevich GK, Kociolek JP (2015b) New species of *Eunotia* (Bacillariophyta) from Lake Baikal with comments on morphology and biogeography of the genus. Phycologia 54(3): 248-260. https://doi.org/doi:10.2216/14-98.1.

Kützing FT (1846) Kurze Mittheilung über einige kieselschalige Diatomeen. Botanische Zeitung 4(247-248.

Lange-Bertalot H (1993) 85 Neue Taxa und über 100 weitere neu definierte Taxa ergänzend zur Süßwasserflora von Mitteleuropa Vol. 2/1-4. – Bibliotheca Diatomologica 27. 454 pp. J. Cramer, Berlin, Stuttgart.

Lange-Bertalot H (1999) Neue Kombinationen von Taxa aus *Achnanthes* Bory (sensu lato). – Iconographia Diatomologica 6: 270-283.

Maddison W, Maddison D (2016).Mesquite: a modular system for evolutionary analysis. Version 3.10 http://mesquiteproject.org. In.

Mann DG (1999) The species concept in diatoms. Phycologia 38(6): 437-495. https://doi.org/10.2216/i0031-8884-38-6-437.1.

Mann DG, Droop SJM (1996) 3. Biodiversity, biogeography and conservation of diatoms. Hydrobiologia 336(1): 19-32. https://doi.org/10.1007/bf00010816.

Medlin LK, 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. https://doi.org/10.2216/i0031-8884-43-3-245.1.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Gateway Computing Environments Workshop (GCE), 2010. 1-8 pp.

Morales EA (2006) Small *Planothidium* Round et Bukhtiyarova (Bacillariophyceae) taxa related to *P. daui* (Foged) Lange-Bertalot from the United States. Diatom Research 21(2): 325-342. https://doi.org/10.1080/0269249X.2006.9705673.

Moser G, Lange-Bertalot H, Metzeltin D (1998) Insel der Endemiten: Geobotanisches Phänomen Neukaledonien. Bibliotheca Diatomologica 38(464 p.

Moss MO, Carter JR (1982) The resurrection of *Achnanthes rostrata* Østrup. Bacillaria 5(157-164.

Müller J, Müller K, Neinhuis C, Quandt D (2010) PhyDE®–Phylogenetic Data Editor. Computer program.

N'Guessan KR, Wetzel CE, Ector L, Coste M, Cocquyt C, Van de Vijver B, Yao SS, Ouattara A, Kouamelan EP, Tison-Rosebery J (2014) *Planothidium comperei* sp. nov. (Bacillariophyta), a new diatom species from Ivory Coast. Plant Ecology and Evolution 147(3): 455-462. https://doi.org/10.5091/plecevo.2014.981.

Novis P, 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.

Patrick R, Reimer C (1966) The diatoms of the United States. Exclusive of Alaska and Hawaii. Volume 1: Fragilariaceae, Eunotiaceae, Achnanthaceae, Naviculaceae. – Monographs of the Academy of Natural Sciences of Philadelphia 1-688.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14(9): 817-834.

Potapova M (2010a) *Planothidium frequentissimum*. In Diatoms of the United States. Retrieved July 13, 2016, from http://westerndiatoms.colorado.edu/taxa/species/planothidium_frequentissimum.

Potapova M (2010b) *Planothidium lanceolatum*. In Diatoms of the United States. Retrieved July 13, 2016, from http://westerndiatoms.colorado.edu/taxa/species/planothidium_lanceolatum.

Potapova M (2011) *Planothidium dubium*. In Diatoms of the United States. Retrieved July 13, 2016, from http://westerndiatoms.colorado.edu/taxa/species/planothidium_dubium.

Poulícková A, Vesela J, Neustupa J, Skaloud P (2010) Pseudocryptic diversity versus cosmopolitanism in diatoms: a case study on *Navicula cryptocephala* Kutz. (Bacillariophyceae) and morphologically similar taxa. Protist 161(3): 353-369. https://doi.org/10.1016/j.protis.2009.12.003.

Rambaut A (2008) Figtree v1.2.2. Insitute of Evolutionary Biology, University of Edinburgh. Available at http://tree.bio.ed.ac.uk/software/figtree.

Romero OE, Jahn R (2013) Typification of *Cocconeis lineata* and *Cocconeis euglypta* (Bacillariophyta). Diatom Research 28(2): 175-184. https://doi.org/10.1080/0269249x.2013.770801.

Souffreau C, Verbruggen H, Wolfe AP, Vanormelingen P, Siver PA, Cox EJ, Mann DG, Vijver BVd, Sabbe K, Vyverman W (2011) A time-calibrated multi-gene phylogeny of the diatom genus *Pinnularia*. Molecular Phylogenetics and Evolution 61(3): 866-879. https://doi.org/10.1016/j.ympev.2011.08.031.

Spaulding SA, Edlund MB, Metzeltin D (2008) *Planothidium*. In Diatoms of the United States. Retrieved March 03, 2015, from http://westerndiatoms.colorado.edu/taxa/genus/Planothidium.

Stamatakis A (2006) "RAxML-VI-HPC: Maximum Likelihood-based Phylogenetic Analyses with Thousands of Taxa and Mixed Models". Bioinformatics 22(21): 2688-2690.

Stamatakis A (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. Bioinformatics 30(9): 1312-1313. https://doi.org/10.1093/bioinformatics/btu033.

Stamatakis A, Hoover P, Rougemont J (2008) "A Fast Bootstrapping Algorithm for the RAxML Web-Servers". Systematic Biology 57(5): 758-771.

Straub F (1985) Variabilité comparée d'*Achnanthes lanceolata* (Bréb.) Grun. et d'*Achnanthes rostrata* Östrup (Bacillariophyceae) dans huit populations naturelles du Jura suisse I: aproche morphologique. Bulletin de la Société Neuchâteloise des Sciences Naturelles 108(135-150.

Straub F (1990) Compared variability of *Achnanthes lanceolata* (Bréb.) Grunow. 2. Biometrical approach of several races of the sub-species *frequentissima* Lange-Bertalot. In: Ricard M (Ed) Ouvrage dedie à la memoire du Professeur Henry Germain (1903-1989). Koeltz Scientific Books, Königstein, Germany; Champaign, Illinois, USA, 243-250.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Molecular Biology and Evolution 30(12): 2725-2729. https://doi.org/10.1093/molbev/mst197.

Tavaré S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences*, **17**, 57–86, 1986. American Mathematical Society.

Van de Vijver B, Wetzel CE, Kopalová K, Zidarova R, Ector L (2013) Analysis of the type material of *Achnanthidium lanceolatum* Brébisson ex Kützing (Bacillariophyta) with the description of two new *Planothidium* species from the Antarctic Region. Fottea 13(2): 105-117.

Wetzel CE, De Vijver BV, Hoffmann L, Ector L (2013) *Planothidium incuriatum* sp. nov. a widely distributed diatom species (Bacillariophyta) and type analysis of *Planothidium biporomum*. Phytotaxa 138(1): 43-57. https://doi.org/10.11646/phytotaxa.138.1.6.

Wetzel CE, Ector L (2014) Taxonomy, distribution and autecology of *Planothidium bagualensis* sp. nov. (Bacillariophyta) a common monoraphid species from southern Brazilian rivers. Phytotaxa 156(4): 201-210. https://doi.org/10.11646/phytotaxa.156.4.2.

Wetzel CE, Van de Vijver B, Kopalová K, Hoffmann L, Pfister L, Ector L (2014) Type analysis of the South American diatom *Achnanthes haynaldii* (Bacillariophyta) and description of *Planothidium amphibium* sp. nov., from aerial and aquatic environments in Oregon (USA). Plant Ecology and Evolution 147(3): 439-454. https://doi.org/10.5091/plecevo.2014.1058.

Zimmermann J, Abarca N, Enke 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. https://doi.org/10.1371/journal.pone.0108793.

Zimmermann J, Jahn R, Gemeinholzer B (2011) Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. Organisms Diversity & Evolution 11(3): 173-192. https://doi.org/10.1007/s13127-011-0050-6.

# 6 General conclusions and outlook

The results presented in this dissertation demonstrate that adopting an integrative approach improves the diversity assessments of epilithic diatoms. In this study, morphological, ecological, molecular and phylogenetic methods were applied, along with diatom cultivation techniques. The number of integrated approaches in the analysis is not the only factor that counts to improve diversity evaluations but also the resolution offered by the methods and their critical assessment.

Regarding morphological methods, light microscopy (LM) is the most widely used technique for the identification of diatoms in ecological assessments (Morales et al. 2001). LM provides low resolution of the intricate morphology of the diatom cell walls, resulting in lumping of different species into one or a few, therefore underestimating diversity, such as in the *Planothidium lanceolatum*/*P. frequentissimum* species complexes, previously reported to have worldwide distributions due to clustering of several species into one (Jahn et al. 2017). Following thorough LM observations accompanied by SEM examinations led to the identification of the largest diatom diversity reported for Mexico to date, leading to the description of new species and therefore rejecting the inherent cosmopolitanism that previous studies suggested for the Mexican diatom flora (Mora et al. 2017). The diversity observed by microscopy is also illustrated in the form of an identification guide (Chapter 3), setting the baseline for identification in future studies using diatoms for the evaluation of the ecological integrity of freshwaters of the Lerma-Chapala Basin. This research topic, though well established in several countries (Kelly et al. 2009), is a field of research that has been poorly studied in Mexico, but that can bring new outcomes in the assessment and monitoring in a country facing serious challenges to protect its freshwater resources (Sedeño-Díaz and López-López 2007).

Another approach to diversity analysis is by the use of the DNA barcoding principle (Hebert et al. 2003) in environmental DNA (eDNA) metabarcoding (Taberlet et al. 2012). Although metabarcoding holds an enormous potential for diatom diversity assessments of freshwaters when compared to light microscopy analysis (Zimmermann et al. 2015), it also has shortcomings, with its success of identification at the species level heavily relying on the barcode marker and the barcode reference library. The eDNA metabarcoding approach conducted in this dissertation retrieved a larger diversity than the microscopy analyses, with one quarter of the taxa assigned to species level due to the herein assembled regional reference library. Even though eDNA metabarcoding retrieved a larger diversity than microscopy, it did not catch all the diversity identified by morphology, highlighting the complementarity of both methods in assessing the diversity of environmental freshwater samples. The comparison of abundance data retrieved from microscopy and High-Throughput Sequencing

showed large disparities. This can prove a major drawback in eDNA metabarcoding monitoring studies, overestimating the abundance of a taxon (Vasselon et al. 2017). But this can be overcomed by calculating correction factors for cell size and gene copies of the barcoding marker. To add up, amplicon PCR is another source of uncertainty in abundance data, due to differential primer efficiency, specificity and template competition, a major challenge to be solved because primer efficiency differs between species (Kermarrec et al. 2013). Although bioinformatics procedures are progressing at an accelerating rate, underestimation or overestimation of diversity is another issue that was faced during this study, which is dependent on the bioinformatics procedures followed when establishing species boundaries through sequence similarity. As implemented in this study, this issue can be ameliorated to a certain extent by the use of tree-based approaches implementing evolutionary models in the delimitation of species boundaries (Monaghan et al. 2009; Zimmermann et al. 2015).

The results presented in this dissertation emphasize that adopting an integrative approach improves the diversity assessment of epilithic diatoms, with methods complementing each other. Since an accurate and comprehensive estimation of diversity is one of the foundations in the application of diatoms to evaluate ecological integrity in freshwaters, the here presented studies provide the baseline for epilithic diatom diversity from the Lerma-Chapala Basin for future biomonitoring studies.

Hebert PD, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc Biol Sci 270(1512): 313-321. https://doi.org/10.1098/rspb.2002.2218.

Jahn R, Abarca N, Gemeinholzer B, Mora D, Skibbe O, Kulikovskiy M, Gusev E, Kusber WH, Zimmermann J (2017) *Planothidium lanceolatum* and *Planothidium frequentissimum* reinvestigated with molecular methods and morphology: four new species and the taxonomic importance of the sinus and cavum. Diatom Research 32(1): 75-107. https://doi.org/10.1080/0269249x.2017.1312548.

Kelly M, Bennett C, Coste M, Delgado C, Delmas F, Denys L, Ector L, Fauville C, Ferreol M, Golub M, Jarlman A, Kahlert M, Lucey J, Ni Chathain B, Pardo I, Pfister P, Picinska-Faltynowicz J, Rosebery J, Schranz C, Schaumburg J, van Dam H, Vilbaste S (2009) A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. Hydrobiologia 621(1): 169-182. https://doi.org/10.1007/s10750-008-9641-4.

Kermarrec L, Franc A, Rimet F, Chaumeil P, Humbert JF, Bouchez A (2013) Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: a test for freshwater diatoms. Mol Ecol Resour 13(4): 607-619. https://doi.org/10.1111/1755-0998.12105.

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJ, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Syst Biol 58(3): 298-311. https://doi.org/10.1093/sysbio/syp027.

Mora D, Carmona J, Jahn R, Zimmermann J, Abarca N (2017) Epilithic diatom communities of selected streams from the Lerma-Chapala Basin, Central Mexico, with the description of two new species. PhytoKeys 88(39-69. https://doi.org/10.3897/phytokeys.88.14612.

Morales EA, Siver PA, Trainor FR (2001) Identification of diatoms (Bacillariophyceae) during ecological assessments: Comparison between Light Microscopy and Scanning Electron Microscopy techniques. Proceedings of the Academy of Natural Sciences of Philadelphia 151(1): 95-103. https://doi.org/10.1635/0097-3157(2001)151[0095:Iodbde]2.0.Co;2.

Sedeño-Díaz JE, López-López E (2007) Water Quality in the Río Lerma, Mexico: An Overview of the Last Quarter of the Twentieth Century. Water Resources Management 21(10): 1797-1812. https://doi.org/10.1007/s11269-006-9128-x.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol 21(8): 2045-2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x.

Vasselon V, Rimet F, Tapolczai K, Bouchez A (2017) Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France). Ecological Indicators 82(1-12. https://doi.org/10.1016/j.ecolind.2017.06.024.

Zimmermann J, Glöckner G, Jahn R, Enke N, Gemeinholzer B (2015) Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Mol Ecol Resour 15(3): 526-542. https://doi.org/10.1111/1755-0998.12336.

# 7 Appendices

## 7.1 Appendix 1

Supplementary material 1. Diatom taxa list from the Lerma–Chapala Basin, Central Mexico. Taxa names are accompanied by the identification reference as well as the plate and figure(s) matching our observations in the references. Online identification references from the Diatoms of the United States website do not include plate or figure number. The complete citation of the identification reference is given at the end of the taxa list.

* Indicates new record for the Lerma–Chapala Basin.

| **Taxa** | **Identification reference** | |
|---|---|---|
| *Achnanthes inflata* var. *inflata* (Kützing) Grunow | Metzeltin et al. 2005 | Pl. 28: 2 |
| *Achnanthidium* aff. *catenatum* (J.Bílý et Marvan) Lange-Bertalot | Hofmann et al. 2013 | Pl. 23: 69–72 |
| *Achnanthidium exiguum* (Grunow) Czarnecki | Rumrich et al. 2000 | Pl. 27: 14–15 |
| ** Achnanthidium exile* (Kützing) Round et Bukhtiyarova | Hofmann et al. 2013 | Pl. 23: 14 |
| *Achnanthidium minutissimum* (Kützing) Czarnecki | Hofmann et al. 2013 | Pl. 23: 18–19 |
| *Achnanthidium* sp. 1 | | |
| *Achnanthidium* sp. 2 | | |
| *Achnanthidium* sp. 3 | | |
| *Achnanthidium* sp. 4 | | |
| *Achnanthidium* sp. 5 | | |
| *Achnanthidium* sp. 6 | | |
| *Amphora pediculus* (Kützing) Grunow | Levkov 2009 | Pl. 78: 40–47 |
| ** Brachysira altepetlensis* D. Mora, R. Jahn et N. Abarca | This study | |
| ** Brachysira brebissonii* Ross | Lange-Bertalot and Moser 1994 | Pl. 41: 17 |
| ** Brachysira microcephala* (Grunow) Compère | Hamilton 2010 | |
| *Brachysira* sp. 1 | | |
| *Brachysira* sp. 2 | | |
| *Caloneis bacillum* (Grunow) Cleve | Hofmann et al. 2013 | Pl. 67: 28 |

| Taxa | Identification reference | |
|---|---|---|
| * *Caloneis clevei* var. *uruguayensis* Frenguelli | Metzeltin et al. 2005 | Pl. 154: 7–8 |
| *Caloneis schumanniana* (Grunow) Cleve | Stancheva et al. 2009 | Fig. 24 |
| * *Caloneis silicula* (Ehrenberg) Cleve | Metzeltin et al. 2005 | Pl. 154: 23 |
| *Caloneis* cf. *silicula var. elliptica* Frenguelli | Metzeltin et al. 2005 | Pl. 155: 21 |
| *Caloneis* sp. 1 | | |
| *Caloneis* sp. 2 | | |
| *Caloneis* sp. 3 | | |
| *Caloneis* sp. 4 | | |
| * *Chamaepinnularia submuscicola* (Krasske) Lange-Bertalot | Hofmann et al. 2013 | Pl. 50: 35–36 |
| cf. *Chamaepinnularia* sp. | | |
| *Cocconeis pediculus* Ehrenberg | Jahn et al. 2009 | P. 281: 10 and 15 |
| *Cocconeis* sp. 1 | | |
| *Cocconeis* sp. 2 | | |
| *Craticula accomoda* (Hustedt) D.G. Mann | Lange-Bertalot 2001 | Pl. 93: 3 |
| * *Craticula acidoclinata* Lange-Bertalot et Metzeltin | Lange-Bertalot 2001 | Pl. 87: 1 |
| *Craticula ambigua* (Ehrenberg) D.G. Mann | Lange-Bertalot 2001 | Pl. 82: 4 |
| * *Craticula buderi* (Hustedt) Lange-Bertalot | Lange-Bertalot 2001 | Pl. 90: 22 |
| *Craticula molestiformis* (Hustedt) Mayama | Krammer and Lange-Bertalot 1997a | Pl. 45: 7–9 |
| *Craticula* cf. *pumilio* Lange-Bertalot et U.Rumrich | Metzeltin et al. 2005 | Pl. 96: 7 |
| *Craticula subminuscula* (Manguin) C.E. Wetzel et Ector | Rumrich et al. 2000 | Pl. 72: 8–9 |
| * *Craticula submolesta* (Hustedt) Lange-Bertalot | Lange-Bertalot 2001 | Pl. 93: 31–34 |
| *Craticula* sp. 1 | | |
| *Craticula* sp. 2 | | |
| *Cyclostephanos invisitatus* (Hohn et Hellermann) Theriot, Stoermer et Håkasson | Håkansson 2002 | Fig. 221 |
| *Cyclotella atomus* Hustedt | Krammer and Lange-Bertalot 1991a | Pl. 51: Fig. 19 |
| *Cyclotella meneghiniana* Kützing | Krammer and Lange-Bertalot 1991a | Pl. 44: 1–2, 4 and 9 |
| *Cymbella kolbei* Hustedt | Krammer 2002 | Pl. 14: 19–21 |
| * *Cymbella tropica* Krammer et Metzeltin | Krammer 2002 | Pl. 44: 2 |
| * *Cymbopleura naviculiformis* (Auerswald) Krammer | Hofmann et al. 2013 | Pl. 83: 20–23 |
| *Diadesmis confervacea* Kützing | Metzeltin et al. 2005 | Pl. 68: 12–16 |
| * *Discostella stelligera* (Cleve et Grunow) Houk et Klee | Krammer and Lange-Bertalot 1991a | Pl. 49: 3 |
| * *Encyonema brevicapitatum* Krammer | Krammer 1997a | Pl. 34: 1–7 |
| *Encyonema* cf. *hebridiforme* Krammer | Krammer 1997a | Pl. 11: 16 |

| Taxa | Identification reference | |
|---|---|---|
| * *Encyonema jemtlandicum* Krammer | Krammer 1997a | Pl. 35: 1–5 |
| * *Encyonema jemtlandicum* var. *venezolanum* Krammer | Krammer 1997a | Pl. 14: 4 |
| * *Encyonema minutiforme* Krammer | Krammer 1997a | Pl. 18: 12 and 14 |
| *Encyonema* cf. *minutiforme* Krammer | Krammer 1997a | Pl. 18: 12 and 14 |
| * *Encyonema minutum* (Hilse) D.G. Mann | Krammer 1997a | Pl. 25: 5–10, 16–19 |
| * *Encyonema pergracile* Krammer | Krammer 1997a | Pl. 88: 6 |
| *Encyonema silesiacum* (Bleisch) D.G. Mann | Krammer 1997a | Pl. 4: |
| *Encyonema triangulum* (Ehrenberg) Kützing | Krammer 1997a | Pl. 78: 1 and 5 |
| *Encyonema* sp. 1 | | |
| * *Encyonopsis subminuta* Krammer et E.Reichardt | Krammer 1997b | Pl. 144: 6–11 |
| *Encyonopsis* cf. *thienemannii* (Hustedt) Krammer | Krammer 1997b | Pl. 149: 28–29 |
| *Encyonopsis* sp. 1 | | |
| *Eolimna* sp. 1 | | |
| *Eolimna* sp. 2 | | |
| *Eolimna* sp. 3 | | |
| *Eolimna* sp. 4 | | |
| *Epithemia adnata* (Kützing) Brébisson | Metzeltin et al. 2005 | Pl. 190: 8–11 |
| *Epithemia sorex* Kützing | Levkov et al. 2007 | Pl. 198: 13–15 |
| *Epithemia turgida* (Ehrenberg) Kützing | Hofmann et al. 2013 | Pl. 120: 1 |
| * *Eunotia bidens* Ehrenberg | Lange-Bertalot et al. 2011 | Pl. 80: 2 |
| *Eunotia* cf. *bigibba* var. *pumila* Grunow | Metzeltin and Lange-Bertalot 2007 | Pl. 90: 12 |
| * *Eunotia kruegeri* Lange-Bertalot | Werum and Lange-Bertalot 2004 | Pl. 5: 8 |
| *Eunotia major* var. *major* (W. Smith) Rabenhorst | Metzeltin et al. 2005 | Pl. 17: 4–5 |
| *Eunotia* cf. *meridiana* Metzeltin et Lange-Bertalot | Metzeltin et al. 2005 | Pl. 20: 22–28 |
| * *Eunotia metamonodon* Lange-Bertalot | Lange-Bertalot et al. 2011 | Pl. 217: 4–5 |
| * *Eunotia minor* (Kützing) Ehrenberg | Hofmann et al. 2013 | Pl. 15: 5–6 |
| *Eunotia* cf. *monodon* Ehrenberg | Metzeltin et al. 2005 | Pl. 19: 18 |
| * *Eunotia mucophila* (Lange-Bertalot, Nörpel-Schempp et Alles) Lange-Bertalot | Lange-Bertalot et al. 2011 | Pl. 23: 21–23 |
| * *Eunotia tridentula* Ehrenberg | Metzeltin et al. 2005 | Pl. 21: 2–6 |
| *Eunotia* sp. 1 | | |
| *Eunotia* sp. 2 | | |
| *Eunotia* sp. 3 | | |
| *Eunotia* sp. 4 | | |

| Taxa | Identification reference | |
|---|---|---|
| *Eunotia* sp. 5 | | |
| *Eunotia* sp. 6 | | |
| *Fallacia pygmaea* (Kützing) D.G. Mann | Metzeltin et al. 2005 | Pl. 51: 15–18 |
| *Fistulifera saprophila* (Lange-Bertalot et Bonik) Lange-Bertalot | Lange-Bertalot 2001 | Pl. 111: 1 |
| * *Fragilaria austriaca* (Grunow) Lange-Bertalot | Hofmann et al. 2013 | Pl. 7: 33–34 |
| *Fragilaria bidens* Heiberg | Krammer and Lange-Bertalot 1991a | Pl. 111: 21 |
| * *Fragilaria pectinalis* (O.F. Müller) Lyngbye | Wetzel and Ector 2015 | P. 280: 78–80 |
| * *Fragilaria rumpens* (Kützing) Carlson | Hofmann et al. 2013 | Pl. 16–17 |
| * *Fragilaria tenera* (W. Smith) Lange-Bertalot | Hofmann et al. 2013 | Pl. 7: 18–19 |
| * *Frustulia crassinervia* (Brébisson) Lange-Bertalot et Krammer | Metzeltin and Lange-Bertalot 1998 | Pl. 117: 10 and 13 |
| *Frustulia neomundana* Lange-Bertalot et Rumrich | Rumrich et al. 2000 | Pl. 97: 5–6 |
| *Frustulia* cf. *spicula* ssp. *spicula* Amossé | Lange-Bertalot 2001 | Pl. 138: 3–4 |
| *Frustulia* cf. *undosa* Melzeltin et Lange-Bertalot | Metzeltin and Lange-Bertalot 1998 | Pl. 117: 6–7 |
| *Frustulia vulgaris* (Thwaites) De Toni | Rumrich et al. 2000 | Pl. 96: 7 |
| *Geissleria decussis* (Østrup) Lange-Bertalot et Metzeltin | Metzeltin et al. 2005 | Pl. 91: 24–25 |
| *Gomphonema acuminatum* Ehrenberg | Levkov et al. 2016 | Pl. 3: 1 |
| *Gomphonema affine* var. *affine* Kützing | Reichardt 1999 | Pl. 7: 2–3 |
| *Gomphonema exilissimum* (Grunow) Lange-Bertalot et E.Reichardt | Levkov et al. 2016 | Pl. 127: 1–28 |
| * *Gomphonema graciledictum* E. Reichardt | Levkov et al. 2016 | Pl. 44: 20 and 24 |
| *Gomphonema* aff. *graciledictum* E. Reichardt | Levkov et al. 2016 | Pl. 44: 20 and 24 |
| *Gomphonema innocens* E. Reichardt | Levkov et al. 2016 | Pl. 123: 11–32 |
| *Gomphonema lagenula* Kützing | Levkov et al. 2016 | Pl. 102: 39–47 |
| *Gomphonema* cf. *lagenula* Kützing | Levkov et al. 2016 | |
| Gomphonema cf. lippertii E. Reichardt et Lange-Bertalot | Reichardt 1999 | Pl. 21: 6–7 |
| *Gomphonema* aff. *mariovense* Levkov et Tofilovska | Levkov et al. 2016 | Pl. 110: 10 |
| *Gomphonema minusculum* Krasske | Levkov et al. 2016 | Pl. 161: 1–25 |
| * *Gomphonema naviculoides* W. Smith | Reichardt 2015 | Fig. 26–27 |
| *Gomphonema* cf. *naviculoides* W. Smith | Reichardt 2015 | Fig. 26–27 |
| *Gomphonema parvulum* (Kützing) Kützing | Levkov et al. 2016 | Pl. 102: 1–19 |
| * *Gomphonema parvuliforme* Levkov, Mitic-Kopanja et E.Reichardt | Levkov et al. 2016 | Pl. 105: 1–34 |
| *Gomphonema* cf. *parvuliforme* Levkov, Mitic-Kopanja et E. Reichardt | Levkov et al. 2016 | Pl. 105: 1–4 |
| *Gomphonema* aff. *parvulius* (Lange-Bertalot et E. Reichardt) Lange-Bertalot et E. Reichardt | Levkov et al. 2016 | Pl. 124: 42–44 |
| *Gomphonema pumilum* (Grunow) E. Reichardt et Lange-Bertalot | Levkov et al. 2016 | Pl. 151: 36–39 |

| Taxa | Identification reference | |
|---|---|---|
| *Gomphonema salae* Lange-Bertalot et E. Reichardt | Metzeltin and Lange-Bertalot 1998 | Pl. 157: 4 |
| *Gomphonema* aff. *sarcophagus* W. Gregory | Levkov et al. 2016 | Pl. 94: 13 |
| *Gomphonema stonei* E. Reichardt | Reichardt 1999 | Pl. 13: 12 |
| *Gomphonema subclavatum* (Grunow) Grunow | Hofmann et al. 2013 | Pl. 95: 23–24 |
| *Gomphonema* sp. 1 | | |
| *Gomphonema* sp. 2 | | |
| *Gomphonema* sp. 3 | | |
| *Gomphonema* sp. 4 | | |
| *Halamphora montana* (Krasske) Levkov | Levkov 2009 | Pl. 93: 10–12 |
| *Halamphora* cf. *pseudomontana* (Cholnoky) Levkov | Levkov 2009 | Pl. 104: 20 |
| *Halamphora veneta* (Kützing) Levkov | Levkov 2009 | Pl. 94: 17–19 |
| *Hantzschia abundans* Lange-Bertalot | Metzeltin et al. 2005 | Pl. 212: 5–6 |
| *Hantzschia amphioxys* (Ehrenberg) Grunow | Metzeltin et al. 2005 | Pl. 212: 10–12 |
| *Hantzschia* sp. 1 | | |
| * *Humidophila contenta* (Grunow) Lowe, Kociolek, Johansen, Van de Vijver, Lange-Bertalot et Kopalová | Metzeltin et al. 2005 | Pl. 57: 13–14 |
| *Humidophila* sp. | | |
| *Luticola goeppertiana* (Bleisch) D.G. Mann | Hofmann et al. 2013 | Pl. 45: 22 |
| *Luticola kotschyi* (Grunow) D.G. Mann | Rumrich et al. 2000 | Pl. 60: 14 |
| *Luticola mutica* (Kützing) D.G. Mann | Hofmann et al. 2013 | Pl. 45: 36–37 and 39 |
| *Luticola* cf. *tomesii* Gerd Moser, Lange-Bertalot et Metzeltin | Levkov et al. 2013 | Pl. 25: 5–7 |
| * *Luticola tropica* Levkov, Metzeltin et A. Pavlov | Levkov et al. 2013 | Pl. 196: 12 |
| * *Luticola undulata* (Hilse) D.G. Mann | Levkov et al. 2013 | Pl. 181: 43–46 |
| *Luticola ventricosa* (Kützing) D.G. Mann | Hofmann et al. 2013 | Pl. 45: 41–42 |
| *Luticola* sp. 1 | | |
| *Luticola* sp. 2 | | |
| *Mayamaea* cf. *crassistriata* Lange-Bertalot, Cavacini, Tagliaventi et Alfinito | Lange-Bertalot et al. 2003 | Pl. 17: 4–5 |
| *Mayamaea permitis* (Hustedt) Bruder et Medlin | Lange-Bertalot 2001 | Pl. 104: 7–13 |
| *Mayamaea* sp. 1 | | |
| *Navicula angusta* Grunow | Lange-Bertalot 2001 | Pl. 2: 5 |
| *Navicula antonii* Lange-Bertalot | Lange-Bertalot 2001 | Pl. 13: 1 and 8 |
| *Navicula capitatoradiata* H. Germain | Lange-Bertalot 2001 | Pl. 29: 17 |
| *Navicula cryptocephala* Kützing | Lange-Bertalot 2001 | Pl. 17: 4–5 |
| *Navicula* cf. *cryptocephala* Kützing | Hofmann et al. 2013 | Pl. 31: 8 |

| **Taxa** | **Identification reference** | |
|---|---|---|
| *Navicula cryptotenella* Lange-Bertalot | Lange-Bertalot 2001 | Pl. 26: 25 |
| *Navicula erifuga* Lange-Bertalot | Lange-Bertalot 2001 | Pl. 35: 14–16 |
| *Navicula germainii* Wallace | Lange-Bertalot 2001 | Pl. 35: 7–9 |
| *Navicula gregaria* Donkin | Metzeltin et al. 2005 | Pl. 45: 16–17 |
| *Navicula* cf. *hoffmanniae* Lange-Bertalot | Lange-Bertalot 2001 | Pl. 20: 21 |
| *Navicula libonensis* Schoeman | Lange-Bertalot 2001 | Pl. 43: 7–10 |
| * *Navicula notha* Wallace | Lange-Bertalot 2001 | Pl. 40: 22–28 |
| * *Navicula reichardtiana* Lange-Bertalot | Lange-Bertalot 2001 | Pl. 13: 25–35 |
| *Navicula riediana* Lange-Bertalot | Lange-Bertalot 2001 | Pl. 34: 1 |
| *Navicula rostellata* Kützing | Lange-Bertalot 2001 | Pl. 35: 1–6 |
| *Navicula symmetrica* R.M. Patrick | Lange-Bertalot 2001 | Pl. 39: 9–12 |
| *Navicula* cf. *tenelloides* Hustedt | Lange-Bertalot 2001 | Pl. 32: 1 |
| *Navicula trivialis* Lange-Bertalot | Lange-Bertalot 2001 | Pl. 29: 1–7 |
| *Navicula veneta* Kützing | Lange-Bertalot 2001 | Pl. 14: 23–30 |
| *Navigiolum uruguayense* Metzeltin, Lange-Bertalot et García-Rodríguez | Metzeltin et al. 2005 | Pl. 44: 9–10 |
| *Neidium* cf. *affine* (Ehrenberg) Pfitzer | Metzeltin et al. 2009 | Pl. 97: 12 |
| *Neidium ampliatum* (Ehrenberg) Krammer | Metzeltin and Lange-Bertalot 2007 | Pl. 189: 3 |
| * *Neidium amphigomphus* (Ehrenberg) Pfitzer | Metzeltin and Lange-Bertalot 2007 | Pl. 189: 2 |
| * *Neidium longiceps* (W. Gregory) R. Ross | Hofmann et al. 2013 | Pl. 53: 7 |
| *Neidium* sp. 1 | | |
| *Neidium* sp. 2 | | |
| *Nitzschia acicularis* (Kützing) W. Smith | Metzeltin et al. 2005 | Pl. 204: 12–13 |
| *Nitzschia amphibia* Grunow | Metzeltin et al. 2005 | Pl. 207: 20–33 |
| *Nitzschia* cf. *bacillum* Hustedt | Krammer and Lange-Bertalot 1997b | Pl. 78: 12 |
| *Nitzschia balcanica* Hustedt | Kociolek 2011a | |
| *Nitzchia clausii* Hantzch | Krammer and Lange-Bertalot 1997b | Pl. 19: 3and 5 |
| *Nitzschia communis* Rabenhorst | Krammer and Lange-Bertalot 1997b | Pl. 79: 3–4 |
| *Nitzschia desertorum* Hustedt | Krammer and Lange-Bertalot 1997b | Pl. 70: 10–11 |
| *Nitzschia dissipata* var. *dissipata* (Kützing) Grunow | Krammer and Lange-Bertalot 1997b | Pl. 11: 2 |
| *Nitzschia fonticola* Grunow | Krammer and Lange-Bertalot 1997b | Pl. 75: 9, 13–15 |
| *Nitzschia frustulum* (Kützing) Grunow | Krammer and Lange-Bertalot 1997b | Pl. 68: 1–4 |
| *Nitzschia gracilis* Hantzsch | Krammer and Lange-Bertalot 1997b | Pl. 66: 7 |
| *Nitzschia* cf. *hantzschiana* Rabenhorst | Krammer and Lange-Bertalot 1997b | Pl. 73: 16–17 |
| *Nitzschia inconspicua* Grunow | Krammer and Lange-Bertalot 1997b | Pl. 69: 4 |

| Taxa | Identification reference | |
|---|---|---|
| *Nitzschia intermedia* Hantzsch | Krammer and Lange-Bertalot 1997b | Pl. 61: 6–8 |
| *Nitzschia lanceolata* W. Smith | Krammer and Lange-Bertalot 1997b | Pl. 16: 1 |
| *Nitzschia linearis* (Agardh) W. Smith | Metzeltin et al. 2005 | Pl. 206: 3–4 |
| *Nitzschia palea* (Kützing) W. Smith | Hofmann et al. 2013 | Pl. 111: 1–9 |
| *Nitzschia palea* var. *debilis* (Kützing) Grunow | Kociolek 2011b | |
| *Nitzschia palea* var. *tenuirostris* Grunow | Kociolek 2011c | |
| *Nitzschia paleacea* Grunow | Krammer and Lange-Bertalot 1997b | Pl. 81: 2–5 |
| * *Nitzschia perminuta* (Grunow) Peragallo | Krammer and Lange-Bertalot 1997b | Pl. 72: 3–4 |
| *Nitzschia recta* Hantzsch | Krammer and Lange-Bertalot 1997b | Pl. 12: 9 |
| *Nitzschia semirobusta* Lange-Bertalot | Lange-Bertalot 1993 | Pl. 120: 4–8 |
| *Nitzschia sinuata* var. *delognei* (Grunow) Lange-Bertalot | Krammer and Lange-Bertalot 1997b | Pl. 40: 8 |
| *Nitzschia solita* Hustedt | Krammer and Lange-Bertalot 1997b | Pl. 71: 3 |
| * *Nitzschia sublinearis* Hustedt | Krammer and Lange-Bertalot 1997b | Pl. 58: 13 |
| *Nitzschia supralitorea* Lange-Bertalot | Krammer and Lange-Bertalot 1997b | Pl. 70: 17–19 |
| * *Nitzschia tubicola* Grunow | Krammer and Lange-Bertalot 1997b | Pl. 63: 10 |
| *Nitzschia umbonata* (Ehrenberg) Lange-Bertalot | Krammer and Lange-Bertalot 1997b | Pl. 51: 1–2 |
| *Nitzschia* sp. 1 | | |
| *Nitzschia* sp. 2 | | |
| *Nitzschia* sp. 3 | | |
| *Nitzschia* sp. 4 | | |
| *Nitzschia* sp. 5 | | |
| *Nitzschia* sp. 6 | | |
| * *Nupela praecipua* (E. Reichardt) E. Reichardt | Rumrich et al. 2000 | Pl. 33: 12 |
| * *Nupela wellneri* (Lange-Bertalot) Lange-Bertalot | Rumrich et al. 2000 | Pl. 35: 1–3 |
| * *Pinnularia acrosphaeria* var. *parva* Krammer | Metzeltin and Lange-Bertalot 2007 | Pl. 276: 12 |
| * *Pinnularia anglica* morphodeme 1 Krammer | Krammer 2000 | Pl. 80: 7 and 11 |
| *Pinnularia anglica* morphodeme 2 Krammer | Krammer 2000 | Pl. 87: 3–5 |
| *Pinnularia borealis* var. *borealis* Ehrenberg | Krammer 2000 | Pl. 7: 8 and 13 |
| * *Pinnularia borealis* var. *scalaris* (Ehrenberg) Rabenhorst | Krammer 2001 | Pl. 8: 11 |
| *Pinnularia* cf. *brebissonii* var. *acuta* Cleve-Euler | Krammer 2000 | Pl. 47: 5 |
| *Pinnularia divergens* W. Smith | Krammer 2000 | Pl. 29: 3–4 |
| * *Pinnularia divergentissima* var. *divergentissima* Grunow | Krammer 2000 | Pl. 11: 7 |
| *Pinnularia gibba* Ehrenberg | Rumrich et al. 2000 | Pl. 140: 11 |
| * *Pinnularia mayeri* Krammer | Krammer 1992 | Pl. 42: 2 |

| Taxa | Identification reference | |
|---|---|---|
| *Pinnularia* cf. *meridiana* var. *parallela* Metzeltin et Krammer | Metzeltin and Lange-Bertalot 1998 | Pl. 181: 3 |
| * *Pinnularia parvulissima* Krammer | Krammer 2000 | Pl. 69: 10 |
| * *Pinnularia saprophila* Lange-Bertalot, H. Kobayasi et Krammer | Krammer 2000 | Pl. 85: 14–18 |
| * *Pinnularia subbrevistriata* Krammer | Krammer 2000 | Pl. 70: 7–8 |
| *Pinnularia* cf. *subcapitata* var. *elongata* Krammer | Krammer 1992 | Pl. 39: 2–3 |
| *Pinnularia viridiformis* Krammer | Krammer 2000 | Pl. 161: 1 |
| *Pinnularia* sp. 1 | | |
| *Pinnularia* sp. 2 | | |
| *Pinnularia* sp. 3 | | |
| *Pinnularia* sp. 4 | | |
| *Pinnularia* sp. 5 | | |
| *Placoneis* cf. *constans* (Hustedt) E.J. Cox | Hofmann et al. 2013 | Pl. 48: 2 |
| *Placoneis undulata* (Østrup) Lange-Bertalot | Hofmann et al. 2013 | Pl. 47: 26 |
| *Planothidium incuriatum* C.E. Wetzel, Van de Vijver et Ector | Wetzel et al. 2013 | P. 49: 62–67, 71–73 |
| *Planothidium cryptolanceolatum* R. Jahn et N. Abarca | Jahn et al. 2017 | Figs 122–146 |
| *Planothidium rostratum* (Østrup) Lange-Bertalot | Krammer and Lange-Bertalot 1991b | Pl. 43: 9–11 |
| *Planothidium victori* Novis, Braidwood et Kilroy | Jahn et al. 2017 | Figs 272–277 |
| *Pseudofallacia monoculata* (Hustedt) Liu, Kociolek et Wang | Metzeltin et al. 2005 | Pl. 61: 12 and 14 |
| *Reimeria sinuata* (W. Gregory) Kociolek et Stoermer | Rumrich et al. 2000 | Pl. 117: 14 |
| *Rhopalodia gibba* (Ehrenberg) O. Müller | Metzeltin et al. 2005 | 192: 3–4 |
| *Rhopalodia operculata* (C. Agardh) Håkansson | Krammer and Lange-Bertalot 1997b | Pl. 115: 10 |
| *Sellaphora atomoides* (Grunow) C.E. Wetzel et Van de Vijver | Wetzel et al. 2015 | P. 220: 205–235 |
| *Sellaphora bacilloides* Hustedt | Metzeltin et al. 2005 | Pl. 66: 19–21 |
| * *Sellaphora blackfordensis* D.G. Mann et S. Droop | Hofmann et al. 2013 | Pl. 41: 11 |
| *Sellaphora cosmopolitana* (Lange-Bertalot) C.E. Wetzel et Ector | Rumrich et al. 2000 | Pl. 77: 36–38 |
| *Sellaphora* cf. *elorantana* (Lange-Bertalot) C.E. Wetzel | Lange-Bertalot and Metzeltin 1996 | Pl. 28: 15–17 |
| * *Sellaphora indistincta* Kociolek | Kociolek et al. 2014 | Pl. 54: 17 |
| *Sellaphora laevissima* (Kützing) D.G. Mann | Hofmann et al. 2013 | Pl. 41: 24 |
| * *Sellaphora madida* (Kociolek) C.E. Wetzel | Kociolek et al. 2014 | Pl. 28: 31–34 |
| *Sellaphora nigri* (De Notaris) C.E. Wetzel et Ector | Wetzel et al. 2015 | P. 220: 319–393 |
| *Sellaphora pupula* (Kützing) Mereschkowsky | Hofmann et al. 2013 | Pl. 41: 6–10 |
| * *Sellaphora queretana* D. Mora, N. Abarca et J. Carmona | This study | |
| *Sellaphora* cf. *rectangularis* (W. Gregory) Lange-Bertalot et Metzeltin | Lange-Bertalot and Metzeltin 1996 | Pl. 25: 10 |
| *Sellaphora saugerresii* (Desmazières) C.E. Wetzel et D.G. Mann | Wetzel et al. 2015 | P. 214: 112–127 |

| **Taxa** | **Identification reference** | |
|---|---|---|
| * *Sellaphora stauroneioides* Lange-Bertalot | Lange-Bertalot and Metzeltin 1996 | Pl. 109: 24 |
| * *Sellaphora wallacei* (Reimer) Potapova et Ponader | Potapova and Ponader 2008 | P. 173, fig. 1C |
| *Sellaphora* sp. 1 | | |
| *Sellaphora* sp. 2 | | |
| *Sellaphora* sp. 3 | | |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | Hofmann et al. 2013 | Pl. 117: 47 |
| * *Stauroneis reichardtii* Lange-Bertalot, Cavacini, Tagliaventi et Alfinito | Bahls 2010 | P. 127, fig. 130401 |
| *Stauroneis* cf. *acidoclinatopsis* Van de Vijver and Lange-Bertalot | Bahls 2010 | P. 29, fig. 452601 (1) |
| *Stauroneis* cf. *schmidiae* R. Jahn et N. Abarca | Zimmermann et al. 2014 | Fig. 4.4: h |
| *Stauroneis* sp. 1 | | |
| *Stauroneis* sp. 2 | | |
| * *Stenopterobia delicatissima* (Lewis) Van Heurck | Krammer and Lange-Bertalot 1997b | Pl. 174: 6 and 10 |
| *Surirella angusta* Kützing | Metzeltin et al. 2005 | Pl. 221: 1–7 |
| * *Surirella apiculata* var. *panduriformis* Frenguelli | Metzeltin et al. 2005 | Pl. 222: 7–9 |
| *Surirella ovalis* Brébisson | Metzeltin et al. 2005 | Pl. 220: 1–2 |
| *Surirella* cf. *pseudolinearis* Krasske | Metzeltin et al. 2005 | Pl. 225: 9 |
| *Surirella* sp. 1 | | |
| *Tryblionella apiculata* W. Gregory | Krammer and Lange-Bertalot 1997b | Pl. 35: 5 |
| *Tryblionella calida* (Grunow) D.G. Mann | Metzeltin et al. 2005 | Pl. 194: 5–6 |
| *Tryblionella hungarica* (Grunow) D.G. Mann | Metzeltin et al. 2005 | Pl. 194: 10–12 |
| *Ulnaria acus* (Kützing) Aboal | Hofmann et al. 2013 | Pl. 5: 3 |
| *Ulnaria ulna* (Nitzsch) Compère | Hofmann et al. 2013 | Pl. 5: 9 |

Bahls L (2010) *Stauroneis* in the Northern Rockies: 50 species of *Stauroneis* sensu stricto from western Montana, northern Idaho, northeastern Washington and southwestern Alberta, including 16 species described as new. Northwest Diatoms 4: 1–172.

Håkansson H (2002) A compilation and evaluation of species in the general *Stephanodiscus*, *Cyclostephanos* and *Cyclotella* with a new genus in the family Stephanodiscaceae. Diatom Research 17(1): 1–139. doi: 10.1080/0269249X.2002.9705534

Hamilton P (2010) *Brachysira microcephala*. Diatoms of the United States. October 31, 2016, from http://westerndiatoms.colorado.edu/taxa/species/brachysira_microcephala [accessed 31.10.2016]

Hofmann G, Lange–Bertalot H, Werum M (2013) Diatomeen im Süßwasser–Benthos von Mitteleuropa. Bestimmungsflora Kieselalgen für die ökologische Praxis. Über 700 der häufigsten Arten und ihre Ökologie. Koeltz Scientific Books, Königstein, 1–908.

Jahn R, Kusber W–H, Romero OE (2009) *Cocconeis pediculus* Ehrenberg and *C. placentula* Ehrenberg var. *placentula* (Bacillariophyta): Typification and taxonomy. Fottea 9(2): 275–288. doi: 10.5507/fot.2009.027

Jahn R, Abarca N, Gemeinholzer B, Mora D, Skibbe O, Kulikovskiy M, Gusev E, Kusber W–H, Zimmermann J (2017) *Planothidium lanceolatum* and *Planothidium frequentissimum* reinvestigated with molecular methods and morphology: four new species and the taxonomic importance of the sinus and cavum. Diatom Research 32(1) 75–107. doi: 10.1080/0269249X.2017.1312548

Kociolek P (2011a) *Nitzschia balcanica*. In Diatoms of the United States. http://westerndiatoms.colorado.edu/taxa/species/nitzschia_balcanica [accessed 01.09.2017]

Kociolek P (2011b) *Nitzschia palea*. In Diatoms of the United States. http://westerndiatoms.colorado.edu/taxa/species/nitzschia_palea_var._debilis [accessed 24.10.2016]

Kociolek P (2011c) *Nitzschia palea*. In Diatoms of the United States. http://westerndiatoms.colorado.edu/taxa/species/nitzschia_palea_var._tenuirostris [accessed 24.10.2016]

Kociolek JP, Laslandes B, Bennett D, Thomas E, Brady M, Graeff C (2014) Diatoms of the United States 1: taxonomy, ultrastructure and descriptions of new species and other rarely reported taxa from lake sediments in the western U.S.A. Bibliotheca Diatomologica 61. J. Cramer, Stuttgart, 1–188.

Krammer K (1992) *Pinnularia*, eine Monographie der europäischen Taxa. Bibliotheca Diatomologica 26. J. Cramer, Berlin–Stuttgart, 1–353.

Krammer K (1997a) Die cymbelloiden Diatomeen. Eine Monographie del weltweit bekannten Taxa. Teil 1. Allgemeines und *Encyonema* Part. Bibliotheca Diatomologica 36. J. Cramer, Berlin–Stuttgart, 1–382.

Krammer K (1997b) Die cymbelloiden Diatomeen. Eine Monographie del weltweit bekannten Taxa. Teil 2. *Encyonema* part., *Encyonopsis* and *Cymbellopsis*. – Bibliotheca Diatomologica 37. J. Cramer, Berlin–Stuttgart, 1–469.

Krammer K (2000) Diatoms of the European inland waters and comparable habitats. The genus *Pinnularia*. Diatoms of Europe 1. A.R.G. Gantner Verlag K.G., Ruggell. 1–703.

Krammer K (2002) Diatoms of the European inland waters and comparable habitats. *Cymbella*. Diatoms of Europe 3. A.R.G. Gantner Verlag K.G., Ruggell, 1–584.

Krammer K, Lange–Bertalot H (1997a) Bacillariophyceae. 1. Teil: Naviculaceae. In: Ettl, H, Gerloff J, Heynig H, Mollenhauer D (Eds) Süßwasserflora von Mitteleuropa. 2/1. Gustav Fischer, Jena, 1–876.

Krammer K, Lange–Bertalot H (1997b) Bacillariophyceae. 2. Teil: Bacillariaceae, Epithemiaceae, Surirellaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (Eds) Süßwasserflora von Mitteleuropa. 2/2. Gustav Fischer, Jena. 1–610.

Krammer K, Lange–Bertalot H (1991a) Bacillariophyceae. 3. Teil: Centrales, Fragilariaceae, Eunotiaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (Eds) Süßwasserflora von Mitteleuropa. 2/3. Spektrum Akademischer Verlag, Heidelberg–Berlin. 1–598.

Krammer K, Lange–Bertalot H (1991b) Bacillariophyceae. 4. Teil: Achnanthaceae. Kritische Ergänzungen zu *Navicula* (Lineolatae) und *Gomphonema* Gesamtliteraturverzeichnis Teil 1–4. In: Ettl H, Gärtner G, Gerloff J, Heynig H, Mollenhauer D (Eds) Süßwasserflora von Mitteleuropa. 2/4. Gustav Fischer Verlag, Stuttgart–Jena. 1–4347.

Lange–Bertalot H (1993) 85 Neue Taxa und über 100 weitere neu definierte Taxa ergänzend zur Süβwasserflora von Mitteleuropa Vol. 2/1–4. Bibliotheca Diatomologica 27. J. Cramer, Berlin–Stuttgart, 1– 454.

Lange–Bertalot H (2001) Diatoms of the European inland waters and comparable habitats. *Navicula* sensu stricto. 10 genera separated from *Navicula* sensu lato. *Frustulia*. Diatoms of Europe 2. A.R.G. Gantner Verlag K.G., Ruggell, 1–526.

Lange–Bertalot H, Metzeltin D (1996) Indicators of oligotrophy: 800 taxa representative of three ecologically distinct lake types: carbonate buffered – oligodystrophic – weakly buffered soft water. Iconographia Diatomologica 2. Koeltz Scientific Books, Königstein, 1–390.

Lange–Bertalot H, Moser G (1994) *Brachysira*. Monographie der Gattung. Wichtige Indikator–Species für das Gewässer Monitoring und *Naviculadicta* nov. gen. Ein Lösungsvorschlag zu dem

Problem *Navicula* sensu lato ohne *Navicula* sensu stricto. Bibliotheca Diatomologica 29. J. Cramer, Berlin–Stuttgart, 1–212.

Lange–Bertalot H, Cavacini P, Tagliaventi N, Alfinito S (2003) Diatoms of Sardinia. Rare and 76 new species in rock pools and other ephemeral waters. Iconographia Diatomologica 12. A.R.G. Gantner Verlag K.G., Ruggell, 1–438.

Lange–Bertalot H, Bąk M, Witkowski A, Tagliaventi N (2011) Diatoms of the European inland waters and comparable habitats. *Eunotia* and some related genera. Diatoms of Europe 6. A.R.G. Gantner Verlag K.G., Ruggell, 1–747.

Levkov Z (2009) Diatoms of the European inland waters and comparable habitats. *Amphora* sensu lato. Diatoms of Europe 5. A.R.G. Gantner Verlag K.G., Ruggell, : 1–916.

Levkov Z, Krstic S, Metzeltin D, Nakov T (2007) Diatoms of lakes Prespa and Ohrid. About 500 taxa from ancient lake system. Iconographia Diatomologica 16. A.R.G. Gantner Verlag K.G., Ruggell, 1–613.

Levkov Z, Metzeltin D, Pavlov A (2013) Diatoms of the European inland waters and comparable habitats. *Luticola* and *Luticolopsis*. Diatoms of Europe 7. Koeltz Scientific Books, Königstein, 1–698.

Levkov Z, Mitić–Kopanja D, Reichardt E (2016) Diatoms of the European inland waters and comparable habitats. The diatom genus *Gomphonema* in the Republic of Macedonia. Diatoms of Europe 8. Koeltz Botanical Books, Oberreifenberg, 1–552.

Metzeltin D, Lange–Bertalot H (1998) Tropical diatoms of South America I. About 700 predominantly rarely known or new taxa representative of the neotropical flora. Iconographia Diatomologica 5. Koeltz Scientific Books, Königstein, 1–695.

Metzeltin D, Lange–Bertalot H (2007) Tropical diatoms of South America II. Special remarks on biogeographic disjunction. Iconographia Diatomologica 18. A.R.G. Gantner Verlag K.G., Ruggell, 1–877.

Metzeltin D, Lange–Bertalot H, García–Rodríguez F (2005) Diatoms of Uruguay. Compared with other taxa from South America and elsewhere. Iconographia Diatomologica 15. A.R.G. Gantner Verlag K.G., Ruggell, 1–736.

Metzeltin D, Lange–Bertalot H, Nergui S, Yanling L (2009) Diatoms in Mongolia. Iconographia Diatomologica 20. A.R.G. Gantner Verlag K.G., Ruggell, 1–686.

Potapova MG, Ponader KC (2008) New species and combinations in the diatom genus *Sellaphora* (Sellaphoraceae) from southeastern United States. Harvard Papers in Botany 13(1): 171–181. doi: 10.3100/1043-4534(2008)13[171:NSACIT]2.0.CO;2

Reichardt E (1999) Zur Revision der Gattung *Gomphonema*. Die Arten um *G. affine*/*insigne*, *G. angustatum*/*micropus*, *G. acuminatum* sowie gomphonemoide Diatomeen aus dem

Oberoligozän in Böhmen. Iconographia Diatomologica 8. A.R.G. Gantner Verlag K.G., Ruggell. 1–203.

Reichardt E (2015) *Gomphonema gracile* Ehrenberg sensu stricto et sensu auct. (Bacillariophyceae): A taxonomic revision. Nova Hedwigia 101(3/4): 367–393. doi: 10.1127/nova_hedwigia/2015/0275

Rumrich U., Lange–Bertalot H, Rumrich M (2000) Diatomeen der Anden. Von Venezuela bis Patagonien/Feuerland. Und zwei weitere Beiträge. Iconographia Diatomologica 9. A.R.G. Gantner Verlag K.G., Ruggell, 1–673.

Stancheva R, Manoylov K, Gillett N (2009) Morphological variation of the *Caloneis schumanniana* species complex (Bacillariophyceae) from different environmental conditions in North American streams. Hydrobiologia 635(1): 157–170. doi 10.1007/s10750-009-9908-4

Werum M, Lange–Bertalot H, Reichardt E (2004) Diatoms in springs from Central Europe and elsewhere under the influence of hydrogeology and anthropogenic impacts: Diatomeen in Quellen unter hydrogeologischen und anthropogenen Einflüssen in Mitteleuropa und anderen Regionen. Iconographia Diatomologica 13. A.R.G. Gantner Verlag K.G., Ruggell, 1–480.

Wetzel CE, Ector L (2015) Taxonomy and ecology of *Fragilaria microvaucheriae* sp. nov. and comparison with the type materials of *F. uliginosa* and *F. vaucheriae*. Cryptogamie, Algologie, 36(3): 271–289. doi/10.7872/crya/v36.iss3.2015.271

Wetzel CE, Ector L, van de Vijver B, Compère P, Mann DG (2015) Morphology, typification and critical analysis of some ecologically important small naviculoid species (Bacillariophyta). Fottea 15(2): 203–234. doi: 10.5507/fot.2015.020

Wetzel C E, van De Vijver B, Hoffmann L, Ector L (2013) *Planothidium incuriatum* sp. nov. a widely distributed datom species (Bacillaryophyta) and type analysis of *Planothidium biporomum*. Phytotaxa 138(1): 43–57. doi: 10.11646/phytotaxa.138.1.6

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. doi:10.1371/journal.pone.0108793

## 7.2 Appendix 2

Supplementary material 1. Taxonomic reference library of epilithic diatoms of streams from the Lerma-Chapala Basin, Central Mexico. Taxon names are accompanied by strain number, INSDC accession number, locality and date of collection, collector and isolator. * Indicates taxa that were obtained as part of this project but have already been published.

| Taxon name | Strain | Locality | Date | Collector | Isolation |
|---|---|---|---|---|---|
| *Achnanthes inflata* var. *inflata* (Kützing) Grunow | D112_028 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Achnanthes inflata* var. *inflata* (Kützing) Grunow | D112_028_mix | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Achnanthes inflata* var. *inflata* (Kützing) Grunow | D112_028_1_mix | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Achnanthidium* aff. *catenatum* (J.Bílý et Marvan) Lange-Bertalot | D107_005 | La Laborcilla 2 | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Achnanthidium* aff. *catenatum* (J.Bílý et Marvan) Lange-Bertalot | D107_013 | La Laborcilla 2 | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Achnanthidium* aff. *catenatum* (J.Bílý et Marvan) Lange-Bertalot | D110_017 | Los Ailes 1 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Achnanthidium minutissimum* (Kützing) Czarnecki | D62_014 | Los Ailes 1 | 18.09.2013 | Demetrio Mora | Jana Bansemer |
| *Achnanthidium minutissimum* (Kützing) Czarnecki | D112_021 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Jana Bansemer |
| *Achnanthidium* aff. *minutissimum* (Kützing) Czarnecki | D62_018 | Los Ailes 1 | 18.09.2013 | Demetrio Mora | Jana Bansemer |
| *Achnanthidium* sp. 1 | D105_007 | Paredones | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Achnanthidium* sp. 1 | D105_011 | Paredones | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Achnanthidium* sp. 1 | D105_014 | Paredones | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Achnanthidium* sp. 1 | D105_016 | Paredones | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Achnanthidium* sp. 1 | D106_003 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Amphora* cf. *pediculus* (Kützing) Grunow | D102_005 | Ojo de Agua de Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Amphora* cf. *pediculus* (Kützing) Grunow | D102_034 | Ojo de Agua de Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Amphora* cf. *pediculus* (Kützing) Grunow | D109_004 | Guanajuatito | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Amphora* cf. *pediculus* (Kützing) Grunow | D109_022 | Guanajuatito | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Brachysira altepetlensis* D. Mora, R. Jahn et N. Abarca | D105_001 | Paredones | 07.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Brachysira altepetlensis* D. Mora, R. Jahn et N. Abarca | D105_002 | Paredones | 07.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Caloneis* sp. 1 | D60_020 | El Membrillo | 17.09.2013 | Demetrio Mora | Jana Bansemer |
| *Caloneis* sp. 2 | D89_051 | Paredones | 09.02.2014 | Demetrio Mora | Demetrio Mora |
| *Caloneis clevei* var. *uruguayensis* Frenguelli | D109_017 | Guanajuatito | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Caloneis clevei* var. *uruguayensis* Frenguelli | D109_031 | Guanajuatito | 07.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Cocconeis* sp. | D86-Coc1b | Ojo de Agua de Calvillo | 08.02.2014 | Demetrio Mora | Oliver Skibbe |
| *Diadesmis confervacea* Kützing | D103_010 | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Diadesmis confervacea* Kützing | D103_015 | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |

| | | | | | |
|---|---|---|---|---|---|
| *Diadesmis confervacea* Kützing | D104_008 | San Martín | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Diadesmis confervacea* Kützing | D104_030 | San Martín | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Diploneis* sp. | D63_112 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Demetrio Mora |
| *Diploneis* sp. | D63_120 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Demetrio Mora |
| *Encyonema minutiforme* Krammer | D113_001 | Los Ailes 2 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Encyonema minutiforme* Krammer | D113_004b | Los Ailes 2 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Encyonema minutiforme* Krammer | D113_004c | Los Ailes 2 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Encyonema minutum* (Hilse) D.G. Mann | D60_172 | El Membrillo | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Encyonema silesiacum* (Bleisch) D.G. Mann | D103_006 | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Encyonema silesiacum* (Bleisch) D.G. Mann | D103_008 | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Encyonema silesiacum* (Bleisch) D.G. Mann | D103_009 | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Encyonema silesiacum* (Bleisch) D.G. Mann | D112_014 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Jana Bansemer |
| *Encyonema silesiacum* (Bleisch) D.G. Mann | D110_011 | Los Ailes 1 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Encyonema silesiacum* (Bleisch) D.G. Mann | D110_018 | Los Ailes 1 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Eunotia minor* (Kützing) Ehrenberg | D111_006 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Eunotia minor* (Kützing) Ehrenberg | D111_015 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Eunotia minor* (Kützing) Ehrenberg | D112_024 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Jana Bansemer |
| *Fistulifera saprophila* (Lange-Bertalot et Bonik) Lange-Bertalot | D101_013 | Calvillo | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| *Fragilaria pectinalis* (O.F. Müller) Lyngbye | D106_009 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Fragilaria pectinalis* (O.F. Müller) Lyngbye | D106_014 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Fragilaria pectinalis* (O.F. Müller) Lyngbye | D106_019 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Fragilaria pectinalis* (O.F. Müller) Lyngbye | D106_021 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Fragilaria pectinalis* (O.F. Müller) Lyngbye | D111_019 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Fragilaria pectinalis* (O.F. Müller) Lyngbye | D111_020 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Fragilaria* cf. *tenera* (W. Smith) Lange-Bertalot | D106_012 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Fragilaria* cf. *tenera* (W. Smith) Lange-Bertalot | D106_015 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Fragilaria* cf. *tenera* (W. Smith) Lange-Bertalot | D106_015a | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Fragilaria* cf. *tenera* (W. Smith) Lange-Bertalot | D110_005 | Los Ailes 1 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Fragilaria* cf. *tenera* (W. Smith) Lange-Bertalot | D111_009 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Fragilaria* sp. | D113_006 | Los Ailes 2 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* cf. *exilissimum* (Grunow) Lange-Bertalot et E.Reichardt | D103_003b | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* cf. *exilissimum* (Grunow) Lange-Bertalot et E.Reichardt | D103_003a | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |

| | | | | | |
|---|---|---|---|---|---|
| *Gomphonema* cf. *exilissimum* (Grunow) Lange-Bertalot et E.Reichardt | D103_011 | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* cf. *exilissimum* (Grunow) Lange-Bertalot et E.Reichardt | D107_004 | La Laborcilla 2 | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Gomphonema* cf. *gracile* Ehrenberg | D111_001 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* cf. *gracile* Ehrenberg | D111_002 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* cf. *gracile* Ehrenberg | D111_003 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* cf. *gracile* Ehrenberg | D111_004 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* cf. *gracile* Ehrenberg | D111_011 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* cf. *gracile* Ehrenberg | D111_012 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* cf. *gracile* Ehrenberg | D111_014 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema lagenula* Kützing | D109_003 | Guanajuatito | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Gomphonema lagenula* Kützing | D100_012 | La Mesa | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Gomphonema lagenula* Kützing | D101_008 | Calvillo | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| *Gomphonema lagenula* Kützing | D101_015 | Calvillo | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| *Gomphonema parvulum* (Kützing) Kützing | D103_004 | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema parvulum* (Kützing) Kützing | D103_014 | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema parvulum* (Kützing) Kützing | D104_005 | San Martín | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Gomphonema parvulum* (Kützing) Kützing | D112_025 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Jana Bansemer |
| *Gomphonema parvulum* (Kützing) Kützing | D110_013 | Los Ailes 1 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Gomphonema parvulum* (Kützing) Kützing | D105_018 | Paredones | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Gomphonema parvulum* (Kützing) Kützing | D105_18a | Paredones | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Gomphonema parvulum* (Kützing) Kützing | D105_019 | Paredones | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Gomphonema parvulum* (Kützing) Kützing | D63_037 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Jana Bansemer |
| *Gomphonema parvulum* (Kützing) Kützing | D63_061 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Jana Bansemer |
| *Gomphonema parvulum* (Kützing) Kützing | D63_062 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Jana Bansemer |
| *Gomphonema parvulum* (Kützing) Kützing | D105_009 | Paredones | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Gomphonema parvulum* (Kützing) Kützing | D61_041 | Guanajuatito | 17.09.2013 | Demetrio Mora | Jana Bansemer |
| *Gomphonema parvulum* (Kützing) Kützing | D63_020 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Jana Bansemer |
| *Gomphonema parvulum* (Kützing) Kützing | D63_063 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Jana Bansemer |
| *Gomphonema parvulum* (Kützing) Kützing | D106_004 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Gomphonema* sp. 1 | D62_022 | Los Ailes 1 | 18.09.2013 | Demetrio Mora | Jana Bansemer |
| *Gomphonema* sp. 2 | D112_008 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Jana Bansemer |
| *Gomphonema* sp. 3 | D61_010 | Guanajuatito | 17.09.2013 | Demetrio Mora | Jana Bansemer |

| | | | | | |
|---|---|---|---|---|---|
| *Mayamaea permitis* (Hustedt) Bruder et Medlin | D101_010 | Calvillo | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| *Mayamaea* cf. *permitis* (Hustedt) Bruder et Medlin | D84_013 | La Mesa | 08.02.2014 | Demetrio Mora | Demetrio Mora |
| *Mayamaea* cf. *permitis* (Hustedt) Bruder et Medlin | D84_026 | La Mesa | 08.02.2014 | Demetrio Mora | Demetrio Mora |
| *Mayamaea* sp. 1 | D61_035 | Guanajuatito | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Mayamaea* sp. 2 | D104_007 | San Martín | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula* cf. *erifuga* Lange-Bertalot | D61_045 | Guanajuatito | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Navicula* cf. *erifuga* Lange-Bertalot | D61_068 | Guanajuatito | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Navicula* aff. *reichardtiana* Lange-Bertalot | D101_006 | Calvillo | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| *Navicula* aff. *reichardtiana* Lange-Bertalot | D101_019 | Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula symmetrica* R.M. Patrick | D113_002 | Los Ailes 2 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Navicula veneta* Kützing | D61_006 | Guanajuatito | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D61_029 | Guanajuatito | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D61_059 | Guanajuatito | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D102_014 | Ojo de Agua de Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D109_019 | Guanajuatito | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D101_031 | Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D101_048 | Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D101_049 | Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D109_021 | Guanajuatito | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D109_023 | Guanajuatito | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D109_025 | Guanajuatito | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Navicula veneta* Kützing | D109_032 | Guanajuatito | 07.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Nitzschia* cf. *amphibia* Grunow | D100_030 | La Mesa | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Nitzschia* cf. *amphibia* Grunow | D100_026 | La Mesa | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Nitzschia* cf. *amphibia* Grunow | D100_031 | La Mesa | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Nitzschia balcanica* Hustedt | D61_008 | Guanajuatito | 17.09.2013 | Demetrio Mora | Jana Bansemer |
| *Nitzschia linearis* (Agardh) W. Smith | D61_032 | Guanajuatito | 17.09.2013 | Demetrio Mora | Jana Bansemer |
| *Nitzschia palea* (Kützing) W. Smith | D60_009 | El Membrillo | 17.09.2013 | Demetrio Mora | Jana Bansemer |
| *Nitzschia palea* (Kützing) W. Smith | D108_005 | El Membrillo | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Nitzschia palea* (Kützing) W. Smith | D104_014 | San Martín | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Nitzschia palea* (Kützing) W. Smith | D104_023 | San Martín | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| *Nitzschia palea* (Kützing) W. Smith | D100_029 | La Mesa | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Nitzschia palea* (Kützing) W. Smith | D104_018 | San Martín | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Nitzschia palea* (Kützing) W. Smith | D104_020 | San Martín | 06.09.2014 | Demetrio Mora | Demetrio Mora |

| | | | | | |
|---|---|---|---|---|---|
| *Nitzschia* cf. *perminuta* (Grunow) Peragallo | D111_018 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Nitzschia* cf. *umbonata* (Ehrenberg) Lange-Bertalot | D61_049 | Guanajuatito | 17.09.2013 | Demetrio Mora | Jana Bansemer |
| *Nitzschia* cf. *umbonata* (Ehrenberg) Lange-Bertalot | D61_067 | Guanajuatito | 17.09.2013 | Demetrio Mora | Jana Bansemer |
| *Nitzschia* sp. 1 | D106_006 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Nitzschia* sp. 1 | D106_006 mix | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Nitzschia* sp. 1 | D106_020 | La Laborcilla 1 | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Nitzschia* sp. 2 | D101_037 | Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Nupela wellneri* (Lange-Bertalot) Lange-Bertalot | D112_019 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Pinnularia divergens* W. Smith | D90-Nav1 | La Laborcilla 1 | 09.02.2014 | Demetrio Mora | Oliver Skibbe |
| *Pinnularia divergens* W. Smith | D96_003 | Laguna de Servín 2 | 03.02.2014 | Demetrio Mora | Demetrio Mora |
| *Pinnularia divergens* W. Smith | D96_005 | Laguna de Servín 2 | 03.02.2014 | Demetrio Mora | Demetrio Mora |
| *Pinnularia divergens* W. Smith | D112_002 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Pinnularia* aff. *gibba* Ehrenberg | D107_002 | La Laborcilla 2 | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Pinnularia* cf. *viridiformis* Krammer | D90-Pin1 | La Laborcilla 1 | 09.02.2014 | Demetrio Mora | Oliver Skibbe |
| *Pinnularia* sp. 1 | D103_002a | Peña Colorada | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Pinnularia* sp. 1 | D103_002b | Peña Colorada | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Pinnularia* sp. 1 | D103_012 | Peña Colorada | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Pinnularia* sp. 2 | D60_016 | El Membrillo | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| **Planothidium cryptolanceolatum* R. Jahn et N. Abarca | D108_021 | El Membrillo | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| **Planothidium victori* Novis, Braidwood et Kilroy | D100_015 | La Mesa | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| **Planothidium victori* Novis, Braidwood et Kilroy | D101_022 | Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| **Planothidium victori* Novis, Braidwood et Kilroy | D109_018 | Guanajuatito | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| **Planothidium victori* Novis, Braidwood et Kilroy | D109_020 | Guanajuatito | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Sellaphora auldreekie* D.G. Mann et S.M. McDonald | D108_015 | El Membrillo | 07.09.2014 | Demetrio Mora | Jana Bansemer |
| *Sellaphora cosmopolitana* (Lange-Bertalot) C.E. Wetzel et Ector | D88_012a 1/5 Ag | San Martín | 09.02.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora cosmopolitana* (Lange-Bertalot) C.E. Wetzel et Ector | D88_013a 1/5 Ag | San Martín | 09.02.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora cosmopolitana* (Lange-Bertalot) C.E. Wetzel et Ector | D105_008 | Paredones | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora pupula* (Kützing) Mereschkowsky | D61_063 | Guanajuatito | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Sellaphora* cf. *saugerresii* (Desmazières) C.E. Wetzel et D.G. Mann | D101_005 | Calvillo | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| *Sellaphora* cf. *saugerresii* (Desmazières) C.E. Wetzel et D.G. Mann | D102_024 | Ojo de Agua de Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora* aff. *saugerresii* (Desmazières) C.E. Wetzel et D.G. Mann | D96_008 | Laguna de Servín 2 | 03.02.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora* aff. *saugerresii* (Desmazières) C.E. Wetzel et D.G. Mann | D96_009 | Laguna de Servín 2 | 03.02.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora* sp. 1 | D101_023 | Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora* sp. 2 | D104_004 | San Martín | 06.09.2014 | Demetrio Mora | Demetrio Mora |

| | | | | | |
|---|---|---|---|---|---|
| *Sellaphora* sp. 3 | D88_001 | San Martín | 09.02.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora* sp. 3 | D88_002 | San Martín | 09.02.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora* sp. 4 | D100_022 | La Mesa | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| *Sellaphora* sp. 5 | D110_015 | Los Ailes 1 | 05.09.2014 | Demetrio Mora | Demetrio Mora |
| *Sellaphora* sp. 6 | D108_025 | El Membrillo | 07.09.2014 | Demetrio Mora | Demetrio Mora |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | D84_020 | La Mesa | 08.02.2014 | Demetrio Mora | Demetrio Mora |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | D84_024 | La Mesa | 08.02.2014 | Demetrio Mora | Demetrio Mora |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | D85_001 | Calvillo | 08.02.2014 | Demetrio Mora | Demetrio Mora |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | D85_007 | Calvillo | 08.02.2014 | Demetrio Mora | Demetrio Mora |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | D85_009 | Calvillo | 08.02.2014 | Demetrio Mora | Demetrio Mora |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | D102_004 | Ojo de Agua de Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | D102_008 | Ojo de Agua de Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | D102_017 | Ojo de Agua de Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Simonsenia* cf. *delognei* (Grunow) Lange-Bertalot | D102_019 | Ojo de Agua de Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Stauroneis* cf. *acidoclinatopsis* Van de Vijver et Lange-Bertalot | D103_013 | Peña Colorada | 06.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Surirella angusta* Kützing | D101_017 | Calvillo | 06.09.2014 | Demetrio Mora | Jana Bansemer |
| *Surirella angusta* Kützing | D101_041 | Calvillo | 06.09.2014 | Demetrio Mora | Demetrio Mora |
| *Surirella angusta* Kützing | D113_005 | Los Ailes 2 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Surirella ovalis* Brébisson | D61_030 | Guanajuatito | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Tryblionella calida* (Grunow) D.G. Mann | D61_034 | Guanajuatito | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Tryblionella hungarica* (Grunow) D.G. Mann | D60_025 | El Membrillo | 17.09.2013 | Demetrio Mora | Demetrio Mora |
| *Ulnaria* cf. *acus* (Kützing) Aboal | D90-Syn1 | La Laborcilla 1 | 09.02.2014 | Demetrio Mora | Oliver Skibbe |
| *Ulnaria* cf. *ulna* (Nitzsch) Compère | D63_032 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Jana Bansemer |
| *Ulnaria* cf. *ulna* (Nitzsch) Compère | D63_039 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Demetrio Mora |
| *Ulnaria* cf. *ulna* (Nitzsch) Compère | D63_051 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Demetrio Mora |
| *Ulnaria* cf. *ulna* (Nitzsch) Compère | D63_059 | Laguna de Servín 1 | 18.09.2013 | Demetrio Mora | Demetrio Mora |
| *Ulnaria* cf. *ulna* (Nitzsch) Compère | D111_007 | Laguna de Servín 1 | 05.09.2014 | Demetrio Mora | Oliver Skibbe |
| *Ulnaria* cf. *ulna* (Nitzsch) Compère | D112_003 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Jana Bansemer |
| *Ulnaria* cf. *ulna* (Nitzsch) Compère | D112_018 | Laguna de Servín 2 | 05.09.2014 | Demetrio Mora | Demetrio Mora |

Supplementary material 2. Diatom taxa list from the Lerma-Chapala Basin, Central Mexico, identified by morphology and HTS. +Indicates taxa that were observed under LM/SEM after the counts. *Indicates taxa that were observed in cultures but not in the environmental samples. " Refers to taxa assigned only from sequences of the BGBM Diatom Sequence Reference Database or from the NCBI nucleotide database, with 100% identity, but since the taxa were not observed morphologically in our samples, are referred to as closely related (cf.).

| **Taxon** | **Morphology** | **HTS** |
|---|---|---|
| *Achnanthes inflata var. inflata* | + | |
| *Achnanthidium aff. catenatum* (D107_005) | 6,277 | 0,630 |
| *Achnanthidium exiguum* | 0,020 | |
| *Achnanthidium exile* | 0,053 | |
| *Achnanthidium minutissimum* | 0,298 | 1,202 |
| *Achnanthidium* sp. 1 (D105_007) | 9,493 | 1,441 |
| *Achnanthidium* sp. 2 | + | |
| *Achnanthidium* sp. 3 | + | |
| *Achnanthidium* sp. 4 | 0,049 | |
| *Achnanthidium* sp. 5 | 18,978 | |
| *Achnanthidium* sp. 6 | 0,193 | |
| *" Amphora* cf. *commutata* | | 0,030 |
| *Amphora pediculus* | 0,250 | |
| *" Amphora* cf. *pediculus* | * | 0,021 |
| *" Anomoeoneis* cf. *sphaerophora* | | 0,022 |
| *Brachysira altepetlensis* (D105_001) | 1,642 | 0,115 |
| *Brachysira brebissonii* | 0,010 | |
| *Brachysira microcephala* | + | |
| *Brachysira* sp. 1 | 0,009 | |
| *Brachysira* sp. 2 | + | |
| *Caloneis bacillum* | 0,010 | |
| *Caloneis schumanniana* | 0,061 | |
| *Caloneis* sp. 1 | + | |
| *Caloneis* sp. 2 | 0,010 | |
| *Caloneis* sp. 3 (D60_020) | 0,010 | 0,057 |
| cf. *Chamaepinnularia* sp. | 0,020 | |
| *Chamaepinnularia submuscicola* | 0,021 | |
| *" Cocconeis* cf. *euglypta* | | 0,151 |
| *Cocconeis pediculus* | 0,452 | 0,019 |
| *Cocconeis* sp. 1 | + | 0,009 |
| *Cocconeis* sp. 2 (D86-Coc1b) | 7,219 | 8,856 |
| *Craticula accomoda* | + | 0,003 |

| | | |
|---|---|---|
| *Craticula* cf. *pumilio* | 0,062 | |
| *" Craticula* cf. *cuspidata* | | 0,027 |
| *Craticula molestiformis* | 0,329 | 0,107 |
| *Craticula* sp. 1 | 0,171 | |
| *Craticula subminuscula* | 2,217 | |
| *Cyclostephanos invisitatus* | 0,010 | |
| *Cyclotella atomus* | 0,113 | |
| *" Cyclotella* cf. *cryptica* | | 1,035 |
| *" Cyclotella* cf. *gamma* | | 0,008 |
| *Cyclotella meneghiniana* | 0,071 | 0,086 |
| *Cymbella* cf. *lanceolata* | | 0,034 |
| *Cymbella tropica* | + | |
| *Cymbopleura naviculiformis* | 0,038 | 0,148 |
| *Diadesmis confervacea* (D103_010) | 0,173 | 0,038 |
| *" Diatoma* cf. *hyemalis* | | 0,077 |
| *Encyonema brevicapitatum* | 0,041 | |
| *Encyonema* cf. *hebridiforme* | 0,030 | |
| *Encyonema jemtlandicum* | 0,288 | |
| *Encyonema jemtlandicum* var. *venezolanum* | 0,031 | |
| *Encyonema minutiforme* | 0,096 | |
| *Encyonema minutum* | 0,410 | 0,008 |
| *Encyonema pergracile* | 0,170 | |
| *Encyonema silesiacum* (D103_006) | 0,071 | 0,067 |
| *" Encyonema* sp. aff. *humile* | | 0,010 |
| *Encyonema triangulum* | + | |
| *Encyonopsis subminuta* | 0,065 | |
| *Encyonopsis* cf. *thienemannii* | + | |
| *Encyonopsis sp.* | 0,049 | |
| *Eolimna* sp. 1 | 0,053 | |
| *Eolimna* sp. 2 | 0,030 | |
| *Eolimna* sp. 3 | 0,020 | |
| *Epithemia adnata* | 0,052 | |
| *Epithemia sorex* | 0,339 | |
| *Epithemia turgida* | 0,031 | 0,059 |
| *Eunotia bidens* | + | |
| *" Eunotia* cf. *formica* | | 0,160 |
| *Eunotia* cf. *meridiana* | 0,204 | |
| *Eunotia major* var. *major* | + | |
| *Eunotia metamonodon* | + | |
| *Eunotia minor* (D111_006) | 0,400 | 0,208 |
| *Eunotia mucophila* | + | |
| *Eunotia tridentula* | + | |
| *Eunotia* sp. 1 | 1,886 | |

| | | |
|---|---|---|
| *Eunotia* sp. 2 | + | |
| *Eunotia* sp. 3 | + | |
| *Eunotia* sp. 4 | 0,020 | |
| *Eunotia* sp. 5 | + | |
| *Fistulifera saprophila* | 0,144 | 0,753 |
| *Fragilaria austriaca* | 2,647 | |
| *Fragilaria bidens* | 0,074 | |
| *Fragilaria pectinalis* (D106_009) | * | 0,191 |
| *Fragilaria rumpens* | + | 0,021 |
| *Fragilaria tenera* (D106_012) | 0,144 | 0,031 |
| *" Fragilaria* cf. *vaucheriae* | | 0,016 |
| *Fragilaria* sp. (D113_006) | * | 0,097 |
| *Frustulia crassinervia* | 0,130 | |
| *Frustulia neomundana* | + | |
| *Frustulia* cf. *spicula* ssp. *spicula* | + | |
| *Frustulia* cf. *undosa* | 0,041 | |
| *Geissleria decussis* | 0,053 | 0,002 |
| *Gomphonema acuminatum* | + | 0,058 |
| *Gomphonema affine* var. *affine* | 0,020 | |
| *" Gomphonema* cf. *angustum* | | 0,152 |
| *" Gomphonema* cf. *angustatum/bourbonense* | | 1,409 |
| *Gomphonema exilissimum* | 1,400 | |
| *Gomphonema* cf. *gracile* | * | |
| *Gomphonema graciledictum* | + | |
| *Gomphonema innocens* | 0,345 | |
| *Gomphonema lagenula* | 2,352 | |
| *Gomphonema* cf. *lagenula* | 0,130 | |
| *Gomphonema minusculum* | 0,154 | |
| *Gomphonema naviculoides* | 0,193 | |
| *Gomphonema* cf. *naviculoides* | + | |
| *Gomphonema parvuliforme* | 0,061 | |
| *Gomphonema* aff. *parvulius* | 0,344 | |
| *Gomphonema parvulum* | 8,574 | 17,330 |
| *" Gomphonema* cf. *productum* | | 0,021 |
| *Gomphonema pumilum* | 0,452 | 0,379 |
| *Gomphonema salae* | 0,010 | |
| *Gomphonema* aff. *sarcophagus* | 0,541 | |
| *Gomphonema subclavatum* | 0,183 | 0,156 |
| *" Gomphonema* cf. *truncatum* | | 0,443 |
| *Gomphonema* sp. 1 | 0,030 | |
| *Gomphonema* sp. 2 | + | |
| *Gomphonema* sp. 3 | + | |
| *Gomphonema* sp. 4 (D112_008) | * | 0,062 |

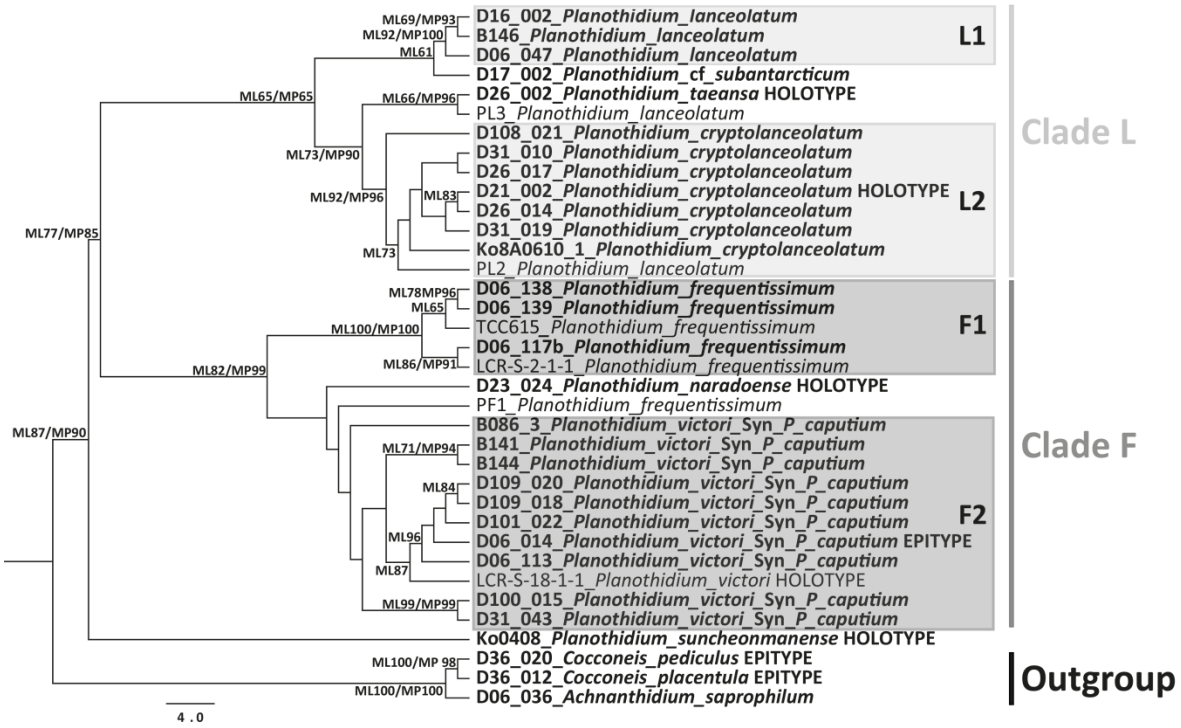
| | | |
|---|---|---|
| *Halamphora montana* | 0,187 | |
| *Halamphora* cf. *pseudomontana* | 0,020 | |
| *Halamphora veneta* | 0,082 | |
| *Iconella delicatissima* | + | 0,042 |
| *Humidophila contenta* | 0,021 | |
| *Luticola goeppertiana* | 0,010 | |
| *Luticola mutica* | 0,031 | |
| *Luticola* cf. *tomesii* | 0,020 | |
| *Luticola undulata* | + | |
| *Luticola ventricosa* | + | |
| *Mayamaea* cf. *crassistriata* | 0,011 | |
| *Mayamaea permitis* | 3,891 | 0,556 |
| *Mayamaea* cf. *permitis* | * | |
| *Mayamaea* sp. 1 | 0,042 | |
| *" Melosira* cf. *varians* | | 0,080 |
| *Navicula angusta* | 0,010 | |
| *Navicula capitatoradiata* | 0,082 | 0,003 |
| *Navicula cryptocephala* | 0,225 | 1,143 |
| *Navicula* cf. *cryptocephala* | 0,132 | |
| *Navicula cryptotenella* | + | |
| *Navicula* cf. *erifuga* (D61_045) | 0,175 | 1,314 |
| *Navicula gregaria* | 0,298 | 0,514 |
| *Navicula libonensis* | 0,081 | |
| *Navicula notha* | 1,638 | 9,099 |
| *" Navicula* cf. *oblonga* | | 0,191 |
| *Navicula reichardtiana* | 0,900 | |
| *Navicula riediana* | 0,010 | |
| *" Navicula* cf. *reinhardtii* | | 0,044 |
| *Navicula rostellata* | 0,120 | |
| *Navicula symmetrica* | 0,165 | 0,641 |
| *Navicula trivialis* | 0,041 | |
| *Navicula veneta* | 0,052 | 1,374 |
| *Navicula* sp. (D101_006) | * | 0,140 |
| *Navigiolum uruguayense* | 0,010 | |
| *Neidium amphigomphus* | + | |
| *Neidium ampliatum* | + | |
| *Neidium* sp. | + | |
| *Nitzschia acicularis* | 0,145 | 1,477 |
| *Nitzschia amphibia* | | 0,008 |
| *Nitzschia* cf. *amphibia (D100_030)* | 0,534 | 0,052 |
| *Nitzschia balcanica* | * | 0,610 |
| *Nitzchia clausii* | + | |
| *Nitzschia desertorum* | + | |

| | | |
|---|---|---|
| *Nitzschia dissipata* | + | 0,039 |
| *Nitzschia fonticola* | 0,091 | |
| *Nitzschia* cf. *gracilis* | 0,165 | 0,268 |
| *Nitzschia* cf. *hantzschiana* | 0,061 | |
| *Nitzschia inconspicua* | 0,314 | 0,052 |
| *Nitzschia intermedia* | 0,010 | |
| *Nitzschia lanceolata* | 0,031 | |
| *Nitzschia linearis* | 0,142 | 6,190 |
| *Nitzschia palea* | 0,713 | 7,974 |
| *Nitzschia palea var. debilis* | 0,737 | |
| *Nitzschia palea var. tenuirostris* | 0,693 | |
| *Nitzschia paleacea* | 0,207 | 0,004 |
| *Nitzschia perminuta* | 0,257 | 0,035 |
| *" Nitzschia* cf. *pusilla* | | 0,031 |
| *Nitzschia recta* | 0,021 | |
| *Nitzschia semirobusta* | 0,221 | 0,062 |
| *Nitzschia sinuata var. delognei* | 0,031 | |
| *Nitzschia sublinearis* | + | |
| *Nitzschia supralitorea* | 0,183 | 0,023 |
| *" Nitzschia* cf. *tenuis* | | 0,019 |
| *Nitzschia tubicola* | 0,010 | |
| *Nitzschia* cf. *umbonata* (D61_049) | * | 0,057 |
| *Nitzschia* sp. 1 | 0,286 | |
| *Nitzschia* sp. 2 | + | |
| *Nitzschia* sp. 3 | 0,020 | |
| *Nitzschia* sp. 4 | 0,010 | |
| *Nitzschia* sp. 5 | + | |
| *Nitzschia* sp. 6 | 0,010 | |
| *Nitzschia* sp. 7 | * | |
| *" Nitzschia* sp. 8 (D06_043) | | 0,062 |
| *Nupela praecipua* | + | |
| *Nupela wellneri* | 0,123 | |
| *" Pinnularia* cf. *acrosphaeria* | | 0,004 |
| *Pinnularia anglica* | 0,010 | |
| *Pinnularia borealis* var. *borealis* | + | |
| *Pinnularia borealis* var. *scalaris* | + | |
| *Pinnularia* cf. *subcapitata* var. *elongata* | + | |
| *Pinnularia divergens* | 0,020 | |
| *Pinnularia divergentissima* var. *divergentissima* | 0,041 | |
| *Pinnularia parvulissima* | 0,010 | |
| *Pinnularia saprophila* (D60_016) | 0,041 | 0,017 |
| *Pinnularia subbrevistriata* | + | |
| *Pinnularia viridiformis* | 0,010 | |

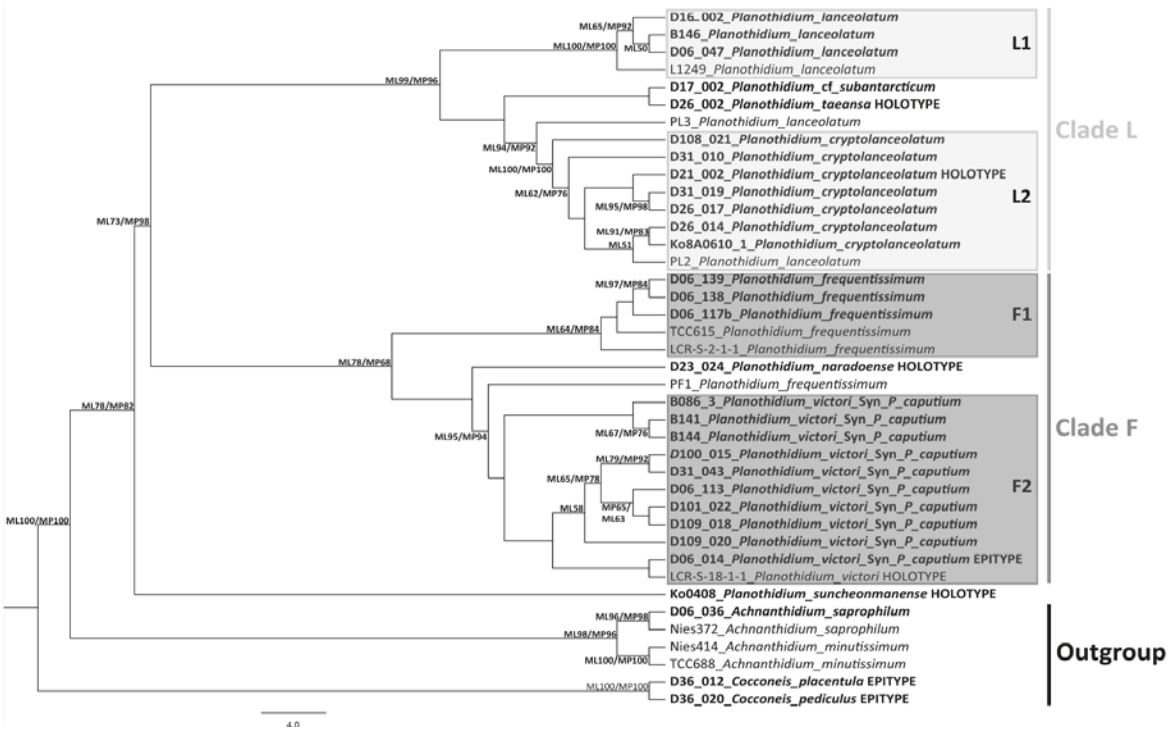
| | | |
|---|---|---|
| *Pinnularia* sp. 1 | * | |
| *Pinnularia* sp. 2 | + | |
| *Pinnularia* sp. 3 | + | |
| *Pinnularia* sp. 4 | + | |
| *Planothidium cryptolanceolatum* (D108_021) | 1,988 | 0,434 |
| *Planothidium incuriatum* | 1,313 | |
| *" Planothidium* cf. *lanceolatum* | | 0,462 |
| *Planothidium rostratum* | 0,159 | |
| *Planothidium victori* | 2,533 | 0,147 |
| *Pseudofallacia monoculata* | 0,020 | |
| *Reimeria sinuata* | 2,204 | 0,648 |
| *Rhopalodia gibba* | 0,011 | 0,047 |
| *Sellaphora atomoides* | 1,131 | |
| *Sellaphora auldreekie* | + | 0,006 |
| *Sellaphora blackfordensis* | 0,010 | |
| *Sellaphora cosmopolitana* (D88_012a) | 0,399 | 0,058 |
| *Sellaphora* cf. *elorantana* | 0,072 | |
| *Sellaphora laevissima* | | |
| *Sellaphora madida* | 0,092 | |
| *Sellaphora nigri* | 1,362 | 0,162 |
| *Sellaphora pupula* | 0,020 | |
| *Sellaphora queretana* | 0,431 | |
| *Sellaphora* cf. *saugerresii* (D101_005) | 1,884 | 0,022 |
| *Sellaphora* aff. *saugerresii* | * | |
| *Sellaphora stauroneioides* | + | |
| *Sellaphora wallacei* | + | |
| *Sellaphora* sp. 1 | 0,061 | |
| *Sellaphora* sp. 2 (D100_022) | * | 0,046 |
| *Sellaphora* sp. 3 | 0,943 | |
| *Sellaphora* sp. 4 | * | |
| *Sellaphora* sp. 5 | * | |
| *Simonsenia* cf. *delognei* | * | |
| *Stauroneis* cf. *acidoclinatopsis* | + | |
| *Stauroneis reichardtii* | + | |
| *Stauroneis* sp. 1 | 0,021 | |
| *" Stephanodiscus* cf. *minutulus* | | 0,036 |
| *Surirella angusta* | 0,237 | |
| *Surirella apiculata* var. *panduriformis* | + | |
| *Surirella* cf. *pseudolinearis* | + | |
| *Surirella ovalis* | | 0,103 |
| *Thalassiosira weissflogii* | + | 0,014 |
| *Tryblionella hungarica* (D60_025) | | 0,023 |
| *" Tryblionella* sp. 1 | | 0,013 |

| | | |
|---|---|---|
| *" Ulnaria acus* | | 0,340 |
| *Ulnaria* cf. *acus* | 0,010 | 0,188 |
| *Ulnaria* cf. *ulna* (D63_032) | 1,125 | 6,602 |
| Unassigned | | 22,775 |
| Σ | 100,00 | 100,00 |

## 7.3 Appendix 3




Supplementary Fig. 1: Strict Consensus Tree of the rbcL Dataset with the Results of Bootstrap Statistics (>50) for ML (LB) and MP (PB). Bold: strains cultured by the authors




Supplementary Fig. 2: Strict Consensus Tree of the 18S Dataset with the Results of Bootstrap Statistics (>50) for ML (LB) and MP (PB). Bold: strains cultured by the authors