UNIVERSITY OF SZCZECIN

INSTITUTE OF MARINE AND ENVIRONMENTAL SCIENCES

POLAND

DIATOMACEOUS (BACILLARIOPHYTA) ASSEMBLAGES OF THE EXTREME ENVIRONMENT FORM - ANATOLIAN SODA LAKES: ECOLOGY, TAXONOMY AND MOLECULAR PHYLOGENY

BY

ELIF YILMAZ

DOCTORAL DISSERTATION

SUPERVISOR:

PROF. DR DAVID MANN

CO- SUPERVISOR:

DR ROMAIN GASTINEAU

SZCZECIN 2024

OŚWIADCZENIE

Elif Yilmaz

Instytut Nauk o Morzu i Środowisku

Uniwersytet Szczeciński

Oświadczam, że moja praca pt.: **„Diatomaceous (Bacillariophyta) Assemblages of the Extreme Environment Form - Anatolian Soda Lakes: Ecology, Taxonomy and Molecular Phylogeny."**

a) została napisana przeze mnie samodzielnie,
b) nie narusza praw autorskich w rozumieniu ustawy z dnia 14 lutego 1994 roku o prawie autorskim i prawach pokrewnych (Dz.U. 2019 r. poz. 1231) oraz dóbr osobistych chronionych prawem,
c) nie zawiera danych i informacji, które uzyskałem / uzyskałam w sposób niedozwolony,
d) nie była podstawą nadania tytułu naukowego lub zawodowego ani mnie ani innej osobie.

Ponadto oświadczam, że treść pracy przedstawionej przeze mnie do obrony, zawarta na przekazanym nośniku elektronicznym jest identyczna z jej wersją drukowaną.

Szczecin, dn:

podpis doktoranta:

OŚWIADCZENIE

Wyrażam zgodę na udostępnienie mojej pracy doktorskiej pt.: **„Diatomaceous (Bacillariophyta) Assemblages of the Extreme Environment Form - Anatolian Soda Lakes: Ecology, Taxonomy and Molecular Phylogeny."**

Szczecin, dn.

podpis doktoranta

OŚWIADCZENIE

Akceptuję ostateczną wersję pracy pt.: **„Diatomaceous (Bacillariophyta) Assemblages of the Extreme Environment Form Anatolian Soda Lakes: Ecology, Taxonomy and Molecular Phylogeny."**

Szczecin, dn.

podpis

Prof. Dr. David Mann

OŚWIADCZENIE

Akceptuję ostateczną wersję pracy pt.: **„Diatomaceous (Bacillariophyta) Assemblages of the Extreme Environment Form - Anatolian Soda Lakes: Ecology, Taxonomy and Molecular Phylogeny."**

Szczecin, dn.

podpis

Dr. Romain Gastineau

PUBLICATIONS

1. **Yılmaz, E.,** Mann, D. G., Gastineau, R., Trobajo, R., Solak, C. N., Górecka, E., Turmel, M., Lemieux, C., Ertorun, N., & Witkowski, A. (2024). Description of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny. *PhytoKe*ys, 241, 27. doi: 10.3897/phytokeys.241.118903

   Impact Factor: 1.4

   Lista czasopism MEiN: 100

   Contribution – 70% (percent): I conceptualized the study, performed single-cell isolation to establish the monoculture, and carried out the laboratory work, including DNA extraction and microscopy imaging (light and electron). I interpreted the data and literature to establish that the species is new to science, prepared the figures and tables, drafted the original manuscript, revised it based on co-authors' suggestions, led the revision process following peer review, and submitted the final version.

   Detailed description: The present study investigated *Navicula vanseea* sp. nov., a newly discovered species of diatom from Lake Van, which is characterized by its highly alkaline environment in Eastern Anatolia, Türkiye. The identification process involved light and scanning electron microscopy on two distinct monoclonal cultures. Both strains underwent sequencing for their complete nuclear rRNA clusters and plastid genomes, while one strain also had its complete mitogenome sequenced. Noteworthy findings include the potential loss of the functional ycf35 gene in the plastomes of both strains. Furthermore, each strain possesses two IB4 group I introns within their rrl, which encode for putative LAGLIDADG homing endonucleases; notably, the first IB4 intron (L1917) is reported for diatoms. The phylogenetic analysis using Maximum Likelihood methods, based on a concatenated dataset of *18S*, *rbcL*, and *psbC*, effectively distinguishes *N. vanseea* sp. nov. from closely related species, including *Navicula cincta* and *Navicula microdigitoradiata*.

2. **Yılmaz, E.,** Gastineau R., Solak C. N., Górecka E., Trobajo R., Turmel M., Lemieux C., Otis C., Witkowski A., & Mann D. G., (2024). Morphological and molecular characterization of *Halamphora vantushpaensis* (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye) *PhytoKe*ys (in press, online ISSN 1314-2003). doi:10.3897/phytokeys.@@@.133205

Impact Factor: 1.4

Lista czasopism MEiN: 100

Contribution – 70% (percent): I conceptualized the study, performed single-cell isolation to establish the monoculture, and obtained light and electron microscopy images, extracted and processed the DNA for characterization, and interpreted the data in relation to the literature to establish that the species is new to science. I prepared the figures and tables, drafted the original manuscript, revised it following the peer review, and submitted the final version.

Detailed description: This research introduces *Halamphora vantushpaensis* sp. nov., a newly identified diatom species from the highly alkaline environment of Lake Van in Eastern Türkiye. The species' morphological features were analyzed using both light and scanning electron microscopy on wild and lab-cultivated samples. A genome-skimming approach was applied to two monoclonal cultures, providing full sequences of their nuclear rRNA clusters, along with mitochondrial and plastid genomes. Genomic analysis revealed high similarity between the strains, with only a few mutations noted, some of which affected protein-coding genes and were non-silent. The mitochondrial *cox1* gene was found to have a group II intron showing significant polymorphism. Additionally, the plastome exhibited a unique feature an extended inverted repeat, reducing the size of the two single-copy regions, distinguishing it from other *Halamphora* species. Phylogenetic analysis using maximum likelihood, based on a dataset combining three genes (*18S*, *psbC,* and *rbcL*), placed the species in Clade K of *Halamphora*, known for its association with hypersaline to freshwater habitats.

3. **Yılmaz, E.,** Gastineau R., Górecka E., Solak C. N., Trobajo R., Peszek Ł., & Mann D. G., (2025). *Halamphora witkowskii* sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye. Nova Hedwigia (in press online ISSN 2363-7188),

Impact Factor: 1.0

Lista czasopism MEiN: 40

Contribution – 75% (percent): I conceptualized the study, performed single-cell isolation to establish the monoculture, and obtained light and electron microscopy images. I extracted the DNA, conducted PCR analysis, interpreted the data and literature to establish that the species is new to science, prepared the figures and tables, drafted the original manuscript, revised it following peer review, and submitted the final version.
Detailed description: This study introduces *Halamphora witkowskii* sp. nov., a newly described species from Lake Van, the world's largest soda lake, known for its high alkalinity. Detailed morphological and morphometric analyses using light and scanning electron microscopy effectively distinguish this new species from other closely related taxa. Additionally, culturing studies on live cells led to the establishment of two separate monocultures, which were used for phylogenetic analysis with the *rbcL* molecular marker. These analyses further confirmed the distinctiveness of *Halamphora witkowskii* compared to other known species. A Maximum Likelihood tree, constructed from the alignment of 260 *rbcL* sequences of diatoms and rooted with *Triparma pacifica,* identified a clade of *Halamphora* species, with two strains of *Halamphora witkowskii* clearly positioned within this group.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my late supervisor and mentor, Prof. Dr. hab. Andrzej Witkowski. His vision, wisdom, and dedication laid the foundation for my work and inspired me throughout this journey. After his passing, I am deeply thankful to my current supervisor, Prof. David Mann, for his ongoing guidance and invaluable support.

I would like to express my heartfelt thanks to my co-supervisor, Dr. Romain Gastineau, whose support and profound expertise were critical at every stage of my thesis. His patient guidance in bioinformatics and molecular work greatly shaped the outcome of this research. I am truly grateful for the countless hours he dedicated to assisting me. I would also like to extend my deepest gratitude to Dr. Ewa Górecka for her unwavering support throughout this research. Her invaluable assistance, from fieldwork to laboratory work, and her patience and encouragement have been greatly appreciated.

My sincere appreciation goes to Prof. Cüneyt Nadir Solak for his generous support, particularly for his assistance with electron microscopy. I would also like to express my deepest gratitude to Dr. Rosa Trobajo, who not only welcomed me during my time in Spain but also played an integral role in guiding me through critical aspects of my thesis. Her contributions have left a lasting impact on my work, and I am grateful for her hospitality, insights, and mentorship.

I also wish to acknowledge Dr. Ana Car for her kind hospitality during my stay in Croatia and for assisting with my research. Additionally, I extend my sincere thanks to Monique Turmel from Laval University, whose expertise and assistance in bioinformatics have been invaluable to this thesis. I owe a debt of gratitude to Przemysław Dąbek, Agnieszka Kierzek, and Renata Ciaś for their dedicated assistance with laboratory work. Their technical support was essential to the successful completion of many experimental processes. I would also like to express my gratitude to Prof. Teresa Radziejewska for her insightful advice and support throughout my studies.

Finally, I would like to extend my heartfelt thanks to the Doctoral School at the University of Szczecin for providing the platform and resources necessary to begin and complete my scientific research.

University of Szczecin

Institute of Marine and Environmental Sciences

Elif Yilmaz, M.Sc.

Tytuł rozprawy doktorskiej:

Diatomaceous (Bacillariophyta) Assemblages of the Extreme Environment Form - Anatolian Soda Lakes: Ecology, Taxonomy and Molecular Phylogeny

Supervisor: Prof. dr **David Mann**

Co- Supervisor**:** dr **Romain Gastineau**

STRESZCZENIE ROZPRAWY DOKTORSKIEJ

Niniejsza praca skupia się na Jeziorze Van, największym jeziorze sodowym na świecie, i analizuje różnorodność biologiczną, taksonomię oraz molekularną filogenezę zespołów okrzemkowych (Bacillariophyta) w sodowych jeziorach Anatolii. Okrzemki, jako pierwotni producenci w ekosystemach wodnych, odgrywają kluczową rolę w cyklach biogeochemicznych oraz służą jako wskaźniki ekologiczne. Głównym celem badań była identyfikacja nowych gatunków okrzemek występujących w Jeziorze Van, charakterystyka ich różnorodności molekularnej oraz zrozumienie ich adaptacji ekologicznych do ekstremalnych, zasadowych warunków jeziora.

W ramach badań pobrano próbki z siedmiu lokalizacji wokół Jeziora Van. W celu uzyskania monoklonalnych kultur przeprowadzono izolację pojedynczych komórek. Wyizolowane kultury okrzemek hodowano w optymalnych warunkach świetlnych i temperaturowych, aby zapewnić ich zdrowy wzrost. Dodatkowo, zastosowano mikroskopię świetlną oraz skaningową mikroskopię elektronową, aby kompleksowo scharakteryzować morfologiczne cechy wyizolowanych okrzemek. Oprócz charakterystyki morfologicznej przeprowadzono także izolację DNA oraz amplifikację wybranych genów w celu określenia różnorodności genetycznej oraz powiązań filogenetycznych między gatunkami. Analizy te doprowadziły do odkrycia trzech nowych gatunków okrzemek należących do rodzajów *Navicula* i *Halamphora*. Badania

te przyczyniły się do lepszego zrozumienia taksonomii okrzemek oraz ich relacji ewolucyjnych.

Podsumowując, badania te nie tylko poszerzają wiedzę taksonomiczną na temat okrzemek żyjących w ekstremalnych siedliskach, ale także pogłębiają nasze zrozumienie ich charakterystyki genomowej. Biorąc pod uwagę ograniczone informacje na temat Jeziora Van i jego specyficznej flory okrzemek, niniejsza praca stanowi istotny wkład w badania nad różnorodnością okrzemek w ekstremalnych środowiskach.

Słowa kluczowe: okrzemki, Jezioro Van, ekstremalne środowisko, filogeneza, jezioro sodowe

ABSTRACT

This study focuses on Lake Van, known as the largest soda lake in the world, and investigates the biodiversity, taxonomy, and molecular phylogeny of diatom (Bacillariophyta) communities in the soda lakes of Anatolia. As primary producers in aquatic ecosystems, diatoms play a critical role in ecosystem cycles and act as environmental bioindicators. The primary goal of this research was to identify new diatom species in Lake Van, characterize their molecular diversity, and understand their ecological adaptations to the lake's highly alkaline conditions.

To achieve this goal, samples were collected from seven stations around Lake Van. Single-cell isolations were performed to establish monoclonal cultures. The isolated diatom cultures were grown under specially optimized light and temperature conditions to ensure healthy growth, resulting in robust and viable cultures. Additionally, detailed light and scanning electron microscopy were used to comprehensively document the morphological characteristics of the isolated diatoms. In addition to morphological characterization, DNA extraction and amplification of selected genes were performed to reveal genetic diversity and phylogenetic relationships among diatom species, followed by molecular analyses. As a result of these studies, three new diatom species belonging to the genera *Navicula* and *Halamphora* were discovered. The study has contributed to a broader understanding of diatom taxonomy and evolutionary relationships.

In conclusion, this research contributes to enhancing the taxonomic knowledge of diatoms living in extreme habitats while advancing our understanding of their genomic characterization. Considering the limited information about the unique features of Lake Van and its diatom flora, this study aims to make significant contributions to the diversity of diatoms in extreme environments.

Keywords: diatom, Lake Van, extreme environment, phylogeny, Soda Lake, *Halamphora, Navicula*, NGS

LIST OF CONTENT

ANNEX I

Yılmaz, E., Mann, D. G., Gastineau, R., Trobajo, R., Solak, C. N., Górecka, E., Turmel, M., Lemieux, C., Ertorun, N., & Witkowski, A. (2024). Description of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny. *PhytoKeys*, *241*, 27.

ANNEX II

Yılmaz, E., Gastineau R., Solak C. N., Górecka E., Trobajo R., Turmel M., Lemieux C., Otis C., Witkowski A., & Mann D. G., (2024). Morphological and molecular characterization of *Halamphora vantushpaensis* (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye) *PhytoKe*ys (proof, in press).

ANNEX III

Yılmaz, E., Gastineau R., Górecka E., Solak C. N., Trobajo R., Peszek Ł., & Mann D. G., . *Halamphora witkowskii* sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye. Nova Hedwigia (proof, in press).

LIST OF ABBREVIATIONS

bp base pairs

bv bootstrap value

DNA deoxyribonucleic acid

LM light microscope

LSC large single copy

NGS next-generation sequencing

orf open reading frame

PCG protein coding genes

PCR polymerase chain reaction

psbC photosystem II CP43 protein

rbcL ribulose-1.5-bisphosphate carboxylase/oxygenase large subunit

RNA ribonucleic acid

rRNA ribosomal ribonucleic acid

SEM scanning electron microscope

SSC small single-copy

SSU small subunit ribosomal ribonucleic acid

SZCZ herbarium code of the Diatom Collection at University of Szczecin

tRNA transfer ribonucleic acid

INTRODUCTION

Diatoms

Diatoms are photosynthetic single-celled eukaryotes found in aquatic habitats worldwide. The exact number of species of diatoms is unknown, but is estimated roughly as approximately 100,000 species (Mann & Vanormelingen 2013) with distinctive characteristics, inhabiting marine, saline, and freshwater environments (Round et al. 1990, Smol & Stoermer 2010). New species are frequently being discovered and described. Their cell lengths range from 1 μm to 500 μm or more, and they function as producers in the food chain (Scala & Bowler 2001). Diatoms, as primary producers in aquatic ecosystems, play a crucial role by utilizing solar energy through photosynthesis to produce organic matter, which serves as the foundation of the aquatic food chain. This organic matter supports other organisms, making diatoms essential contributors to ecosystem dynamics. Beyond their role in the food chain, diatoms significantly influence global biogeochemical cycles, impacting the overall functioning of ecosystems. Their exceptional diversity and widespread distribution make them indispensable components of aquatic environments, shaping the intricate structure of life within these systems (Seckbach & Kociolek 2011).

Diatoms are sensitive to many factors that affect biological diversity and ecosystem dynamics in aquatic ecosystems. Environmental variables such as light, temperature, flow rate, salinity, pH, and oxygen influence the growth, reproduction, and distribution of diatoms (B-Béres et al. 2023). In ecological, palaeolimnological, and biological monitoring studies, diatoms are used as model organisms to assess the quality of aquatic ecosystems and understand environmental changes (Anderson et al. 1990). The species composition, abundance, and distribution of diatom assemblages reflect specific environmental conditions, such as nutrient levels and pollution, thereby providing insights into factors affecting the aquatic environment and the presence of other organisms in the ecosystem. Diatoms exhibit notable adaptability to various environmental conditions. This characteristic, together with fast growth, makes them sensitive bioindicators of aquatic ecosystem health. Their ecological preferences and responses to environmental changes are reflected in significant shifts in species composition, abundance, and distribution, providing invaluable insights into ecosystem

dynamics (Dixit et al. 1992, Stevenson et al. 1999, Bąk et al. 2012). For this reason also, diatom fossil records play a significant role in understanding past aquatic conditions in palaeolimnological studies (Stoermer & Smol, 1999). Diatoms serve as sentinel organisms in both ecological and palaeolimnological research, offering valuable information about historical and current environmental conditions. By analyzing diatom communities and fossilized remains, researchers can reconstruct historical environmental changes, monitor anthropogenic impacts, and evaluate the resilience of aquatic ecosystems over time (Smol & Douglas 1996).

When diatoms are used in ecological and paleoenvironmental research, their most important trait is the structure of their external siliceous covering, termed the frustule.

Structure of the diatom frustule

Diatom frustules are formed by the overlapping of two thecae; the upper theca is referred to as the "epitheca" and consists of a larger piece, the "epivalve" together with a set of bands (“copulae”), which together form the “epicingulum”; the lower theca is termed the "hypotheca" and consists of a "hypovalve" and a second set of bands forming the “hypocingulum” (Figure 1, 2). The hypocingulum and epicingulum (often referred to together as the “girdle”) combine to link the two valves. The axis connecting the apices of the valves is referred to as the "apical axis", and its transverse dimension as the "transapical axis". The axis perpendicular to the center of valves surface is the "pervalvar axis". Valves can be radial, linear, lanceolate, or elliptical in shape (Figure 1, 2). The copulae (girdle bands) can be few or many and often possess species-specific ornamentation. The position of the frustule, when the girdle is visible from the front, is referred to as a girdle view (Ross et al. 1979).

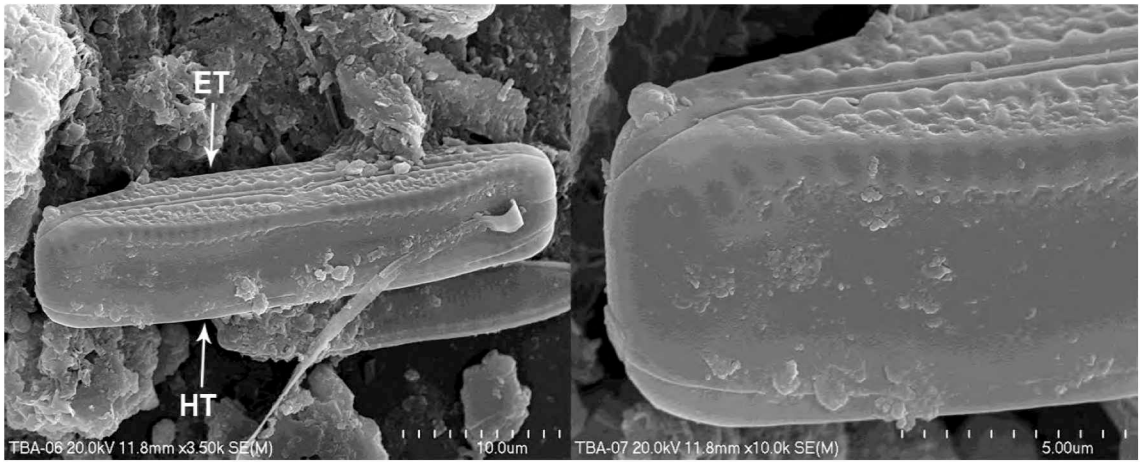


Figure 1. SEM image of a diatom frustule (ET: Epitheca – including the epivalve (uppermost); HT: Hypotheca - the only part visible is the top of the hypovalve, the rest of the hypotheca being hidden by the epitheca).

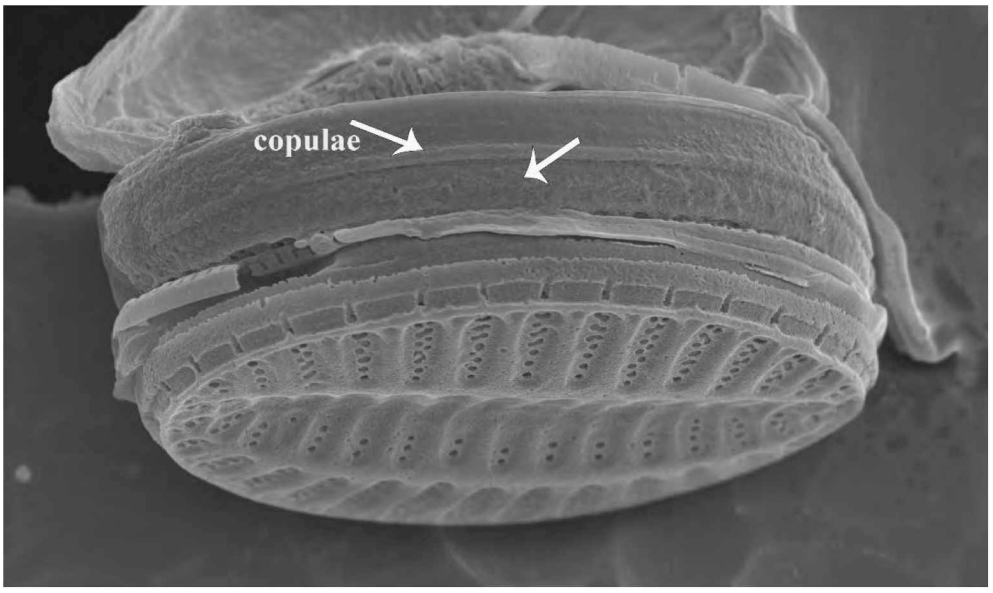


Figure 2. SEM image of a diatom in a girdle view.

"Areolae", which can be either circular or rectangular in shape, are holes present on the surface of the valve. Areolae can either be empty or lined with an internal layer (hymen) that forms different patterns (Riaux-Gobin et al. 2014). The formation of a row of areolae is called a "stria". The stria arrangement on the valve surface is parallel, radiate, or convergent. Usually, the arrangement of striae is uniseriate (single-rowed). In some genera (such as *Planothidium*), the striae arrangement can be biseriate or even triseriate (Potapova 2012).

The slits that divide the cell symmetrically or asymmetrically on the surface of the valve in many diatoms, including the species studied in detail here, are termed the "raphe" and are one of the most important characters in species identification. This trait is characteristic of a group of diatoms called "raphid". The function of the raphe is in locomotion (Round et al. 1990). In some genera, the raphe may be absent completely (centric and araphid diatoms) or may occur on just one valve (monoraphid diatoms). The raphe slit may appear as straight or undulate and terminates in central and polar regions. The endings of the raphe located in the central region of the valves are termed "proximal raphe ends". The endings of the raphe located in the polar regions of the valves are termed "distal raphe ends". Both endings can be straight, curved, or hooked; the distal raphe ends finish internally with a "helictoglossa" (having the shape of a pair of lips or a rolled tongue) but may extend to the valve apex externally (Ross et al. 1979, Round et al. 1990).

In some diatoms, the raphes are simple slits. In contrast, "fibulae" represent a more advanced structure, forming a partially closed channel (raphe canal) on one side of the valve by creating a porous belt under the raphe (Figure 3). The fibulae are found in genera such as *Nitzschia* and *Surirella*. In some genera (e.g., *Epithemia, Rhopalodia*), "costae" are observed. These structures resemble the ribs between the striae but form a complete arch across the valve, parallel to the striae (Figure 3), unlike the fibulae.

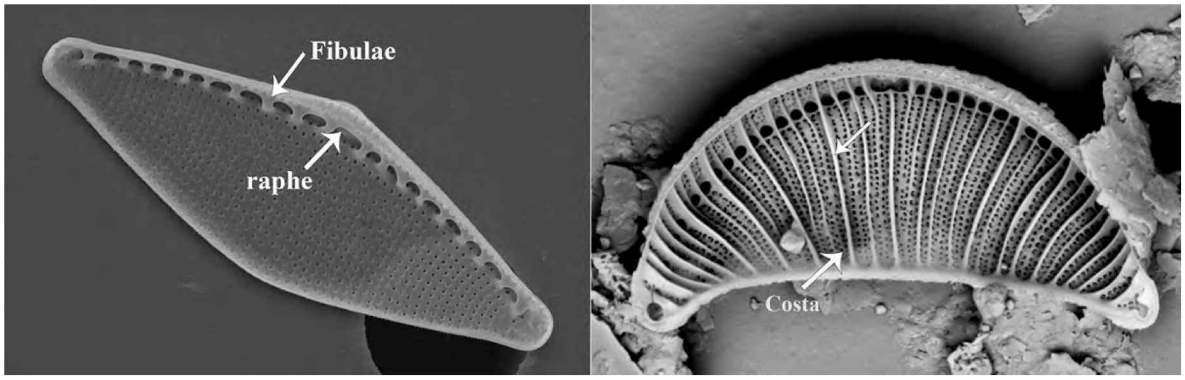


Figure 3. SEM image of an internal view of Nitzschia sp. and Rhopalodia sp. valves.

Morphologically, diatoms are classified into two groups. These are defined as centric and pennate diatoms.

Centric Diatoms

Centric diatoms include the class Coscinodiscophyceae. These diatoms are typically species with a circular or disc-shaped morphology and radial symmetry (Figure 4). Centric diatoms are commonly found in the plankton and are widespread in freshwater and marine habitats. They absorb solar energy and produce organic matter through photosynthesis, making them fundamental components of many ecosystems. In addition, some other centric diatoms, while also having a valve pattern that is radially symmetrical, have a valve outline that is triangular, multiangular or elongate. These are sometimes separated into a separate class, the Mediophyceae (e.g. Sims et al. 2006), which probably evolved from a group within the Coscinodiscophyceae.

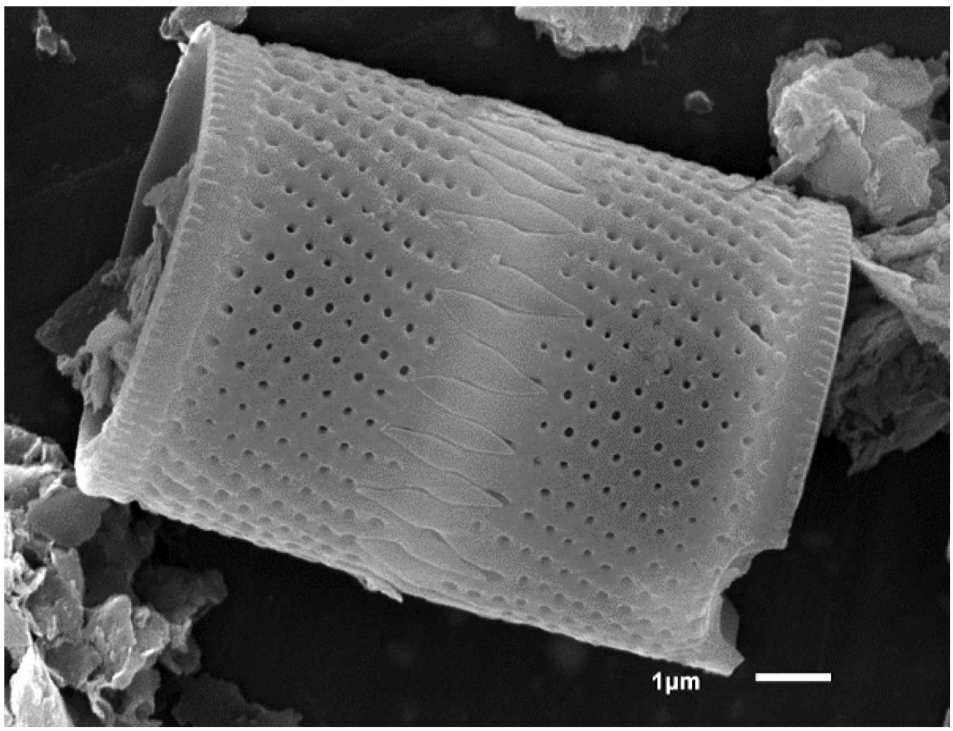


Figure 4. SEM image of siliceous appendages in a centric diatom (Aulacoseira).

Pennate diatoms

Pennate diatoms are further divided into three morphological groups: 1. araphid; 2. monoraphid; and 3. biraphid diatoms (Figure 5). Araphid diatoms (Figure 5A) lack a raphe system. Their valve structure is simpler than in other groups. A characteristic feature of some araphid diatoms are the rimoportulae, which look like openings in the valve and are also present in many centric diatoms. On the internal valve face, the rimoportula opening may have the shape of a pair of lips, while on the external valve face, it can be a simple, round aperture or a tube extending out from the valve. Monoraphid (Figure 5B) diatoms have a raphe on one valve only, while the other valve possesses a simple longitudinal rib (the sternum) instead. The frustules of monoraphid diatom genera show two different ornamentation patterns – this is called heterovalvy: the side of frustule with a raphe is called the raphe valve and the side without a raphe is called the sternum valve. Biraphid diatoms (Figure 5C) are a group in which the raphe is present on both valves; many biraphid diatoms are bilaterally symmetrical (Figure 5C, but not always (Figure 3).

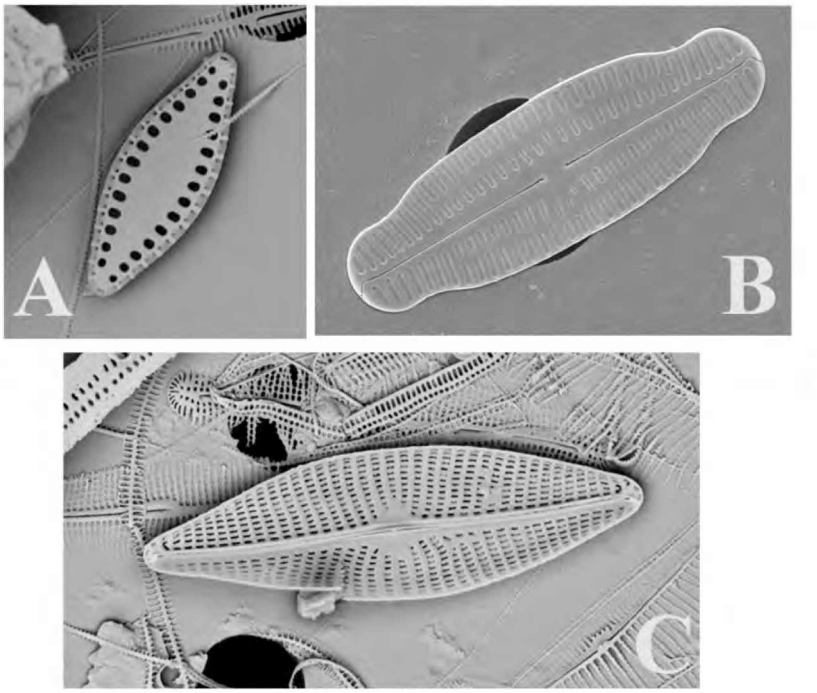


Figure 5. ***A-C****. SEM microphotographs of diatom frustules.* ***A****. Example of araphid diatom.* ***B****. Raphe valve of monoraphid diatom.* ***C****. Internal view of biraphid diatom valve.*

Diatom Ecology

Diatoms are widely distributed in habitats with water and sunlight. They can be found under various physicochemical conditions, in environments ranging from fresh to fully marine water, cave entrances and hot springs (Round 1990, Lauriol et al. 2006, Van de Vijver & Cocquyt 2009). Diatoms spread across seas with different temperatures, from the Arctic to the Antarctic, from the Pacific Ocean to the Baltic Sea (Snoeijs 1993, Poulin 2002, Riaux-Gobin & Poulin 2004a). They can inhabit estuarine areas, salt marshes, and even sandy areas affected by tides (Al-Yamani & Saburova 2011). The proliferation and continuity of diatom populations is controlled by many factors, including light, water temperature and salinity. Studies have shown that species compositions can vary significantly between low and high salinity areas (Sabbe et al. 1995, Stidolph et al. 2012). However, also essential for diatom growth is the presence of nutrients, which is – for most taxa – more important than other factors, e.g., water currents (Ağlaç & Balkıs 2014, Taş 2014). The main nutrients influencing the continuity of diatom populations are phosphorus, nitrogen, and silica. The ratio of these nutrients can be crucial for determining the distribution and ecological classification of species (e.g. Shraoui et al. 2009). Eutrophication, characterized by an excess of nutrients, can also significantly impact diatom populations. When favorable conditions exist, diatom communities may consist of many species, while under unfavorable conditions, the community may be represented by one or a few species only (Hendey 1964).

Diatom groups are divided into two categories based on their habitats. These are "planktonic diatoms," which are found floating in the pelagic zone of oceans, and "benthic diatoms" which are associated with bottom sediments in the littoral zone. Planktonic diatom forms have morphologically distinct structures (e.g. Taş & Hernández-Becceril 2017). Along with dinoflagellates and cyanobacteria, they play a significant role during algal blooms or they may be the sole group responsible for a bloom, for example in spring in temperate latitudes (Taş et al. 2016). On the other hand, benthic diatoms have been found to have a significant impact on grazer animals in shallow habitats (Cox et al. 2020). Epiphytic diatoms grow on the surfaces of plants, while epizoobiontic diatoms live on other organisms (Kaleli 2018). Benthic diatom communities are typically found clustered on the shallow bottom of water bodies. They tend to form dense aggregations where the water is calm.

Some diatoms are notable for their ability to survive in extreme environmental conditions, particularly in ecosystems characterized by high salinity, alkalinity, and low nutrient concentrations. An example of extreme environments are the so-called soda lakes characterized by high pH values (high alkalinity). Diatom communities found in soda lakes can optimize their metabolism to sustain growth and reproduction under such harsh conditions. Their adaptations to extreme environments are supported by genetic diversity and specific morphological traits, making them fundamental components of such ecosystems (Kociolek 2007). An example of an extreme, alkaline environment is Lake Van in Türkiye.

In this context, two genera, *Halamphora* and *Navicula*, exemplify the diatoms that have adapted to extreme environments, and studying them will in future undoubtedly showcase the unique physiological and ecological traits that allow them to thrive under challenging conditions.

Halamphora

The genus *Halamphora,* which is a raphid pennate diatom, was first described by Cleve (1895) as a subgenus of *Amphora,* and later reclassified as a separate genus by Levkov (2009). Species within this genus are typically characterized by valves with a dorsally outwardly curved margin and a ventrally flat or slightly swollen margin. The valves exhibit round or capitate apices and a narrow raphe sternum, and the bands are punctate. The raphe, positioned near the ventral margin, is either straight or slightly curved (Round et al. 1990; Levkov 2009; Stepanek & Kociolek 2018). According to current taxonomic classifications, there are 157 recognized species within the genus *Halamphora* (Guiry & Guiry 2024).

Although species of this genus are commonly found in inland water bodies with high conductivity, which can include both freshwater and brackish environments (Sala et al. 2007), they can also inhabit freshwater and semi-terrestrial environments (Levkov 2009, Van de Vijver et al. 2014, You et al. 2015). However, the majority of *Halamphora* species are primarily found in marine and brackish water habitats (Cleve 1895, Stepanek & Kociolek 2013).

Navicula

The genus *Navicula,* also a representative of the raphid pennate diatoms, is one of the most species-rich genera within the Bacillariophyceae, largely due to its historical use as a "catch-all" category for bilaterally symmetrical raphid diatoms with simple structural characteristics. The genus was first established in 1822 by Bory in his Dictionnaire Classique d'Histoire Naturelle (Bory de Saint-Vincent 1822). The name *Navicula*, meaning "small ship" in Latin, was chosen to describe the cell shape, resembling a weaving shuttle. Cells are generally solitary and motile, though some species may exist in mucilaginous tubes (Millie & Wee 1981). *Navicula* species possess two parietal chloroplasts, one on each side of the girdle, and their valves are symmetrical in both apical and transapical views, with rounded, acute, or capitate ends. The central area of the valves is often distinctly expanded (Patrick 1959, Cox 1979, Round et al. 1990). The areolae of *Navicula* species are elongate (parallel to the apical axis, as they are also in related genera such as *Pseudogomphonema*, *Seminavis*, *Hippodonta* and *Trachyneis* (e.g. Round et al. 1990), and contrast with the rounded areolae present in many other raphid diatom genera.

Navicula is an extremely widespread and ecologically versatile genus, occurring in fresh, brackish and marine waters, in various types of benthic habitat.

AREA OF STUDY

The studies described in this work were conducted in Lake Van (Figure 6). Van, located in the eastern Anatolian plateau of Türkiye, is a notable volcanic lake formed around 600,000 years ago by the eruption of the Nemrut volcanic mountain. Surrounded by towering mountains, it holds the distinction of being the largest soda lake globally, spanning an expansive area of 3764 square kilometres. Positioned at an average altitude of 1646 meters above sea level, Lake Van boasts an average depth of 171 metres, with its deepest known point plunging to 451 metres.

Figure 6. A map of Türkiye (A) with location of Lake Van (red frame, B).

The area around Lake Van has long been a focus of interest for ecologists due to its unique ecological characteristics and considerable biodiversity. Worldwide, there exist 35 biodiversity hotspots, three of which are situated within Türkiye: the Mediterranean, Caucasian, and Iran-Anatolia regions (Marchese 2015). In the latter, Lake Van stands out as an important area, the largest lake in the Iran-Anatolia region.

Despite the lake's prominence, the biodiversity of its animals is surprisingly under-explored (Kempe & Degens 1978). However, there are two known endemic species of fish: *Alburnus tarichi* Güldenstädt, 1814, a species adapted to the salty and alkaline environment that lives near the mouths of the streams flowing into the lake (Danulat & Kempe 1992), and *Oxynoemacheilus ercisianus* Erk'akan & Kuru, 1986, a stone loach

inhabiting the lake's microbialites (see below). The invertebrate biodiversity of Lake Van is still poorly understood as well, though recent studies have identified species new to science (Arslan et al. 2018).

On account of its unique characteristics, Lake Van has fascinated geologists, palaeoecologists and biologists. The fascination with Lake Van has been steadily growing since the groundbreaking research conducted by Degens et al. 1978, followed by the influential work of Landmann et al. in 1996. The significance of Lake Van has been underscored by the International Continental Scientific Drilling Program (ICDP), which, since 2001, has been conducting a project exploring the lake's sedimentary records (Litt et al. 2009). A unique feature of Lake Van is the presence of microbialites. These are sedimentary deposits that can soar to impressive heights of up to 40 meters above the bottom in Van, marking them as some of the largest in the world (Kempe et al. 1991). Microbialites are carbonate structures, like stromatolites, and they also result from the metabolism and activities of cyanobacteria. They are home to a diverse community of alkaliphilic heterotrophic bacteria (López-Garcia et al. 2005). According to the pioneer research by Kempe et al. (1991), the microbialites from Lake Van share similarities with stromatolites from Precambrian oceans and are therefore interesting models to study the development of life in Precambrian times. Akkus et al. (2021), in their study on the fishes living inside microbialites, noted that the microbialites are covered with various algae.

However, despite its prominence, Lake Van remains relatively unexplored in terms of its diatom flora. Approximately 80 years ago, Legler & Krasske (1940) conducted pioneering research on Lake Van, resulting in the description of 24 diatom taxa. Among these, they described 6 new diatom taxa, including *Amphiprora paludosa* var. *densestriata*, *Nitzschia incognita*, *Rhopalodia musculus* var. *suprasemicirculatus*, and *Surirella invicta*. They laid the foundation for diatom research in Lake Van. Subsequently, Lange-Bertalot et al. (1996) further examined and documented some of these species using light microscopy (LM) and scanning electron microscopy (SEM). Since then, there has been a notable absence of taxonomical studies on Lake Van diatoms. Therefore, until now, Legler & Krasske's (1940) pioneering work, along with the subsequent re-evaluations by Lange-Bertalot et al. (1996), constitutes the bulk of the research conducted on diatoms in the lake. Recently, however, Solak et al. (2021) contributed to this body of knowledge by utilizing modern molecular tools to provide evidence for the existence of a cryptic diatom species in Lake Van, *Nitzschia anatoliensis*.

THE AIMS AND HYPOTHESIS

In view of the unique characteristics of Lake Van and the paucity of knowledge on its diatom flora, the overarching objective of the research leading to this thesis is to explore the diatom assemblages of Lake Van to characterize their ecology, taxonomy, and molecular phylogeny. This way, this research study will contribute to the knowledge on diatom diversity, and biodiversity in general, in extreme environments.

In this thesis, this objective has been narrowed down to test the following hypotheses with reference to representatives of two genera (*Halamphora* and *Navicula*):

Hypothesis 1: The Lake Van diatom assemblages contain new, undocumented, diatom species that should contribute to our understanding of diatom diversity in extreme environments.

This hypothesis will be tested by addressing the following research questions:

- Can *Halamphora* and *Navicula* species new to science be found in Lake Van
- How does the knowledge on Lake Van diatoms contribute to our understanding of diatom taxonomy in extreme environments?

Hypothesis 2: The diatoms of Lake Van exhibit unique molecular diversity and reveal evolutionary adaptations to the lake's extreme conditions.

The research questions addressed to test this hypothesis are:

- What is the level of molecular diversity within the Lake Van diatom populations studied?
- How do the new diatom species phylogenetically relate to known species, and what can be inferred about their evolutionary relationships?
- Which molecular traits can be regarded as adaptations to the extreme environments represented by soda lakes?

It is assumed that phylogenetic analyses will assist in clarifying the relationships among these species and in elucidating their responses to environmental stressors, and the study will thus contribute to a broader understanding of diatom adaptation strategies in soda lakes.

Sampling and material collection

Samples were collected from 7 stations distributed around Lake Van on its shores (Figure 7) and located at sites reflecting the diversity of the lake's habitats. Samples were obtained by scraping epilithic material from hard substrates (rock surfaces). The collected samples were placed into sterile 50 mL falcon tubes and transported to the laboratory for diatom cell isolation.

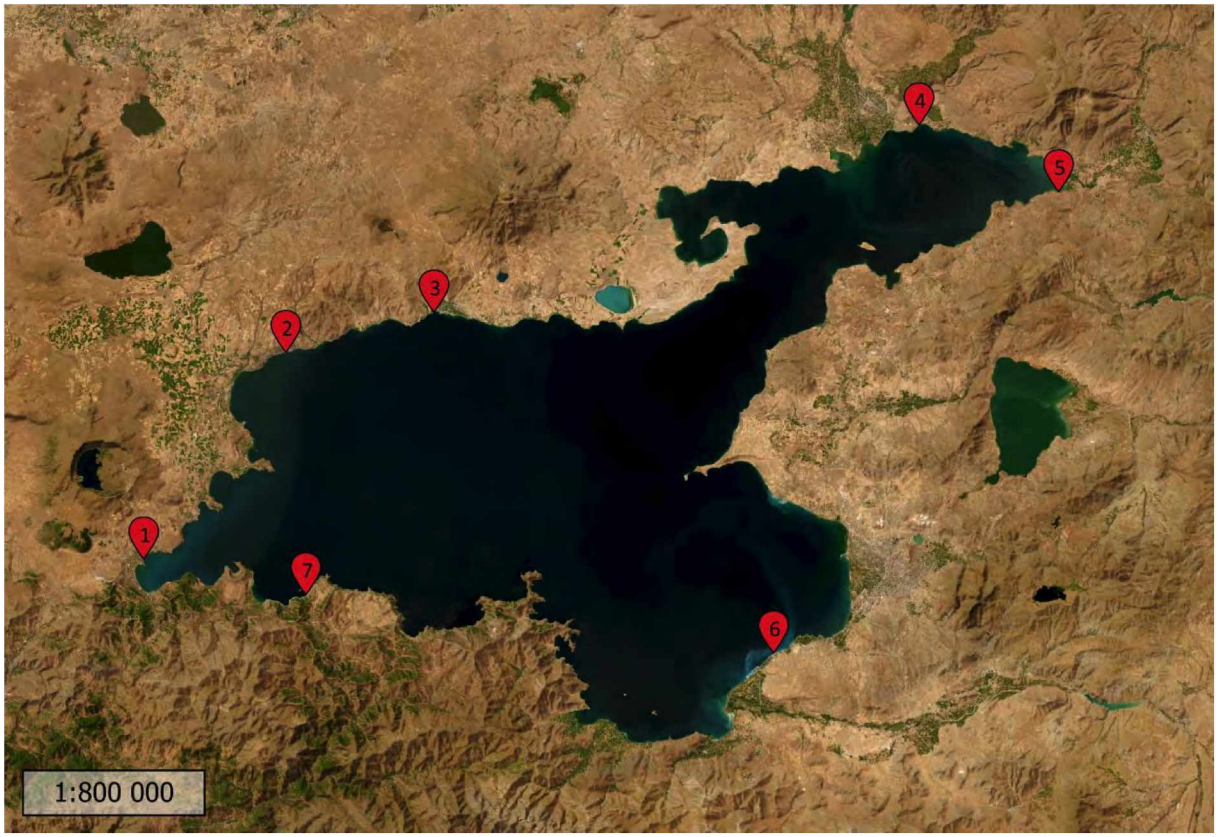


Figure 7. Map showing the seven sampling stations around Lake Van.

Monoclonal culture preparation

Step 1. Isolation of single cells

Water samples were examined under an inverted microscope to look for live diatoms. Whenever live *Halamphora* and *Navicula* were observed, single individuals were isolated with micropipettes (Figure 8) to minimize contamination. Isolations were repeated at least three times.

Figure 8. Detailed view of the micropipette used for single-cell isolation.

Step 2. Inoculation into F/2 Medium

The single cells isolated were inoculated onto F/2 medium (Guillard 1975, modified), a suitable medium containing the nutrients and salts necessary for diatom growth and reproduction. In the original formulation of F/2, seawater was used as the base. For my isolates the salinity level of the F/2 medium was kept at 18 psu, to roughly correspond to the conditions occurring in Lake Van. The medium was sterilized in the autoclave under standard conditions of 120°C for 20 minutes.

Culturing under controlled light and temperature

Monoclonal cultures were cultivated at the laboratory under controlled light and temperature conditions. The cultures were consistently maintained in active growth at 22°C under a light intensity of 60 µmol photons $m^{-2}s^{-1}$ and a photoperiod cycle of 14 hours of light followed by 10 hours of darkness (Figure 9).

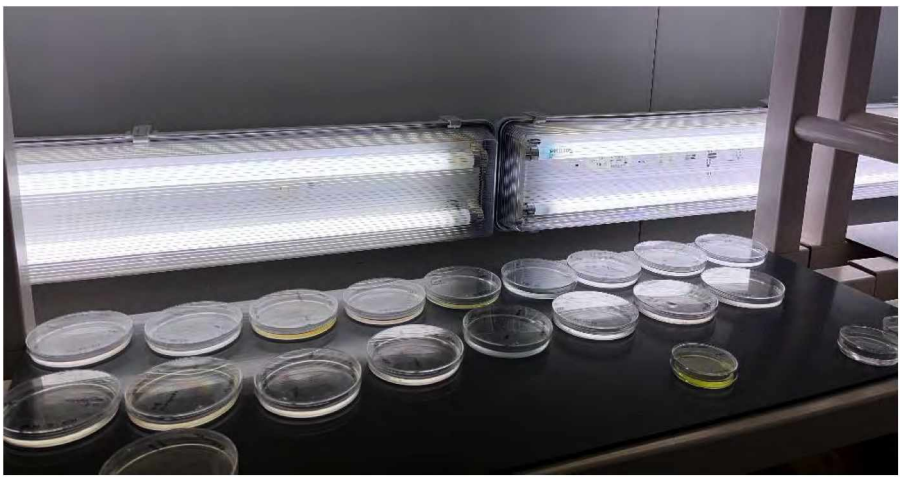


Figure 9. Cultivation setup with controlled light and temperature conditions.

Subsequent Isolation Procedures

Each individual cell that had been isolated underwent a re-isolation process to be utilized for obtaining a pure culture. Throughout this procedure, the multiplication of cells within the culture and the possible contamination by other microorganisms were monitored following each isolation step. The process of isolating single cells was repeated if necessary. A minimum of three successive isolation cycles were carried out, thereby guaranteeing monoclonal diatom cultures with desired levels of purity and homogeneity. No attempt was made to fully eliminate bacteria.

Light (LM) and scanning electron microscopy (SEM)

Diatom cultures were directly transferred onto glass slides and covered with a cover glass for imaging. Living diatom images were captured using light microscope Zeiss Axio Scope A1 (Carl Zeiss, Jena, Germany) with Canon EOS 500D camera and Canon EOS Utility software (Canon, Tokyo, Japan) at a magnification of 1000× .

To prepare cleaned frustules for microscopy, 5 mL of monoclonal cultures were transferred to 20 mL beakers containing 10 mL of 10% hydrochloric acid (HCl). After 24 hours, samples were washed four times with distilled water, then suspended in 30% hydrogen peroxide (H_2O_2) and boiled for approximately four hours. Finally, the samples were washed again four times with distilled water.

For LM, the cleaned material was air-dried on cover glasses and mounted on glass slides with Naphrax (Brunel Microscopes Ltd., Chippenham, UK) solution. Images of cleaned diatom frustules were captured using the Zeiss Axio Scope A1 with the Canon EOS 500D camera and Canon EOS Utility software, as above.

For SEM, the cleaned samples were deposited as a drop onto a Nuclepore Track-Etch Whatman membrane (Maidstone, England). The membranes were air-dried overnight, mounted on aluminum stubs with carbon tape, and coated with gold using a Q150T coater from Quorum Technologies (Laughton, UK). SEM examination was conducted at the Faculty of Chemical Technology and Engineering, Western Pomeranian University of Technology in Szczecin (Poland), using a Hitachi SU8020 microscope (Tokyo, Japan).

Molecular analyses

DNA isolation and PCR protocol for single gene molecular analyses

The genomic DNA extraction procedure for PCR and Sanger sequencing employed Chelex resin (Cat. No. 142-2842-MSDS, Bio-Rad, Hercules, CA, USA). A 10% working solution of Chelex was prepared by dissolving 1 g of resin in 10 mL of distilled water. Subsequently, 150 µL of the working solution was added to 1.5 mL Eppendorf® tubes containing centrifuged diatom biomass. After vortexing and heating the mixture to 95°C, the suspension was incubated for 20 min. After this, centrifugation was carried out for 5 minutes at 4°C and 10,000 rpm (RCF = 6.010g), followed by the transfer of the DNA-containing supernatant into sterile Eppendorf® tubes. The extracted DNA material is stored at -21°C in the Szczecin Diatom Culture Collection (University of Szczecin). In the PCR process, one nuclear coding gene, 18S rRNA (SSU), was amplified alongside two plastid molecular markers, *rbcL* and *psbC*. The three genes have been extensively and successfully used for phylogeny in recent years (e.g. Theriot et al. 2010, Ashworth et al. 2022), whereby reference data are available (Gorecka 2022).

The PCRs were run at the University of Szczecin, Institute of Marine and Environmental Sciences, utilizing the S1000™ Thermal Cycler (Bio-Rad, Hercules, CA, USA). The experimental design was based on protocols adapted from Theriot et al. (2010), and Ashworth et al. (2013). Reaction mixtures were prepared in PCR Eppendorf tubes for a final volume of 25 µL. The mixture consisted of 17.35 µL of nuclease-free water, 2.5 µL of 10x DreamTaq™ reaction buffer containing 20 mM MgCl2 (Cat. No. B65, Thermo Fisher Scientific, Waltham, MA, USA), 1 µL of deoxyribonucleotide triphosphate mix (dNTPs) (Cat. No. E0503, EURx, Gdańsk, Poland), 0.5 µL of each forward and reverse primer, 0.15 µL of DreamTaq™ DNA Polymerase (cat. no. EP0702, Thermo Fisher Scientific), and 3 µL of template DNA isolated prior to PCR amplification. Forward and reverse primers were synthesized by GENEWIZ, Inc. (Leipzig, Germany). Quality control measures were rigorously applied, with two control samples prepared alongside each PCR run: one containing DNA that had been previously amplified with success, to confirm the accuracy of reaction mixture preparation, the other without DNA to rule out contamination. Following PCR amplification, the resultant products were stored at +4°C in the refrigerator until they were dispatched for sequencing, ensuring the preservation of sample integrity and data reliability throughout the experimental process.

To assess the quality and validate the PCR products, electrophoresis was conducted on a 1% agarose gel (Maximus, Łódź, Poland) blended with 5 µL of SimplySafe™ dye (Cat. No. E4600, EURx) for the visualization of nucleic acids. Employing a horizontal gel electrophoresis system (Peqlab, VWR, USA) and liquid 1x TAE buffer (Cat. No. 0220, EURx), the process was carried out for 25 minutes under a voltage of 120 mV. Each PCR product was loaded onto the gel with loading buffer BLUE (Cat. No. E0260, EURx). Post-electrophoresis, the gel was visualized using the G: BOX F3 gel documentation system (Syngene, Frederick, MD, USA) to assess the PCR.

The PCR products were dispatched to GENEWIZ (Leipzig, Germany) for Sanger sequencing. Upon completion of sequencing, one or two fragments of a gene, each ranging from 800 to 1000 bp in length, were obtained. These raw sequencing reads were aligned utilizing the Contig Assembly Program (CAP) within BioEdit software version 7.2.5 (Hall, 1999), and their accuracy was verified manually through chromatograms. Subsequently, a consensus sequence derived from the raw reads was compared against the NCBI nucleotide database utilizing the BLAST Sequence Analysis Tool (Madden, 2013) to identify significant discrepancies in the sequence and ascertain the correct orientation of the sequence.

Next generation sequencing, assembly and annotation of the plastome

DNA from clones SZCZEY2172, SZCZEY2262, SZCZEY2166, SZCZEY2167, SZCZEY2176 and SZCZEY2177 was extracted using a modified version of the protocol of Doyle and Doyle (1990). The cell walls of diatoms were disrupted by grinding them in liquid nitrogen. The broken diatom cells were processed in a lysis buffer with proteinase K and left to incubate at 65°C during 30 minutes. Chloroform was added to the lysate and centrifuged 30 minutes at 3000 rpm in order to remove proteins. The supernatant was transferred into another tube with isopropanol and left 20 minutes at -20°C. Finally, DNA was precipitated by centrifugation at 15,000 rpm during 15 minutes, rinsed with cold 70% ethanol overnight, air dried and finally diluted in TE buffer. Total DNA was sent to the Beijing Genomics Institute (BGI) in Shenzhen (China) to be sequenced on a DNBSEQ platform. Usually, an amount of 60 to 80M clean paired-end reads of 150 bp were obtained from each sample. Reads were assembled using SPAdes (Bankevich et al. 2012) with a stringent k-mer of 125. The contigs files obtained after assembly were datamined by customized blastn queries using organellar genomes or nuclear ribosomal RNA genes

from previous works as references (e.g. Solak et al. 2021). Contigs of interest were extracted and their circularity was verified. Gene detection for the organellar genome was performed using a Perl script developed at the Laval Université in Québec (Gagnon 2004). Annotation was done using Sequin. To draw the genome maps, the portal of OGDRAW (Greiner et al. 2019). was employed.

Molecular dataset preparation

The DNA-based phylogenetic analysis was conducted. The first step of the analysis involved creating a database consisting of full or partial sequences of three selected genes: SSU, *rbcL,* and *psbC.* This reference database was constructed using the GenBank online sequence library. Taxon sampling was based on diatom systematics and the availability of existing sequences. Priority reference taxa were identified, and - to achieve optimal results - sequences containing complete gene sequences and all three genes were given precedence during database construction. Incomplete or partial sequences were treated as missing data.

Phylogenetic analysis

In order to confront results obtained from morphological analyses, DNA sequence-based phylogenetic analyses was performed. Genes chosen for these analyses were the most popular and most accessible for the biggest number of diatom species (Theriot et al. 2010, Witkowski et al. 2016, Dąbek et al. 2017, Stepanek & Kociolek 2019, Gorecka et al. 2022, Ashworth et al. 2022, etc). In this study three phylogenetic trees were calculated, two based on three-gene concatenated datasets (SSU, *rbcL*, and *psbC*) and one based on *rbcL* sequences datasets only. The reference sequences were chosen based on their systematic position in relation to the studied strains, of *Navicula* and *Halamphora*, so that the closest relatives were represented optimally. The reference datasets consisted of sequences available in GenBank, an on-line repository, at the time of manuscript preparation. For each gene a dataset was prepared separately in FASTA format.

All datasets were aligned using MAFFT 7 (Katoh & Standley 2013) and automatically trimmed with trimAl (Capella-Gutiérrez et al. 2009). The best model of evolution for each gene was chosen with ModelTest-NG (Darriba et al. 2020). For the three gene datasets, alignments were concatenated after trimming using Phyutility 2.7.1 (Smith & Dunn 2008), but in case of the *rbcL*-only phylogenetic analysis, this step was not necessary. The

Maximum Likelihood phylogenetic trees were reconstructed using IQ-TREE 2.2.0 (Minh et al. 2020) with 1000 ultrafast bootstrap replicates. The final phylogenetic tree was the ML tree with the highest score, and the trees were visualized using MEGA 11 software (Tamura et al. 2021). The bootstrap value are all indicated at the nodes.

RESULTS

The taxonomic position of six diatom strains was determined using an integrated approach. This approach involved the combination of frustule morphology and ultrastructure observations along with plastid observations and DNA sequence data, including phylogenetic analysis based on single and three-gene concatenated datasets.

As part of the study, three species belonging to diatom genera *Halamphora* (*H. vanseea, H. witkowskii*) and *Navicula* (*N. vantushpaensis*) were described as new to science using methods described above. Additionally, two plastid and two mitochondrial genomes of diatoms *H. vanseea* and *N. vantushpaensis* were assembled, annotated and analyzed. The molecular data that resulted from this study are freely accessible in an on-line repository (NCBI GenBank).

The new species' descriptions are presented below.

Navicula vanseea Yilmaz, Gastineau, Solak & Witkowski, sp. nov

The following chapter summarizes the description of *Navicula vanseea* Yilmaz, Gastineau, Solak & Witkowski, sp. nov, which was the first new species discovered in the course of the PhD. This description has been published in Phytokeys (online ISSN 1314-2003) as Yılmaz E, Mann DG, Gastineau R, Trobajo R, Solak CN, Górecka E, Turmel M, Lemieux C, Ertorun N, Witkowski A (2024) Description of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterization of its organellar genomes and multigene phylogeny. PhytoKeys 241: 27-48. https://doi.org/10.3897/phytokeys.241.118903

Etymology. The name given to the species refers to the German name of Lake Van (Vansee, the Lake of Van) as it was used in the work of Legler and Krasske (1940) and is meant as a tribute to these authors and their work.

Taxonomic description:

Light Microscopy (Figure 10A-Y): Valves: smaller specimens are elliptic, tapering towards cuneately rounded apices, while larger specimens are linear-elliptic-lanceolate, narrowly rounded, with occasionally slightly protracted poles. Valve dimensions (n = 39): length 11.0–28.0 μm, width 5.0–6.5 μm. The raphe is filiform and straight. The central area is small and rounded, with a narrow axial area. Striae are strongly radiate, sometimes

irregularly shortened around the central area, approximately 12–13 in 10 μm. Lineolae are difficult to resolve in LM, approximately 50 in 10 μm.

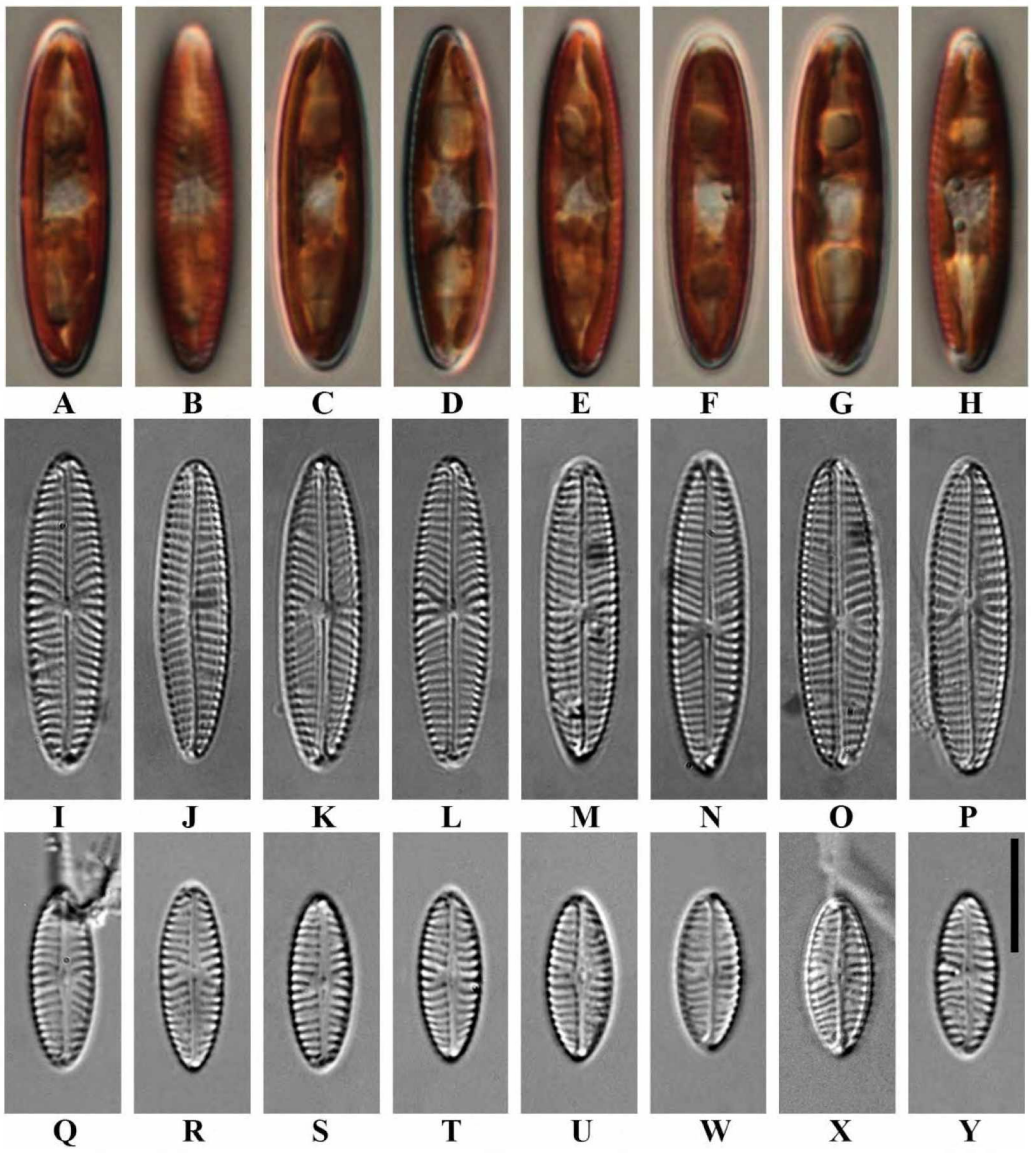


Figure 10. ***A-Y****. Navicula vanseea sp. nov. LM micrographs.* ***A–H****. In vivo pictures of Navicula vanseea sp. nov. SZCZEY2172.* ***I****. LM image of a cleaned valve from wild material.* ***J–P****. Cleaned valves of Navicula vanseea sp. nov. SZCZEY2172.* ***Q–Y****. cleaned valves of Navicula vanseea sp. nov. SZCZEY2262 Scale bar: 10 μm.*

Scanning Electron Microscopy (Figure 11A-H): External valve surface: The valve surface is flat, with apically elongated areolae. The raphe sternum is slightly elevated above the valve face level. The axial area is very narrow, with a very slightly expanded, small, asymmetric central area. Proximal raphe endings are drop-like, slightly deflected unilaterally. Distal raphe endings are strongly hooked in the same direction.

Internal valve surface: The valve surface is slightly arched with transapical striae positioned in relatively deep grooves, bordered by virgae that become thicker towards the center of the valve. The central area is asymmetric, usually only slightly expanded, but sometimes more strongly. The internal lineolae openings are slit-like, narrower than the vimines. Lineolae are occluded by hymens; two isolated lineolae are present at the valve apex. The raphe sternum slightly widens at the center to form a fusiform ridge enclosing the central raphe endings, which are simple, straight, and separated. Distally, the raphe terminates in well-developed helictoglossae.

These characteristics distinguish *Navicula vanseea* as a new species, notable for its specific ecological and morphological features (Yilmaz et al 2024). The differences in size and shape between the two strains of *N. vanseea* are striking (Fig. 10I-Y) but not unusual for raphe pennate diatoms and conform to the 'rules' of shape and size change during the life cycle established by Geitler (1932; see also Woodard et al. 2016).

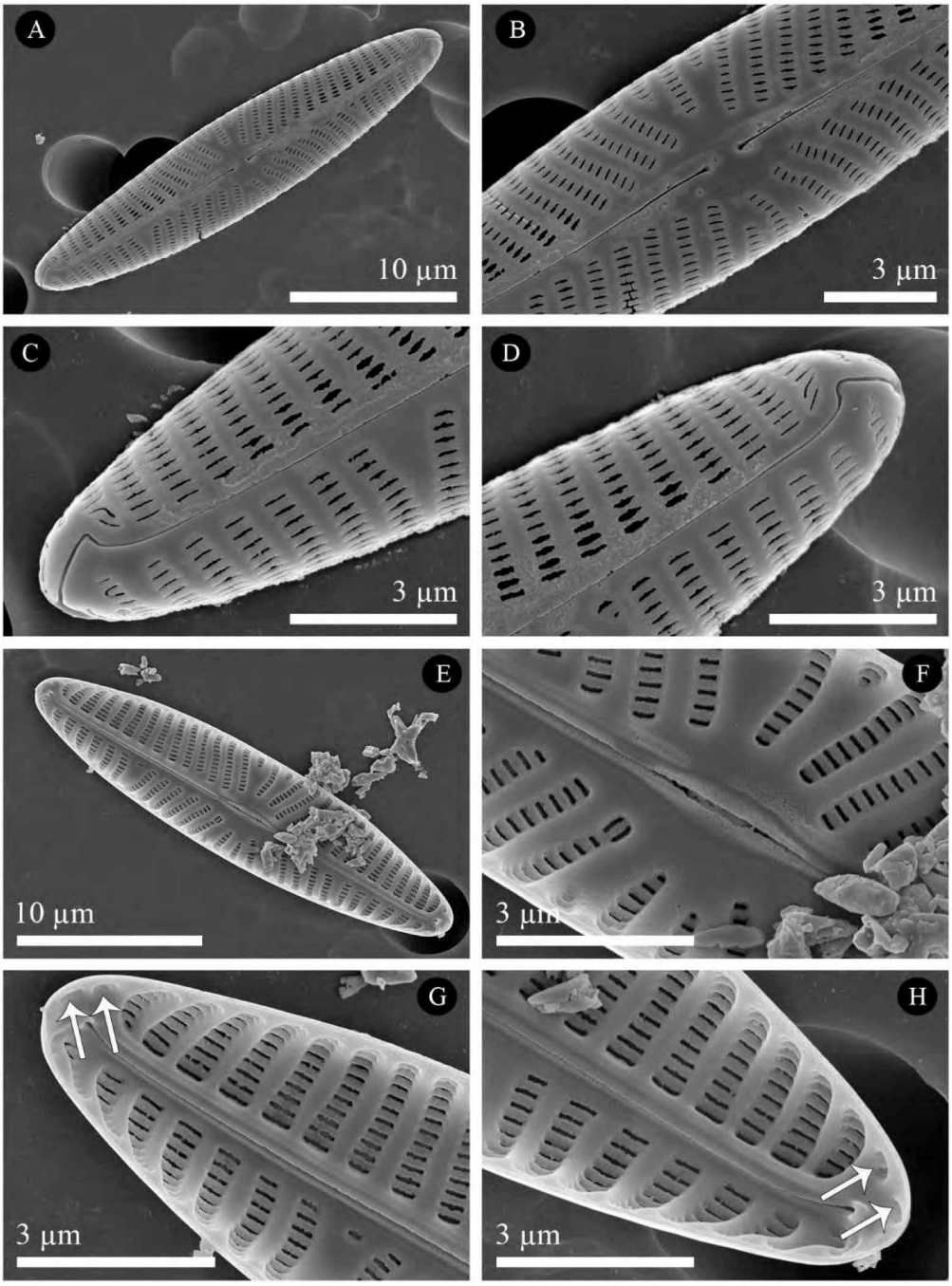


Figure 11. ***A-H****. SEM micrographs of Navicula vanseea sp. nov. SZCZEY2172.* ***A****. External view of the entire valve.* ***B****. Details of central area showing simple, slightly drop-shaped proximal raphe endings and shortened striae.* ***C-D****. Details of the two apices of a single valve.* ***E****. Internal view of the entire valve.* ***F****. Details of central area showing filiform proximal raphe endings in a fusiform expansion of the raphe-sternum.* ***G-H****. Details of apices showing well-developed helictoglossae showing two isolated lineolae (white arrows). Scale bars: 10 µm (A, E); 3 µm (B–D, F–H).*

Genomics and phylogeny

The nuclear rRNA gene cluster

The entire rRNA gene cluster was sequenced for both clones and was subsequently submitted to the NCBI GenBank under accession numbers OR797294 (SZCZEY2172) and OR797293 (SZCZEY2262). The cluster spans a length of 4902 base pairs, with the following distribution: 18S – 1792 bp, ITS1 – 195 bp, 5.8S – 155 bp, ITS2 – 260 bp, 28S – 2500 bp.

Mitochondrial genome

Analysis of the mitochondrial genome revealed a 43997 bp contig corresponding to strain SZCZEY2262. However, due to repeated sequences at its ends, circularization of the contig was not feasible. Nonetheless, for clarity in representation, it is depicted as circular in the map (Figure 12). This mitogenome encompasses 34 protein-coding genes along with the conserved open-reading frame (ORF) orf150 situated between rps11 and mttB, as well as two rRNA genes and 23 tRNA genes (GenBank: OR795084). Notably, the nad11 gene is bifurcated into two distinct subunits, with two protein-coding genes, two rRNA genes, and one tRNA gene interspersed between them. Within the repeated segment of the genome, two copies of the same ORF, orf145, were identified. Additionally, a 767-bp group I intron was found within the rnl gene.

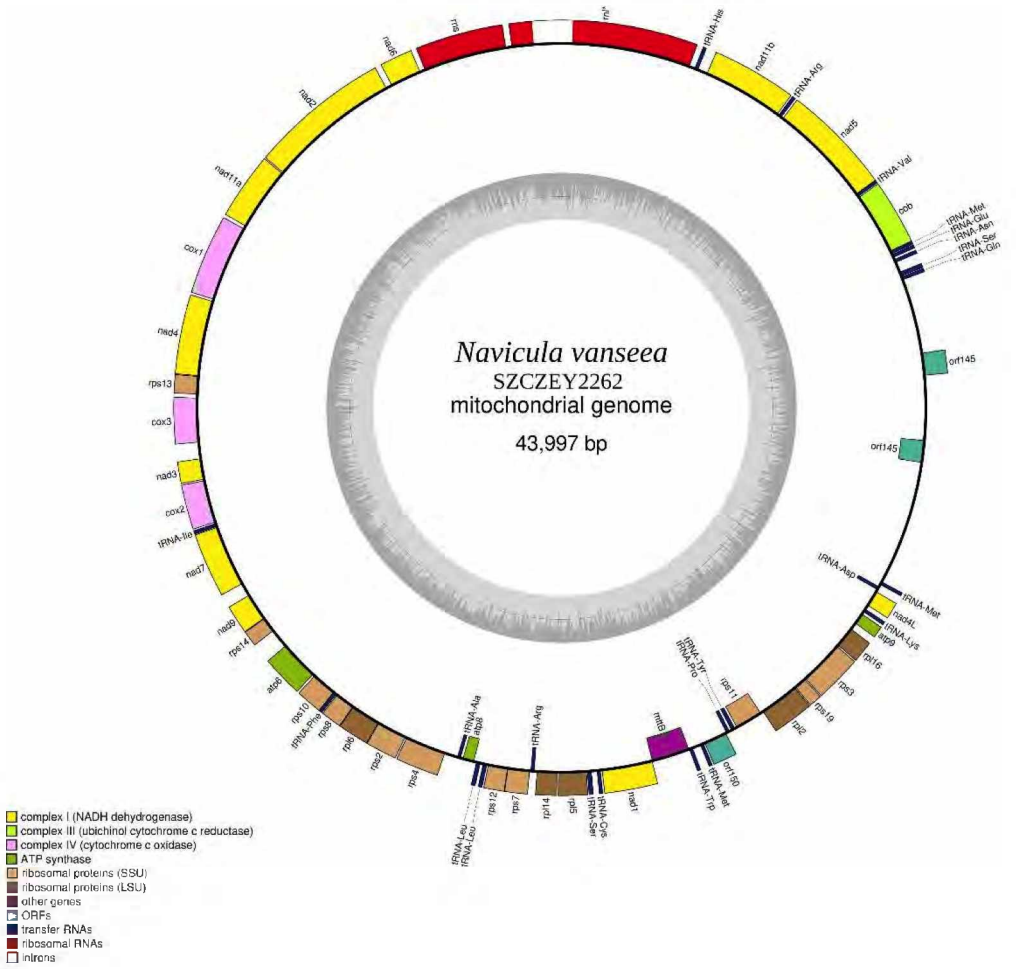


Figure 12. Map of the mitochondrial genome of Navicula vanseea sp. nov. SZCZEY2262.

The plastid genome analysis revealed significant differences between strains SZCZEY2262 and SZCZEY2172. In SZCZEY2262 (GenBank: OR795085), the plastome is 158,005 bp long (Figure 13), comprising a large single-copy (LSC) region of 72,941 bp, a small single-copy (SSC) region of 49,714 bp, and an inverted repeat (IR) of 17,675 bp. The LSC contains 74 conserved protein-coding genes, two non-conserved ORFs, two putative integrase/recombinase xerC genes, and 17 tRNAs. The SSC encodes 51 conserved protein-coding genes, eight tRNAs, and five non-conserved ORFs. The IR contains one conserved protein-coding gene, three rRNAs, six non-conserved ORFs, four tRNAs, and one putative serC gene. Additionally, there are two IB4 group I introns in the

ribosomal rrl gene at positions 1917 and 1931. These introns contain two putative LAGLIDADG homing endonuclease genes referred to as L1917 and L1931.

In SZCZEY2172 (GenBank: OR795086), the plastome is 157,990 bp long (Figure 14), with a gene content similar to that in SZCZEY2262. The LSC is 72,913 bp long, the SSC is 49,727 bp long, and the IR is 17,675 bp long. Both strains contain a 43 AA ORF in their SSC, like the hypothetical chloroplast RF35 encoded by ycf35.

There are notable differences in non-conserved ORFs and several SNPs have also been spotted in protein coding genes. These differences underscore the genomic diversity between these strains.

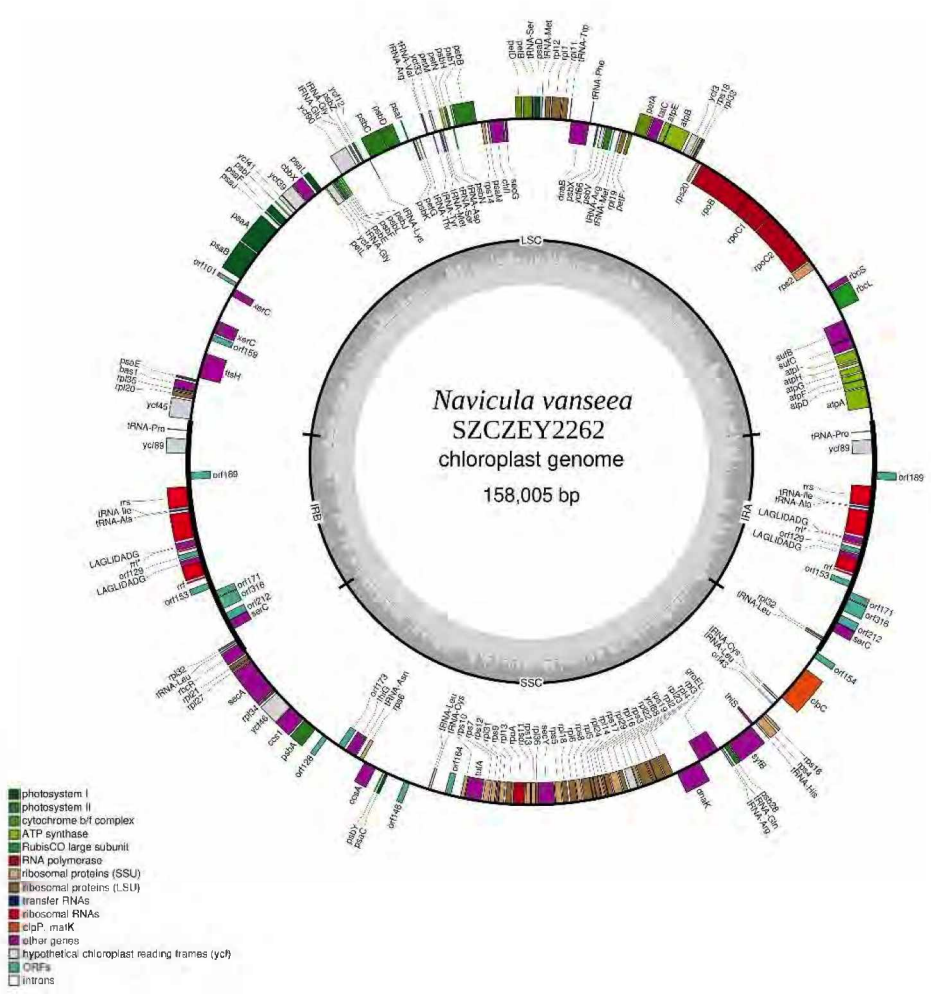


Figure 13. Map of the plastid genome of Navicula vanseea sp. nov. SZCZEY2262.

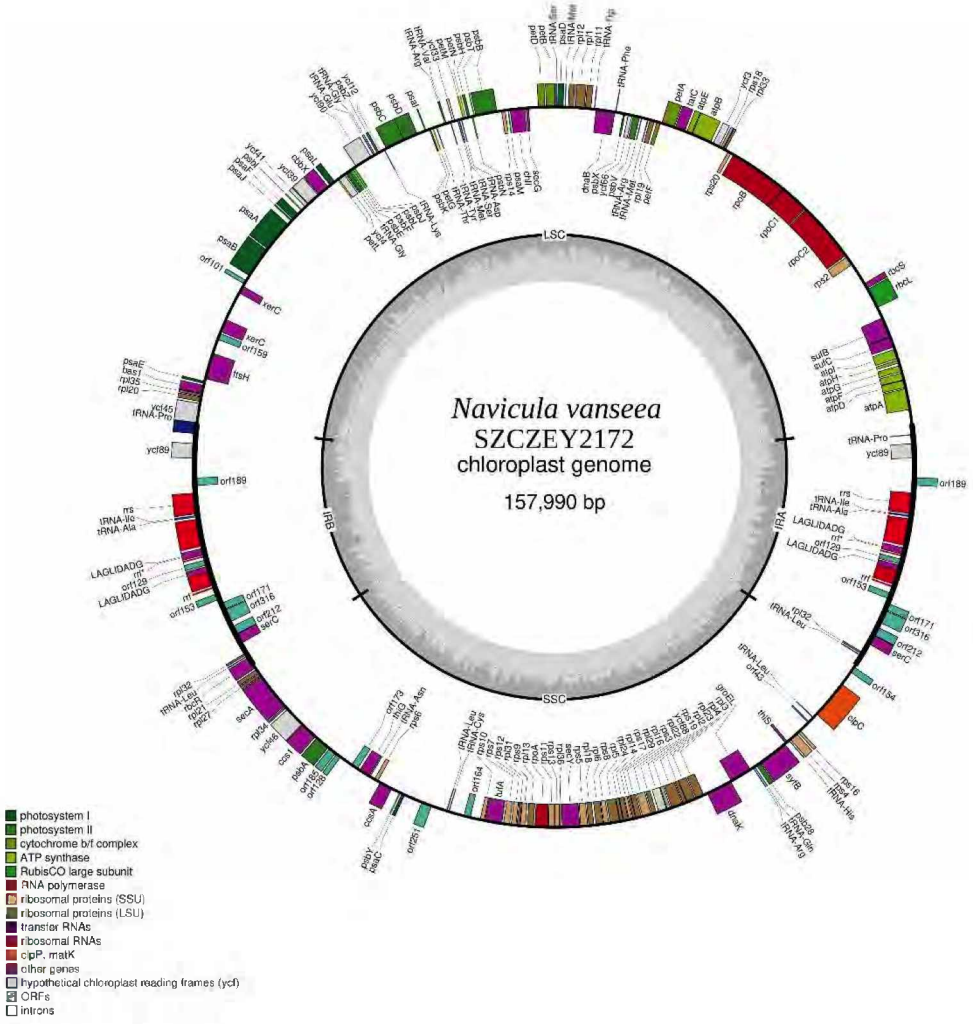


Figure 14. Map of the plastid genome of Navicula vanseea sp. nov. SZCZEY2172.

A Maximum Likelihood phylogenetic tree was inferred from a concatenated alignments of *psbC*, *rbcL*, and 18S sequences. The tree provides insights into the placement of *Navicula vanseea* within the broader phylogenetic context of *Navicula* and the related genera *Rhoikoneis*, *Pseudogomphonema*, *Seminavis* and *Cymatoneis* (Figure 15); the tree indicates that the genus *Navicula* is paraphyletic and will therefore need revision in due course, once more species have been sampled and characterized molecularly. It is noteworthy that the phylogenetic tree strictly distinguishes *N. vanseea* from *Navicula cincta* (Ehrenberg) Ralfs 1861 and *Navicula microdigitoradiata* Lange-Bertalot, 1993, two morphologically similar species with which it might have been easily misidentified, if it were not for the methods employed here.

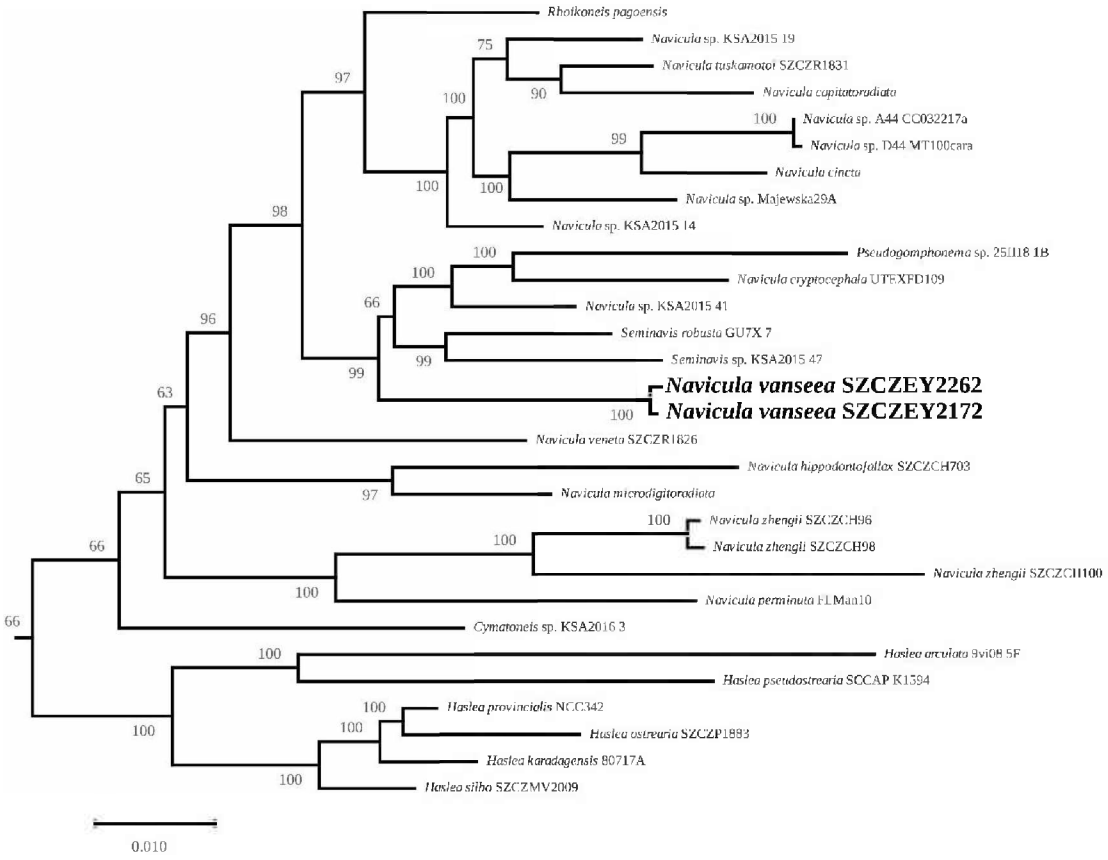


Figure 15. A cut from a Maximum Likelihood phylogenetic tree obtained from concatenated alignments of psbC, rbcL and 18S.

Halamphora witkowskii sp. nov. Yilmaz, Solak, & Gastineau, sp. nov.

The following chapter introduces *Halamphora witkowskii* sp. nov. Yilmaz, Solak, & Gastineau, sp. nov., a species named after our beloved colleague and mentor, Prof. Dr. Hab. Andrzej Witkowski. This article is currently in press in a special issue of Nova Hedwigia dedicated to Prof. Witkowski (online ISSN 2363-7188).

Etymology: The new diatom species is dedicated to the esteemed mentor we have recently lost, Prof. dr. hab. Andrzej Witkowski, a Fellow of the Polish Academy of Sciences.

Taxonomic description:

Light Microscopy (Figure 16A-R): The valves are semi-lanceolate in shape, appearing dorsiventral with a gently arched dorsal margin and nearly straight ventral margin. The valve tips are slightly extended, capped, and exhibit a weak ventral curvature. The valves measure 14–16 μm in length and 4–5 μm in width. The axial area appears narrow, more prominently expressed on the ventral side. The central area expands towards the ventral side. The raphe displays a biarcuate structure with proximal endings slightly bent dorsally. Dorsal striae exhibit distinct punctation, radiating throughout, while ventral striae are challenging to discern with light microscopy, becoming recognizable near the ventral margin, occurring at a density of 20–22 in 10 μm (Yilmaz et al. in prees).

The two strains were very similar morphologically and represent very similar stages in the life cycle of the new species (in contrast to, for example, *Navicula vanseea*). It can be expected that the length range will be found to be much greater once further samples have been obtained from Lake Van.

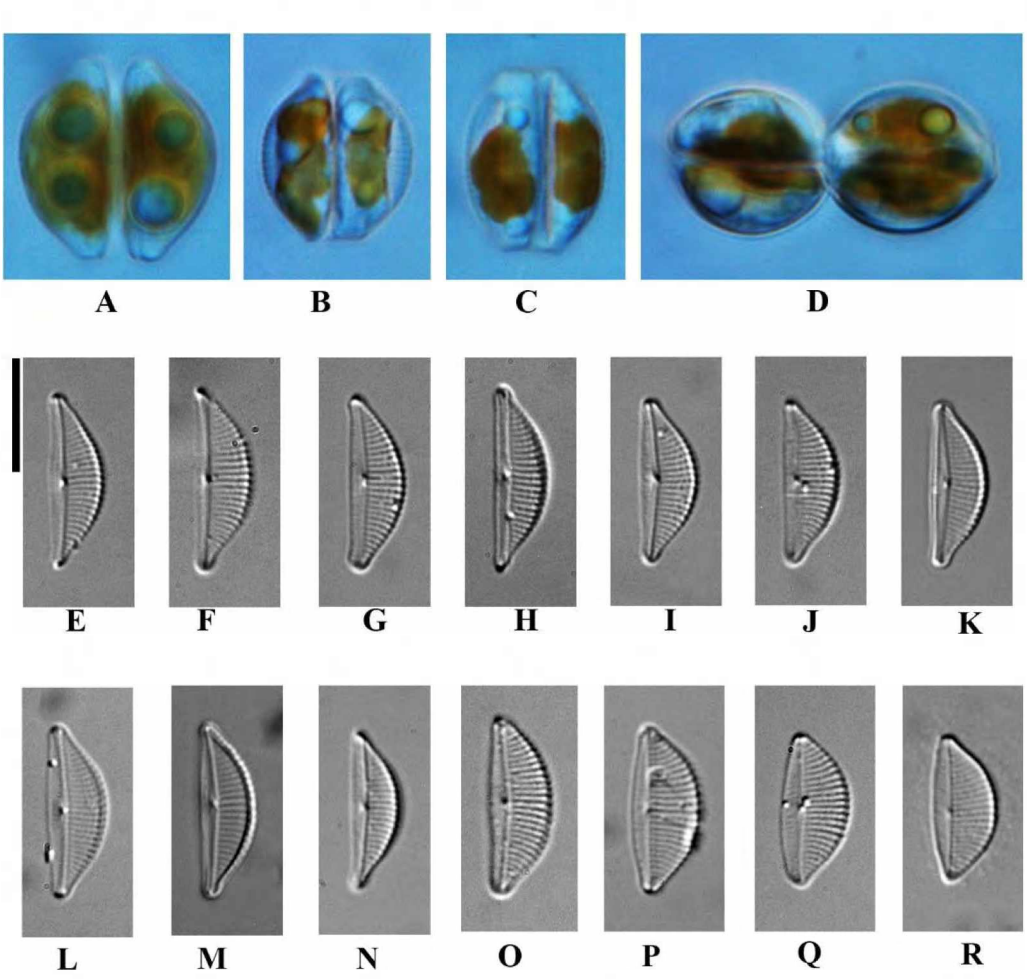


Figure 16. ***A-R****. Halamphora witkowskii sp. nov. LM micrographs.* ***A-D****. In vivo pictures of Halamphora witkowskii sp. nov. SZCZEY2176.* ***E-H****. LM image of a cleaned valve from wild material, I-N. cleaned valves of Halamphora witkowskii sp. nov. SZCZEY2177.* ***O-R****. Cleaned valves of Halamphora witkowskii sp. nov. SZCZEY2176. Scale bar = 10 µm.*

Scanning Electron Microscopy (Figure 17A-F): The valves are oriented dorsiventrally, featuring a convex dorsal margin and nearly straight ventral margin. The raphe ledge appears wide, with crenulations evident on the dorsal side, aligning with adjacent striae. The raphe exhibits a subtle arching, with proximal endings expanding into central pores. Dorsal striae display a biseriate arrangement throughout, occasionally interrupted by longitudinal bars, creating isolated sections near the dorsal margin. Areolae contain small circular foramina arranged in two rows. The biseriate nature of the striae is also observable internally. Ventral striae are uniseriate, composed of areolae with varying lengths, increasing towards the central area. Areolae within the central area appear rounded and less dispersed. Internally, a single row of dorsal areolae near the raphe is

delineated by internal transapical ribs. Distal raphe endings exhibit a slight ventral deflection and terminate with poorly developed helictoglossae. Proximally, the raphe concludes with fused central helictoglossae (Yilmaz et al. in press).

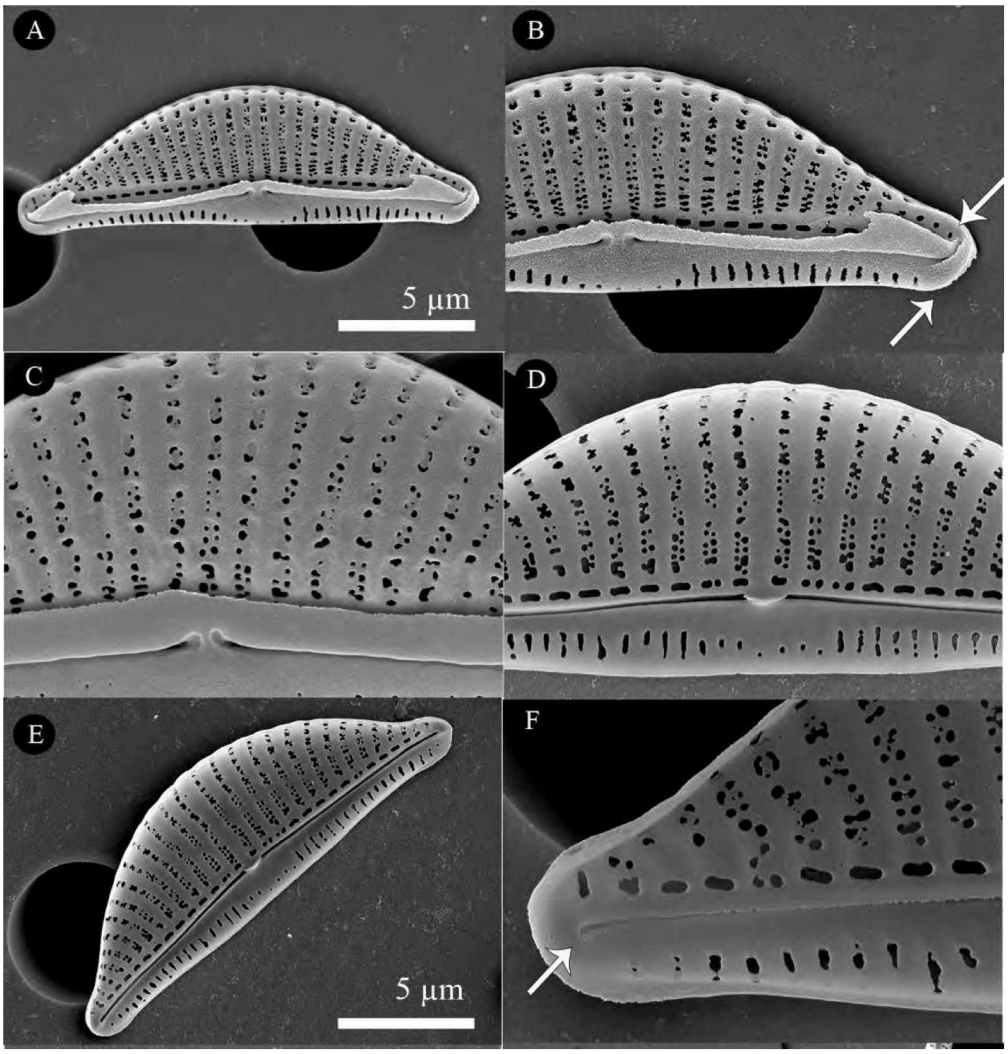


Figure 17. ***A-F****. SEM micrographs of Halamphora witkowskii sp. nov.* ***A****. External view of the entire valve.* ***B****. External detail, showing the expanded central area on the ventral side and dorsal raphe ledge (arrow).* ***C****. External valve view showing central raphe endings.* ***D****. Internal valve view showing central nodule.* ***E****. Internal view of the entire valve.* ***F****. Detail of internal valve apex showing distal raphe ending (helictoglossa: arrow). Scale bars = (A, E) 5 µm, (B–D, F) 2 µm.*

Molecular phylogenetic analysis

In the phylogenetic analysis, both strains of *Halamphora witkowskii* (SZCZEY 2176 and 2177) were identified within a monophyletic group (bootstrap value = 100), situated within a larger *Halamphora* clade encompassing 54 distinct sequences from 45 species.

H. witkowskii was found to be closely related to a clade consisting of *H. incelebrata*, *H. intramaritima, H. calidilacuna*, *H. americana*, *H. siqueirosii,* and *H. bonnevillensis*. The high bootstrap value (>92) for this clade suggests a stable tree topology, reinforcing the distinctiveness of the newly identified species (Figure 18).

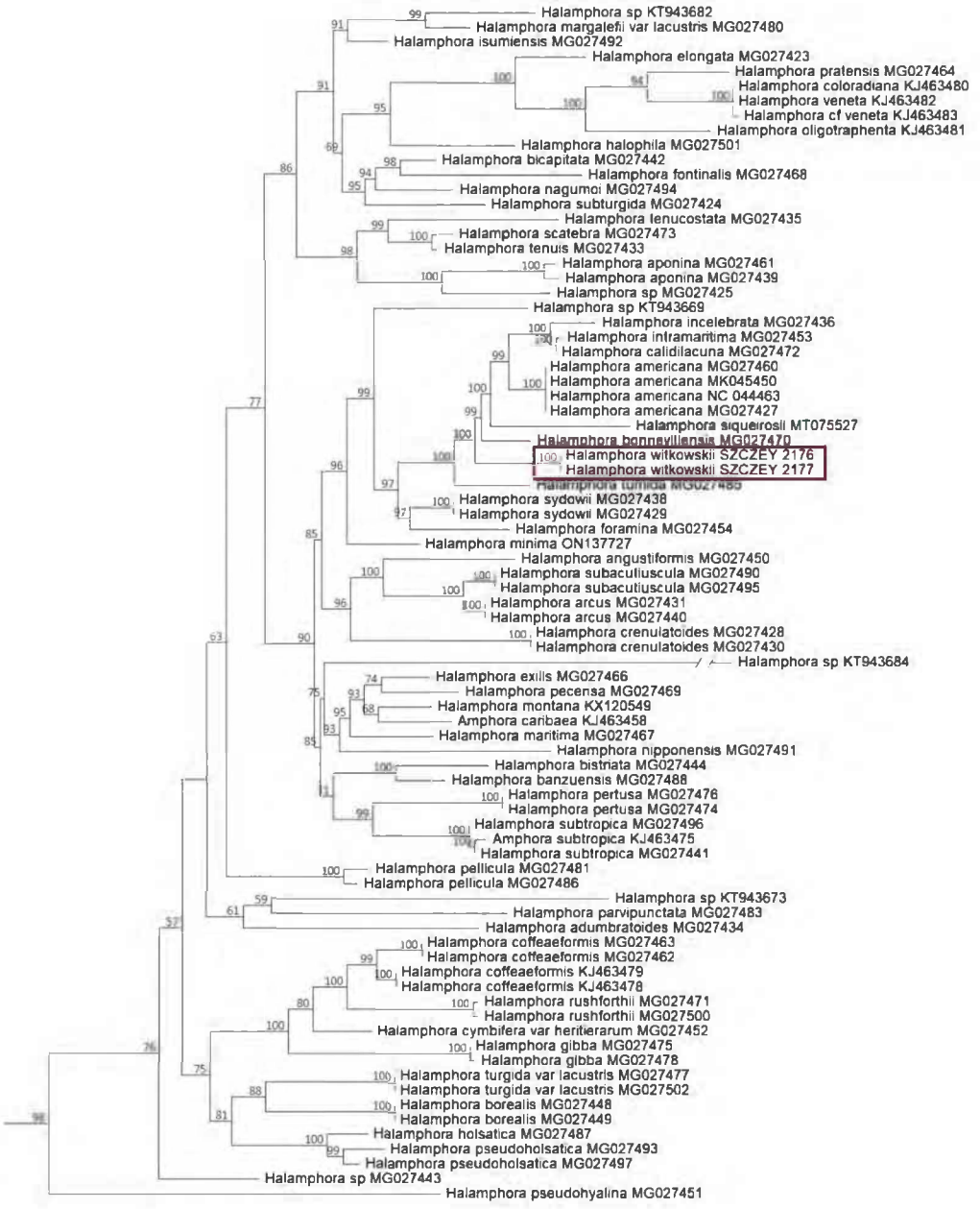


Figure 18. : The clade of Halamphora spp. cut from Maximum Likelihood tree inferred from the alignment of 260 rbcL sequences of diatoms and rooted with 2 Triparma pacifica sequences. Rectangle indicates the position of two strains of Halamphora witkowskii.

Halamphora vantushpaensis sp. nov. Yilmaz, Solak, & Gastineau, sp. nov.

This last chapter is dedicated to *Halamphora vantushpaensis* sp. nov. Yilmaz, Solak, &Gastineau, sp. nov., a second new species from the genus *Halamphora*. This article is currently in press in Phytokeys (online ISSN 1314-2003).

Etymology: The species is named with regard to both Lake Van and the city of Tushpa, capital of the Iron Age kingdom of Urartu, which was located in the vicinity of the lake.

Taxonomic description:

Light Microscopy (Figure 19A-M): Valves semi-lanceolate, dorsiventral with arched dorsal margin and slightly tumid ventral margin. Valve ends protracted, capitate in larger specimens; slightly sub-protracted, not separated from the rest of the valve in smaller specimens; and ventrally bent. Valve length 24.0–42.0 µm, valve width 4.0–5.0 µm (n = 35). Axial area very narrow, wider on ventral valve side. Central area visible in larger specimens; indistinct on dorsal side, semi-lanceolate on ventral side. Raphe ledge almost straight, slightly arched, located near median line of valve. Proximal raphe endings slightly dorsally bent. Striae hard to resolve in LM, dorsally slightly radiate entire valve, 27–32 in 10 µm.

As with the two sequenced isolates of *Navicula vanseea*, the two *Halamphora vantushpaensis* isolates differed considerably in their dimensions and also in shape, the larger-celled isolate having a much more elaborate outline. If the isolates had been seen in isolation and had not been fully characterized with respect to their chloroplast genomes (see below), it might have been thought that they represented different species. However, the differences observed are, as with *N. vanseea*, fully consistent with the changes observed in many other pennate diatoms, in which length changes much more than width during the life cycle, stria density remains little altered, and the valve outline becomes much simpler as the cells get smaller (Geitler 1932, Round et al. 1990).

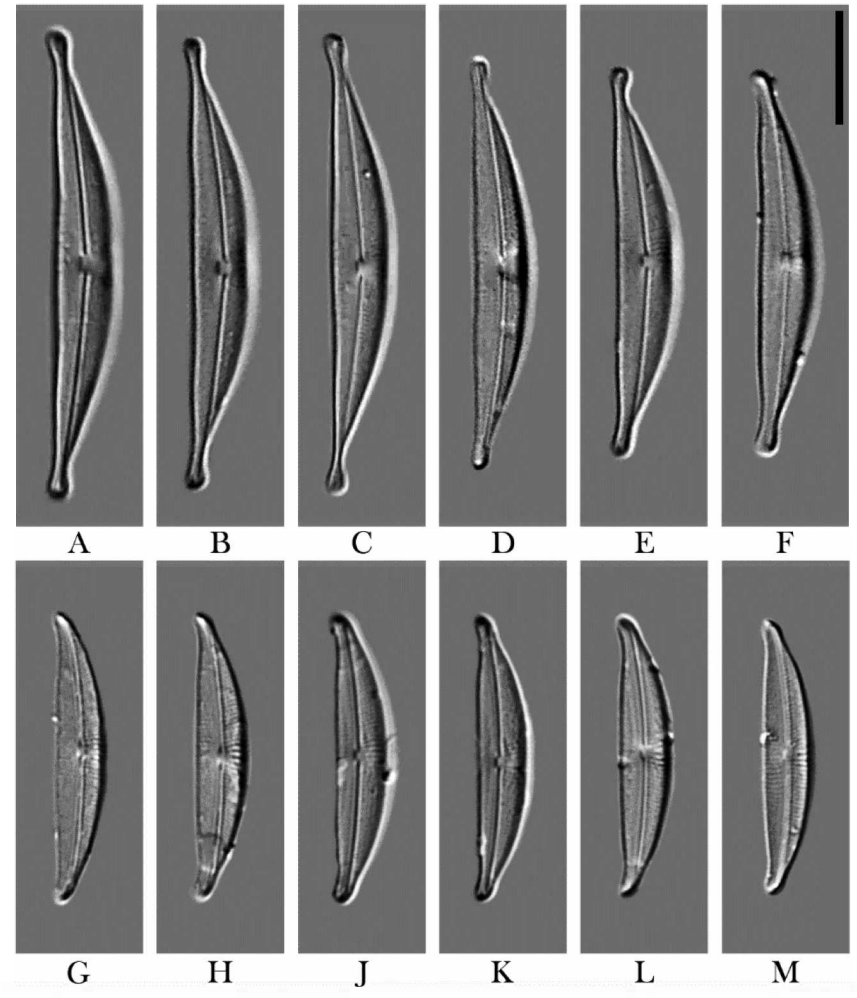


Figure 19. ***A-M****. Halamphora vantushpaensis sp. nov. LM micrographs.* ***A-F****. The images of larger specimens cleaned valves from the culture isolated from Van Lake.* ***G-M****. The images of smaller specimens cleaned valves from the culture isolated from Van Lake. Scale bar = 10 µm.*

Scanning Electron Microscopy (Figure 20A-F): Externally, valves dorsiventral with convex dorsal margin and straight to slightly tumid ventral margin. Raphe ledge narrow and linear, present on dorsal side of valve. Raphe branches slightly arched. Proximal raphe endings slightly expanded into central depressions dorsally deflected with small pores within depression. Distal raphe endings dorsally deflected. Central area is distinct on dorsal side in larger specimens; dorsally absent in smaller specimens and ventrally elliptic large hyaline area in larger specimens; semi-lanceolate with regular shortening of several central striae on ventral side of smaller specimens. Internally, central area symmetrically visible and large hyaline area on dorsal side of larger specimens; very small on both sides in smaller specimens. Proximally, raphe terminates with fused central

helictoglossae. Distal raphe endings deflected ventrally and terminate internally with developed helictoglossae. Striae uniseriate, composed of areolae with elliptical internal openings.(Yilmaz et al. in press).

There was full agreement in valve ultrastructure between the two isolates, supporting the interpretation given above (Light Microscopy), that the two isolates represent different stages in the size-reduction cycle of a single species.

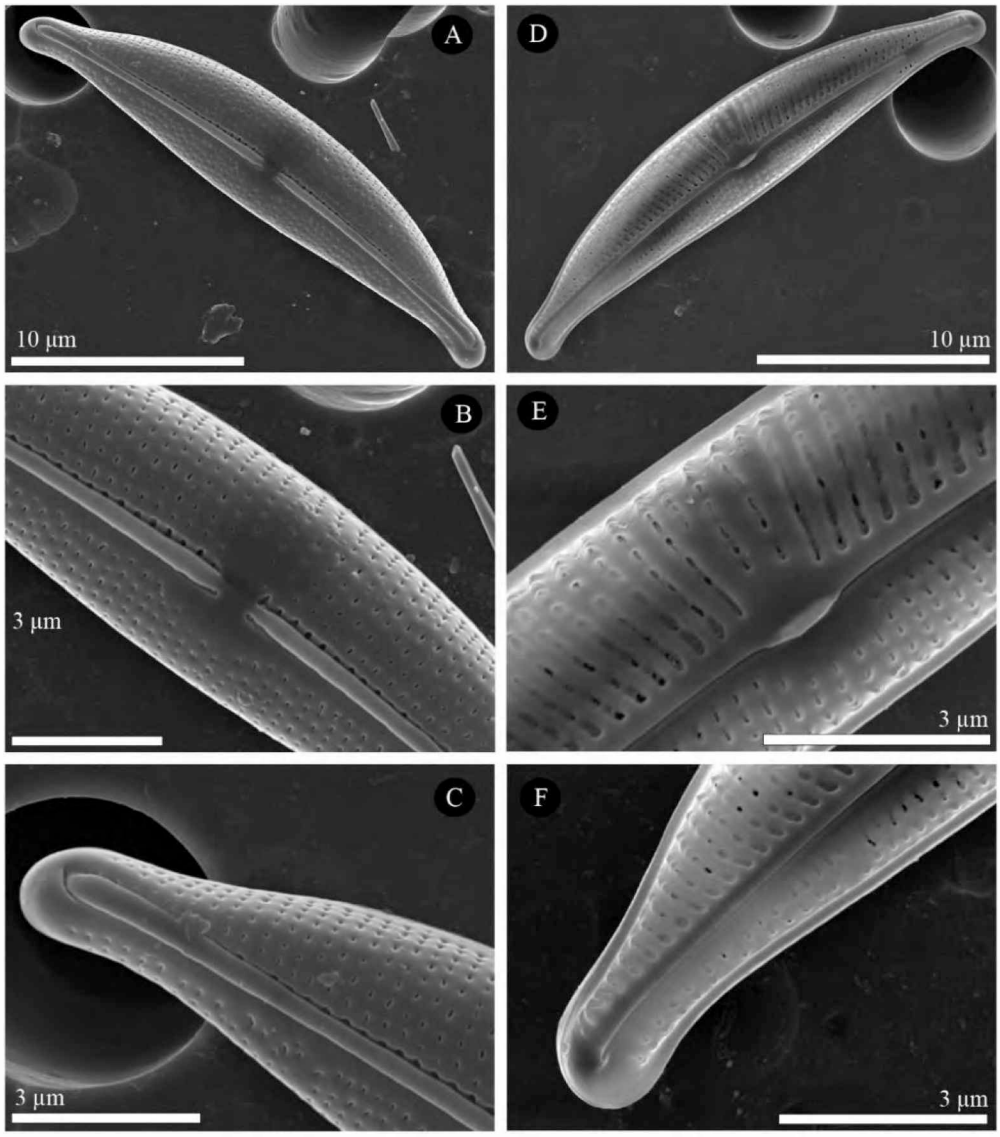


Figure 20. ***A-F****. Halamphora vantushpaensis sp. nov. SEM micrographs of smaller specimens.* ***A****. External view of the entire valve.* ***B****. Details of central area showing simple proximal raphe endings and regular shortened striae.* ***C****. Details of apex showing the terminal fissure.* ***D****. Internal view of the entire valve.* ***E****. Details of central area showing long fused central helictoglossa between the proximal raphe endings.* ***F****. Details of apex showing well-developed helictoglossa. Scale bars = (A, D) 10 µm, (B, C, E, F) 3 µm*

Genomics and phylogeny

Plastid genome and comparison with Halamphora spp.

The plastid genome of *H. vantushpaensis* is 133,852 bp long in SZCZEY2167 and 133,866 bp long in SZCZEY2166. The only difference between them was spotted in intergenic sequences, and for this reason only SZCZEY2167 has been deposited on GenBank (GenBank:OR797294,OR797293). The large single copy (LSC) is 61,691 bp long and has 70 conserved protein coding genes (PCGs) as well as a single non-conserved open reading frame (ORF) and 17 tRNA. The short single copy (SSC) is 39,615 bp long, has 46 PCGs, also a single non conserved ORF and six tRNA. The inverted repeat IR is 16,273 bp long and contains 10 PCGs, three rRNA and four tRNA (2).

Three plastid genomes are available in GenBank for the genus *Halamphora*, all originating from the same study (Hamsher et al. 2019). They represent *Halamphora americana* Kociolek, 2014, *Halamphora calidilacuna* J.G.Stepanek & Kociolek, 2018 and *H. coffeiformis* (C.Agardh) Levkov 2009. In Table 1, the total lengths of these genomes and the lengths of their different compartments are compared.

The complete three-gene tree and the *rbcL*-only tree can be found as indicated in the data availability statement (Figure 21). In the three-gene tree, *H. vantushpaensis* strains appear as a highly supported (99%) long-branched sister group to a larger cluster composed of 18 *Halamphora* species, namely *H. subacutiuscula*, *H. angistiformis*, *H. foramina*, *H. sydowii*, *H. tumida*, *H. witkowskii*, *H. bonnewillensis*, *H. americana*, *H. calidilacuna*, *H. intramaritima*, *H. incelebrata*, *H. banzuensis*, *H. bistriata*, *H. pertusa*, *H. subtropica* plus three unidentified *Halamphora* species.

Table 1. Length of the different compartments of the plastid genomes of four Halamphora species

Species	**Length of the LSC**	**Length of the SSC**	**Length of the IR**	**Total length**
Halamphora calidilacuna	82,227 bp	49,698 bp	9,407 bp	150,739 bp
Halamphora americana	77,289 bp	44,724 bp	10,269 bp	142,551 bp
Halamphora coffeiformis	64,938 bp	41,485 bp	7,752 bp	121,927 bp
Halamphora vantushpaensis	61,691 bp	39,615	16,273 bp	133,852

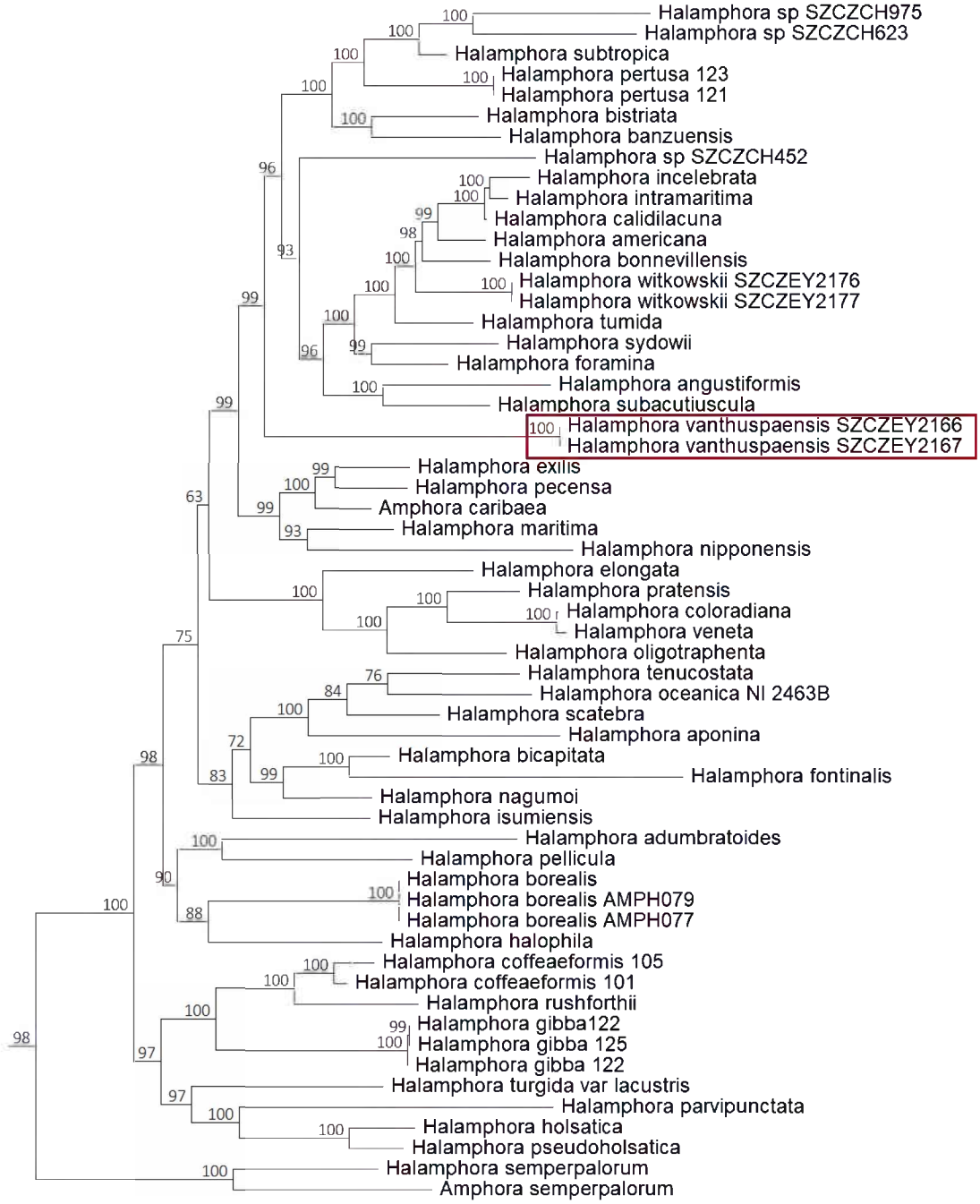


Figure 21. A cut from a Maximum Likelihood phylogenetic tree inferred from concatenated alignments of psbC, rbcL and 18S.

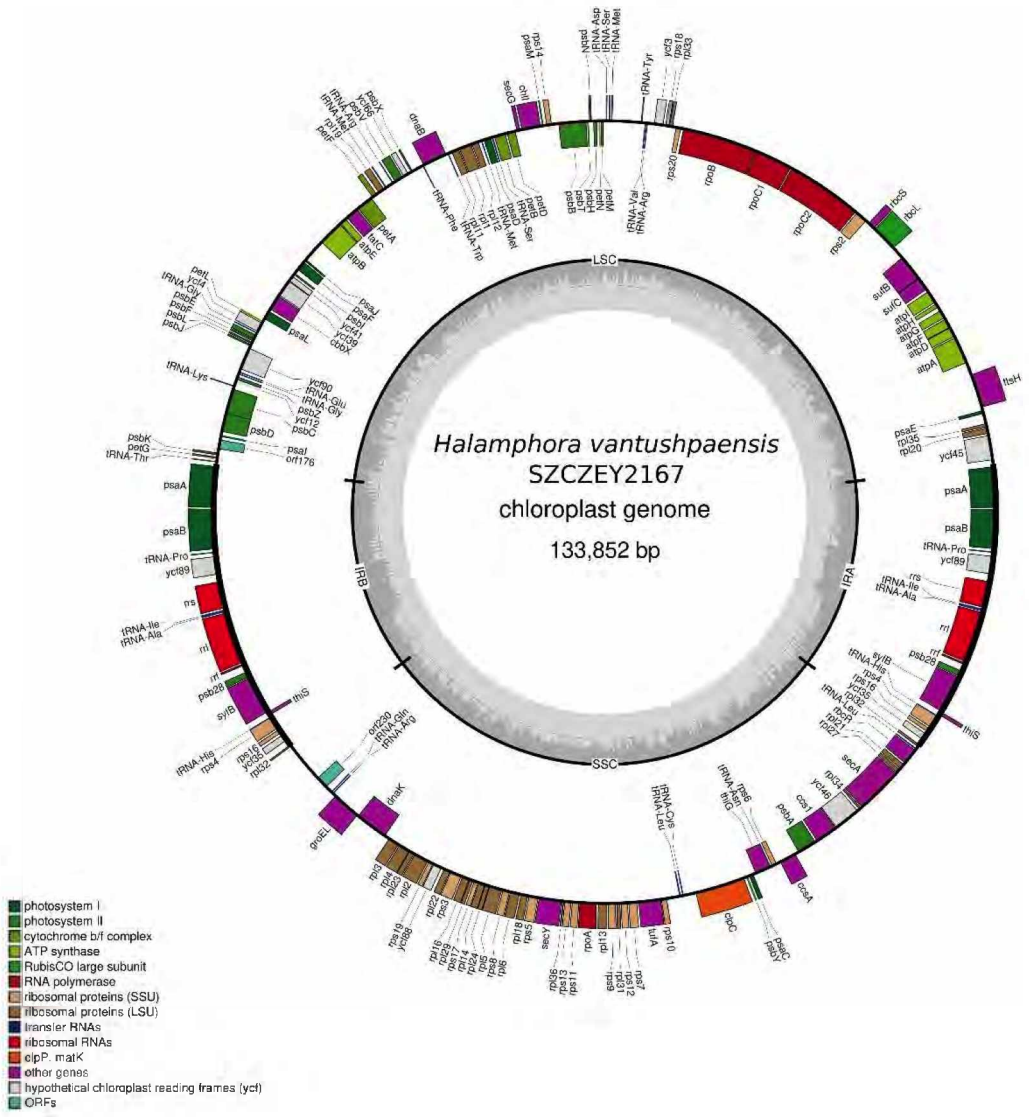


Figure 2. Map of the plastid genome of Halamphora vantushpaensis sp. nov. SZCZEY2167.

Halamphora vantushpaensis has shorter LSC and SSC but its IR is consistently longer when compared to other species. The gene content of the IR is compared for all these species in Table 1. The restricted set of conserved genes found among *H. calidilacuna* or *Halamphora americana* and which consists of a single PCG (*ycf89*), three tRNA and three rRNA seems to be shared by many unrelated species and genera such as *Navicula veneta* Kützing, 1844 or *Tryblionella apiculata* W.Gregory, 1857 (Gastineau et al., 2021a). As with *H. americana*, an extension of the IR may result from the presence of non-conserved ORF or putative genes of plasmid origin, as exemplified by its ORF9 and the putative integrase/recombinase encoded by the gene labelled as *tyrC* by Hamsher et al. (2019).

The case of *H. vantushpaensis* is entirely different in the sense that the extension of the IR is a consequence of the incorporation of several conserved PCGs plus one tRNA. When compared with the gene content of the other species, it appears that this extension has been done at the expense of both the LSC and the SSC, which distinguishes it from species such as *Climaconeis* spp. (Gastineau et al., 2021b) among which the IR seemed to have only taken over the SSC. Indeed, among the other *Halamphora* spp., *psaA* and *psaB* are in the LSC while the other genes belong to the SSC in which they form a well conserved cluster.

SUMMARY AND CONCLUSIONS

The aim of this thesis was to explore the diatom assemblages of Lake Van, a highly alkaline and extreme environment (a soda lake), with modern research tools. Soda lakes are known for their harsh conditions, making the diversity of the organisms inhabiting them and the putative underlying ecological adaptations a significant research focus. By conducting both fieldwork and laboratory analyses, I made an attempt to contribute with new elements regarding these topics. Ecological studies will subsequently provide insights into the relationship of the alkaliphilic diatoms with environmental factors, population dynamics, and the roles of these diatoms in the ecosystem, but first the species present need to be described and characterized.

First of all, it has to be stated that the methodology employed resulted in the isolation and cultivation of several diatom strains from Lake Van. As obvious as this statement might be, it should be reminded that with such an extreme environment, this step constituted a milestone. The protocol employed led to the description of several new species, in addition to a few species that had already been known, but which were previously not reported or accurately identified by Legler & Krasske (1940).

In view of the results obtained, the two hypotheses are confirmed. Indeed, several taxa were not reported by Legler & Krasske (1940). As illustrated by the three articles this thesis is based on, Lake Van has a high potential for the discovery of new species, with putatively several more to investigate (Hypothesis 1). The protocol employed, which associates precise microscopical documentation with molecular methods, proved suitable to conduct these investigations (Hypothesis 2). Moreover, as suggested by Solak et al. (2021) and illustrated here, especially by *N. vanseea* sp. nov., these methods might be absolutely necessary to discriminate between Lake Van taxa and other morphologically similar known species. In future it may be worth considering the use of a combination of morphological studies and metabarcoding methods (e.g. Rimet et al. 2023) to survey the diatom diversity in detail, including for cryptic species. Furthermore, characterization of the whole ribosomal gene cluster, as done here in *Navicula vanseea*, will aid the development of long-read metabarcoding of the whole ribosomal operon (Jamy et al. 2020, 2023), which is likely to become an important way to simultaneously investigate diatom communities and diatom phylogeny.

Taxonomic investigations involved identifying and classifying diatom species in the alkaline lake, while molecular phylogenetic analyses aimed at exploring evolutionary relationships and genetic diversity among these species. I hope that the results of this study will contribute to a better understanding of the biodiversity and complex structure of diatom assemblages in alkaline lakes. Additionally, this research might provide additional data for a better understanding of the evolutionary and ecological processes of organisms adapted to extreme conditions.

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTARY FILES

REFERENCES

Ağlaç, E., & Balkıs, N. (2014). Investigation of the seasonal variation in the phytoplankton in the surface waters of the Gulf of Edremit. *European Journal of Biology*, *73*(2), 31-46.

Akkuş, M., Sarı, M., Ekmekçi, F., & Yoğurtçuoğlu, B. (2021). The discovery of a microbialite-associated freshwater fish in the world's largest saline soda lake, Lake Van (Turkey). *Zoosystematics and Evolution*, *97*(1), 181-189.

Al-Yamani, F. Y., & Saburova, M. A. (2011). Illustrated guide on the benthic diatoms of Kuwait's Marine Environment. Primera ed. Kuwait Institute for Scientific Research, Kuwait. 364.

Anderson, N. J., Rippey, B., & Stevenson, A. C. (1990). Change to a diatom assemblage in a eutrophic lake following point source nutrient re-direction: a palaeolimnological approach. *Freshwater Biology*, *23*(2), 205-217.

Arslan, N., Timm, T., Rojo, V., Vizcaíno, A., Schmelz R. M. (2018). A new species of Enchytraeus (Enchytraeidae, Oligochaeta) from the profundal of Lake Van, the world's largest soda Lake (Turkey, East Anatolia). *Zootaxa*, *4382*(2), 367-380.

Ashworth, M. P., Nakov, T., & Theriot, E. C. (2013). Revisiting Ross and Sims (1971): toward a molecular phylogeny of the Biddulphiaceae and Eupodiscaceae (Bacillariophyceae). *Journal of Phycology*, *49*(6), 1207-1222.

Ashworth, M. P., Majewska, R., Frankovich, T. A., Sullivan, M., Bosak, S., Filek, K., Van de Vijver, B., Arendt, M., Schwenter, J., Nel, R., Robinson, N. J., Gary, M. P., Theriot, E. C., Stacy, N. I., Lam, D. W., Perrault, J. R., Manire, C. A. & Manning, S. R. (2022). Cultivating epizoic diatoms provides insights into the evolution and ecology of both epibionts and hosts. *Scientific Reports 12*, 15116.

Bąk, M., Witkowski, A., Żelazna-Wieczorek, J., Wojtal, A. Z., Szczepocka, E., Szulc, K., & Szulc, B. (2012). Klucz do oznaczania okrzemek w fitobentosie na potrzeby oceny stanu ekologicznego wód powierzchniowych w Polsce. *Biblioteka Monitoringu Środowiska, Warszawa*, 7-452.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... & Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, *19*(5), 455-477.

B-Béres, V., Stenger-Kovács, C., Buczkó, K., Padisák, J., Selmeczy, G. B., Lengyel, E., & Tapolczai, K. (2023). Ecosystem services provided by freshwater and marine diatoms. *Hydrobiologia*, *850*(12), 2707-2733.Benoiston, A. S., Ibarbalz, F. M., Bittner, L., Guidi, L., Jahn, O., Dutkiewicz, S., & Bowler, C. (2017). The evolution of diatoms and their biogeochemical functions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1728), 20160397.

Bory de Saint-Vincent, M. (1822). Dictionnaire classique d'histoire naturelle.

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972-1973.

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

Cox, E. J. (1979). Studies on the diatom genus *Navicula* Bory. *Navicula scopulorum* Bréb. and a further comment on the genus *Berkeleya* Grev. *British Phycological Journal*, *14*(2), 161-174.

Cox, T. E., Cebrian, J., Tabor, M., West, L., & Krause, J. W. (2020). Do diatoms dominate benthic production in shallow systems? A case study from a mixed seagrass bed. *Limnology and Oceanography Letters*, *5*(6), 425-434.

Dąbek, P., Ashworth, M. P., Witkowski, A., Li, C., Bornman, T. G., Gonçalves, V., ... & Khim, J. S. (2017). Towards a multigene phylogeny of the Cymatosiraceae (Bacillariophyta, Mediophyceae) I: novel taxa within the subfamily Cymatosiroideae based on molecular and morphological data. *Journal of Phycology*, *53*(2), 342-360.

Danulat, E., & Kempe, S. (1992). Nitrogenous waste excretion and accumulation of urea and ammonia in *Chalcalburnus tarichi* (Cyprinidae), endemic to the extremely alkaline Lake Van (Eastern Turkey). *Fish Physiology and Biochemistry*, 9(5), 377-386.

Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Molecular biology and evolution*, *37*(1), 291-294.

Degens, E. T., Wong, H. K., Kurtman, F., Finkch, P., Degens, E. T., & Kurtman, F. F. (1978). Geological development of Lake Van. The Geology of Lake Van, 134-146.

Dixit, S. S., Smol, J. P., Kingston, J. C., & Charles, D. F. (1992). Diatoms: powerful indicators of environmental change. *Environmental Science & Technology*, *26*(1), 22-33.

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus (San Francisco, Calif.) *12*, 13–15.

Erk'akan, F., & Kuru, M. (1986). A new Noemacheilinae loach subspecies from Turkey (Osteichthyes, Cobitidae). *Turkish Journal of Biology*, *10*(1), 106-109.

Gagnon, A. G. (Ed.). (2004). *Québec: State and society* (pp. 427-446). Peterborough: Broadview Press.

Gastineau R, Hamedi C, Baba Hamed MB, Abi-Ayad S-ME-A, Bąk M, Lemieux C, Turmel M, Dobosz S, Wróbel RJ, Kierzek A, Lange-Bertalot H, Witkowski A (2021a) Morphological and molecular identification reveals that waters from an isolated oasis in Tamanrasset (extreme South of Algerian Sahara) are colonized by opportunistic and pollution-tolerant diatom species. *Ecological Indicators 121*(6): 107104. 0

Gastineau, R., Davidovich, N. A., Davidovich, O. I., Lemieux, C., Turmel, M., Wróbel, R. J., & Witkowski, A. (2021b). Extreme enlargement of the inverted repeat region in the plastid genomes of diatoms from the genus *Climaconeis*. *International Journal of Molecular Sciences*, *22*(13), 7155.

Geitler, L. (1932). Der Formwechsel der pennaten Diatomeen. *Archiv für Protistenkunde 78*, 1-226.

Górecka, E., Dąbek, P., Davidovich, N., Davidovich, O., Tremblay, R., Belzile, C., ... & Witkowski, A. (2022). Life history of the diatom *Schizostauron trachyderma*: Cell size and lipid accumulation. *Frontiers in Marine Science*, *8*, 793665.

Gorecka, E. (2022). *Taxonomy and molecular phylogeny of monoraphid marine benthic diatoms with an account on their oil production potential*. PhD dissertation. University of Szczecin, Institute of Marine and Environmental Sciences, Szczecin, 93pp.

Greiner, S., Lehwark, P., & Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3. 1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic acids research*, *47*(W1), W59-W64.

Guillard, R. R. L. (1975). Culture of phytoplankton for feeding marine invertebrates, in *Culture of Marine Invertebrate Animals*, edited by W. L. Smith and M. H. Chanley, pp. 26–60, Plenum Press, New York.

Guiry, M. D., & Guiry, G. M. (2024). *AlgaeBase*. World-wide electronic publication, National University of Ireland, Galway. https://www.algaebase.org; searched on 21 March.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, *41*, 95.

Hamsher, S. E., Keepers, K. G., Pogoda, C. S., Stepanek, J. G., Kane, N. C., & Kociolek, J. P. (2019). Extensive chloroplast genome rearrangement amongst three closely related Halamphora spp. (Bacillariophyceae), and evidence for rapid evolution as compared to land plants. *PloS One*, *14*(7), e0217824.

Hendey, N. I. (1964). *An Introductory Account of the Smaller Algae of British Coastal Waters. Part V: Bacillariophyceae (Diatoms).* Fishery Investigations, Series IV. London, England, Her Majesty's Stationery Office, 317.

Kaleli, M. A. (2018). *Caretta caretta* (Linneaus, 1758) karapaksı üzerindeki epibiont diyatomelerin taksonomik ve biyocoğrafik açıdan incelenmesi. PhD dissertation. University of Dumlupinar, Institute of Sciences, Kutahya, 293pp.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, *30*(4), 772-780.

Kempe, S., Kazmierczak, J., Landmann, G., Konuk, T., Reimer, A., & Lipp, A. (1991). Largest known microbialites discovered in Lake Van, Turkey. *Nature, 349*(6310), 605-608.

Kempe, S., & Degens, E. T. (1978). Lake Van varve record: the past 10,420 years. In Degens, E.T., and Kurtman, F. (Eds.), The Geology of Lake Van: The Mineral Research and Exploration Institute of Turkey, Ankara pp. 56–63 (MTA Press)

Kociolek, J. P. (2007). Diatoms: Unique eukaryotic extremophiles providing insights into planetary change. In *Instruments, Methods, and Missions for Astrobiology X* (Vol. 6694, pp. 242-256). SPIE.

Landmann, G., Reimer, A., & Kempe, S. (1996). Climatically induced lake level changes at Lake Van, Turkey, during the Pleistocene/Holocene transition. *Global Biogeochemical Cycles*, *10*(4), 797-808.

Lange-Bertalot, H., Külbs, K., Lauser, T., Nörpel-Schempp, M. & Willmann, M. (1996). Diatom taxa introduced by Georg Krasske. Documentation and revision. Dokumnetation und Revision der von Georg Krasske beschribenen Diatomeen-Taxa. Iconographia Diatomologica 3: 2-358.

Lauriol, B., Prévost, C., & Lacelle, D. (2006). The distribution of diatom flora in ice caves of the northern Yukon Territory, Canada: relationship to air circulation and freezing. *International Journal of Speleology*, *35*(2), 4.

Legler, F., & Krasske, G. (1940). Diatomeen aus dem Vansee (Armenien). Beiträge zur Ökologie der Brackwasserdiatomeen I. *Beihefte zum Botanischen Centralblatt, 60*, 335-345.

Levkov, Z. (2009) *Diatoms of Europe: Diatoms of the European inland waters and comparable habitats. vol. 5*. Amphora sensu lato. Ruggell: A.R.G. Gantner Verlag.

Litt, T., Krastel, S., Sturm, M., Kipfer, R., Örcen, S., Heumann, G., ... & Niessen, F. (2009). 'PALEOVAN', International Continental Scientific Drilling Program (ICDP): site survey results and perspectives. *Quaternary Science Reviews*, *28*(15-16), 1555-1567.

López-García, P., Kazmierczak, J., Benzerara, K., Kempe, S., Guyot, F., & Moreira, D. (2005). Bacterial diversity and carbonate precipitation in the giant microbialites from the highly alkaline Lake Van, Turkey. *Extremophiles*, *9*, 263-274.

Madden, T. (2013). The BLAST sequence analysis tool. *The NCBI handbook*, *2*(5), 425-436.

Mann, D. G., & Vanormelingen, P. (2013). An inordinate fondness? The number, distributions, and origins of diatom species. *Journal of eukaryotic microbiology*, *60*(4), 414-420.

Marchese, C. (2015). Biodiversity hotspots: A shortcut for a more complicated concept. *Global Ecology and Conservation*, *3*, 297-309.

Millie, D. F., & Wee, J. L. (1981). Observations on the tube-dwelling diatom Navicula tripunctata var. schizonemoides (VH) Patr. using light and electron microscopy. In *Proceedings of the Iowa Academy of Science, 88*(2), 74-78).

Minh, B. Q., Hahn, M. W., & Lanfear, R. (2020). New methods to calculate concordance factors for phylogenomic datasets. *Molecular Biology and Evolution*, *37*(9), 2727-2733.

Patrick, R. (1959). New species and nomenclatural changes in the genus Navicula (Bacillariophyceae). *Proceedings of the Academy of Natural Sciences of Philadelphia*, *111*, 91-108.

Potapova, M. G. (2012). New species and combinations in monoraphid diatoms (family Achnanthidiaceae) from North America. *Diatom Research*, *27*(1), 29-42.

Poulin, M. (2002). Lectotypification of previously described diatoms from marine coastal habitats (Quebec, Canada). *Diatom Research, 17*, 465-466.

Riaux-Gobin, C., & Poulin, M. (2004a). Possible symbiosis of *Berkeleya adeliensis* Medlin, *Synedropsis fragilis* (Manguin) Hasle et al. and *Nitzschia lecointei* Van-Heurck (Bacillariophyta) associated with land-fast ice in Adelie Land, Antarctica. *Diatom Research, 19*, 265-274.

Riaux-Gobin, C., Compère, P., Coste, M., Straub, F., & Taxböck, L. (2014). Cocconeis napukensis sp. nov.(Bacillariophyceaea) from Napuka Atoll (South Pacific) and lectotypification of Cocconeis subtilissima Meister. *Fottea*, *14*(2), 209-224.

Rimet, F., Pinseel, E., Bouchez, A., Japoshvili, B., & Mumladze, L. (2023). Diatom endemism and taxonomic turnover: Assessment in high-altitude alpine lakes covering a large geographical range. *Science of the Total Environment*, *871*, 161970.

Ross, R., Cox, E. J., Karayeva, N. I., Mann, D. G., Paddock, T. B. B., Simonsen, R., & Sims, P. A. (1979). An amended terminology for the siliceous components of the diatom cell. *Nova Hedwigia, Beiheft 64*: 513–533

Round, F. E. (1990). *The Diatoms: Biology and Morphology of the Genera*. Cambridge University Press. 747 pp.

Sabbe, K., Witkowski, A., & Vyverman, W. (1995). Taxonomy, morphology and ecology of *Biremis lucens* comb. nov. (Bacillariophyta): a brackish-marine, benthic diatom species comprising different morphological types. *Botanica Marina*, *38*(5), 379-391.

Sala, S. E., Guerrero, J. M., & Coste, M. (2007). Valve morphology of *Amphora chilensis* Hustedt (Bacillariophyceae). *Nova Hedwigia 85*, 353–364.

Scala, S., & Bowler, C. (2001). Molecular insights into the novel aspects of diatom biology. *Cellular and Molecular Life Sciences CMLS*, *58*, 1666-1673.

Seckbach, J., & Kociolek, P. (Eds.). (2011). *The diatom world* (Vol. 19). Springer Science & Business Media.

Shraoui, I., Hlaili, A. S., Mabrouk, H. H., Léger, C., & Bates, S. S. (2009). Blooms of diatom genus *Pseudo-nitzschia* H. Peragallo in Bizerte Lagoon (Tunusia, SW Mediterranean). *Diatom Research, 24*, 175-190.

Smith, S. A., & Dunn, C. W. (2008). Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics*, *24*(5), 715-716.

Sims, P. A., Mann, D. G., & Medlin, L. K. (2006). Evolution of the diatoms: insights from fossil, biological and molecular data. *Phycologia*, *45*(4), 361-402.

Smol, J. P., & Douglas, M. S. (1996). Long-term environmental monitoring in arctic lakes and ponds using diatoms and other biological indicators. *Geoscience Canada* 23, 225–230.

Smol, J. P., & Stoermer, E. F. (Eds.). (2010). *The diatoms: applications for the environmental and earth sciences*. Cambridge University Press.

Snoeijs, P. (1993). Intercalibration and distribution of diatom species in the Baltic Sea. *The Baltic Marine Biologists Publication No. 16a*, 1-129.

Solak, C. N., Gastineau, R., Lemieux, C., Turmel, M., Gorecka, E., Trobajo, R., Rybak, M., Yilmaz, E., & Witkowski, A. (2021). *Nitzschia anatoliensis* sp. nov., a cryptic diatom species from the highly alkaline Van Lake (Turkey*). PeerJ*, 9, e12220.

Stepanek, J. G., & Kociolek, J. P. (2013). Several new species of *Amphora* and *Halamphora* from the western USA. *Diatom Research, 28,* 61–76.

Stepanek, J.G. & Kociolek, J.P. (2018). *Amphora* and *Halamphora* from coastal and inland waters of the United States and Japan, with the description of 33 new species. *Bibliotheca Diatomologica* 66: 1–260.

Stepanek, J. G., & Kociolek, J. P. (2019). Molecular phylogeny of the diatom genera *Amphora* and *Halamphora* (Bacillariophyta) with a focus on morphological and ecological evolution. *Journal of Phycology, 55*, 442–456.

Stevenson, R. J., & Bahls, L. L. (1999). Periphyton protocols. In Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling (eds) *Rapid bioassessment protocols for use in streams and wadeable rivers: Periphyton, benthic macroinvertebrates and fish,* Second Edition, Chapter 6. EPA 841-B-99-002, U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

Stidolph, S. R., Sterrenburg, F. A. S., Smith, K. E. L., & Kraberg, A. (2012). Stuart R. Stidolph Diatom Atlas. *U.S. Geological Survey Open-File Report* 2012-1163, 199.

Stoermer, E. F., & Smol, J. P. (1999). Applications and uses of diatoms: prologue. *The diatoms: applications for the environmental and earth sciences*, pp. 3-10. Cambridge University Press.

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, *38*(7), 3022-3027.

Taş, S. (2014). Phytoplankton composition and abundance in the coastal waters of the Datça and Bozburun Peninsulas, south-eastern Aegean Sea (Turkey). *Mediterranean Marine Science*, *15*, 84-94.

Taş, S., Dursun, F., Aksu, A., & Balkıs, N. (2016). Presence of the diatom genus *Pseudonitzschia* and particulate domoic acid in the Golden Horn Estuary (Sea of Marmara, Turkey). *Diatom Research, 31*, 339-349.

Taş, S., & Hernández-Becceril, D. U. (2017). Diversity and distribution of the planktonic diatom genus *Chaetoceros* (Bacillariophyceae) in the Golden Horn Estuary (Sea of Marmara). *Diatom Research, 32*, 309-323.

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

Van de Vijver, B., & Cocquyt, C. (2009). Four new diatom species from La Calera hot spring in the Peruvian Andes (Colca Canyon). *Diatom Research*, *24*(1), 209-223.

Van de Vijver, B., Kopalová, K., Zidarova, R., & Levkov, Z. (2014). Revision of the genus *Halamphora* (Bacillariophyta) in the Antarctic Region. *Plant Ecology and Evolution, 147*, 374–391.

Woodard, K., Kulichová, J., Poláčková, T., & Neustupa, J. (2016). Morphometric allometry of representatives of three naviculoid genera throughout their life cycle. *Diatom Research*, *31*, 231–242.

Yılmaz, E., Mann, D. G., Gastineau, R., Trobajo, R., Solak, C. N., Górecka, E., Turmel, M., Lemieux, C, Ertorun, N., & Witkowski, A. (2024). Description of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny. *PhytoKeys*, *241*, 27.

Yilmaz, E., Gastineau, R., Górecka, E., Solak, C. N., Trobajo, R., Peszek, Ł., Mann, D. G. (in press) *Halamphora witkowskii* sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye. *Nova Hedwigia*.

Yilmaz E, Gastineau R., Solak, C. N., Górecka, E., Trobajo, R., Turmel, M., Lemieux, C., Otis C., Witkowski A., Mann, D. G. (in press) Morphological and molecular characterization of *Halamphora vantushpaensis* (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye). *PhytoKeys*,

You, Q. M., Kociolek, J. P., & Wang, Q. X. (2015). Taxonomic studies of the diatom genus *Halamphora* (Bacillariophyceae) from the mountainous regions of southwest of China, including the description of two new species. *Phytotaxa 205*, 75–89.

Witkowski, A., Li, C., Zgłobicka, I., Yu, S. X., Ashworth, M., Dąbek, P., ... & Hajrah, N. H. (2016). Multigene assessment of biodiversity of diatom (Bacillariophyceae) assemblages from the littoral zone of the Bohai and Yellow Seas in Yantai Region of Northeast China with some remarks on ubiquitous taxa. *Journal of Coastal Research*, *74,* 166-195.

Woodard, K., Kulichová, J., Poláčková, T., & Neusputa, J. (2016). Morphometric allometry of representatives of three naviculoid genera throughout their life cycle. *Diatom Research, 31*, 231-242.

SUPPLEMENTARY DATA

Supplementary File S 1. Yılmaz, E., Mann, D. G., Gastineau, R., Trobajo, R., Solak, C. N., Górecka, E., ... & Witkowski, A. (2024). Description of Naviculavanseea sp. nov.(Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny. *PhytoKeys*, 241, 27. doi: 10.3897/phytokeys.241.118903

Impact Factor: 1.4

Lista czasopism MEiN: 100

Supplementary File S 2. ***Yılmaz, E.****, Gastineau R., Solak C. N., Górecka E., Trobajo R., Turmel M., Lemieux C., Otis C., Witkowski A., & Mann D. G., (2024). Morphological and molecular characterization of* Halamphora vantushpaensis *(Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye)* PhytoKeys *(proof, in press).*

Impact Factor: 1.4

Lista czasopism MEiN: 100

Supplementary File S 3. Yılmaz, E., Gastineau R., Górecka E., Solak C. N., Trobajo R., Peszek Ł., & Mann D. G., . Halamphora witkowskii *sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye.* Nova Hedwigia *(proof, in press).*

Impact Factor: 1.0

Lista czasopism MEiN: 40

INSTYTUT NAUK O MORZU
I ŚRODOWISKU
UNIWERSYTETU SZCZECIŃSKIEGO

Elif Yilmaz
Institute of Marine and Environmental Sciences
University of Szczecin
Mickiewicza 16a, 70-383 Szczecin, Poland

Declaration

I hereby declare that my contribution in the preparation of the article *"Description of Navicula vanseea sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny,"* published in *PhytoKeys* 241: 27-48, of which I am co-author with Mann D. G., Gastineau R., Trobajo R., Solak C. N., Górecka E., Turmel M., Lemieux C., Ertorun N., and Witkowski A., equals to 65%. For the purpose of the study, I conceptualized the study, performed laboratory work, interpreted the data, prepared figures and tables, drafted the original manuscript, revised it following the co-authors' suggestions, led the revision process after peer review, and submitted the final version.

Elif Yilmaz

ul. Mickiewicza 16
70-383 Szczecin
tel. +48 91 444 24 61
www.nms.usz.edu.pl

Prof. Dr David G. Mann
Research Associate
Royal Botanic Garden Edinburgh
Edinburgh EH3 5LR
Scotland, UK

Declaration

I hereby declare that my contribution in the preparation of the article *"Description of Navicula vanseea sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny,"* published in *PhytoKeys* 241: 27-48, of which I am co-author with Yilmaz E., Gastineau R., Trobajo R., Solak C. N., Górecka E., Turmel M., Lemieux C., Ertorun N., and Witkowski A., equals to 10%. I supervised the investigation and work of PhD student Yilmaz E., and revised the original draft.

David G. Mann

INSTYTUT NAUK O MORZU
I ŚRODOWISKU
UNIWERSYTETU SZCZECIŃSKIEGO

Dr Romain Gastineau
Institute of Marine and Environmental Sciences
University of Szczecin
Mickiewicza 16a, 70-383 Szczecin, Poland

Declaration

I hereby declare that my contribution in the preparation of the article *"Description of Navicula vanseea sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny,"* published in *PhytoKeys* 241: 27-48, of which I am co-author with Yilmaz E., Mann D. G., Trobajo R., Solak C. N., Górecka E., Turmel M., Lemieux C., Ertorun N., and Witkowski A., equals to 15%. For the purpose of the study, I supervised the work PhD student Yilmaz E., and took part in the molecular methodology, bioinformatics development, and revision of the original draft.

Romain Gastineau

ul. Mickiewicza 16
70-383 Szczecin
tel. +48 91 444 24 61
www.nms.usz.edu.pl

Prof. Dr Cüneyt Nadir Solak
Department of Biology,
Faculty of Science & Art, Dumlupınar University,
43000 Kütahya, Türkiye

Declaration

I hereby declare that my contribution in the preparation of the article *"Description of Navicula vanseea sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny,"* published in *PhytoKeys* 241: 27-48, of which I am co-author with Yilmaz E., Mann D. G., Gastineau R., Trobajo R., Turmel M., Lemieux C., Ertorun N., and Witkowski A., equals to 2%. For the purpose of the study, I participated in the taxonomy methodology and reviewed the original draft.

Cüneyt Nadir Solak

Associate professor Nesil ERTORUN

Department of Biology,

Faculty of Science,

Eskisehir Technical University,

26000 Eskisehir, Türkiye

Declaration

I hereby declare that my contribution in the preparation of the article *Description of Navicula vanseea sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny,* published in *PhytoKeys* 241: 27-48, of which I am co-author with Yilmaz E., Mann D. G., Gastineau R., Trobajo R., Solak C. N., Turmel M., Lemieux C., and Witkowski A., equals to 1%. For the purpose of the study, I assisted in taxonomy methodology and scanning electron microscopy studies.

Nesil ERTORUN

N. Ertorun

INSTYTUT NAUK O MORZU
I ŚRODOWISKU
UNIWERSYTETU SZCZECIŃSKIEGO

Elif Yilmaz
Institute of Marine and Environmental Sciences
University of Szczecin
Mickiewicza 16a, 70-383 Szczecin, Poland

Declaration

I hereby declare that my contribution in the preparation of the article "*Morphological and molecular characterization of Halamphora vantushpaensis (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye)*," co-authored with Gastineau R., Solak C. N., Górecka E., Trobajo R., Turmel M., Lemieux C., Otis C., Witkowski A., and Mann D. G., currently in press with *PhytoKeys* (online ISSN 1314-2003), equals 65%. For the purpose of the study, I conceptualized the research, performed the isolation of the species, conducted laboratory experiments, carried out the taxonomic analysis, interpreted the data, prepared figures and tables, drafted the original manuscript, revised it following co-authors' feedback, led the revision process after peer review, and submitted the final version.

Elif Yilmaz

ul. Mickiewicza 16
70-383 Szczecin
tel. +48 91 444 24 61
www.nms.usz.edu.pl

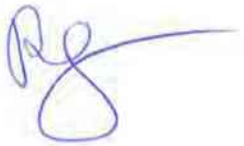
INSTYTUT NAUK O MORZU
I ŚRODOWISKU
UNIWERSYTETU SZCZECIŃSKIEGO

Dr Romain Gastineau
Institute of Marine and Environmental Sciences
University of Szczecin
Mickiewicza 16a, 70-383 Szczecin, Poland

Declaration

I hereby declare that my contribution in the preparation of the article *" Morphological and molecular characterization of Halamphora vantushpaensis (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye),"* co-authored with Yilmaz E., Solak C. N., Górecka E., Trobajo R., Turmel M., Lemieux C., Otis C., Witkowski A., and Mann D. G., currently in press with PhytoKeys (online ISSN 1314-2003), equals 15%. For the purpose of the study, I supervised the work PhD student Yilmaz E., and took part in the molecular methodology, bioinformatics development, and revision of the original draft.

Romain Gastineau



ul. Mickiewicza 16
70-383 Szczecin
tel. +48 91 444 24 61
www.nms.usz.edu.pl

Prof. Dr David G. Mann

Research Associate

Royal Botanic Garden Edinburgh

Edinburgh EH3 5LR

Scotland, UK

Declaration

I hereby declare that my contribution in the preparation of the article "*Morphological and molecular characterization of Halamphora vantushpaensis (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye)*," co-authored with Yilmaz E., Gastineau R., Solak C. N., Górecka E., Trobajo R., Turmel M., Lemieux C., Otis C., and Witkowski A., currently in press with *PhytoKeys* (online ISSN 1314-2003), equals 10%. I supervised the investigation and work of PhD student Yilmaz E., and revised the original draft.

David G. Mann

INSTYTUT NAUK O MORZU
I ŚRODOWISKU
UNIWERSYTETU SZCZECIŃSKIEGO

Dr Ewa Górecka
Institute of Marine and Environmental Sciences
University of Szczecin
Mickiewicza 16a, 70-383 Szczecin, Poland

Declaration

I hereby declare that my contribution in the preparation of the article *"Morphological and molecular characterization of Halamphora vantushpaensis (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye),"* co-authored with Yilmaz E., Gastineau R., Solak C. N., Trobajo R., Turmel M., Lemieux C., Otis C., Witkowski A., and Mann D. G., currently in press with *PhytoKeys* (online ISSN 1314-2003), equals 2%. For the purpose of the study, took part in the molecular methodology, and helped with revision in peer-review process.

Ewa Górecka

ul. Mickiewicza 16
70-383 Szczecin
tel. +48 91 444 24 61
www.nms.usz.edu.pl

Prof. Dr Cüneyt Nadir Solak

Department of Biology,

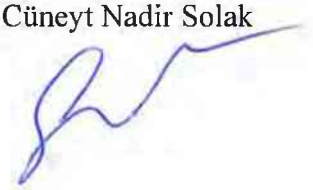
Faculty of Science & Art, Dumlupınar University,

43000 Kütahya, Türkiye

Declaration

I hereby declare that my contribution in the preparation of the article *" Morphological and molecular characterization of Halamphora vantushpaensis (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye),"* co-authored with Yilmaz E., Gastineau R., Górecka E., Trobajo R., Turmel M., Lemieux C., Otis C., Witkowski A., and Mann D. G., currently in press with *PhytoKeys* (online ISSN 1314-2003), equals 2%. For the purpose of the study, I took part in the taxonomy methodology, and helped with scanning electron microscopy in process.

Cüneyt Nadir Solak



INSTYTUT NAUK O MORZU
I ŚRODOWISKU
UNIWERSYTETU SZCZECIŃSKIEGO

Elif Yilmaz
Institute of Marine and Environmental Sciences
University of Szczecin
Mickiewicza 16a, 70-383 Szczecin, Poland

Declaration

I hereby declare that my contribution in the preparation of the article "*Halamphora witkowskii sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye*" co-authored with Gastineau R., Górecka E., Solak C. N., Trobajo R., Peszek Ł., and Mann D. G. to be published in Nova Hedwigia dedicated to Prof. Witkowski (online ISSN 2363-7188), equals 70%. For this study, I was responsible for conceptualizing the research, isolating the species, and performing the taxonomic analysis. Additionally, I conducted the laboratory experiments, analyzed and interpreted the data, prepared the figures and tables, wrote the initial draft of the manuscript, addressed co-authors' suggestions, managed the peer review revisions, and finalized the submission.

Elif Yilmaz

ul. Mickiewicza 16
70-383 Szczecin
tel. +48 91 444 [illegible]
www.oms.usz.edu.pl

Prof. Dr David G. Mann

Research Associate

Royal Botanic Garden Edinburgh

Edinburgh EH3 5LR

Scotland, UK

Declaration

I hereby declare that my contribution in the preparation of the article " *Halamphora witkowskii sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye*," Yilmaz E., Gastineau R., Górecka E., Solak C. N., Trobajo R., and Peszek Ł., currently in press in Nova Hedwigia (dedicated to Prof. Witkowski, online ISSN 2363-7188), equals 10%. I supervised the investigation and work of PhD student Yilmaz E., and revised the original draft.

David G. Mann

INSTYTUT NAUK O MORZU
I ŚRODOWISKU
UNIWERSYTETU SZCZECIŃSKIEGO

Dr Ewa Górecka
Institute of Marine and Environmental Sciences
University of Szczecin
Mickiewicza 16a, 70-383 Szczecin, Poland

Declaration

I hereby declare that my contribution to the preparation of the article "*Halamphora witkowskii sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye*," co-authored with Yilmaz E., Gastineau R., Solak C. N., Trobajo R., Peszek Ł., and Mann D. G., currently in press in Nova Hedwigia (dedicated to Prof. Witkowski, online ISSN 2363-7188), equals 10%. For the purpose of the study, I participated in the conceptualization of the study, the molecular methodology, and assisted with revisions during the peer-review process.

Ewa Górecka

ul. Mickiewicza 16
70-383 Szczecin
tel. +48 91 444 24 61
www.inms.usz.edu.pl

Prof. Dr Cüneyt Nadir Solak

Department of Biology,

Faculty of Science & Art, Dumlupınar University,

43000 Kütahya, Türkiye

Declaration

I hereby declare that my contribution in the preparation of the article *Halamphora witkowskii sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye,"* Yilmaz E., Gastineau R., Górecka E., Trobajo R., Peszek Ł., and Mann D. G. currently in press in Nova Hedwigia (dedicated to Prof. Witkowski, online ISSN 2363-7188), equals 2%. For the purpose of the study, I took part in conceptualization of the study, and assisted with revision in peer-review process.

Cüneyt Nadir Salak

UNIWERSYTET SZCZECIŃSKI

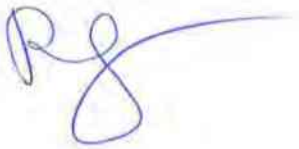
INSTYTUT NAUK O MORZU
I ŚRODOWISKU
UNIWERSYTETU SZCZECINSKIEGO

Dr Romain Gastineau
Institute of Marine and Environmental Sciences
University of Szczecin
Mickiewicza 16a, 70-383 Szczecin, Poland

Declaration

I hereby declare that my contribution in the preparation of the article *Halamphora witkowskii sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye,"* Yilmaz E., Górecka E., Solak C. N., Trobajo R., Peszek Ł., and Mann D. G., currently in press in Nova Hedwigia (dedicated to Prof. Witkowski, online ISSN 2363-7188), equals 10%. For the purpose of the study, I supervised the work PhD student Yilmaz E., and took part in the molecular methodology, revision of the original draft.

Romain Gastineau



ul. Mickiewicza 16
70-383 Szczecin
tel. +48 91 444 24 61
www.nms.usz.edu.pl

PhytoKeys

PhytoKeys 241: 27–48 (2024)
DOI: 10.3897/phytokeys.241.118903

Research Article

Description of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny

Elif Yılmaz[1], David G. Mann[2], Romain Gastineau[1], Rosa Trobajo[3], Cüneyt Nadir Solak[4], Ewa Górecka[1], Monique Turmel[5], Claude Lemieux[5], Nesil Ertorun[6], Andrzej Witkowski[1*]

1 *Institute of Marine and Environmental Sciences, University of Szczecin, Mickiewicza 16A, PL70–383 Poland*
2 *Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, Scotland, UK*
3 *Marine and Continental Waters, Institute for Food and Agricultural Research and Technology (IRTA), Crta de Poble Nou Km 5.5, E-43540 La Ràpita, Catalunya, Spain*
4 *Department of Biology, Faculty of Science & Art, Dumlupınar University, 43000 Kütahya, Türkiye*
5 *Département de biochimie, de microbiologie et de bio-Informatique, Institut de Biologie Intégrative et des Systèmes, Université Laval, Québec, QC, Canada*
6 *Department of Biology, Science Faculty, Eskişehir Technical University, 26000 Eskişehir, Türkiye*

Corresponding author: Elif Yılmaz (elfyilmaz38@gmail.com)

Abstract

The current article describes *Navicula vanseea* **sp. nov.**, a new species of diatom from Lake Van, a highly alkaline lake in Eastern Anatolia (Türkiye). The description is based on light and scanning electron microscopy performed on two monoclonal cultures. The complete nuclear rRNA clusters and plastid genomes have been sequenced for these two strains and the complete mitogenome for one of them. The plastome of both strains shows the probable loss of a functional *ycf35* gene. They also exhibit two IB4 group I introns in their *rrl*, each encoding for a putative LAGLIDADG homing endonuclease, with the first L1917 IB4 intron reported amongst diatoms. The Maximum Likelihood phylogeny inferred from a concatenated alignment of *18S*, *rbcL* and *psbC* distinguishes *N. vanseea* sp. nov. from the morphologically similar species *Navicula cincta* and *Navicula microdigitoradiata*.

Key words: Group I intron, LAGLIDADG, mitogenome, Naviculaceae, plastome, pseudogene, soda lake

Academic editor: Bing Liu
Received: 16 January 2024
Accepted: 19 February 2024
Published: 8 April 2024

Citation: Yilmaz E, Mann DG, Gastineau R, Trobajo R, Solak CN, Górecka E, Turmel M, Lemieux C, Ertorun N, Witkowski A (2024) Description of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny. PhytoKeys 241: 27–48. https://doi.org/10.3897/phytokeys.241.118903

* Deceased author.

Introduction

Lake Van is located in Eastern Anatolia, Turkey (Republic of Türkiye). It is Turkey's largest inland water body and also world's largest soda lake. The lake is surrounded by dormant volcanoes and its formation was a consequence of the eruption of the Nemrut stratovolcano (not to be confused with the Nemrut Mountain, also in Turkey), which is 2247 m above sea level. As a result of the erosion of volcanic rocks in the catchment and evaporation, the lake water is salty (21.4‰) and alkaline (155 m mEq−1, pH 9.81) (Glombitza et al. 2013; Ersoy Omeroglu et al. 2021). The lake is notable for its unusual chemistry,

which results from the constant losses of calcium as carbonate and of magnesium in the form of mineral phases rich in Mg–silica. Thus, the Mg cycle is closely related to the silica cycle, which is itself dependent on the production of biosilica by diatoms, eventually followed by the dissolution of their frustules (Reimer et al. 2009).

The genus *Navicula* is amongst the most species-rich genera of Bacillariophyceae, although this is partly because it was used for a long time as a 'catch-all' for simply structured, bilaterally symmetrical raphid diatoms. It was erected as early as 1822 by Bory in his 'Dictionnaire Classique d'Histoire Naturelle' (Bory de Saint-Vincent 1822). The name chosen by Bory refers to the shape of the cells, similar to the shuttle that was used for weaving. The cells are generally solitary and motile, although some species live in mucilage tubes (Millie and Wee 1981). Cells have two parietal chloroplasts. Their valves are symmetrical both apically and transapically and have rounded, acute or capitate ends. The central area is often distinctly expanded (Patrick 1959; Cox 1979; Round et al. 1990).

The only account ever published on the diatoms from Lake Van was written by Legler and Krasske (1940). Amongst the 24 species they recorded were three taxa of *Navicula*, all of them considered as varieties of *Navicula cryptocephala*, namely *Navicula cryptocephala* Kützing, 1844, *Navicula cryptocephala* var. *intermedia* Grunow 1880 and a taxon noted as *Navicula cryptocephala* var. *veneta* (Kützing) Grunow, which possibly corresponds to *Navicula cryptocephala* var. *veneta* (Kützing) Rabenhorst, 1864. Out of these three taxa, only *N. cryptocephala* Kützing, 1844 is still deemed to be valid. *Navicula cryptocephala* var. *intermedia* is considered a synonym of *Navicula capitatoradiata* H. Germain ex Gasse 1986 and *Navicula cryptocephala* var. *veneta* is now treated as an independent species, *Navicula venet*a Kützing 1844.

Navicula species are rather well documented in inland waters where they are known for their bioindicator potential (Lange-Bertalot 2001). For instance, *Navicula tripunctata* (O.F. Müller) Bory is a good indicator of eutrophic waters with an average to high electrolyte content, while taxa such as *Navicula gregaria* Donkin, 1861, *Navicula meulemansii* A.Mertens, A.Witkowski & Lange-Bertalot 2013 and *N. veneta* are common in brackish to electrolyte rich waters (Cox 1995; Adrienne Mertens et al. 2014).

Preliminary results from a new sampling campaign conducted in 2021 in Lake Van strongly suggested that the biodiversity of diatoms had been underestimated in the previous work of Legler and Krasske (1940). One illustration is a previously undocumented *Nitzschia*, *N. anatoliensis* Górecka, Gastineau & Solak (Solak et al. 2021), which would have probably been overlooked if it had not been for the combined use of microscopic and molecular tools. Amongst several other monoclonal cultures from the 2021 campaign, two contained strains of a *Navicula* species were identified, which is the subject of the present article. Although both strains were noticeably different in size, they were quickly proven to belong to the same new species.

The aim of the following article is to formally describe *Navicula vanseea* sp. nov. from Lake Van. The description combines the use of light microscopy (live specimen and cleaned frustules) and scanning electron microscopy. The complete cluster of nuclear ribosomal RNA genes and the complete plastid genome were obtained for both strains by means of next generation sequencing and

also the mitogenome of one of these strains. These results were included in a multigene Maximum Likelihood phylogeny which unambiguously separated *Navicula vanseea* sp. nov. from morphologically similar known species, whose differences with *Navicula vanseea* sp. nov. are discussed. As it was the first time that a L1917 group I intron with its putative LAGLIDADG homing endonuclease gene had been discovered in the plastid genome of a diatom, special attention has been paid to this feature, with a phylogeny of the putative LAGLIDADG protein being performed.

Materials and methods

Sampling, isolation and cultivation

Epilithic samples were collected by brushing rocks in the littoral of Lake Van in July 2021, in the vicinity of Erciş Municipality (Fig. 1). Samples were re-suspended in surface water from the lake in 50 ml tubes before being brought to the University of Szczecin for subsequent analyses. Samples were then transferred into Petri dishes containing sterile f/2 medium (Guillard 1975) modified to 18‰ salinity. Single cells were isolated by micropipette under an inverted Nikon Eclipse light microscope. Successive re-isolations were performed (at least 3 times) before the culture was considered monoclonal. Strains were later transferred into 250 ml Erlenmeyer flasks with modified f/2 medium. Cultures were maintained in active growth under a light intensity of 60 µmol photons m^{-2} s^{-1} and a photoperiod of 14 h light/10 h darkness). Two morphologically distinct clones with different cell sizes were registered in the Szczecin Culture Collection as SZCZEY2172 and SZCZEY2262.

Light and scanning electron microscopy

Pictures of living diatoms were taken using a Light Microscope (LM) Zeiss Axio Scope A1 (Carl Zeiss, Jena, Germany at a magnification of 1000× by transferring diatom cultures directly on to the glass slide.

To prepare cleaned frustules for microscopy, 5 ml of monoclonal cultures were transferred into 20 ml beakers with 10 ml of 10% hydrochloric acid (HCl). After 24h, samples were washed four times with distilled water then re-suspended in 30% hydrogen peroxide (H_2O_2) and boiled for about four hours. Finally, samples were washed again four times with distilled water. For LM, cleaned material was then air-dried on cover glasses and mounted on glass slide with Naphrax® (Brunel Microscopes Ltd., Chippenham, UK) solution and pictures were taken with the Zeiss Axio Scope A1. For SEM, a drop of cleaned sample was deposited on a Nuclepore Track-Etch membrane from Whatman (Maidstone, England). The membranes were air-dried overnight, mounted on aluminium stubs with carbon tape and coated with gold using a Q150T coater from Quorum Technologies (Laughton, U.K.). SEM observations were made at the Faculty of Chemical Technology and Engineering, Western Pomeranian University of Technology in Szczecin (Poland), using a Hitachi SU8020 (Tokyo, Japan) and Eskişehir Technical University (Türkiye) using a ZEISS Ultra microscope (Oberkochen, Germany).

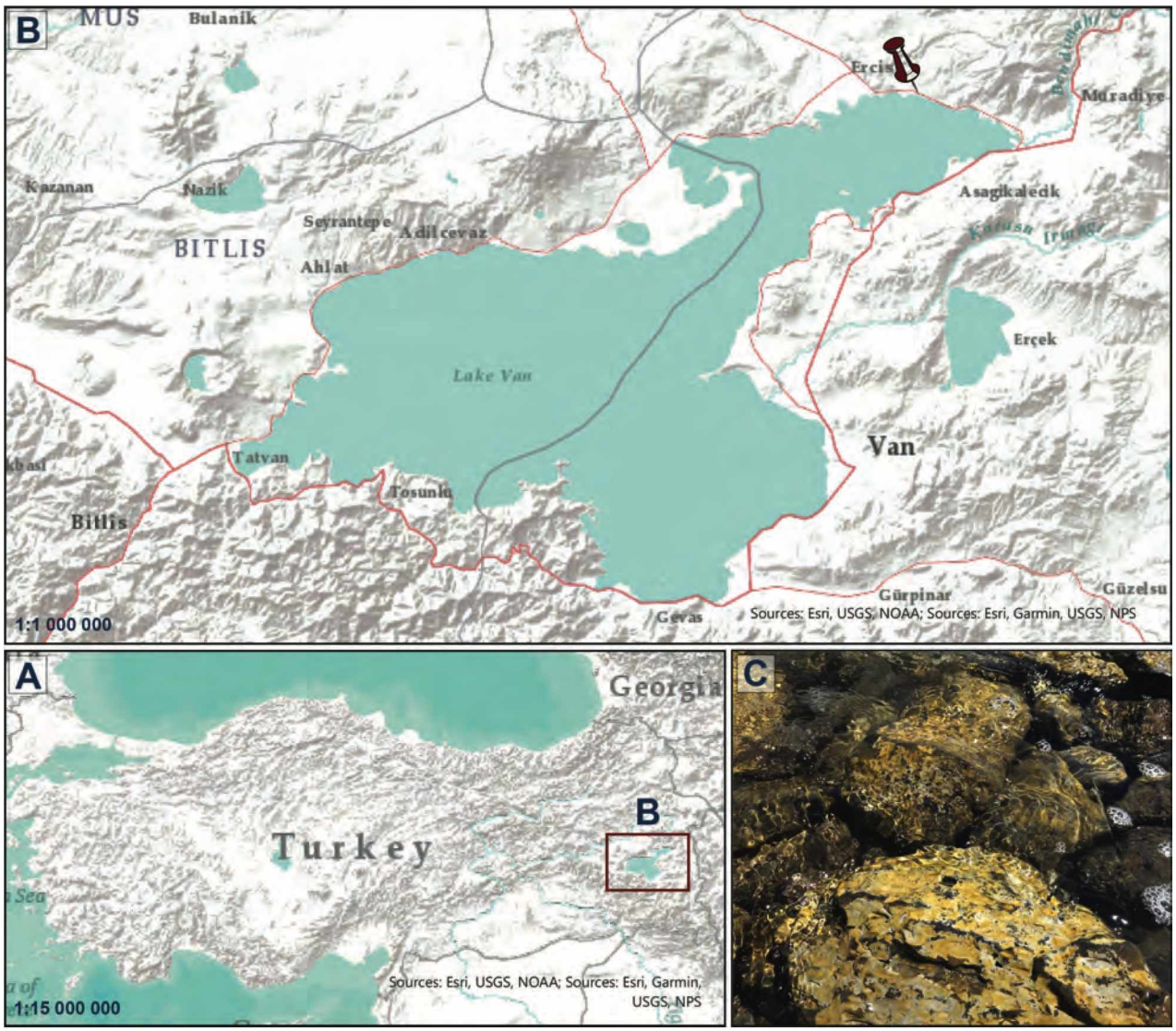


Figure 1. Map of the sampling location **A** location of Lake Van in Turkey. The red frame indicates the position of Lake Van **B** general view of the lake. The pin indicates the position of the sampling area **C** photo of the epilithic sampling area on the rock (Esri. (2023). ArcGIS Pro 3.1.0. Environmental Systems Research Institute).

Next generation sequencing and bioinformatic analyses

DNA was extracted from clones SZCZEY2172 and SZCZEY2262 using the protocol of Doyle and Doyle (1990). Total DNA was then sent to the Beijing Genomics Institute (BGI) in Shenzhen (China) to be sequenced on a DNBSEQ platform. For each clone, a total of ca. 40M clean 150 bp paired-end reads was obtained. Reads were assembled with a k-mer parameter of 125 using SPAdes 3.15.0 (Bankevich et al. 2012). Contigs of interest were retrieved by customised command-line BLASTn analyses as previously described (Dąbek et al. 2022; Gastineau et al. 2022). Consed (Gordon and Green 2013) was used to merge the different subunits of the plastome and when trying to circularise the mitogenome. Annotations were performed using the same tools as described in Gastineau et al. (2022). The maps of the organellar genomes were obtained from the OGDRAW online portal (Lohse et al. 2013). The different parts of the nuclear rRNA gene cluster were identified with the help of Rfam 14 (Kalvari et al. 2021).

Molecular phylogeny

The three gene datasets (*18S*, *rbcL* and *psbC*) already used in previous publications (Dąbek et al. 2017; Li et al. 2020; Górecka et al. 2021a;) were obtained and the corresponding genes from various Naviculaceae and *N. vanseea sp. nov*. appended. Sequences for *N. capitatoradiata*, *N. microdigitoradiata* Lange-Bertalot, 1993 and *N. cincta* (Ehrenberg) Ralfs 1861 were also added. However, it should be noted that these three species were represented in GenBank just by *18S* and *rbcL* (*N. capitatoradiata* and *N. cincta*) or *rbcL* only (*N. microdigitoradiata*). Genes were aligned separately with MAFFT 7 (Katoh and Standley 2013) with the -auto option and trimmed using trimAl (Capella-Gutiérrez et al. 2009) with the -automated1 option. The best model of evolution for each of these genes was selected with ModelTest-NG (Darriba et al. 2020) and were GTR+I+G4 (*psbC*), TIM3+I+G4 (*rbcL*) and TIM1+I+G4 (*18S*). Alignments were then concatenated using Phyutility 2.7.1 (Smith and Dunn 2008) for a final size of 3301 bp. A Maximum Likelihood phylogeny was constructed from the concatenated alignment using IQ-TREE 2.2.0 (Minh et al. 2020) with 1000 ultrafast bootstrap replicates and a dataset partitioned, based on the best models of evolution found for each gene. *Triparma pacifica* (Guillou & Chrétiennot-Dinet) Ichinomiya & Lopes dos Santos, 2016 was used as an outgroup.

For the phylogeny of the putative LAGLIDADG endonuclease proteins, protein sequences found in IB4 - L1917 and - L1931 introns presented in Lucas et al. (2001) and Haugen and Bhattacharya (2004) were downloaded from GenBank. Sequences from the current study, from Turmel et al. (2002), Haugen et al. (2007) and Lemieux et al. (2014) were appended. The IB2 – L1917 LAGLIDADG sequence from *Coxiella burnetii* (Derrick, 1939) Philip, 1948 was used as an outgroup (Raghavan et al. 2007). The phylogeny was conducted in a similar way to the multigene phylogeny, except that alignment was not trimmed. The best model of evolution returned by ModelTest-NG was LG+G4.

Results

Taxonomy

***Navicula vanseea* Yilmaz, Gastineau, Solak & Witkowski, sp. nov.**
Figs 2–4

Type material. *Holotype*: Slide number SZCZEY2172 in the collection of Andrzej Witkowski at the University of Szczecin, Poland. Valves representing the holotype population are illustrated in Fig. 2L.

***Isotype*:** Slide number TR_Erciş_Van_2021 deposited in Kütahya Dumlupınar University (Turkey).

Registration. http://phycobank.org/104542

Type locality. Erciş Van, Türkiye (38°59'47.3"N 43°24'15.3"E) collected by: Elif Yilmaz, 31 July 2021.

Etymology. The name given to the species refers to the German name of Lake Van (Vansee, the sea of Van) as it was used in the work of Legler and Krasske and is meant as a tribute to these authors and their work.

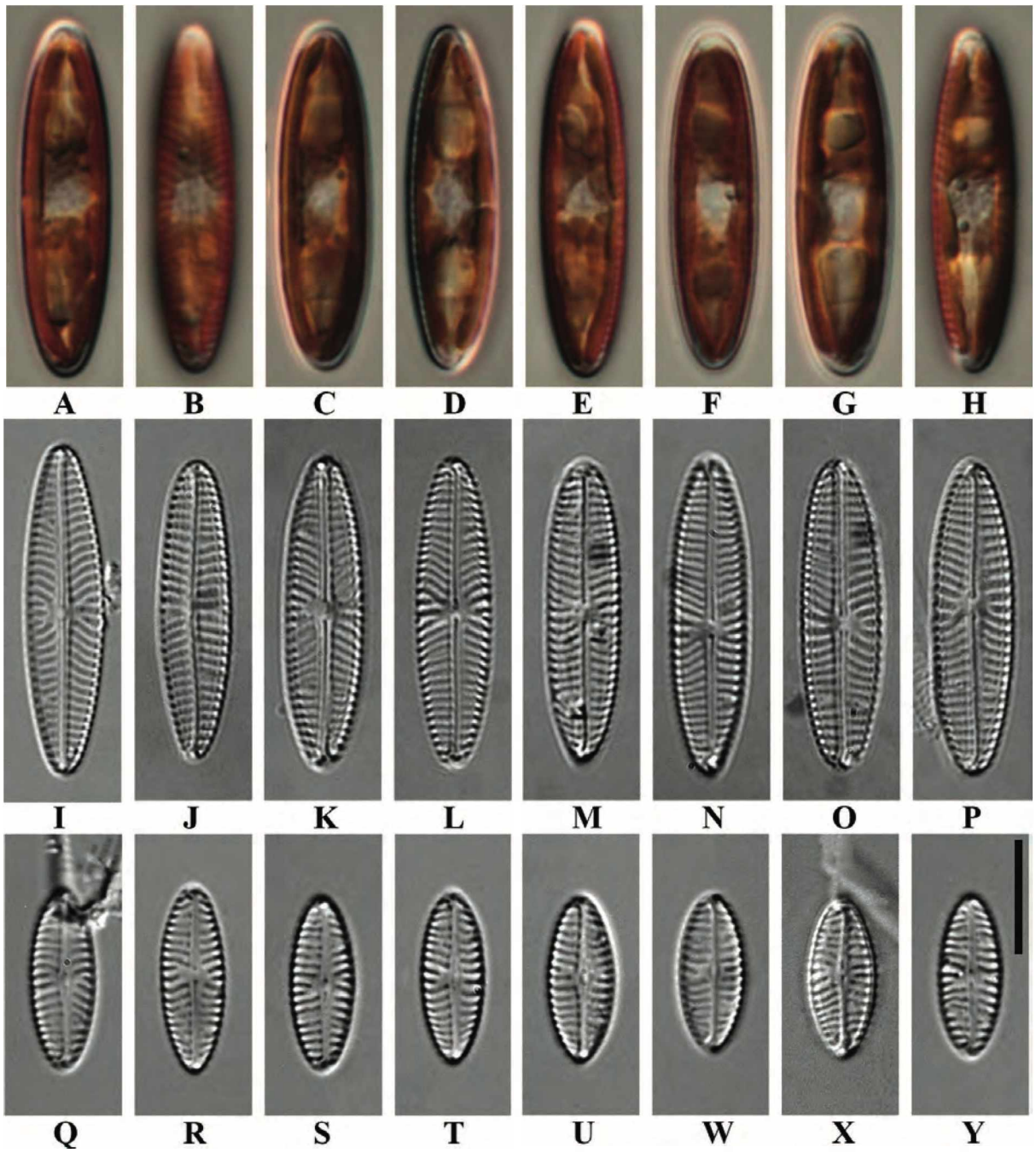


Figure 2. *Navicula vanseea* sp. nov. LM micrographs **A–H** *in vivo* pictures of *Navicula vanseea* sp. nov. SZCZEY2172 **I** LM image of a cleaned valve from wild material **J–P** cleaned valves of *Navicula vanseea* sp. nov. SZCZEY2172 **Q–Y** cleaned valves of *Navicula vanseea* sp. nov. SZCZEY2262 Scale bar: 10 µm.

Distribution and ecology. The taxon was exclusively observed within benthic epilithic assemblages in Lake Van (salinity 21.4‰ and pH 9.5).

Description. ***LM*** (Fig. 2A–Y) Valves: smaller specimens elliptic, tapering towards cuneately rounded apices, larger specimens linear-elliptic-lanceolate narrowly rounded, with narrowly rounded poles, which are occasionally slightly protracted (Fig. 2E, K, N). Valve dimensions (n = 39): length 11.0–28.0 µm, width 5.0–6.5 µm. Raphe filiform, straight. Central area small and rounded, axial area narrow. Striae strongly radiate, sometimes irregularly shortened around the central area, 12–13 in 10 µm, lineolae difficult to resolve in LM, ca. 50 in 10 µm.

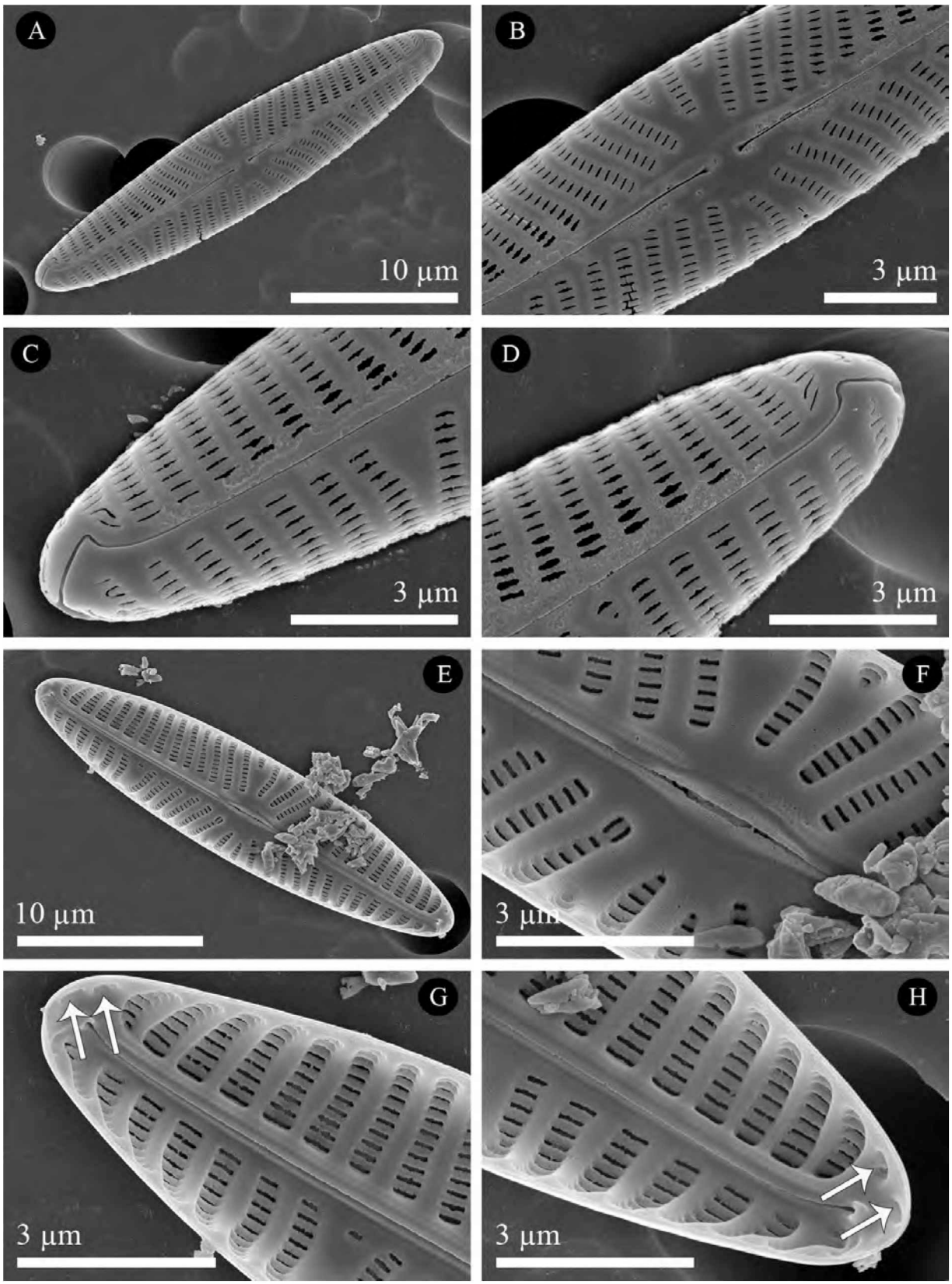


Figure 3. SEM micrographs of *Navicula vanseea* sp. nov. SZCZEY2172 **A** external view of the entire valve **B** details of central area showing simple, slightly drop-shaped proximal raphe endings and shortened striae **C, D** details of the two apices of a single valve showing the terminal fissures **E** internal view of the entire valve **F** details of central area showing filiform proximal raphe endings in a fusiform expansion of the raphe-sternum **G, H** details of apices showing well-developed helictoglossae showing two isolated lineolae (white arrows). Scale bars: 10 µm (**A, E**); 3 µm (**B–D, F–H**).

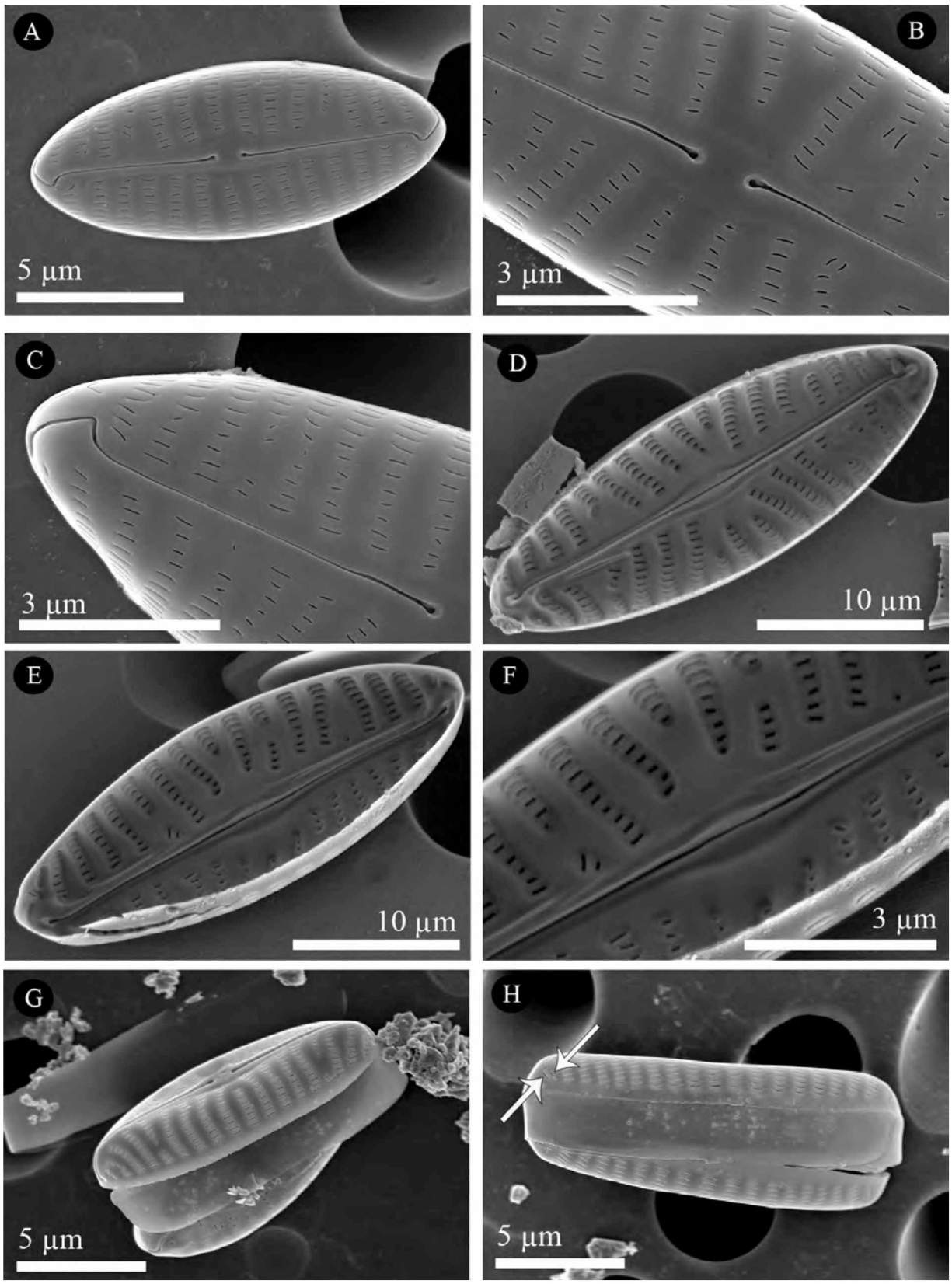


Figure 4. SEM micrographs of *Navicula vanseea* sp. nov. SZCZEY2262 **A** external view of the entire valve **B** details of central area showing simple proximal raphe endings and shortened striae **C** details of apex showing the terminal fissure **D, E** internal view of two entire valves, showing the central area and filiform proximal raphe endings **F** details of apex showing well-developed helictoglossae **G, H** girdle view of valves showing continuous areolation on mantle and two isolated lineolae (white arrows). Scale bars: 5 µm (**A, D, E, G, H**); 3 µm (**B, C, F**).

SEM External valve surface (Figs 3A–D, 4D–F): Valve surface flat (Fig. 3G, H), areolae apically elongated (Figs 3B–D, 4B). Raphe sternum slightly elevated above the valve face level (Fig. 4G). Axial area very narrow, central area very slightly expanded, small, asymmetric (Figs 3B, 4A, B). Proximal raphe endings drop-like, slightly deflected unilaterally (Fig. 4B). Distal raphe endings strongly hooked in the same direction (Fig. 3C, D, which are the two ends of the same valve and Fig. 4A).

SEM Internal valve surface (Figs 3E–H, 4D–F): valve surface slightly arched with transapical striae positioned in relatively deep grooves, bordered by virgae that become thicker towards to the centre of the valve (Figs 3H, 4D, F). Central area asymmetric, usually only slightly expanded (Figs 3E, F, 4F), but sometimes more strongly (Fig. 4D). The internal lineolae openings are slit-like (Fig. 3F–H), narrower than the vimines. Lineolae occluded by hymens (Fig. 4F); two isolated lineolae are present at the valve apex. Raphe sternum slightly widened at the centre to form a fusiform ridge enclosing the central raphe endings, which are simple, straight and separated (Figs 3F, 4F). Distally, the raphe terminates in well-developed helictoglossae (Figs 3G, H).

Genomics and phylogeny

The nuclear rRNA gene cluster

The complete rRNA gene cluster was sequenced for both clones and deposited in NCBI GenBank with accession numbers OR797294 (SZCZEY2172) and OR797293 (SZCZEY2262). The cluster is 4902 bp long, distributed as follows: 18S – 1792 bp, ITS1 – 195 bp, 5.8S – 155 bp, ITS2 – 260 bp, 28S – 2500 bp. Comparing the two clones, there was one single nucleotide polymorphism (SNP) found in the 18S (in the V2 region), three in the ITS1, one in the 5.8S, three in the ITS2 and two in the 28S (both in the D1/D2 region).

Mitochondrial genome

A 43997 bp contig corresponding to the mitochondrial genome was retrieved for strain SZCZEY2262, but could not be circularised because of the presence of repeated sequences at its ends. However, for easier reading, it is displayed as circular on the map (Fig. 5). The mitogenome encodes 34 protein-coding genes plus the conserved open-reading frame (ORF) *orf150* between *rps11* and *mttB* (Pogoda et al. 2019), two rRNA genes and 23 tRNA genes (GenBank: OR795084). The *nad11* gene is split into two distinct subunits, separated from each other by two protein-coding genes, two rRNA and one tRNA. In the repeated part of the genome, there are two copies of the same ORF, *orf145*. There is a 767-bp group I intron in the *rnl* gene.

Despite several attempts, it was impossible to assemble the mitogenome of strain SZCZEY2172. Lowering the k-mer parameter to 75 only allowed the recovery of a short ca. 500 bp fragment with a low coverage. This fragment was used as a seed to try an assembly with NOVOPlasty 4.3.3 (Dierckxsens et al. 2017), using the mitogenome of SZCZEY2262 as reference sequence and a k-mer of 25, but this attempt also failed.

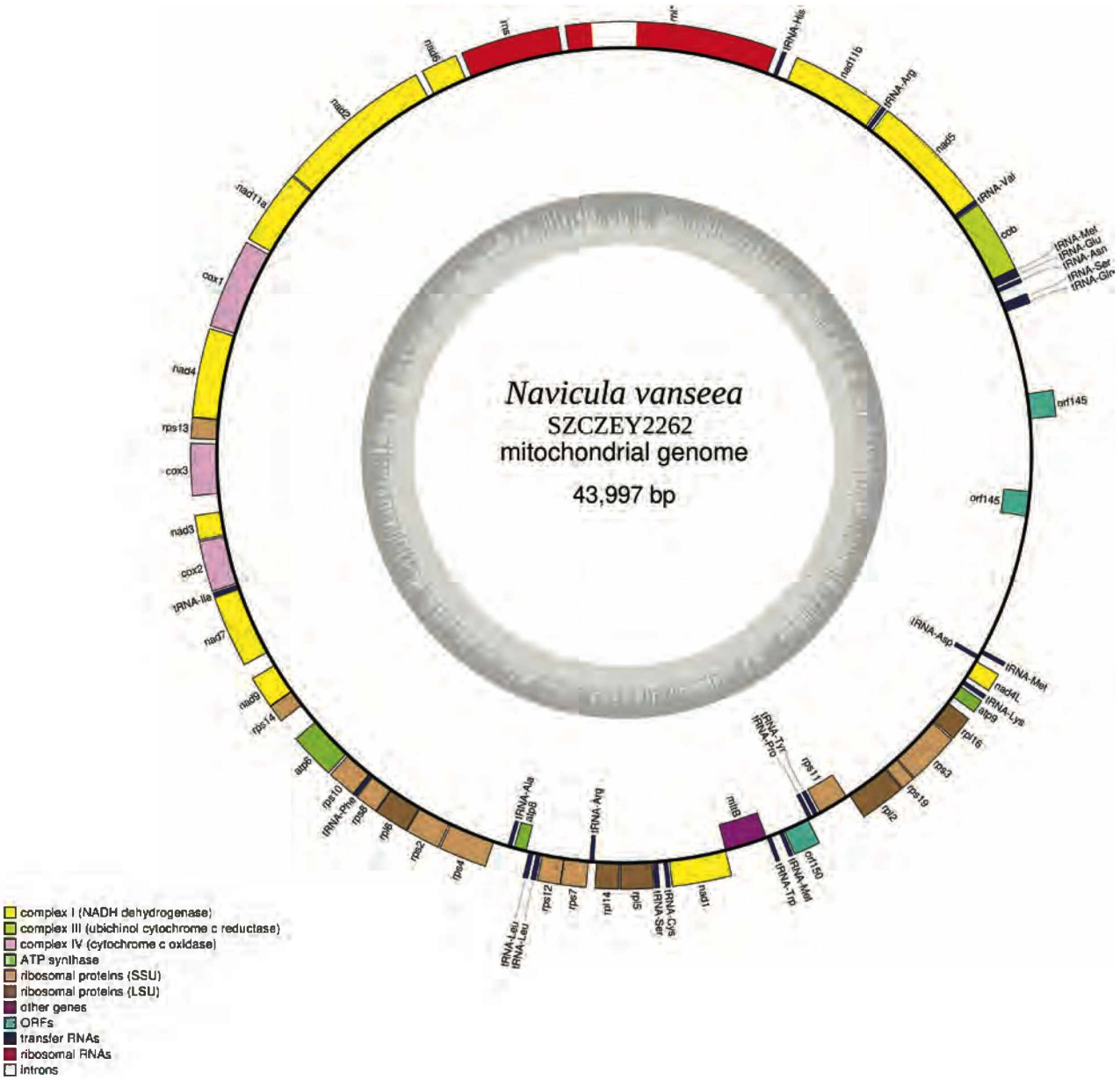


Figure 5. Map of the mitochondrial genome of *Navicula vanseea* sp. nov. SZCZEY2262.

Plastid genome

The plastome is 158,005 bp long for SZCZEY2262 (Fig. 6) and 157,990 bp long for SZCZEY2172 (Fig. 7). For SZCZEY2262 (GenBank: OR795085), the large single-copy (LSC) is 72,941 bp long and encodes 74 conserved protein-coding genes, two non-conserved ORF, two putative integrase/recombinase *xerC* genes and 17 tRNAs. The small single-copy region (SSC) is 49,714 long and encodes 51 conserved protein-coding genes, eight tRNAs, and five non-conserved ORFs of a size higher than 100 amino-acids (AA). The inverted repeat (IR) is 17,675 bp long and encodes one conserved protein-coding gene, three rRNAs, six non-conserved ORFs, four tRNAs and one putative *serC* gene and *rbcR* overlaps the inverted repeat B (IRB) and the SSC. There are two IB4 group I introns in the *rrl* gene at positions 1917 and 1931 (based on the reference sequence U00096 from *Escherichia coli* T. Escherich, 1885 str. K-12 substr. MG1655),

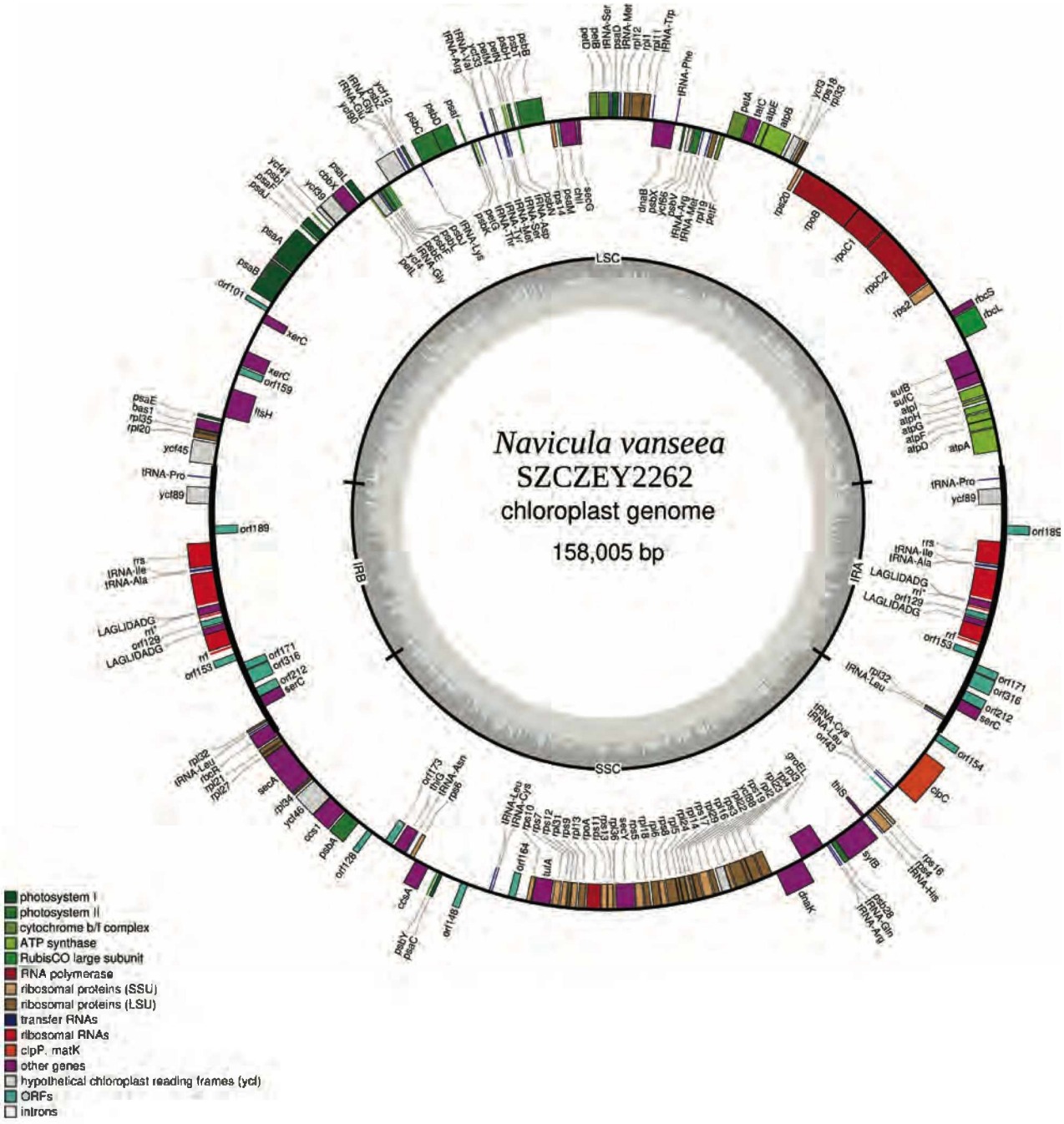


Figure 6. Map of the plastid genome of *Navicula vanseea* sp. nov. SZCZEY2262.

both containing two putative *LAGLIDADG* homing endonuclease genes. They will be referred to as L1917 and L1931.

For SZCZEY2172 (GenBank: OR795086), the LSC is 72,913 bp long and has an identical gene content compared to SZCZEY2262, the SSC is 49727 bp long and encodes 51 conserved protein-coding genes, eight tRNAs and six non-conserved ORFs of more than 100 AA. The IR is 17,675 bp long and has an identical gene and intron content to SZCZEY2262, with the same overlap of *rbcR* between the IRB and the SSC.

Both genomes contain a 43 AA ORF in their SSC that cannot be extended because of the presence of stop-codons. This ORF shows similarities to the hypothetical chloroplast RF35 encoded by *ycf35*, which is missing between both

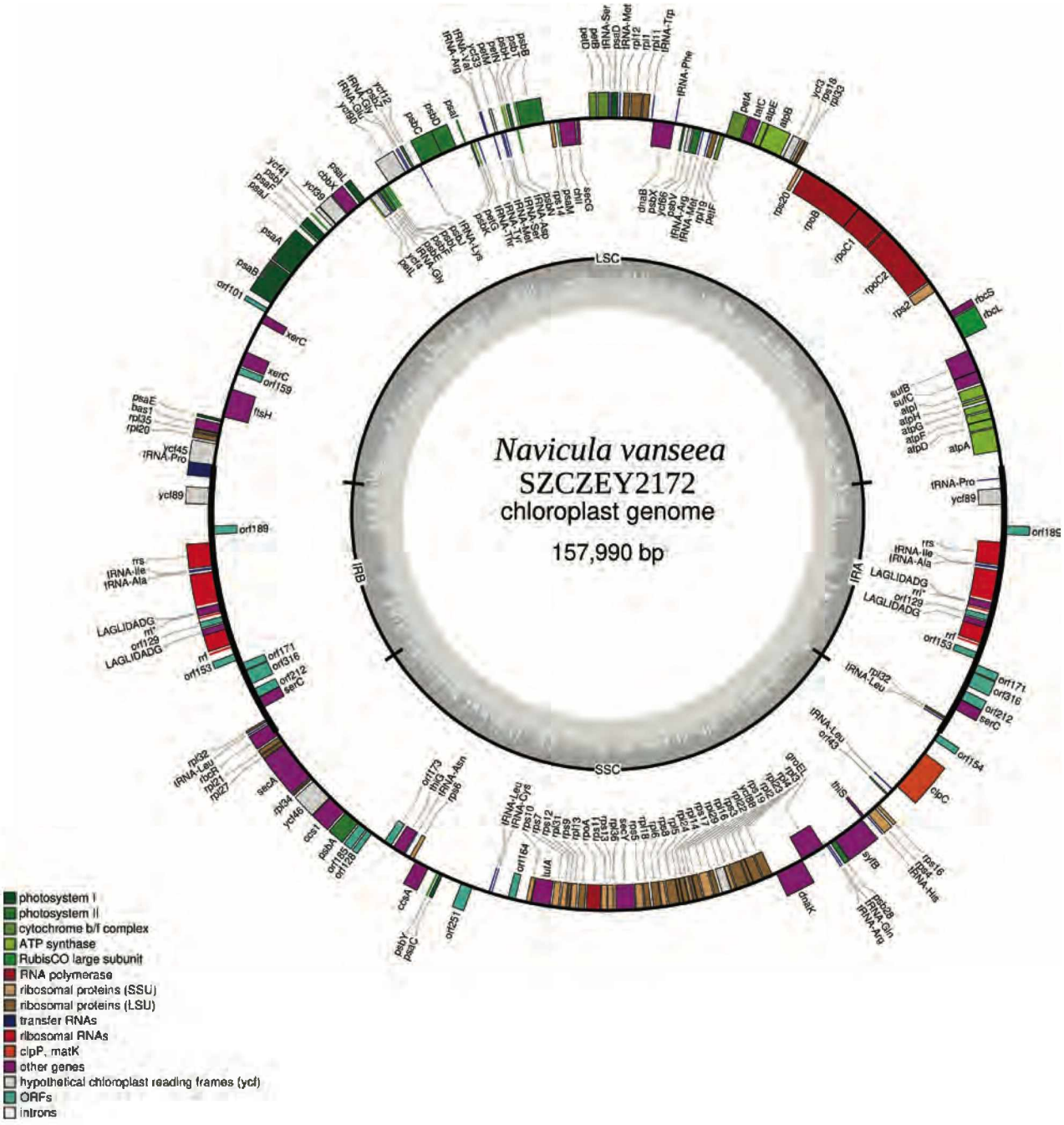


Figure 7. Map of the plastid genome of *Navicula vanseea* sp. nov. SZCZEY2172.

strains. The position of this ORF also corresponds to the position of *ycf35* in *Navicula veneta*, between *clpC* and *rps13* (Gastineau et al. 2021a).

It is worth noting that, in addition to the differences in length and content in the non-conserved ORFs, there is a slight degree of extra polymorphism in the two strains, the extent depending on the part of the genome considered. There were only two SNPs in the inverted repeat (one in the spacer between *ycf45* and *tRNA-Pro* and the other inside *rbcR*). On the other hand, a gene such as *psbC* displayed three SNPs, two of them silent, but one leading to a phenylalanine–leucine substitution. The two *xerC* genes, although present in both strains, differed in length.

Multigene phylogeny

The subtree containing Naviculaceae (Fig. 8) has been extracted from the complete multigene tree (downloadable as explained in the data availability statement). *Navicula vanseea* sp. nov. appears in a well-supported (99%) clade that also contains the freshwater diatom *Navicula cryptocephala* UTEX FD109 (sometimes indexed as *Navicula cryptocephala var. veneta* UTEX FD109), a specimen isolated by the late David B. Czarnecki from North Dakota, USA (Theriot et al. 2010) and the marine *Navicula* sp. KSA2015 41, which originates from the vicinity of Rabigh on the Red Sea, Saudi Arabia (Sabir et al. 2018). The clade also contains *Pseudogomphonema* sp. and two species of *Seminavis* spp. The phylogeny unambiguously separates *N. vansee* sp. nov. from the two morphologically similar species *N. cincta* and *N. microdigitoradiata*. *Navicula cincta* appears in a different, strongly-supported clade (97%) that contains *N. capitatoradiata*, *Navicula tsukamotoi* (Sterrenburg & Hinz) Yuhang Li & Kuidong Xi 2017, several unnamed species of *Navicula* spp. and *Rhoikoneis pagoensis* Lobban, 2015. *Navicula microdigitoradiata* is also easily distinguished and appears as sister to *Navicula hippodontofallax* Witkowski & Chulian Li 2016.

Phylogeny of the putative homing endonuclease LAGLIDADG proteins

Once rooted with sequence ABR25263, the phylogenetic tree of LAGLIDADG proteins (Fig. 9) distinguished the two groups. The tree associates sequences from *N. vanseea* SZCZEY2172 and SZCZEY2262 with those in other species that are of the same type and occupy the same positions. For example, the L1931 LAGLIDADG sequences of the two *N. vanseea* clones were found to be sister to a L1931 LAGLIDADG in the plastid genome of the diatom *Schizostauron trachyderma* (F. Meister) Górecka, Riaux-Gobin & Witkowski, 2021 (Górecka et al. 2021b) and then to the green algae *Pterosperma cristatum* Schiller, 1925 (Prasinophyceae) and *Pedinomonas tuberculata* (Vischer) Gams, 1947 (Pedinophyceae), a synonym of *Chlorochytridion tuberculatum* Vischer 1945 (both from plastid genomes). In contrast to the topology of the L1931 clade, in the L1917 tree, the *N. vanseea* LAGLIDADG sequences appeared at the base of the clade with maximum support. In this clade, sequences from the plastomes of various Viridiplantae form a strong clade, separated from *N. vanseea* by two Prokaryota, namely the heterotrophic bacteria *Pseudothermotoga thermarum* (Windberger et al. 1992) Bhandari and Gupta 2014 and the cyanobacterium *Synechococcus* sp. C9.

Discussion

Morphological comparison with similar taxa

Navicula vanseea sp. nov. has elliptic valves that taper towards cuneately rounded apices in smaller specimens and linear-elliptic-lanceolate with narrowly rounded, protracted endings in larger specimens (Table 1). *Navicula cincta*, *Navicula dealpina* Lange-Bert., *N. microdigitoradiata*, *N. meulemansii* and *Navicula veronensis* Lange-Bertalot and Cantonati are the most similar taxa. Regarding valve outline, *Navicula dealpina* is very similar, being linear-lanceolate to almost elliptic.

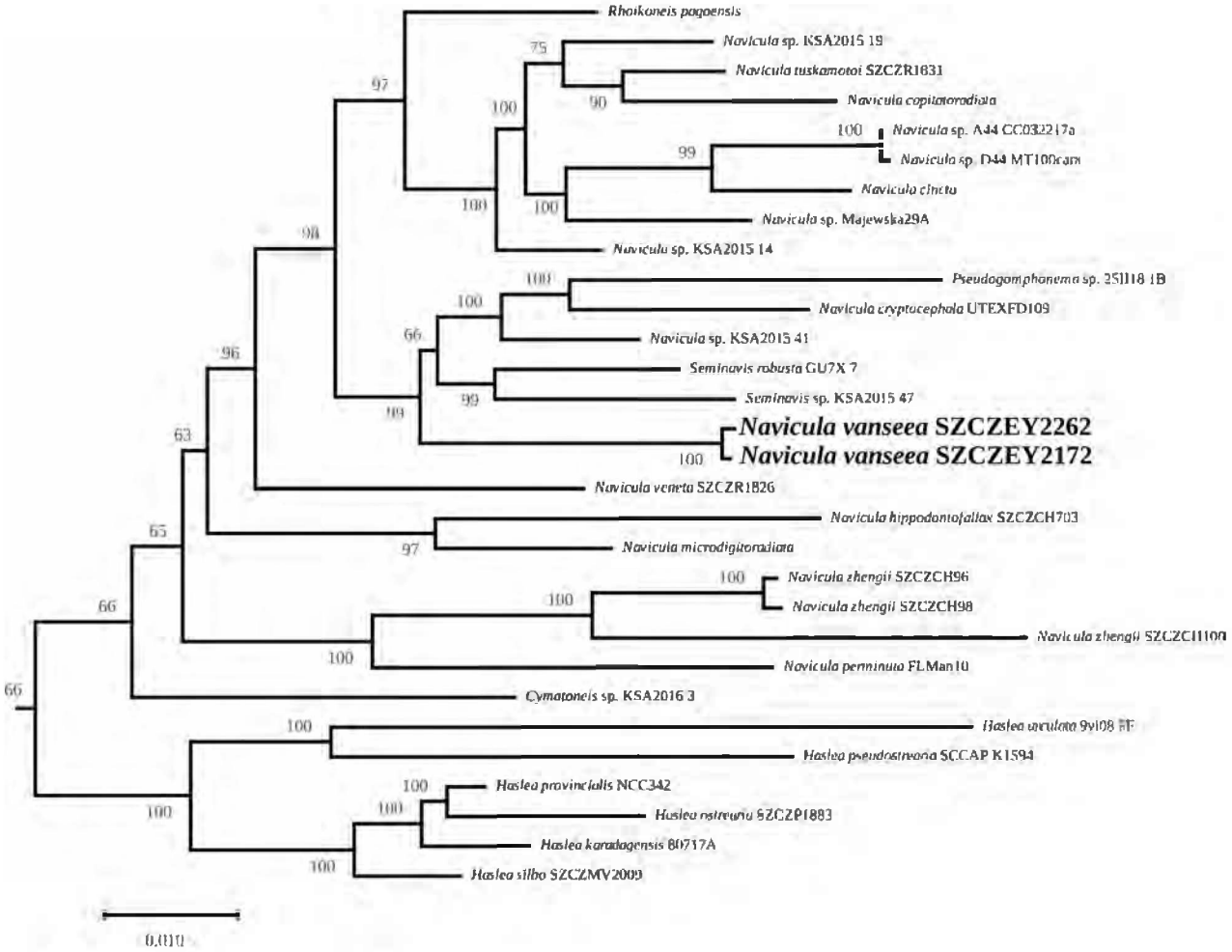


Figure 8. Maximum Likelihood phylogenetic tree obtained from concatenated alignments of *psbC*, *rbcL* and *18S*.

However, the apices of *N. dealpina* are much narrower compared to the bluntly rounded apices of *N. vanseea*. Moreover, *N. dealpina* is larger (26–86 µm length, 8–12 µm width) and has a lower stria density (8–10 striae in 10 µm) than *N. vanseea* (11–28 µm length, 5–6 µm width and 12–13 striae in 10 µm) and it has a transversely rectangular central area (not elliptic and small in *N. vanseea*). Regarding dimensions, *N. veronensis* and *N. meulemansii* are quite similar. *Navicula veronensis* has similarly linear-lanceolate valve outline, but *N. vanseea* is more elliptic. Moreover, *N. veronensis* has relatively larger and more visible central area with gradually wedge-shaped, finally obtusely rounded apices, while *N. vanseea* has an indistinctive central area with narrowly rounded apices. In SEM, *N. veronensis* has a thicker sternum structure than *N. vanseea*, while *N. vanseea* has more strongly-hooked distal raphe endings. Similarly, *N. meulemansii* have thicker sternum structure in SEM. Additionally, *N. meulemansii* has elliptic-lanceolate valve outline with cuneate apices and higher striae density than *N. vanseea* (15–21 in 10 µm vs. 12–13 in 10 µm, respectively). Amongst the similar taxa, *Navicula microdigitoradiata* and *N. cincta* have slightly larger sizes (up to 40–45 µm length and 7–8 µm width) and lower striae density (max. 10 in 10 µm for *N. cincta* and 11 in 10 µm for *N. microdigitoradiata*) than *N. vanseea* (max. 28 µm length, 6 µm width and 13 striae in 10 µm). *N. meulemansii* has an almost similar length with narrower specimens (3–5 µm width) and higher striae density (15–21 stri-

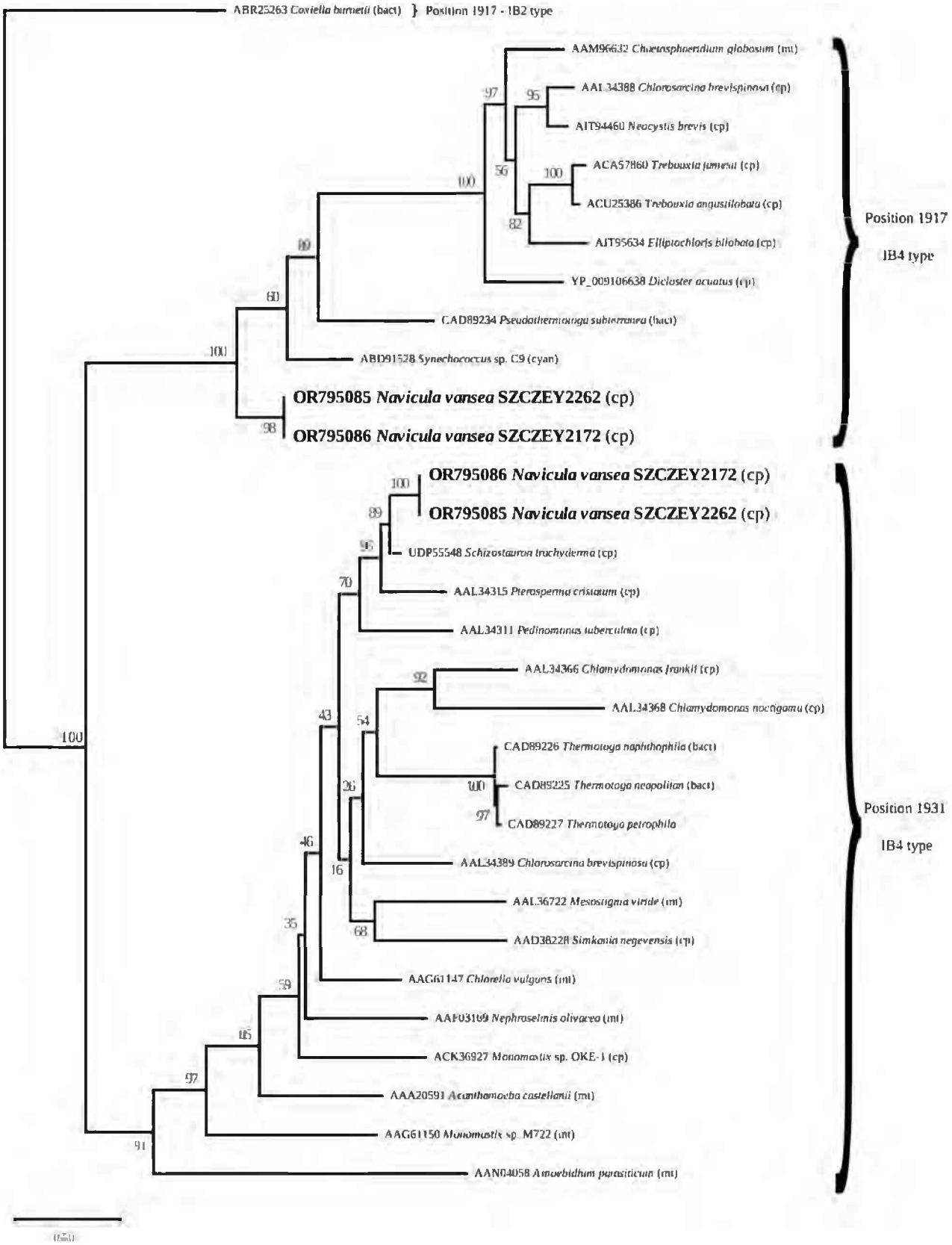


Figure 9. Maximum Likelihood phylogenetic tree inferred from the alignment of the putative LAGLIDADG endonuclease proteins found in the group I introns of *Navicula vanseea* sp. nov. and other taxa. The type of genome is indicated between brackets: cp – plastome, mt – mitogenome, bact – bacteria, cyan – cyanobacteria.

ae in 10 µm). The valve outline is relatively more elliptic with cuneately rounded apices in *N. meulemansii*, while *N. vanseea* has a more linear-lanceolate outline with narrowly-rounded apices (Table 1). It is worth mentioning that *N. cincta*, *N. veronensis* and *N. dealpina* are also all species suspected to prefer alkaline water (Cantonati et al. 2016), so it is especially important to clarify the distinctions between them and *N. vanseea*.

In relation to the three taxa mentioned in the introduction as having been found in Van Lake by Legler and Krasske (1940) – *N. cryptocephala*, *N. capitatoradiata* and *N. veneta* – *N. vansee* is easily distinguished in LM by the shape of its apices, which are rounded, while all the others have narrow, protracted apices. Molecular phylogeny, despite the limitations in the sampling of taxa, also concurs to discriminate *N. vanseea* sp. nov. from these three species (Fig. 8).

Genetic polymorphisms and genome evolution

The organellar genomes, especially the plastomes, show some interesting features. For example, introns are not considered to be conserved genetic elements and are known to vary amongst isolates of a single species (e.g. Gastineau et al. (2021b)), but they were fully conserved between our two isolates, whereas protein-coding genes displayed non-silent polymorphisms. In addition, all three markers commonly used for phylogeny reconstruction in diatoms (18S, *rbcL*

Table 1. Comparison of *Navicula vanseea* sp. nov. and similar species.

	***Navicula vanseea* sp. nov.**	***Navicula meulemansii***	***Navicula microdigitoradiata***	***Navicula cincta***	***Navicula cariocincta***	***Navicula veronensis***	***Navicula dealpina***
Valve length (µm)	11.5–28.5	12–30	15–40	14–45	30–50	19–40	26–86
Valve width (µm)	5–6	3.5–5.5	5–7	5.5–8	5.5–7	4–7	8–12
Stria density (in 10 µm)	12–13	15–21	10–11	8–10	10–12	11–13	8–10
Striae around central area	irregularly shortened	irregularly shortened	single shortened on either side	irregularly shortened	irregularly shortened	2–4 striae distinctly shortened on either side	irregularly shortened
Lineolae density (in 10 µm)	ca. 50	ca. 44–51	–	ca. 40	ca. 30	ca. 50	ca. 26 (LM)
Valve shape	smaller elliptic, larger linear-elliptic-lanceolate	elliptic-lanceolate	elliptic-lanceolate to linear-lanceolate	elliptic to lanceolate to linear-elliptic-lanceolate	linear- elliptic -lanceolate	linear-lanceolate	linear-lanceolate to almost elliptical
Central area	elliptic and small	very small and asymmetric	elliptic and very small	small	large, transversely rectangular to elliptical	broadly rectangular or transapically elliptic	almost symmetrical, transversally rectangular
Valve apices	narrowly rounded	cuneately rounded	obtusely rounded	obtusely rounded	narrowed to a wedge, obtusely-rounded, never protracted	gradually wedge-shaped, finally obtusely rounded	obtusely wedge-shaped
Raphe	filiform	weakly lateral; central pores very close together	weakly lateral, central pores very close	filiform	strongly radiate	filiform	weakly to strongly lateral, lying outside the median area towards the central pores
References	this study	Mertens et al. (2014)	Lange-Bertalot (2001)	Lange-Bertalot (2001)	Tsarenko et al. (2000)	Cantonati et al. (2016)	Lange-Bertalot (2001)

and *psbC*) exhibited single nucleotide polymorphisms in *N. vanseea*; this was especially surprising for the nuclear 18S gene, which generally exhibits very few differences between closely-related species (Evans et al. 2007). The polymorphism in *18S* appeared to be in the variable V2 region, while the V4 or V9 regions, which are often used in metabarcoding studies, were found fully conserved.

In the plastome of *N. vanseea*, the *ycf35* gene has seemingly been turned to a pseudogene, which would be the first time to our knowledge that this has been observed in diatoms, although *ycf35* pseudogenes have been observed in Rhodophyta (Costa et al. 2016). This gene seems to be lost altogether amongst other taxa, such as *Rhizosolenia imbricata* Brightwell, 1858 (Sabir et al. 2014) or *Proboscia* sp. and *Licmophora* sp. (Yu et al. 2018). It is not clear if the gene has been completely lost in these species or if it has been transferred to the nuclear genome, which is known to have happened with the plastid *petF* gene in *Thalassiosira oceanica* (Lommer et al. 2010). In *N. vanseea*, the *ycf35* gene is likely no longer functional. The size of the Ycf35 protein amongst diatoms is ca. 1130 AA long. Its origin can be traced back to Cyanobacteria. Its function is unknown, but it has been suggested, based on experiments conducted on *Synechocystis*, to participate in CO_2 capture (Jiang et al. 2015). The results obtained from *N. vansee* sp. nov. suggest that it is not necessary to its metabolism and survival, unless *ycf35* is also already present in the nucleus.

Our study also illustrates the added value that next generation sequencing provides when describing new species, in three ways. First, it is a convenient way to gather data for multigene phylogenies, whatever the species considered. Second, in the current case with *N. vanseea*, it made it possible to find SNPs in supposedly conserved genes of two sympatric strains of the same species. This needs to be taken into account in interpreting phylogenetic and metabarcoding analyses. Third, serendipitous discoveries can occur that increase our knowledge of the organellar genomes of diatoms and other stramenopiles, such as the loss of a functional *ycf35* gene here or the first documented L1917 intron found in a stramenopile.

Acknowledgements

This work started under the supervision of our beloved colleague and mentor, the late Prof. Dr. Hab. Andrzej Witkowski, who passed away on 17 September 2023. His guidance is already deeply missed.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Funding

This work was co-financed by the Minister of Science under the "Regional Excellence Initiative" Program (2024–2027). Claude Lemieux and Monique Turmel were supported by grant RGPIN-2017-04506 from the Natural Sciences and Engineering Research Council

of Canada (NSERC). We also acknowledge support from the CERCA Programme/Generalitat de Catalunya. The Royal Botanic Garden Edinburgh is supported by the Scottish Government's Rural and Environment Science and Analytical Services Division.

Author contributions

Conceptualisation: EY. Data curation: EY. Funding acquisition: AW. Investigation: EY, RG, CNS, EG, CL, MT, NE. Methodology: AW, RT, DGM. Project administration: CNS, AW. Supervision: AW, DGM, RT, CNS, RG. Visualisation: EY, NE, CNS, EG. Writing - original draft: EY. Writing - review and editing: DGM, RT, CL, MT, CNS, RG.

Author ORCIDs

Elif Yılmaz https://orcid.org/0000-0001-7939-1814
David G. Mann https://orcid.org/0000-0003-0522-6802
Romain Gastineau https://orcid.org/0000-0001-8661-5118
Rosa Trobajo https://orcid.org/0000-0001-9498-3797
Cüneyt Nadir Solak https://orcid.org/0000-0003-2334-4271
Ewa Górecka https://orcid.org/0000-0003-0590-7480
Monique Turmel https://orcid.org/0000-0001-7060-035X
Claude Lemieux https://orcid.org/0000-0001-9580-8042
Nesil Ertorun https://orcid.org/0000-0001-6224-7314
Andrzej Witkowski https://orcid.org/0000-0003-1714-218X

Data availability

All sequences have been deposited and are already available on GenBank with accession numbers OR797293, OR797294, OR795084, OR795085, OR795086. The alignment, partition file and complete tree in Newick format can be obtained from Zenodo following this link: https://doi.org/10.5281/zenodo.10518939.

References

Adrienne Mertens AW, Lange-Bertalot H, Ribeiro L, Rhiel E (2014) *Navicula meulemansii* sp. nov., (Bacillariophyaceae) from brackish waters in Europe and the USA. Nova Hedwigia 98(1–2): 201–212. https://doi.org/10.1127/0029-5035/2013/0152

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19(5): 455–477. https://doi.org/10.1089/cmb.2012.0021

Bory de Saint-Vincent JBGM (1822) Bacillariées. Dictionnaire classique d'histoire naturelle, 2. Rey et Gravier, Libraries-Editeurs, Paris, 127–129.

Cantonati M, Angeli N, Spitale D, Lange-Bertalot H (2016) A new *Navicula* (Bacillariophyta) species from low-elevation carbonate springs affected by anthropogenic disturbance. Fottea 16(2): 255–265. https://doi.org/10.5507/fot.2016.013

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics (Oxford, England) 25(15): 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Costa JF, Lin SM, Macaya EC, Fernández-García C, Verbruggen H (2016) Chloroplast genomes as a tool to resolve red algal phylogenies: A case study in the Nemaliales. BMC Evolutionary Biology 16(1): 205. https://doi.org/10.1186/s12862-016-0772-3

Cox EJ (1979) Taxonomic studies on the diatom genus *Navicula* Bory. The typification of the genus. Bacillaria 2: 137–153.

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

Dąbek P, Ashworth MP, Witkowski A, Li C, Bornman TG, Gonçalves V, Park J, Khim JS (2017) Towards a multigene phylogeny of the Cymatosiraceae (Bacillariophyta, Mediophyceae) I: Novel taxa within the subfamily cymatosiroideae based on molecular and morphological data. Journal of Phycology 53(2): 342–360. https://doi.org/10.1111/jpy.12501

Dąbek P, Gastineau R, Bornman TG, Lemieux C, Turmel M, Hallegraeff G, Mouget J-L, Witkowski A (2022) The blue diatom *Haslea ostrearia* from the Indian Ocean coast of South Africa, with comparative analysis of *Haslea* organellar genomes. Frontiers in Marine Science 9: 950716. https://doi.org/10.3389/fmars.2022.950716

Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T (2020) ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. Molecular Biology and Evolution 37(1): 291–294. https://doi.org/10.1093/molbev/msz189

Dierckxsens N, Mardulyn P, Smits G (2017) NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Research 45(4): e18. https://doi.org/10.1093/nar/gkw955

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus (San Francisco, Calif.) 12: 13–15.

Ersoy Omeroglu E, Sudagidan M, Yurt MNZ, Tasbasi BB, Acar EE, Ozalp VC (2021) Microbial community of soda Lake Van as obtained from direct and enriched water, sediment and fish samples. Scientific Reports 11(1): 18364. https://doi.org/10.1038/s41598-021-97980-3

Evans KM, Wortley AH, Mann DG (2007) An assessment of potential diatom "barcode" genes (*cox1*, *rbcL*, *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

Gastineau R, Hamedi C, Baba Hamed MB, Abi-Ayad S-ME-A, Bąk M, Lemieux C, Turmel M, Dobosz S, Wróbel RJ, Kierzek A, Lange-Bertalot H, Witkowski A (2021a) Morphological and molecular identification reveals that waters from an isolated oasis in Tamanrasset (extreme South of Algerian Sahara) are colonized by opportunistic and pollution-tolerant diatom species. Ecological Indicators 121(6): 107104. https://doi.org/10.1016/j.ecolind.2020.107104

Gastineau R, Hansen G, Poulin M, Lemieux C, Turmel M, Bardeau JF, Leignel V, Hardivillier Y, Morançais M, Fleurence J, Gaudin P, Méléder V, Cox EJ, Davidovich NA, Davidovich OI, Witkowski A, Kaczmarska I, Ehrman JM, Soler Onís E, Quintana AM, Mucko M, Mordret S, Sarno D, Jacquette B, Falaise C, Séveno J, Lindquist NL, Kemp Jr PS, Eker-Develi E, Konucu M, Mouget JL (2021b) *Haslea silbo*, A Novel Cosmopolitan Species of Blue Diatoms. Biology (Basel) 10(4): 328. https://doi.org/10.3390/biology10040328

Gastineau R, Li C, Ashworth MP, Witkowski A, Turmel M, Górecka E, Frankovich TA, Wachnicka A, Lobban CS, Theriot EC, Otis C, Dąbek P, Binczewska A, Lemieux C (2022) A New Insight Into *Amicula*, a Genus of Tiny Marine Benthic Diatoms With the Description of Two New Tropical Species and the Largest Mitogenome Known for a Stramenopile. Frontiers in Marine Science 9: 941506. https://doi.org/10.3389/fmars.2022.941506

Glombitza C, Stockhecke M, Schubert CJ, Vette A, Kallmeyer J (2013) Sulfate reduction controlled by organic matter availability in deep sediment cores from the saline, alkaline Lake Van (Eastern Anatolia, Turkey). Frontiers in Microbiology 4: 209. https://doi.org/10.3389/fmicb.2013.00209

Gordon D, Green P (2013) Consed: A graphical editor for next-generation sequencing. Bioinformatics (Oxford, England) 29(22): 2936–2937. https://doi.org/10.1093/bioinformatics/btt515

Górecka E, Ashworth MP, Davidovich N, Davidovich O, Dąbek P, Sabir JSM, Witkowski A (2021a) Multigene phylogenetic data place monoraphid diatoms *Schizostauron* and *Astartiella* along with other fistula-bearing genera in the Stauroneidaceae. Journal of Phycology 57(5): 1472–1491. https://doi.org/10.1111/jpy.13192

Górecka E, Gastineau R, Davidovich NA, Davidovich OI, Ashworth MP, Sabir JSM, Lemieux C, Turmel M, Witkowski A (2021b) Mitochondrial and Plastid Genomes of the Monoraphid Diatom *Schizostauron trachyderma*. International Journal of Molecular Sciences 22(20): 11139. https://doi.org/10.3390/ijms222011139

Guillard RRL (1975) Culture of phytoplankton for feeding marine invertebrates. In: Smith WL, Chanley MH (Eds) Culture of Marine Invertebrate Animals. Plenum Press, New York, 26–60. https://doi.org/10.1007/978-1-4615-8714-9_3

Haugen P, Bhattacharya D (2004) The spread of LAGLIDADG homing endonuclease genes in rDNA. Nucleic Acids Research 32(6): 2049–2057. https://doi.org/10.1093/nar/gkh520

Haugen P, Bhattacharya D, Palmer JD, Turner S, Lewis LA, Pryer KM (2007) Cyanobacterial ribosomal RNA genes with multiple, endonuclease-encoding group I introns. BMC Evolutionary Biology 7(1): 159. https://doi.org/10.1186/1471-2148-7-159

Jiang HB, Song WY, Cheng HM, Qiu BS (2015) The hypothetical protein Ycf46 is involved in regulation of CO_2 utilization in the cyanobacterium *Synechocystis* sp. PCC 6803. Planta 241(1): 145–155. https://doi.org/10.1007/s00425-014-2169-0

Kalvari I, Nawrocki EP, Ontiveros-Palacios N, Argasinska J, Lamkiewicz K, Marz M, Griffiths-Jones S, Toffano-Nioche C, Gautheret D, Weinberg Z, Rivas E, Eddy SR, Finn RD, Bateman A, Petrov AI (2021) Rfam 14: Expanded coverage of metagenomic, viral and microRNA families. Nucleic Acids Research 49(D1): D192–D200. https://doi.org/10.1093/nar/gkaa1047

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. https://doi.org/10.1093/molbev/mst010

Lange-Bertalot H (2001) *Navicula sensu stricto* 10 genera separated from *Navicula sensu lato*, *Frustulia*. A.R.G. Gantner Verlag K.G, Ruggell.

Legler F, Krasske G (1940) Diatomeen aus dem Vansee (Armenien). Beihefte zur Botanischen Zentralblatt 60, Abteilung B, 335–345.

Lemieux C, Otis C, Turmel M (2014) Chloroplast phylogenomic analysis resolves deep-level relationships within the green algal class Trebouxiophyceae. BMC Evolutionary Biology 14(1): 211. https://doi.org/10.1186/s12862-014-0211-2

Li C, Ashworth MP, Mackiewicz P, Dąbek P, Witkowski J, Górecka E, Krzywda M, Witkowski A (2020) Morphology, phylogeny, and molecular dating in Plagiogrammaceae family focused on *Plagiogramma-Dimeregramma* complex (Urneidophycidae, Bacillariophyceae). Molecular Phylogenetics and Evolution 148: 106808. https://doi.org/10.1016/j.ympev.2020.106808

Lohse M, Drechsel O, Kahlau S, Bock R (2013) OrganellarGenomeDRAW – A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing

expression data sets. Nucleic Acids Research 41(W1): W575–W581. https://doi.org/10.1093/nar/gkt289

Lommer M, Roy AS, Schilhabel M, Schreiber S, Rosenstiel P, LaRoche J (2010) Recent transfer of an iron-regulated gene from the plastid to the nuclear genome in an oceanic diatom adapted to chronic iron limitation. BMC Genomics 11(1): 718. https://doi.org/10.1186/1471-2164-11-718

Lucas P, Otis C, Mercier JP, Turmel M, Lemieux C (2001) Rapid evolution of the DNA-binding site in LAGLIDADG homing endonucleases. Nucleic Acids Research 29(4): 960–969. https://doi.org/10.1093/nar/29.4.960

Millie DF, Wee JL (1981) Observations on the tube-dwelling diatom *Navicula tripunctata* var. *schizonemoides* (VH) Patr. using light and electron microscopy. Proceedings of the Iowa Academy of Science 88(2): 74–78. https://scholarworks.uni.edu/pias/vol88/iss2/8

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution 37(5): 1530–1534. https://doi.org/10.1093/molbev/msaa015

Patrick R (1959) New species and nomenclatural changes in the genus *Navicula* (Bacillariophyceae). Proceedings of the Academy of Natural Sciences of Philadelphia 111: 91–108. https://www.jstor.org/stable/4064508

Pogoda CS, Keepers KG, Hamsher SE, Stepanek JG, Kane NC, Kociolek JP (2019) Comparative analysis of the mitochondrial genomes of six newly sequenced diatoms reveals group II introns in the barcoding region of *cox1*. Mitochondrial DNA, Part A, DNA Mapping, Sequencing, and Analysis 30(1): 43–51. https://doi.org/10.1080/24701394.2018.1450397

Raghavan R, Miller SR, Hicks LD, Minnick MF (2007) The unusual 23S rRNA gene of *Coxiella burnetii*: Two self-splicing group I introns flank a 34-base-pair exon, and one element lacks the canonical omegaG. Journal of Bacteriology 189(18): 6572–6579. https://doi.org/10.1128/JB.00812-07

Reimer A, Landmann G, Kempe S (2009) Lake Van, eastern Anatolia, hydrochemistry and history. Aquatic Geochemistry 15(1): 195–222. https://doi.org/10.1007/s10498-008-9049-9

Round FE, Crawford RM, Mann DG (1990) The Diatoms. Biology & morphology of the genera. Cambridge Univ. Press, Cambridge, 747 pp.

Sabir JS, Yu M, Ashworth MP, Baeshen NA, Baeshen MN, Bahieldin A, Theriot EC, Jansen RK (2014) Conserved gene order and expanded inverted repeats characterize plastid genomes of Thalassiosirales. PLOS ONE 9(9): e107854. https://doi.org/10.1371/journal.pone.0107854

Sabir JSM, Theriot EC, Manning SR, Al-Malki AL, Khiyami MA, Al-Ghamdi AK, Sabir MJ, Romanovicz DK, Hajrah NH, El Omri A, Jansen RK, Ashworth MP (2018) Phylogenetic analysis and a review of the history of the accidental phytoplankter, Phaeodactylum tricornutum Bohlin (Bacillariophyta). PLOS ONE 13(6): e0196744. https://doi.org/10.1371/journal.pone.0196744

Smith SA, Dunn CW (2008) Phyutility: A phyloinformatics tool for trees, alignments and molecular data. Bioinformatics (Oxford, England) 24(5): 715–716. https://doi.org/10.1093/bioinformatics/btm619

Solak CN, Gastineau R, Lemieux C, Turmel M, Gorecka E, Trobajo R, Rybak M, Yılmaz E, Witkowski A (2021) *Nitzschia anatoliensis* sp. nov., a cryptic diatom species from the highly alkaline Van Lake (Turkey). PeerJ 9: e12220. https://doi.org/10.7717/peerj.12220

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

Tsarenko PM, Lange-Bertalot H, Stupina VV, Wasser SP (2000) Bacillariophyta. Biodiversity of cyanoprocaryotes, algae and fungi of Israel: Cyanoprocaryotes and algae of continental Israel. A.R.G. Gantner Verlag K.G, Ruggell.

Turmel M, Otis C, Lemieux C (2002) The chloroplast and mitochondrial genome sequences of the charophyte *Chaetosphaeridium globosum*: Insights into the timing of the events that restructured organelle DNAs within the green algal lineage that led to land plants. Proceedings of the National Academy of Sciences of the United States of America 99(17): 11275–11280. https://doi.org/10.1073/pnas.162203299

Yu M, Ashworth MP, Hajrah NH, Khiyami MA, Sabir MJ, Alhebshi AM, Al-Malki AL, Sabir JSM, Theriot EC, Jansen RK (2018) Chapter five – evolution of the plastid genomes in diatoms. Advances in Botanical Research 85: 129–155. https://doi.org/10.1016/bs.abr.2017.11.009

PhytoKeys @@@: #–## (2024)
DOI: 10.3897/phytokeys.@@@.133205

Research Article

Morphological and molecular characterization of *Halamphora vantushpaensis* (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye)

Elif Yılmaz[1], Romain Gastineau[1], Cüneyt Nadir Solak[2], Ewa Górecka[1], Rosa Trobajo[3], Monique Turmel[4], Claude Lemieux[4], Christian Otis[5], Andrzej Witkowski[1†], David G. Mann[3,6]

1 *Institute of Marine and Environmental Sciences, University of Szczecin, Mickiewicza 16A, PL70–383 Poland*
2 *Department of Biology, Faculty of Science & Art, Dumlupınar University, 43000 Kütahya, Turkiye*
3 *Marine and Continental Waters, Institute for Food and Agricultural Research and Technology (IRTA), Crta de Poble Nou Km 5.5, E-43540 La Ràpita, Catalunya, Spain*
4 *Département de biochimie, de microbiologie et de bio-Informatique, Institut de Biologie Intégrativeet des Systèmes, Université Laval, Québec, QC, Canada*
5 *Plateforme d'Analyse Génomique, Institut de biologie intégrative et des systèmes, Université Laval, Québec, QC, Canada*
6 *Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, Scotland, UK*

Corresponding authors: Elif Yılmaz (elfyilmaz38@gmail.com); Romain Gastineau (romain.gastineau@usz.edu.pl)

Abstract

In this study, we describe *Halamphora vantushpaensis* **sp. nov.**, a newly identified diatom species found in the highly alkaline Lake Van in Eastern Turkey (Türkiye). This new species is characterized morphologically by light and scanning electron microscopy, performed on both wild and cultivated samples. Two monoclonal cultures were submitted to a genome-skimming approach, giving access to the complete sequence of their nuclear rRNA cluster of genes, mitochondrial and plastid genomes. Both strains were highly similar from the genomic point of view, with few mutations noted, although in organellar genomes some of them concerned protein coding genes and were non-silent. Also, the group II intron in the mitochondrial *cox1* gene was found to display a relatively high number of polymorphisms. The plastome also distinguishes itself from other *Halamphora* spp. by the extension of its inverted repeat at the expense of the two single copy regions of the genome. Maximum likelihood molecular phylogeny inferred from a concatenated three genes dataset (*18S*, *psbC* and *rbcL*) positions this species within the K clade, which is known to contain hypersaline to freshwater species.

Key words: Alkaline lake, group II introns, inverted repeat, mitochondrial genome, multigene phylogeny, nuclear rRNA genes, plastid genome

Academic editor: Kalina Manoylov
Received: 26 July 2024
Accepted: 4 September 2024
Published: ## @@@ 2024

Citation: Yilmaz E, Gastineau R, Solak CN, Górecka E, Trobajo R, Turmel M, Lemieux C, Otis C, Witkowski A, Mann DG (2024) Morphological and molecular characterization of *Halamphora vantushpaensis* (Bacillariophyceae, Amphipleuraceae), a new diatom species widely dispersed on the shores of the soda Lake Van (Türkiye). PhytoKeys @@@: #–##. https://doi.org/10.3897/phytokeys.@@@.133205

Introduction

Soda lakes are among the rarest and most geochemically distinctive wetlands on Earth. They are characterized by their alkaline waters containing high levels of carbonate and bicarbonate ions, typically resulting in elevated pH levels. Lake Van is the largest soda lake in the world (Glombitza et al.

† Deceased author.

2013), with water that is both saline (21.4‰) and alkaline (155 m mEq^{-1}, pH 9.81) (Aydar et al. 2003; Kempe et al. 1991). The lake has existed for 600,000 years, spanning multiple glacial–interglacial cycles (Stockhecke et al. 2014; North et al. 2018) and hosts endemic species of fishes (e.g. Akku et al. 2021). However, studies on the phytoplanktonic flora of the lake, and especially diatoms, have been rather scarce.

For a long time, investigations into the contemporary diatom flora from Lake Van were restricted to a single study by Legler and Krasske (1940). These authors worked on samples brought to them in Germany from Lake Van and, based on light microscopy, they found 24 diatom taxa. Among them, two belonged to the genus *Amphora*. The first one was described as '*Amphora coffeiformis* Ag.' (more correctly referred to as *Amphora coffeiformis* (C.Agardh) Kützing, 1844). This is now a non-accepted synonym of *Halamphora coffeiformis* (C.Agardh) Levkov following the recent revisions of the amphoroid diatoms in which Cleve's sect. *Halamphora* has been recognized at the genus level (Cleve 1895; Levkov 2009; Stepanek and Kociolek 2013, 2014, 2019). The second was '*Amphora commutata* Grun.', published in Van Heurck (1880), considered a valid species as of today. Both taxa are known for being brackish species and both have wide distributions (Guiry and Guiry 2024).

In the last few years, new investigations have been conducted on Lake Van diatoms using an integrative approach that combines light/scanning electron microscopy (LM/SEM) and molecular phylogenies derived from next generation sequencing results. With these data, three new species already have been discovered and described, namely *Nitzschia anatoliensis* Górecka, Gastineau and Solak (Solak et al. 2021), *Navicula vanseea* Yılmaz, Gastineau, Solak and Witkowski (Yilmaz et al. 2024) and *Halamphora witkowskii* (Yilmaz et al. in press). Lake Van can be divided into four different basins of various depths: the shallow Erçis basin (northeast), the Van basin (southeast), the Ahlat basin (northwest) and, at the centre of the lake, a deeper fourth basin (Kaden et al. 2010). *Navicula vanseea* and *Ni. anatoliensis* were both described from material from the Erciş basin but have not been observed yet in other parts of the lake. *Halamphora witkowskii* is known so far only from the Ahlat basin (Yilmaz et al. in press).

In the current article, we describe another *Halamphora* species, *Halamphora vantushpaensis* sp. nov., using the same integrative approach and tools previously used for *Na. vanseea* and *Ni. anatoliensis*.

Material and methods

Sampling, isolation and cultivation

Live samples were collected by scraping stones at four different stations around Lake Van: Ahlat, Bitlis (38°75'45.748"N, 42°50'71.257"E); Erciş, Van (39°00'07.9"N, 43°25'40.4"E); Adilcevaz, Bitlis (38°79'83"N, 42°72'16"E); and Edremit, Van (38°42'07.09"N, 43°23'74.39"E) (Fig. 1). Individual diatom cells were isolated from the Ahlat samples using a micropipette under a Nikon TS100 inverted microscope (NIKON, Tokyo, Japan). The strains were subsequently moved into 250 mL Erlenmeyer flasks containing F/2 medium (Guillard, 1975), which had been adjusted to 18‰ salinity. The cultures were maintained in conditions promoting active growth, with a light intensity of 60 µmol photons m^{-2} s^{-1},

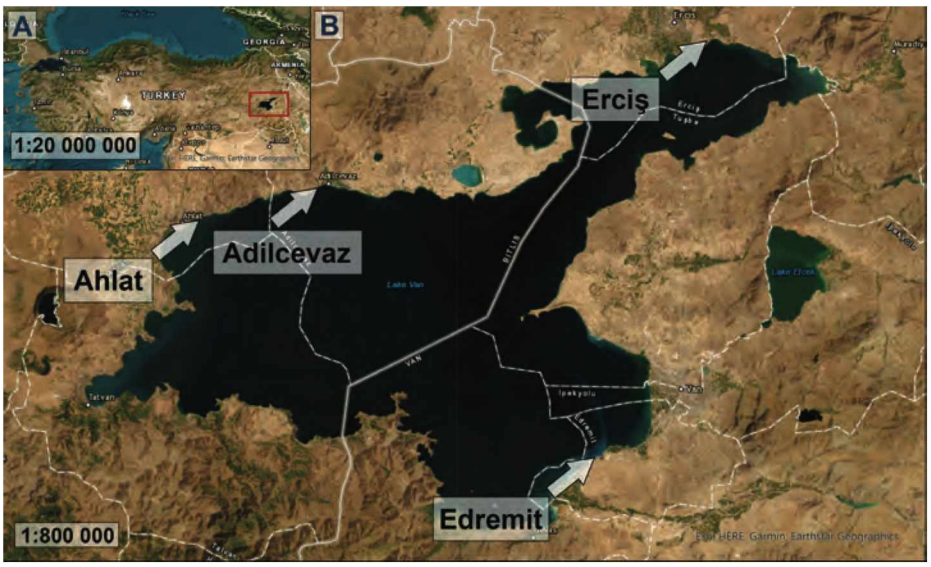


Figure 1. Map of the sampling location **A** location of Lake Van in Turkey. The red frame indicates the position of Lake Van **B** general view of the lake. The areas indicated by arrows are the sampling stations. (Esri. (2023). ArcGIS Pro 3.1.0. Environmental Systems Research Institute)

and a photoperiod of 14 hours of light and 10 hours of darkness at a temperature of 18 °C. The two monoclonal cultures were obtained and registered in the Szczecin Diatom Culture Collection as SZCZEY 2166 and SZCZEY 2167.

Light and scanning electron microscopy

Light microscopy (LM) documentation was obtained at the University of Szczecin with a Zeiss Axio Scope A1 (Carl Zeiss, Jena, Germany) using a Canon EOS 500D camera and Canon EOS Utility software (Canon, Tokyo, Japan). Images were obtained using a 100× Plan Apochromat oil immersion objective (numerical aperture = 1.4).

For the preparation of diatom frustules for both LM and scanning electron microscope (SEM) observations, samples (pellets of cells from monoclonal cultures or wild samples) were moved into 20 mL beakers and 10 mL of 10% HCl added. Over a 24-hour period, the samples were washed with distilled water four times, allowing the material to sediment naturally between washes. Next, the samples were re-suspended in 30% H_2O_2 and boiled on a hotplate at 170 °C for approximately four hours. The final step involved washing the samples four times with distilled water in 24 h, as before. For LM, the material was air-dried on cover glasses and then affixed to a glass slide using Naphrax (Brunel Microscopes, Chippenham, UK). LM measurements were done on a total of 91 valves. For SEM, a drop of cleaned sample was placed on a Nuclepore Track-Etch membrane (Whatman, Maidstone, U.K.). Following air-drying overnight, the membranes were mounted on aluminium stubs using carbon tape and coated with gold using a Q150T coater (Quorum Technologies, Laughton, UK). SEM imaging was conducted at the Faculty of Chemical Technology and Engineering, Western Pomeranian University of Technology in Szczecin (Poland) on a Hitachi SU8020 field emission microscope (Tokyo, Japan). The imaging was conducted with an accelerating voltage of 5kV and a working distance of 8500–8600 µm.

Next generation sequencing and bioinformatic analyses

Clones SZCZEY2166 and SZCZEY2167 were harvested by centrifugation and DNA was extracted following Doyle and Doyle (1990). Total DNA was sent to the Beijing genomics Institute (BGI) in Shenzhen (China) where they were sequenced on a DNBSEQ platform for a total for each clone of ca. 40M 150 bp paired-end clean reads. Assemblies were conducted using SPAdes 3.15.0 (Bankevich et al. 2012) with a k-mer parameter of 125. The contigs corresponding to the plastid and mitochondrial genomes or nuclear rRNA clusters were data-mined by standalone blastn queries. The subunits of the plastid genome were merged with each other with the help of Consed (Gordon and Green 2013). Annotation of protein coding genes was done as described in Gastineau et al. (2023).

Molecular phylogeny

Maximum likelihood phylogenies were inferred from two different datasets. One contained a concatenated alignment of *18S*, *rbcL* and *psbC* genes representing 214 taxa downloaded from GenBank appended with those obtained in the course of this study. Two strains of *Triparma pacifica* (Guillou and Chrétiennot-Dinet) Ichinomiya and Lopes Santos were used as an outgroup. Among the diatom taxa, two were lacking *rbcL* data, 22 strains lacked *psbC* and 21 strains lacked 18S. A *rbcL*-only tree was built as well, in order to compare the tree topologies. Sequences were aligned using MAFFT 7 (Katoh and Standley 2013) and trimmed automatically with trimAl (Capella-Gutiérrez et al. 2009). The best model of evolution was selected separately for each gene with ModelTest-NG (Darriba et al. 2020). In case of three-gene dataset, trimmed alignments of *18S*, *rbcL* and *psbC* were concatenated with Phyutility 2.7.1 (Smith and Dunn 2008). Maximum Likelihood phylogenes were computed using IQ-TREE 2.2.0 (Minh et al. 2020) with 1000 ultrafast bootstrap replicates; the dataset was partitioned based on the best models of evolution found for each gene. The trees were visualised with MEGA11 (Tamura et al. 2021). Lists of the sequences with their corresponding accession numbers can be accessed as explained in the data availability statement.

Results

Taxonomy

Phylum Bacillariophyta Haeckel
Class Bacillariophyceae Haeckel
Family Amphipleuraceae Grunow
Genus *Halamphora* (Cleve) Levkov

***Halamphora vantushpaensis* Yilmaz, Solak & Gastineau, sp. nov.**
Figs 2–4

LM (Figs 2A–M). Valves semi-lanceolate, dorsiventral with arched dorsal margin and slightly tumid ventral margin. Valve ends protracted and capitate in larger specimens (Figs 2A–F); but less protracted and not clearly separated

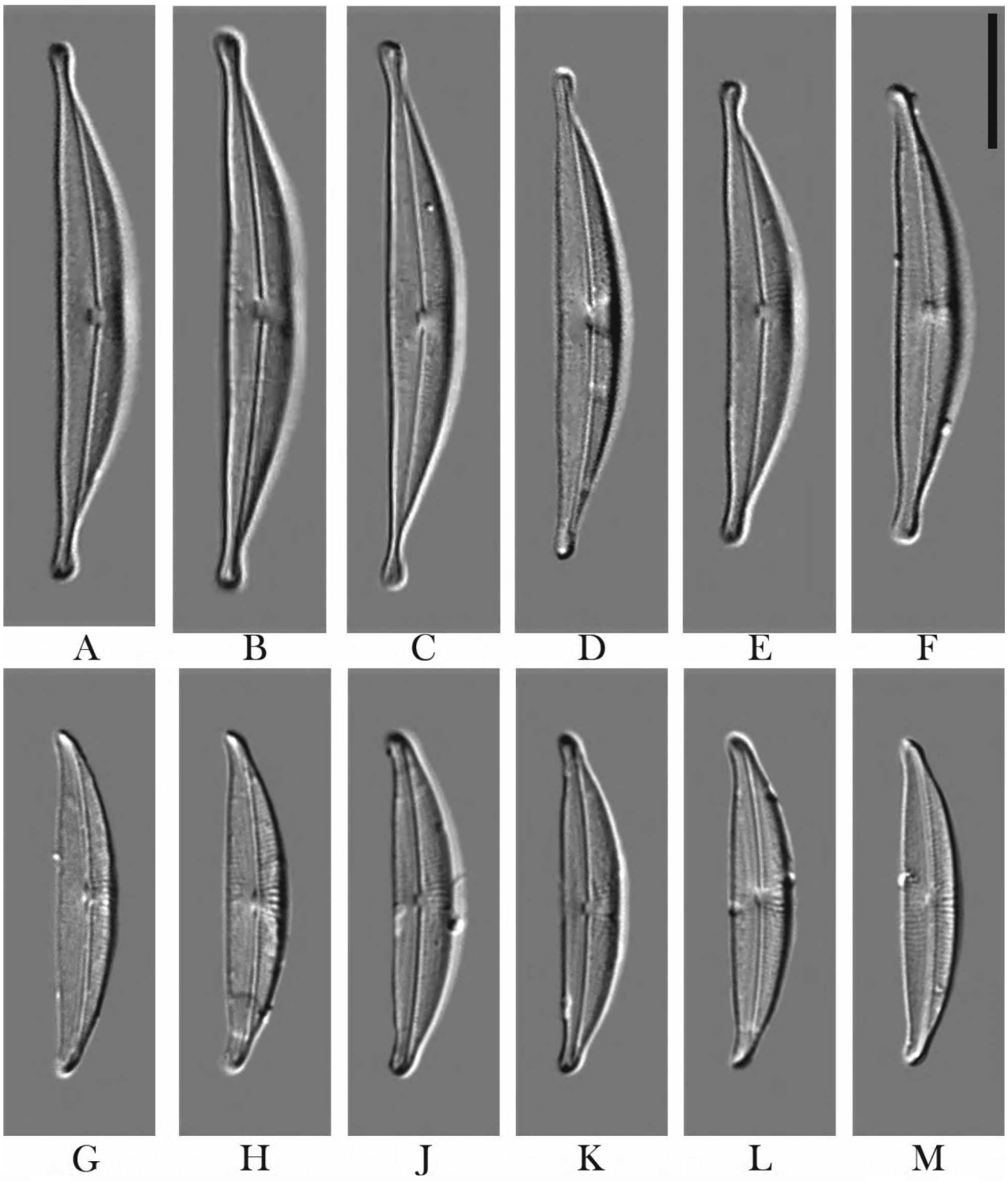


Figure 2. A–M *Halamphora vantushpaensis* sp. nov. LM micrographs **A–F** cleaned valves of the larger specimen (SZCZEY2167) **G–M** cleaned valves of the smaller specimen (SZCZEY2166). Scale bar: 10 µm

from the rest of the valve in smaller specimens (Figs 2G–M), ventrally bent. Valve length 24.0–42.0 µm, valve width 4.0–5.0 µm (n = 35). Axial area very narrow, wider on the ventral side. Central area visible in larger specimens: indistinct on the dorsal side, semi-lanceolate on the ventral side. Raphe almost straight, slightly arched, appearing to be located near the median line of the valve or slightly dorsal in valve view (Fig. 2). Sometimes the proximal raphe

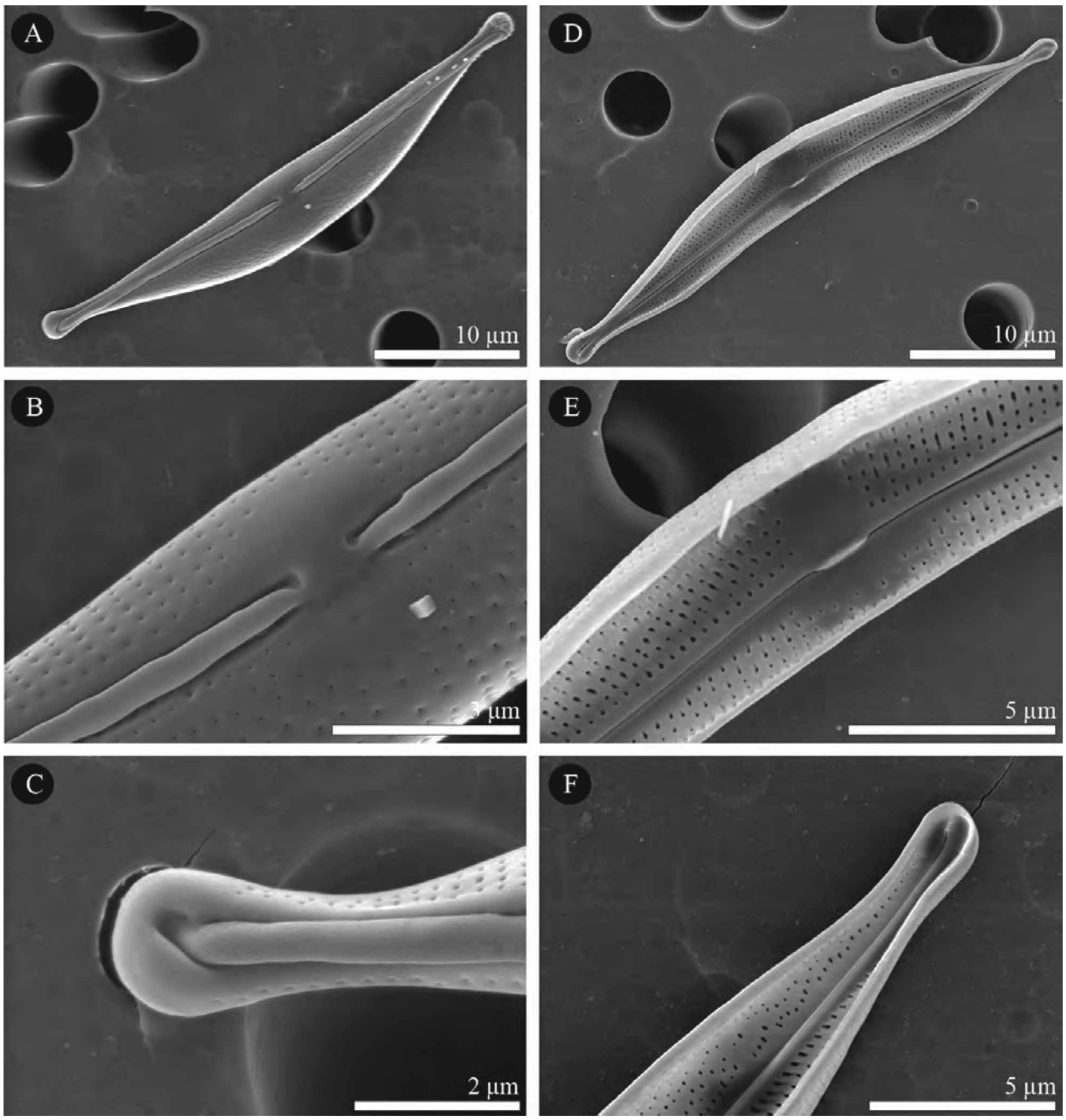


Figure 3. A–F *Halamphora vantushpaensis* sp. nov. SEM micrographs of strain SZCZEY2167 **A** External view of the entire valve **B** details of central area showing simple proximal raphe endings and regular shortened striae **C** details of apex showing the terminal fissure **D** internal view of the entire valve **E**details of central area showing fused central helictoglossae in proximal raphe endings **F** details of apex showing well-developed helictoglossae. Scale bars: 10 µm (**A, D**); 5 µm (**E, F**); 3 µm (**B**); 2 µm (**C**).

endings can be seen to be slightly dorsally bent (Fig. 2B). Striae hard to resolve in LM, dorsally slightly radiate over the entire valve (see SEM images for clearer demonstration of this), 27–32 in 10 µm.

SEM (Figs 3A–F, 4A–F, 5A–J). Externally, the valve face is arched, merging gently into the mantles (Figs 3B, 4B, C, 5H). Raphe ledge narrow and linear, present on the dorsal side of the raphe, with a prominent groove separating it from the valve face. The proximal raphe endings are slightly expanded into central depressions

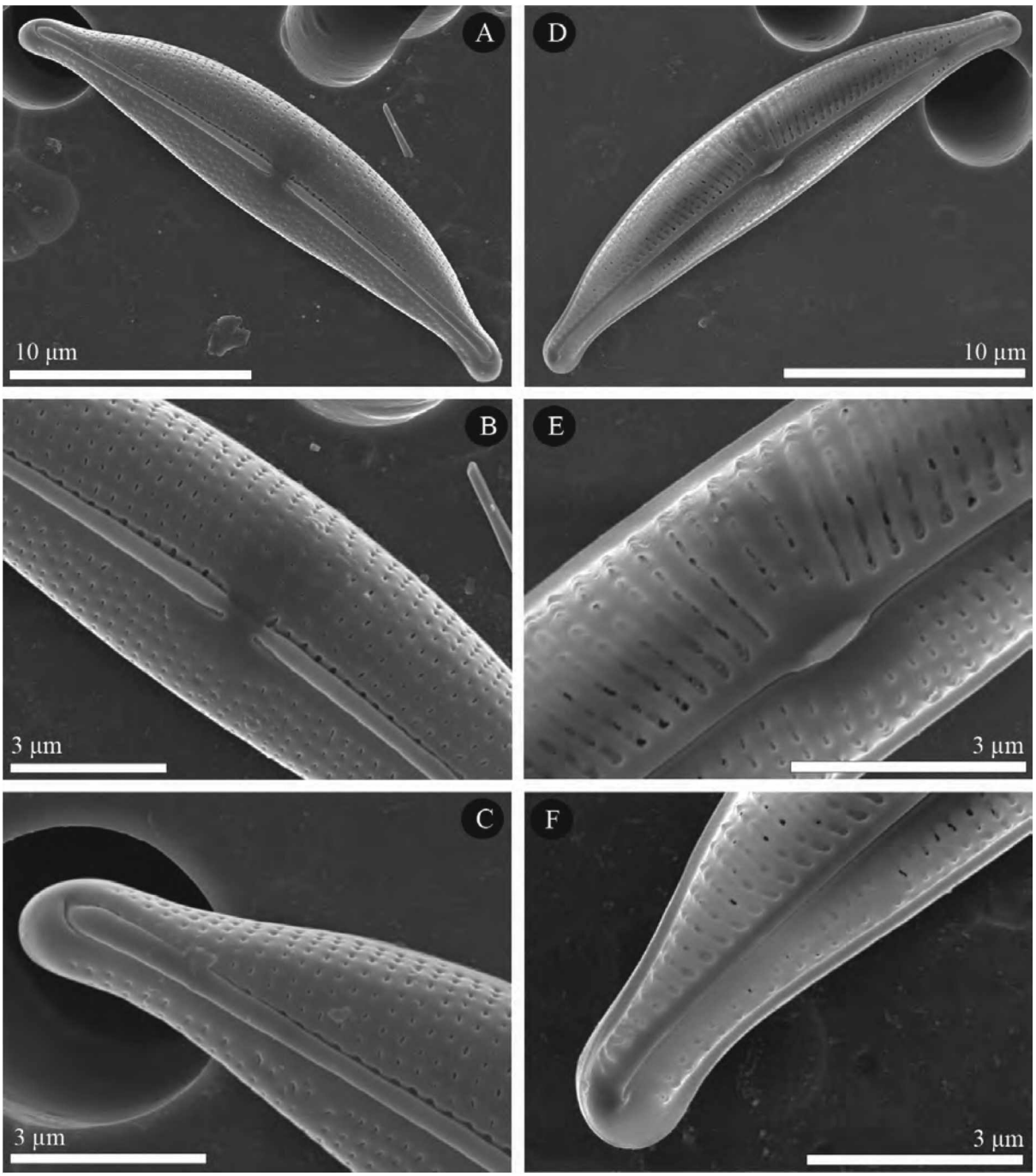


Figure 4. A–F *Halamphora vantushpaensis* sp. nov. SEM micrographs of strain SZCZEY2166 **A** external view of the entire valve **B** details of central area showing simple proximal raphe endings and regular shortened striae **C** details of apex showing the terminal fissure **D** internal view of the entire valve **E** details of central area showing fused central helictoglossae in proximal raphe endings **F** details of apex showing well-developed helictoglossae. Scale bars: 10 µm (**A, D**); 3 µm (**B, C, E, F**).

and are dorsally deflected (Figs 3B, 4B, 5H). The distal raphe endings are dorsally deflected and hook around to link with the groove bordering the raphe ledge (Figs 3C, 4C, 5J). The striae are simple and uniseriate, containing small round or slightly elongate poroids (Figs 3B, 4B, C, 5H), which are somewhat irregularly spaced (Fig. 4B, 5H and see also the internal views in Figs 3E, 4E, 5B, E).

The internal view of the valve shows the overall structure (Figs 3D, 5A). The central area is easier to detect than the external area and appears symmetrical

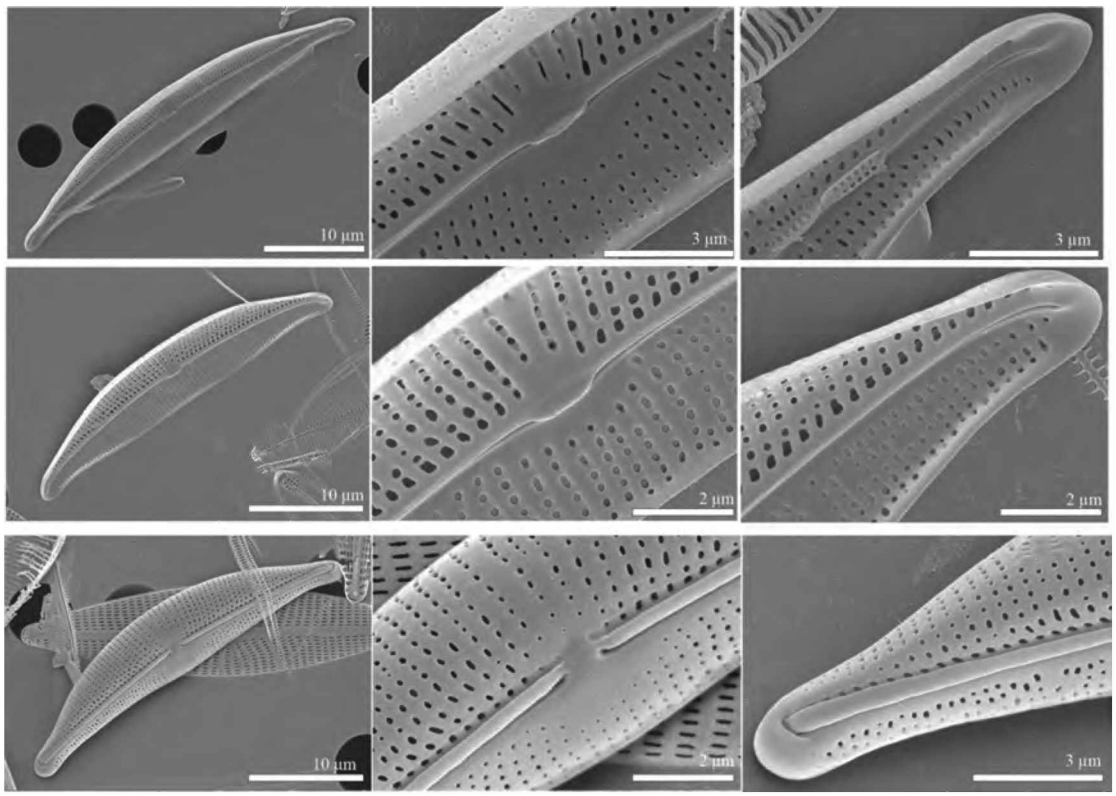


Figure 5. A–J SEM images of a cleaned valve from wild material **A** external view of the entire valve of the larger specimens **B** details of central area showing simple proximal raphe endings and regular shortened striae **C** details of apex showing the terminal fissure **D** internal view of the entire valve of the smaller specimens **E** details of central area showing fused central helictoglossae in proximal raphe endings **F** details of apex showing well-developed helictoglossae **G** external view of the entire valve of the smaller specimens **H** details of central area showing simple proximal raphe endings and regular shortened striae **J** details of apex showing the terminal fissure. Scale bars: 10 µm (**A, D, G**); 3 µm (**B, C, J**); 2 µm (**E, H, F**).

and large on the dorsal side in larger specimens (Fig. 3E, 5B); but very small on both sides in smaller specimens (Fig. 4E, 5E). Proximally, the raphe terminates within a fused central helictoglossa (Figs 3E, 4E, 5E). The distal raphe endings are slightly deflected ventrally and terminate in well-developed helictoglossae (Figs 3F, 4F, 5C, F). Internally, the poroids have round to elliptical internal openings (Figs 3E, 4E, 5B, E). These characteristics are summarized and compared with those of similar species in Table 1.

Phycobank. http://phycobank.org/104935.

Holotype. Slides number SZCZEY2167 in the collection of Andrzej Witkowski at the University of Szczecin, Poland. Valves representing the holotype population here illustrated in Fig. 2D.

Isotype. Slide number TR_Erciş_Van_2021 deposited in Kütahya Dumlupınar University (Türkiye).

Type locality. Erciş Van, Turkey (39°00'07.9"N, 43°25'40.4"E) leg. Elif Yılmaz, 31 July 2021.

Etymology. The species is named with regard to both Lake Van and the city of Tushpa, capital of the Iron Age kingdom of Urartu, which was located in the vicinity of the lake.

Distribution. The presence of this taxon has been assessed and confirmed at four different stations around Lake Van: Ahlat (North West of the lake), Adilcevaz (North), Erciş (North East), and Edremit (South East).

Table 1. Morphological characteristics of *Halamphora vantushpaensis* sp. nov. and similar *Halamphora* (-- represents no information) (for *H. vantushpaensis* measurements, n = 35).

	***Halamphora vantushpaensis* sp. nov.**	***H. atacamana***	***H. borealis***	***H. gasseae***	***H. salinicola***	***H. sardiniensis***	***H. thermalis***
Valve length (µm)	**24.0–42.0**	29.0–45.0	19.0–40.0	19.0–35.0	20.0–34.0	13.0–27.5	18.0–40.0
Valve width (µm)	**4.0–5.0**	4.5–8.0	3.0–4.0	3.5–4.5	2.5–3.7	3.0–4.5	4.0–6.5
Stria density (in 10 µm)	**27–32**	25–28	20–24	20–24	21–26	36–42	26–36
Valve outline	**semi-lanceolate with arched dorsal margin, slightly tumid ventral margin**	semi-lanceolate, arched dorsal margin, concave or straight to weakly tumid ventral margin	semi-lanceolate, arched dorsal margin, straight to weakly tumid ventral margin	semi-lanceolate, smoothly arched dorsal margin, straight to weakly concave ventral margin	semi-lanceolate, smoothly arched dorsal margin, straight to weakly concave ventral margin	semi-lanceolate, strongly arched dorsal margin and straight to slightly concave ventral margin	semi-lanceolate to lanceolate, smoothly arched dorsal margin, straight to slightly tumid ventral margin
Valve ending	**subprotracted in smaller specimens; protracted, capitate in larger specimens and ventrally bent**	slightly subprotracted and ventrally bent	protracted, capitate and slightly ventrally bent	shortly protracted and capitate	shortly protracted and capitate	shortly protracted and capitate	attenuate
Raphe ledge	**narrow, arched with equal width throughout**	narrow, equal width throughout and dorsally elevated from the valve face	narrow, linear	--	narrow, expanded on both valve sides	narrow, expanded on both valve sides	narrow, equal width throughout
Axial area	**narrow, widening ventrally**	narrow, widening ventrally	narrow, widening ventrally	narrow, widening ventrally	narrow, widening ventrally	narrow, widening ventrally	narrow, slightly dorsally bent
References	**in this study**	Levkov 2009	Levkov 2009	Levkov 2009	Levkov 2009	Levkov 2009	Levkov 2009

Genomics - the cluster of nuclear ribosomal RNA genes

Complete clusters of the rRNA genes, containing *18S*, internal transcribed spacer 1 (ITS1), *5.8S*, internal transcribed spacer 2 (ITS2) and *28S*, were obtained for both strains and registered with GenBank accession numbers PP726705 and PP726703 for SZCZEY2166 and SZCZEY2167 respectively. Their sizes and sequences were identical except for a single T/G SNP in the ITS1. The sizes of the different parts of the cluster are indicated in Table 2 and compared with results obtained on the same set of species as in Hamsher et al. (2019), which concern *Halamphora americana* Kociolek, 2014, *Halamphora calidilacuna* Stepanek & Kociolek, 2018 and *H. coffeiformis*. Lengths of *18S*, *5.8S* and *28S* were very conserved among species, except for *H. americana*, which has a group II intron in its *18S* that also contains on ORF coding for a putative reverse transcriptase protein. *Halamphora vantushpaensis* sp. nov. displays a longer ITS1 when compared to other species.

Genomics - mitochondrial genome

Complete mitogenomes were obtained on both strains and registered with GenBank accession numbers PP962256 (SZCZEY2166) (Fig. 6) and PP962257 (SZCZEY2167) (Fig. 7). The genomes are 42,659 bp and 43,152 bp long (SZCZEY2166 and SZCZEY2167 respectively). The genomes both contain 35 conserved protein coding genes (PCGs), two rRNA and 26 tRNA genes. The mitogenomes

Table 2. Length (in bp) of the different portions of the nuclear rRNA cluster for four species of *Halamphora* spp. The length of the *18S* gene of *H. americana* is indicated with and without the intron.

Species	*Halamphora vantushpaensis*	*Halamphora calidilacuna*	*Halamphora americana*	*Halamphora coffeiformis*
Accession number	PP726705 and PP726703	MH810165	MH810166	MH810167
Total length	5932	5764	9254	5938
18S	1767	1769	1783 (5241 with intron)	1767
ITS1	431	223	229	347
5.8S	156	155	155	154
ITS2	368	405	419	454
28S	3210	3212	3210	3217

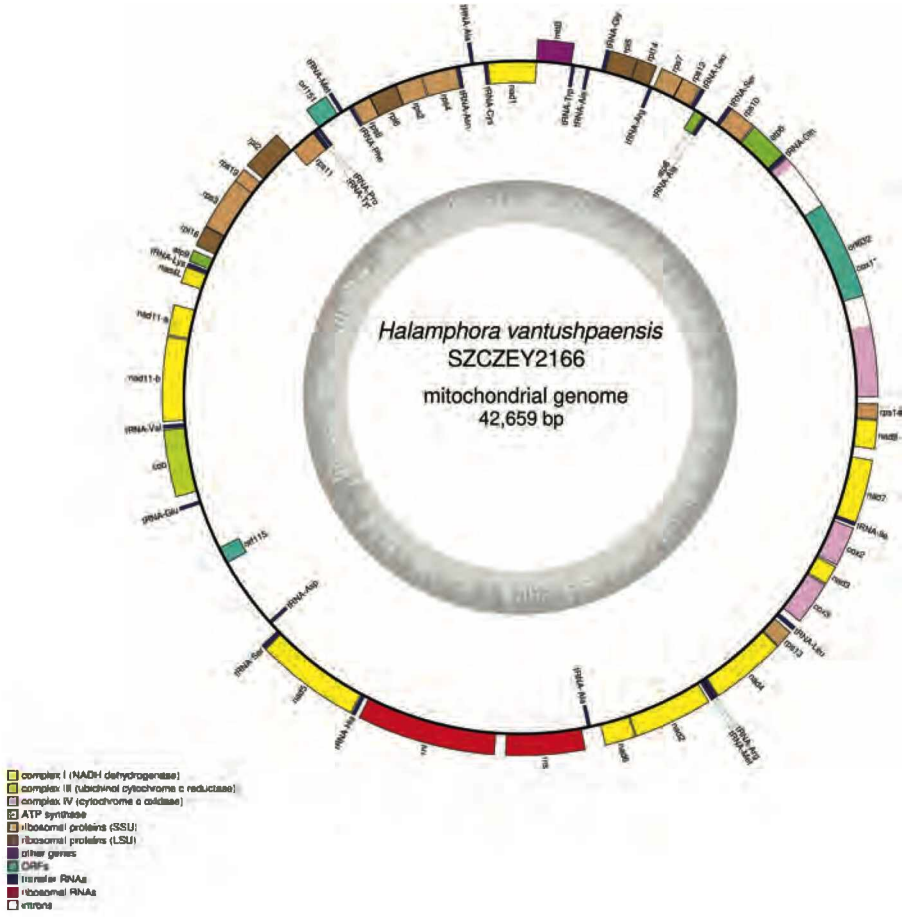


Figure 6. Map of the mitochondrial genome of *Halamphora vantushpaensis* sp. nov. SZCZEY2166

harbour the conserved open reading frame (ORF) orf151, although its position differs from most know species among which it is included in a conserved cluster of genes together with *rps11* and *mttb/tatC* (Pogoda et al. 2019) whereas here it is located between *rps11* and *rps8*. There is also a non-conserved ORF (orf115) between *cob* and *nad5*. The *cox1* gene contains a group II intron with an ORF coding for a putative reverse transcriptase. There are discrepancies in the length of this putative protein which is 632 amino-acid long in SZCZEY2166

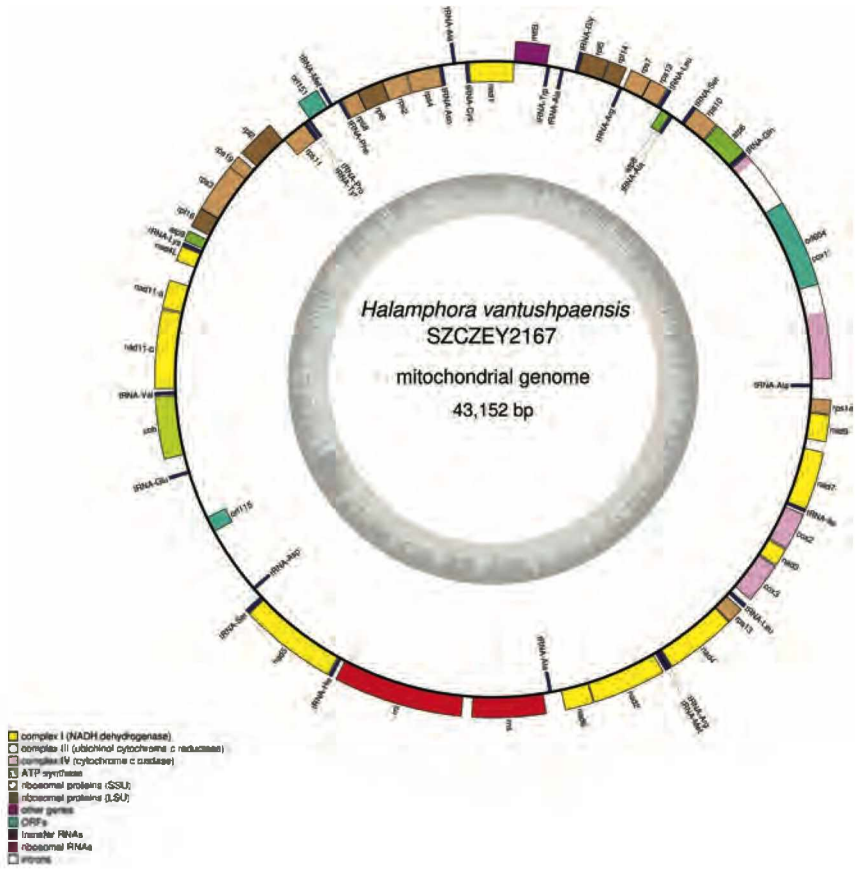


Figure 7. Map of the mitochondrial genome of *Halamphora vantushpaensis* sp. nov. SZCZEY2167

while it is 604 amino-acid long in SZCZEY2167, the extra-length being at the C-terminal part of the putative protein entirely. The polymorphisms between both strains mostly occurred in intergenic parts, hence the slight differences in lengths of the mitogenomes. The conserved protein coding genes were strongly conserved with some of them completely identical, although a certain number of polymorphisms could still be spotted in PCGs in the following genes, with the number of SNPs/lengths indicated between brackets: *cob* (6/1287), *nad2* (1/1536), *nad4* (1/1473), *rpl2* (1/810), *rpl6* (1/573), *rps3* (2/1038), *rps10* (1/540). These mutations were silent in *nad4* and *rpl2*, but led for each of the other encoded protein to one amino-acid substitution. Several variations were otherwise found in the *cox1* intron, whose size varied because of indels (four in total). It otherwise displays 17 SNPs for a total length of 3433/3435 bp, with nine of them being found in the 1815 bp shared between the two putative reverse transcriptase encoding ORF, leading to seven amino-acid substitutions.

Genomics - plastid genome

Both plastome were also obtained. Their lengths are 133,866 bp long for strain SZCZEY2166 (GenBank: PP962255) and 133,852 bp long for strain SZCZEY2167 (GenBank: PP727123). The two plastomes came out as different isomers after assembly, hence the difference of strand of the large single copy region (LSC) that can be observed between SZCZEY2166 (Fig. 8) and SZCZEY2167 (Fig. 9). There were slight differences of lengths for LSC and SSC.

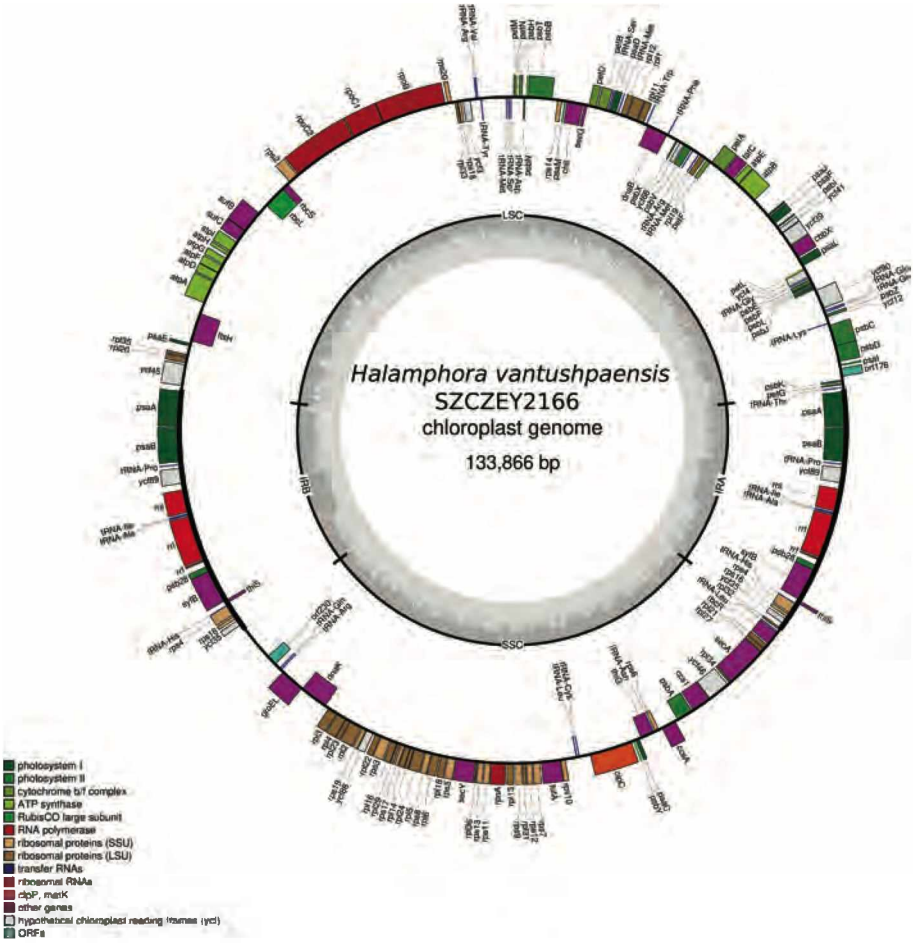


Figure 8. Map of the plastid genome of *Halamphora vantushpaensis* sp. nov. SZCZEY2166

The LSC is 61,711/61,691 bp long and display 26 indels and six SNPs. Out of these six SNPs, five were found in PCGs (*psbC*, *ycf90*, *petB*, *rps20* and *rpoC2*) and were silent only in *petB* and *rpoC2*. The short single copy (SSC) is 39,615 bp long for both strains, with no indels and 17 SNPs, all located in intergenic area. The inverted repeat had identical lengths and displayed six consecutive polymorphisms in the intergene between *rpl32* and *ycf35*. The LSC contains 70 PCGs, a single non-conserved open reading frame (ORF), and 17 tRNAs. The SSC encodes for 46 PCGs, also a single non-conserved ORF and six tRNAs. The inverted repeat IR contains 10 PCGs, three rRNA and four tRNA.

Three plastid genomes are available in GenBank for the genus *Halamphora*, all originating from the same study (Hamsher et al. 2019). In Table 3, the total lengths of these genomes and the lengths of their different compartments are compared.

Halamphora vantushpaensis has shorter LSC and SSC but its IR is consequently longer when compared to other species. The gene content of the IR is compared for all these species in Table 4. The restricted set of conserved genes found among *H. calidilacuna* or *H. americana* and which consists of a single PCG (*ycf89*), three tRNA and three rRNA seems to be shared by many unrelated species and genera such as *Navicula veneta* Kützing 1844 or *Tryblionella apiculata* Gregory 1857 (Gastineau et al. 2021a). As with *H. americana*, an extension of the IR may result from the presence of non-conserved ORF or putative genes of plasmid origin, as exemplified by its ORF9 and the putative

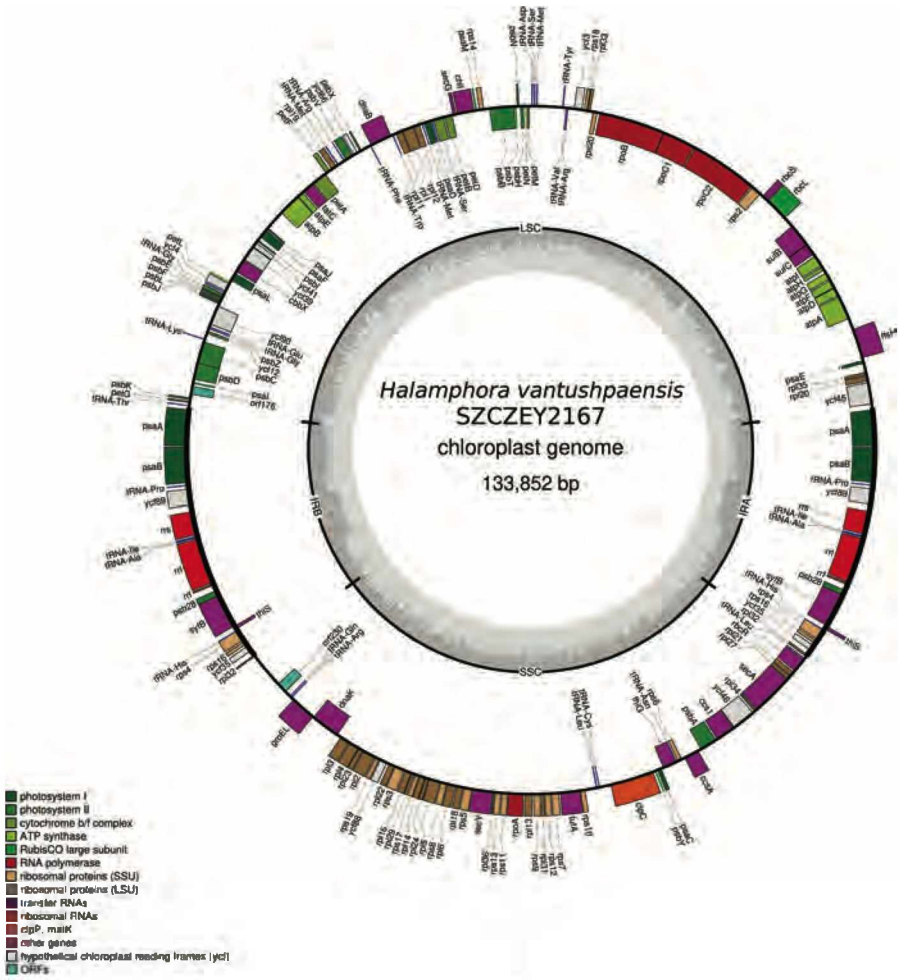


Figure 9. Map of the plastid genome of *Halamphora vantushpaensis* sp. nov. SZCZEY2167

Table 3. Lengths (in bp) of the different compartments of the plastid genomes of four species of *Halamphora* spp.

Species	Length of the LSC	Length of the SSC	Length of the IR	Total length
Halamphora calidilacuna	82,227	49,698	9,407	150,739
Halamphora americana	77,289	44,724	10,269	142,551
Halamphora coffeiformis	64,938	41,485	7,752	121,927
Halamphora vantushpaensis	61,705/61,691	39,615/39,615	16,273	133,866/133,852

Table 4. Gene composition of the inverted repeats of the plastid genomes of four species of *Halamphora* spp. Genes highlighted in bold for *Halamphora vantushpaensis* sp. nov. are genes found in the LSC in other species. Genes in bold italic concern genes usually found in the SSC. Genes marked by an asterisk are non-conserved genes of probable plasmidic origin.

Species	Gene composition of the IR
Halamphora calidilacuna	*tRNA-Pro, ycf89, rrs, tRNA-Ile, tRNA-Ala, rrl, rrf*
Halamphora americana	*tRNA-Pro, ycf89, ORF9*, rrs, tRNA-Ile, tRNA-Ala, rrl, rrf, tyrC**
Halamphora coffeiformis	*tRNA-Pro, ycf89, rrs, tRNA-Ile, tRNA-Ala, rrl, rrf, ycf35*
Halamphora vantushpaensis	**psaA, psaB**, *tRNA-Pro, ycf89, rrs, tRNA-Ile, tRNA-Ala, rrl, rrf,* ***psb28, syfb, thiS, tRNA-His, rps4, rps16, ycf35, rpl32***

integrase/recombinase encoded by the gene labelled as *tyrC* by Hamsher et al. (2019). The case of *H. vantushpaensis* is entirely different in the sense that the extension of the IR is a consequence of the incorporation of several conserved PCGs plus one tRNA. When compared with the gene content of the other species, it appears that this extension has been done at the expense of both the LSC and the SSC, which distinguishes it from species like *Climaconeis* spp. (Gastineau et al. 2021b) among which the IR seemed to have only taken over the SSC. Indeed, among the other *Halamphora* spp., *psaA* and *psaB* are located in the LSC while the other genes belong to the SSC in which they form a well-conserved cluster.

Molecular maximum likelihood phylogeny

Fig. 10 presents the *Halamphora* clade as a sub-tree derived from the three-gene inferred phylogeny. The complete three-gene tree and the *rbcL*-only tree can be found as indicated in the data availability statement. In the three-gene tree, *H. vantushpaensis* strains appear as a highly supported (99%) long-branched sister group to a larger cluster composed of 18 *Halamphora* species, namely *H. subacutiuscula*, *H. angistiformis*, *H. foramina*, *H. sydowii*, *H. tumida*, *H. witkowskii*, *H. bonnewillensis*, *H. americana*, *H. calidilacuna*, *H. intramaritima*, *H. incelebrata*, *H. banzuensis*, *H. bistriata*, *H. pertusa*, *H. subtropica* plus three unidentified *Halamphora* species. The topology of the *rbcL*-inferred tree slightly differs regarding the species sister to *H. vantushpaensis*, which are, in this case, *H. angustiformis* (bv = 93) and *H. subacutiuscula* (bv = 96). These strains are further nested in a clade with *H. maritima*, *H. pecensa*, *"Amphora" caribeana*, *H. exilis*, *H. subtropica*, *H. pertusa*, *H. banzuensis and H. bistriata* with low support (bv < 50) and, together with these, sister to *H. tumida*, *H. witkowskii*, *H. bonnebillensis*, *H. americana*, *H. calidilacuna*, *H. intramaritima*, *H. incelebrata*, *H. foramina*, *H. sydowii* and *Halamphora* sp. SZCZCH45

Discussion

Morphological comparison with similar taxa

Halamphora vantushpaensis sp. nov. is a new species, characterized through the extensive study of two distinct cultivated clones as well as examination of wild samples. The findings indicate that the morphological characteristics of *H. vantushpaensis* can strongly vary and that LM observations might not be sufficient. *Halamphora atacamana* (Patrick) Levkov, *H. borealis* (Kützing) Levkov, *H. gasseae* Levkov, and *H. salinicola* Levkov and Diaz have been identified as the most similar species. In terms of outline, *H. borealis* exhibits a semi-lanceolate shape similar to *H. vantushpaensis* (Table 1). However, distinguishing features of *H. vantushpaensis* such as the larger ventral side and the indistinct striae with a higher density (more than 27 striae in 10 µm) set it apart. Additionally, SEM images reveal differences in striae composition between *H. vantushpaensis* and *H. borealis*, further supporting their taxonomic differentiation. *Halamphora atacamana* exhibits a similar outline, especially to smaller specimens of *H. vantushpaensis*, with slightly protracted valve endings; however, larger specimens of *H. vantushpaensis* have elongated valve ends. *Halamphora atacamana* also

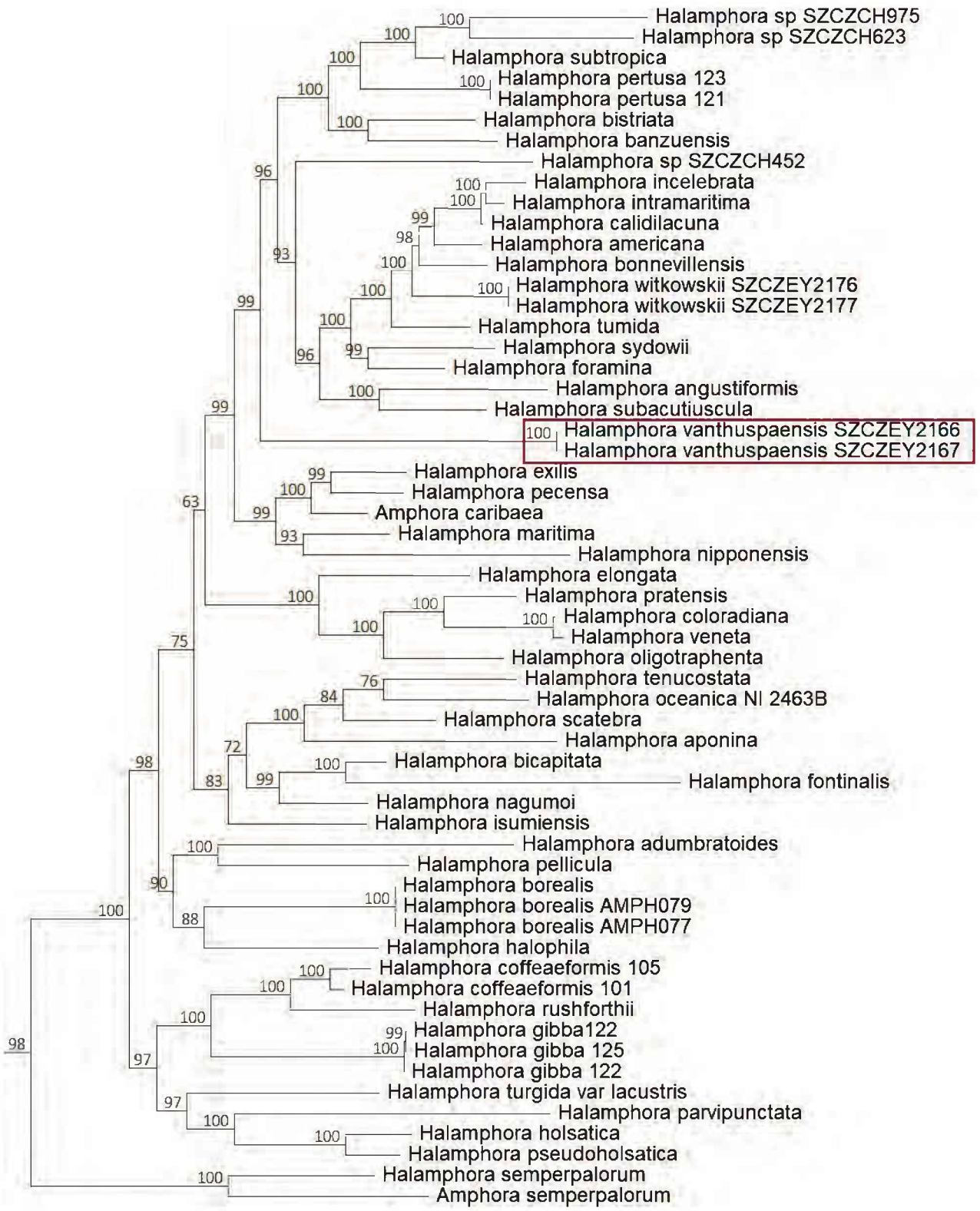


Figure 10. Maximum Likelihood phylogenetic tree inferred from concatenated alignments of *psbC*, *rbcL* and *18S*. The figure represents the sub-tree that contains the *Halamphora* clade.

tends to have lower stria densities (< 28 in 10 µm). *Halamphora gassae* and *H. salinicola* are further similar taxa, but both have smaller valves (< 35 µm in length, < 4.5 µm in width), lower stria densities (< 27 in 10 µm), and smoothly arched dorsal margins. Moreover, we observed that *H. salinicola* has larger

areolae on the dorsal side, one elongate areolae on the ventral side, and a raphe ledge that extends continuously over the entire length of the valve. Among other species that might be compared with *H. vantushpaensis*, *H. sardiniensis* has smaller valves and a higher stria density (> 36 striae in 10 µm) and strongly arched dorsal margin. Additionally, *H. sardiniensis* has elongate areolae on the dorsal side, one elongate areola on the ventral side, and a large central area on the ventral side (visible in SEM: Levkov 2009, pl.245, fig. 4). *Halamphora thermalis* is similar to smaller specimens of *H. vantushpaensis*. However, *H. thermalis* has a smoothly arched dorsal margin and a more visibly tumid ventral margin. In SEM, *H. thermalis* has larger irregularly rounded elongate areolae on dorsal side and rounded areolae on ventral side and areolae become smaller toward the central area ventrally. Also, the proximal raphe endings open into larger depressions (Levkov 2009, pl. 230, figs 1–6).

Genomic results and phylogenies

Initially, when comparing SZCZEY2166 and SZCZEY2167 by the means of LM, it was unclear whether or not they belonged to identical or different species, particularly because of the differences of shape of their apices. Of course, SEM brought supplementary clues of their identity, but in the end, molecular methods provided the decisive argument. With regards to this, it should be noted that within a single round of short-reads sequencing, complete nuclear rRNA clusters, mitochondrial genomes and plastid genomes were obtained, which allowed to derive accurate phylogenies. The *rbcL*-inferred phylogeny strictly positions *H. vantushpaensis* within a clade of species previously described as 'K clade' (Stepanek and Kociolek 2019). This is a noteworthy result, because this clade is known to contain species with extremely different preferences for what regards salinity, ranging from freshwater to hypersaline. As such, this clade is regarded as an interesting model to study transition between habitats, although it is noteworthy that such a transition seemed to have occurred repeatedly and independently among the genus Halamphora. The genomic approach we employed, sometimes described as 'genome-skimming', has reliably provided results on diatoms (for recent examples, see Gastineau et al. 2021c; Solak et al. 2021 ; Dąbek et al. 2022 ; Yilmaz et al. 2024). Aside from their interest in molecular phylogeny, the availability of full-length RNA operon reference sequences from duly identified organisms could become increasingly valuable with the development of long-read metabarcoding (e.g. Jamy et al. 2020). In the current case, it was interesting to see that the only polymorphism between both strains of *H. vantushpaensis* was located in one of the internal transcribed spacers, a portion that does not participate to the final 3D structure of the ribosome and as such, is more likely to display variations.

When comparing the plastomes, the low number of polymorphisms slightly misled us at first into thinking that these SNPs might only have concerned non-coding parts. Surprisingly, it was not the case, and although the number of SNPs is rather low, interestingly, several among them were not silent. However, this variability between the two strains is consequently lower than what was observed with the previously published Lake Van-species *N. vanseea* (Yılmaz et al. 2024). A gene such as *psbC* displayed three times more polymorphisms between the two strains studied at that time when compared to what was unveiled

between both strains of *H. vantushpaensis*. At the time *N. vanseea* was being investigated, it was possible to sequence the mitogenome for only one of the strains, for reasons that remain unknown but might be related to the amount of bacterial contamination in the DNA pool. This time, with *H. vantushpaensis*, sequencing of both strains was successful. We note that the *cox1* gene, which seems to be a sensitive marker for the study of diatoms at the subspecific level (Gastineau et al. 2013, 2021c; Dąbek et al. 2022) was entirely conserved in its exonic parts. The position of the *cox1* intron was perfectly conserved, unlike previous reports on other species (Gastineau et al. 2021c), but there were noticeable differences in its sequence. This is exemplified by the changes in length and sequence of the putative reverse-transcriptase it contains. Among land plants, introns in organellar genomes have been documented to be relevant as population markers (e.g. Spaniolas et al. 2010; Grosser et al. 2023), but no such work seems to exist for diatoms.

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Funding

This work was co-financed by the Minister of Science under the "Regional Excellence Initiative" Program for 2024–2027 (RID/SP/0045/2024/01). Claude Lemieux and Monique Turmel were supported by grant RGPIN-2017-04506 from the Natural Sciences and Engineering Research Council of Canada (NSERC). We also acknowledge support from the CERCA Programme/Generalitat de Catalunya. The Royal Botanic Garden Edinburgh is supported by the Scottish Government's Rural and Environment Science and Analytical Services Division.

Author contributions

Conceptualisation: EY. Data curation: EY. Funding acquisition: AW. Investigation: EY, RG, CNS, EG, CL, MT, CO. Methodology: AW, RT, DGM. Project administration: CNS, AW. Supervision: AW, DGM, RT, CNS, RG. Visualisation: EY, CNS, EG. Writing - original draft: EY. Writing - review and editing: DGM, RT, CL, MT, CO, CNS, RG.

Author ORCIDs

Elif Yılmaz https://orcid.org/0000-0001-7939-1814
Romain Gastineau https://orcid.org/0000-0001-8661-5118
Cüneyt Nadir Solak https://orcid.org/0000-0003-2334-4271
Ewa Górecka https://orcid.org/0000-0003-0590-7480
Rosa Trobajo https://orcid.org/0000-0001-9498-3797
Monique Turmel https://orcid.org/0000-0001-7060-035X
Claude Lemieux https://orcid.org/0000-0001-9580-8042
Christian Otis https://orcid.org/0000-0001-9680-5863
Andrzej Witkowski https://orcid.org/0000-0003-1714-218X
David G. Mann https://orcid.org/0000-0003-0522-6802

Data availability

All sequences have been deposited and are already available on GenBank with accession numbers PP726705, PP726703, PP962256, PP962257, PP962255, PP727123. The complete 3-gene tree, the rbcL-inferred trees and the lists of sequences used for phylogeny can be found on Zenodo following this link: https://doi.org/10.5281/zenodo.12963401.

References

Akku M, Sar M, Ekmekçi F, Yoğurtçuoğlu B (2021) The discovery of a microbialite-associated freshwater fish in the world's largest saline soda lake, Lake Van (Turkey). Zoosystematics and Evolution 97: 181–189. https://doi.org/10.3897/zse.97.62120

Aydar E, Gourgaud A, Ulusoy I, Digonnet F, Labazuy P, Şen E, Bayhan HG, Kurttas T, Tolluoğlu AÜ (2003) Morphological analysis of active Mount Nemrut stratovolcano, eastern Turkey: evidences and possible impact areas of future eruption. Journal of Volcanology and Geothermal Research 123: 301–312. https://doi.org/10.1016/S0377-0273(03)00002-7

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19(5): 455–477. https://doi.org/10.1089/cmb.2012.0021

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics (Oxford, England) 25(15): 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Cleve PT (1895) Synopsis of the naviculoid diatoms, part II. Kongliga Svenska-Vetenskaps Akademiens Handlingar 27: 1–219.

Dąbek P, Gastineau R, Bornman TG, Lemieux C, Turmel M, Hallegraeff G, Mouget J-L and Witkowski A (2022) The blue diatom *Haslea ostrearia* from the Indian Ocean coast of South Africa, with comparative analysis of *Haslea* organellar genomes. Frontiers in Marine Science 9: 950716. https://doi.org/10.3389/fmars.2022.950716

Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T (2020) ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. Molecular Biology and Evolution 37(1): 291–294. https://doi.org/10.1093/molbev/msz189

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12: 13–15. https://doi.org/10.2307/2419362

Gastineau R, Leignel V, Jacquette B, Hardivillier Y, Wulff A, Gaudin P, Bendahmane D, Davidovich NA, Kaczmarska I, Mouget JL (2013) Inheritance of mitochondrial DNA in the Pennate diatom Haslea ostrearia (Naviculaceae) during auxosporulation suggests a uniparental transmission. Protist 164(3): 340–351. https://doi.org/10.1016/j.protis.2013.01.001

Gastineau R, Hamedi C, Baba Hamed MB, Abi-Ayad S-ME-A, Bąk M, Lemieux C, Turmel M, Dobosz S, Wróbel RJ, Kierzek A, Lange-Bertalot H, Witkowski A (2021a) Morphological and molecular identification reveals that waters from an isolated oasis in Tamanrasset (extreme South of Algerian Sahara) are colonized by opportunistic and pollution-tolerant diatom species. Ecological Indicators 121(6): 107104. https://doi.org/10.1016/j.ecolind.2020.107104

Gastineau R, Davidovich NA, Davidovich OI, Lemieux C, Turmel M, Wróbel RJ, Witkowski A (2021b) Extreme Enlargement of the Inverted Repeat Region in the Plastid Genomes of Diatoms from the Genus *Climaconeis*. International journal of molecular sciences 22(13): 7155. https://doi.org/10.3390/ijms22137155

Gastineau R, Hansen G, Poulin M, Lemieux C, Turmel M, Bardeau JF, Leignel V, Hardivillier Y, Morançais M, Fleurence J, Gaudin P, Méléder V, Cox EJ, Davidovich NA, Davidovich OI, Witkowski A, Kaczmarska I, Ehrman JM, Soler Onís E, Quintana AM, Mucko M, Mordret S, Sarno D, Jacquette B, Falaise C, Séveno J, Lindquist NL, Kemp PS, Jr.; Eker-Develi E, Konucu M, Mouget JL (2021c) *Haslea silbo*, A Novel Cosmopolitan Species of Blue Diatoms. Biology 10: 328. https://doi.org/10.3390/biology10040328

Gastineau R, Harðardóttir S, Guilmette C, Lemieux C, Turmel M, Otis C, Boyle B, Levesque RC, Gauthier J, Potvin M, Lovejoy C (2023) Mitochondrial genome sequence of the protist *Ancyromonas sigmoides* Kent, 1881 (Ancyromonadida) from the Sugluk Inlet, Hudson Strait, Nunavik, Québec. Frontiers in Microbiology 14: 1275665. https://doi.org/10.3389/fmicb.2023.1275665

Glombitza C, Stockhecke M, Schubert CJ, Vetter A, Kallmeyer J (2013) Sulfate reduction controlled by organic matter availability in deep sediment cores from the saline, alkaline Lake Van (Eastern Anatolia, Turkey). Frontiers in Microbiology 4: 209. https://doi.org/10.3389/fmicb.2013.00209

Gordon D, Green P (2013) Consed: A graphical editor for next-generation sequencing. Bioinformatics 29(22): 2936–2937. https://doi.org/10.1093/bioinformatics/btt515

Grosser MR, Sites SK, Murata MM, Lopez Y, Chamusco KC, Love Harriage K, Grosser JW, Graham JH, Gmitter Jr FG, Chase CD (2023). Plant mitochondrial introns as genetic markers - conservation and variation. Frontiers in plant science 14: 1116851. https://doi.org/10.3389/fpls.2023.1116851

Guillard RRL (1975) Culture of phytoplankton for feeding marine invertebrates. In: Smith WL, Chanley MH (Eds) Culture of Marine Invertebrate Animals. Plenum Press, New York, 26–60. https://doi.org/10.1007/978-1-4615-8714-9_3

Guiry MD, Guiry GM 30 May 2024. *AlgaeBase*. World-wide electronic publication, National University of Ireland, Galway. https://www.algaebase.org [searched on 02 April 2024]

Hamsher SE, Keepers KG, Pogoda CS, Stepanek JG, Kane NC, Kociolek JP (2019) Extensive chloroplast genome rearrangement amongst three closely related *Halamphora* spp. (Bacillariophyceae), and evidence for rapid evolution as compared to land plants. PloS ONE 14(7): e0217824. https://doi.org/10.1371/journal.pone.0217824

Jamy M, Foster R, Barbera P, Czech L, Kozlov A, Stamatakis A, Bending G, Hilton S, Bass D, Burki F (2020) Long-read metabarcoding of the eukaryotic rDNA operon to phylogenetically and taxonomically resolve environmental diversity. Molecular Ecology Resources 20(2): 429–443. https://doi.org/10.1111/1755-0998.13117

Kaden H, Peeters F, Lorke A, Kipfer R, Tomonaga Y, Karabiyikoglu M (2010) Impact of lake level change on deep-water renewal and oxic conditions in deep saline Lake Van, Turkey. Water Resources Research 46(11): W11508. https://doi.org/10.1029/2009WR008555

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. https://doi.org/10.1093/molbev/mst010

Kempe S, Kazmierczak J, Landmann G, Konuk T, Reimer A, Lipp A (1991) Largest known microbialites discovered in Lake Van, Turkey. Nature 349: 605–608. https://doi.org/10.1038/349605a0

Kützing FT (1844) Die kieselschaligen Bacillarien oder Diatomeen. Nordhausen, 1–152. https://doi.org/10.5962/bhl.title.64360

Legler F, Krasske G (1940) Diatomeen aus dem Vansee (Armenien). Beihefte zur Botanischen Zentralblatt 60, Abteilung B, 335–345.

Levkov Z (2009) *Amphora sensu lato*. Diatoms of Europe 5: 1–916.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution 37(5): 1530–1534. https://doi.org/10.1093/molbev/msaa015

North SM, Stockhecke M, Tomonaga Y Mackay AW (2018) Analysis of a fragmentary diatom record from Lake Van (Turkey) reveals substantial lake-level variability during previous interglacials MIS7 and MIS5e. Journal of Paleolimnology 59: 119–133. https://doi.org/10.1007/s10933-017-9973-z

Pogoda CS, Keepers KG, Hamsher SE, Stepanek JG, Kane NC, Kociolek JP (2019) Comparative analysis of the mitochondrial genomes of six newly sequenced diatoms reveals group II introns in the barcoding region of cox1. Mitochondrial DNA. Part A, DNA mapping, sequencing, and analysis, 30(1): 43–51. https://doi.org/10.1080/24701394.2018.1450397

Smith SA, Dunn CW (2008) Phyutility: a phyloinformatics tool for trees, alignments and molecular data. Bioinformatics (Oxford, England) 24(5): 715–716. https://doi.org/10.1093/bioinformatics/btm619

Solak CN, Gastineau R, Lemieux C, Turmel M, Gorecka E, Trobajo R, Rybak M, Yılmaz E, Witkowski A (2021) *Nitzschia anatoliensis* sp. nov., a cryptic diatom species from the highly alkaline Van Lake (Turkey). PeerJ 9: e12220. https://doi.org/10.7717/peerj.12220

Spaniolas S, Bazakos C, Spano T, Zoghby C, Kalaitzis P (2010) The potential of plastid trnL (UAA) intron polymorphisms for the identification of the botanical origin of plant oils. Food Chemistry 122(3): 850–856. https://doi.org/10.1016/j.foodchem.2010.02.039

Stepanek JG, Kociolek JP (2013) Several new species of *Amphora* and *Halamphora* from the western USA. Diatom Research 28(1): 61–76. https://doi.org/10.1080/0269249X.2012.735205

Stepanek JG, Kociolek JP (2014) Molecular phylogeny of Amphora sensu lato (Bacillariophyta): an investigation into the monophyly and classification of the amphoroid diatoms. Protist 165(2): 177–195. https://doi.org/10.1016/j.protis.2014.02.002

Stepanek JG, Kociolek JP (2019) Molecular phylogeny of the diatom genera *Amphora* and *Halamphora* (Bacillariophyta) with a focus on morphological and ecological evolution. Journal of phycology 55(2): 442–456. https://doi.org/10.1111/jpy.12836

Stockhecke M, Kwiecien O, Vigliotti L, Anselmetti FS, Beer J, Çağatay MN, Channell JE, Kipfer R, Lachner J, Litt T, Pickarski N, Sturm M (2014) Chronostratigraphy of the 600,000 year old continental record of Lake Van (Turkey). Quaternary Science Reviews 104: 8–17. https://doi.org/10.1016/j.quascirev.2014.04.008

Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis version 11. Molecular Biology and Evolution 38(7): 3022–3027. https://doi.org/10.1093/molbev/msab120

Van Heurck HF (1880) Synopsis des Diatomées de Belgique. Ducaju et Cie., 1–696. https://doi.org/10.5962/bhl.title.1990

Yilmaz E, Gastineau R, Górecka E, Solak CN, Trobajo R, Peszek Ł, Mann DG (in press) *Halamphora witkowskii* sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye. Nova Hedwigia.

Yilmaz E, Mann DG, Gastineau R, Trobajo R, Solak CN, Górecka E, Turmel M, Lemieux C, Ertorun N, Witkowski A (2024) Description of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the highly alkaline Lake Van (Republic of Türkiye) with complete characterisation of its organellar genomes and multigene phylogeny. PhytoKeys 241: 27–48. https://doi.org/10.3897/phytokeys.241.118903

Halamphora witkowskii sp. nov. (Catenulaceae, Bacillariophyta), a new diatom species from the alkaline waters of Lake Van, Republic of Türkiye

Elif Yılmaz[1], Romain Gastineau[1], Ewa Górecka[1], Cüneyt Nadir Solak[2], Rosa Trobajo[3], Łukasz Peszek[4], David G. Mann[5]

[1]*Institute of Marine and Environmental Sciences, University of Szczecin, Mickiewicza 16A, PL70–383 Poland*

[2]*Department of Biology, Faculty of Science & Art, Dumlupınar University, 43000 Kütahya, Türkiye*

[3]*Marine and Continental Waters, Institute for Food and Agricultural Research and Technology (IRTA), Crta de Poble Nou Km 5.5, E-43540 La Ràpita, Catalunya, Spain*

[4]*Department of Agroecology and Forest Utilization, University of Rzeszów, ul. Ćwiklinskiej 1A, 35-601 Rzeszów, Poland*

[5]*Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, Scotland, UK*

Abstract

This study introduces *Halamphora witkowskii* sp. nov., a newly discovered diatom species from Lake Van, the world's largest soda lake known for its high alkalinity. Detailed morphological and morphometric analyses using light and scanning electron microscopy effectively distinguished this new species from closely related taxa, including *Halamphora minima*, *H. coffeaeformis*, and *H. tumida*. Additionally, culturing studies from live cells led to the establishment of two separate monocultures, which were utilized for phylogenetic analysis using the *rbcL* molecular marker, further confirming the distinctiveness of *Halamphora witkowskii* from other known species.

Keywords: *Halamphora*, alkaline, soda lake, *rbcL*, phylogeny, taxonomy

Introduction

Lake Van, located in the eastern Anatolian region of Turkey, is the world's largest soda lake and is notable for its saline water composition. Situated between the borders of Bitlis and Van provinces, this volcanic lake was formed by the accumulation of waters in the crater

resulting from the eruption of Mount Nemrut. The lake spans 70 km in the east-west direction, covers an area of 3,522 km², and has a volume of 607 km³. Its shoreline extends 430 km overland (Kadıoğlu et al. 1997). Lake Van is situated at an average altitude of 1,648 meters above sea level, with an average depth of 171 meters and a maximum depth of 460 meters (Özalp et al. 2016). The perimeter of the lake features volcanic rocks to the north and west, metamorphic rocks to the south, and predominantly sedimentary rocks to the east. The geological diversity and saline composition of Lake Van's water have drawn significant research interest, particularly regarding the chemical parameters of the lake's water across different periods (Degens et al. 1984).

More than 80 years ago, Legler & Krasske (1940) conducted significant research on the diatoms of Lake Van, describing six new diatom taxa, including *Amphiprora paludosa var. densestriata* Krasske, *Nitzschia incognita* Legler & Krasske, *Rhopalodia musculus* var. *suprasemicirculatus* Krasske, and *Surirella invicta* Krasske. Subsequently, some of these species underwent further examination and documentation using light microscopy (LM) and scanning electron microscopy (SEM) by Lange-Bertalot et al. (1996). More recently, Solak et al. (2021) and Yılmaz et al. (2024) contributed to this body of knowledge by utilizing modern molecular tools to describe two new species of diatoms from Lake Van: *Nitzschia anatoliensis* Górecka, Gastineau & Solak and *Navicula vanseea* Yılmaz, Gastineau, Solak & Witkowski.

The genus *Halamphora* was first described by Cleve (1895) as a subgenus of *Amphora*, which was recently separated as a genus by Levkov (2009). The morphological characters of *Halamphora* species are outwardly curvedon the dorsal side and shallow and flat on the ventral side, generally characterized as having round or capitate-tipped valves and punctate bands valves, a straight or slightly curved raphe near the ventral edge, and a narrow raphe sternum (e.g. Round et al. 1990; Levkov 2009; Stepanek & Kociolek 2018). Taxonomically, there are currently 157 accepted *Halamphora* species (Guiry & Guiry 2024). Although this genus is common in inland water bodies with high conductivity (e.g. Sala et al. 2007) and can also be found in freshwaters and semiterrestrial habitats (e.g. Levkov 2009, Van de Vijver et al. 2014, You et al. 2015), most species occur in marine and brackish water habitats (e.g. Cleve 1895, Stepanek & Kociolek 2013). Ongoing research on the species richness of *Halamphora* with a specific focus on studies of cultured isolates, are continuously leading to the discovery of new species and expansion of our knowledge regarding diversity within

the genus (Jiang et al. 2015, Olivares-Rubio et al. 2017, Zhang et al. 2019, López-Fuerte et al. 2020, An et al. 2022).

In the current study, two strains of *Halamphora* from Lake Van have been isolated, cultivated and characterized. These cultures, together with a specimen found in environmental sample, are hereby proposed as a new species, based on morphological characteristics observed using light and scanning electron microscopy (LM and SEM), as well as the molecular characteristics of the *rbcL* gene.

Materials and methods

Sample collection and culture

Benthic samples were collected on July 31st 2021, from the Ahlat shore of Lake Van, using a toothbrush to scrape epilithic material from submerged rocks (Fig. 1). Single diatom cells were isolated using glass micropipette, under a Nikon Eclipse TS100 inverted microscope (Nikon, Tokyo, Japan), into sterile, plastic Petri dishes containing sterile F/2 culture medium (Guillard 1975), modified to have 18 ppt salinity. The medium inoculated with isolated single cells was carefully cultivated under controlled light and temperature conditions within the laboratory; the temperature was 18°C. Light intensity was 60 µmol photons $m^{-2}s^{-1}$ and a photoperiod cycle of 14 hours of light followed by 10 hours of darkness. A minimum of three successive isolation procedures were carried out, thereby guaranteeing that cultures were monoclonal. Strains were reinoculated into newly prepared, sterile flasks with culture medium at four-week intervals. Monoclonal cultures of the studied taxon are currently being kept (April 2024) in the Szczecin Diatom Culture Collection at University of Szczecin, Institute of Marine and Environmental Sciences under numbers SZCZ EY2176 and EY2177.

Observations

For LM and SEM observations, approximately 5 mL of dense monoclonal culture was transferred into 20 mL beakers, followed by the addition of 10 mL of 10% HCl. Afterward, the acidic residues were subjected to four successive washes with distilled water to aid in their removal. The sedimentation method was employed to separate the diatoms from the water, wherein the mixture was allowed to stand undisturbed, allowing the denser diatoms to settle to the bottom while the water was removed. The samples were then resuspended in 30% H_2O_2 and boiled for four hours in 170°C. Finally, the samples underwent another four washes with distilled water. For LM, the material was air-dried on cover glasses, then

mounted on microscope glass with Naphrax (Brunel Microscopes Ltd., Chippenham, UK) solution.

Light microscopy documentation was taken at the University of Szczecin with a Zeiss Axio Scope A1 (Carl Zeiss, Jena, Germany) with immersion lens (Zeiss Plan Apochromat 100×) using a Canon EOS 500D camera with the Canon EOS Utility software (Canon, Tokyo, Japan). Images were taken at 1000× magnification using a 100× Plan Apochromat oil immersion objective (numerical aperture (NA) = 1.4). During the observations, we examined a total of 62 valves under the light microscope.

For SEM analysis, a drop of cleaned sample was applied to a Nuclepore Track-Etch membrane (Whatman, Maidstone, UK). After air-drying the membranes overnight, they were affixed to aluminum stubs using carbon tape and subsequently gold-coated using a Q150T coater (Quorum Technologies, Laughton, U.K.). SEM observations were conducted at the University of Rzeszów (Poland), using a Hitachi SU8010 microscope (Hitachi, Tokyo, Japan) with accelerating voltage = 5 kV and working distance = 8500–8600 µm for cultured specimen, and a Phenom ProX (PIK Instruments sp. z o.o., Piaseczno, Polska) for environmental sample specimen.

Phylogenetic analysis

Genomic DNA was isolated from monospecific, non-axenic cultures of diatoms using Chelex® 100 resin (cat. no. 142-2842-MSDS, Bio-Rad, Hercules, CA, USA; Walsh et al. 1991). 150 µL of 10% Chelex® working solution was added into 1.5 mL Eppendorf tubes containing pelleted diatom biomass harvested from monocultures and heated in 95°C, after 20 min the samples were centrifuged at 10 000 rpm for 5 min. The resultant supernatant containing diatom DNA was transferred into new, sterile 1.5 mL Eppendorf tubes.

The primers and PCR protocol for the plastid-encoded *rbcL* molecular marker followed Theriot et al. (2010). Sanger sequencing was performed by GENEWIZ® (Azenta Life Sciences, Burlington, Massachusetts, USA; https://www.genewiz.com). Maximum Likelihood (ML) analysis was conducted with 260 diatom strains and two strains of *Triparma pacifica* (Guillou & Chrétiennot-Dinet) Ichinomiya & Lopes Santos that served as an outgroup. The sequence dataset was constructed using sequences available in GenBank before 02.2024. The sequences were aligned using MAFFT 7 (Katoh & Standley 2016) and trimmed with the -automated1 option with trimAl (Capella-Gutiérrez et al. 2009) into 1449 bp long alignment. Prior to the ML analysis the alignment underwent evaluation for the best

evolutionary model using ModelTest-NG (Flouri et al. 2015, Darriba et al. 2020), which selected TIM3+I+G4. Phylogenetic analysis was performed using IQ-TREE 2.2.0 (Nguyen et al. 2015, Minh et al. 2020) with 1000 bootstrap replicates using ultra-fast bootstrap replications; the resulting tree was then visualized using MEGA 11 (Tamura et al. 2021), and bootstrap values (bv) were added at the nodes if bigger than 50%.

Results

Class: Bacillariophyceae Haeckel

Order: Thalassiophysales D.G.Mann

Family: Catenulaceae Mereschkowsky

Genus: *Halamphora* (Cleve) Levkov

***Halamphora witkowskii* Yilmaz, Solak & Gastineau sp. nov.**

Description: LM (Figs 2E-R) Valves semi-lanceolate and dorsiventral with strongly convex dorsal margin and a ventral margin that is almost straight (Fig. 2H, J) or weakly convex (Fig. 2E, L) or more convex in smaller valves (Fig. 2R). Valve ends shortly protracted, rostrate to almost subcapitate in the largest valves and weakly bent towards the ventral side. Valve length 14–31 µm, width 4–6 µm. Axial area narrow, more expressed on ventral side (Figs 2E-R). Central area is enlarged in the middle on the ventral side. Raphe biarcuate with slightly dorsally bent proximal endings, though this feature is barely visible in LM. Dorsal striae finely punctate, radiate throughout; 20–22 in 10 µm; ventral striae hard to resolve sometimes faintly recognizable near the ventral margin (Figs 2E-R). One H-shaped chloroplast per cell appressed to the ventral side, daughter cells often remaining attached to each other for some hours after division (Figs 2A-D) while the chloroplast rearranges within the cell.

Description: SEM (Figs 3A-G) A raphe ledge is present on the dorsal side of the raphe (Fig. 3A). The raphe is weakly arched with closely positioned proximal endings, slightly dorsally bent and expanded into central pores. The distal raphe endings are slightly deflected towards the dorsal side (Fig. 3C). Dorsal striae biseriate throughout (except immediately adjacent to the raphe-sternum), but interrupted by longitudinal bars (vimines), which isolate sections corresponding to the 'puncta' visible in LM (Fig. 3E). The bars become more frequent towards the dorsal margin. Dorsally the areolae are small and circular, except for the first row of areola in each stria, which is actually located under the raphe ledge and usually a

156 single apically elongate poroid, separated from the remainder of the stria by a wide space;
157 here, therefore, the stria is uniseriate and only occasionally biseriate (Fig. 3C). Ventral striae
158 appear to diminish in length towards the central area, where they become very short and
159 almost rounded, leaving a narrow, semi- lanceolate central area (Fig. 3B). In the centre itself,
160 however, the ventral areolae are rounded and positioned close to the margin, creating a
161 narrow but noticeable lanceolate central area (Fig. 3B). Distally the internal raphe fissure
162 end in poorly developed helictoglossa; proximally, there is a prominent fused central
163 helictoglossa (Fig. 3G).

164 It was not possible to study the girdle in detail because the frustules became disassembled
165 during specimen preparation. In addition, although the use of nucleopore filters is very
166 helpful for removing fine debris (e.g. clay particles in natural samples, broken valves, etc),
167 facilitating the examination of clean valves, one disadvantage is that thin elements like the
168 delicate girdle bands of *Halamphora* pass through the filter and are lost. Nevertheless, the
169 residual bands present in our material showed that the bands are open and that some possess
170 one row of simple round poroids, others two (Figs 3D-F). We did not see any longitudinal
171 ribbing on the girdle bands like that present in *H. oligotraphenta* (Lange-Bertalot) Levkov
172 (Levkov 2009, pl. 234).

173 **Type:** Ahlat Bitlis, Türkiye (38°75'45.748"N, 42°50'71.257"E) collected by: Elif Yilmaz,
174 July 31st, 2021. (holotype SZCZ 27591 University of Szczecin, Poland). Valves representing
175 the holotype population here illustrated in Fig. 2H; isotype Slide number
176 TR_Ahlat_Van_2021 deposited in Kütahya Dumlupınar University (Türkiye).

177 **Etymology:** The new diatom species is dedicated to the esteemed mentor we recently lost,
178 Prof. dr. hab. Andrzej Witkowski, fellow member of the Polish Academy of Sciences. This
179 choice of name celebrates Professor Witkowski's profound impact on diatom research and
180 his enduring legacy in the field, a legacy we humbly hope to maintain alive.

181 **Registration**. http://phycobank.org/xxx

182 **DNA sequences**: The sequences are deposited at Genbank with accession numbers
183 PP476484 (SZCZ EY2176) and PP476485 (SZCZ EY2177).

184 **Distribution:** So far observed only from the type locality.

185 **Ecology:** Lake Van is a saline lake exhibiting a distinct soda chemistry defined by the fact
186 that alkali cations, in particular sodium and potassium, maintain the charge balance of

bicarbonate and carbonate ions in addition to alkaline earth ions (Kazmierczak & Kempe 2003). Soda lakes show large salinity and pH ranges and, especially in closed lakes with high salt contents and alkalinities above 100 mEq/L, the pH may rise above 10 (Eugster & Hardie 1978, Lerman & Stumm 1989, Kempe et al. 1989). Lake Van has pH 9.5–9.9, 21–24 ppt salinity and 155 mEq/L alkalinity.

Molecular phylogenetic analysis

The single-gene phylogenetic analysis based on *rbcL* molecular marker included pennate diatoms belonging to the Bacillariales, Mastogloiales, Naviculales, Rhopalodiales, Surrirelales and Thalassiophysales. The full phylogenetic tree, the list of sequences used and their corresponding accession numbers, as well as the alignment are available at https://doi.org/10.5281/zenodo.13693584. The strains chosen as an outgroup belong to Parmales, the closest diatom relative. The phylogenetic analysis (clade containing *Halamphora* spp. presented in Fig. 4) retrieved both *Halamphora witkowskii* strains (SZCZ EY2176 and EY2177) in a single clade nested inside a large *Halamphora* clade, consisting of 80 *Halamphora* sequences belonging to 60 *Halamphora* species. *Halamphora witkowskii* is sister to clade composed of *H. incelebrata*, *H. intramaritima*, *H. calidilacuna*, *H. americana*, *H. siqueirosii* and *H. bonnevillensis*, and further to *H. sydowii*, and H. *foramina*. The statistical support (bootstrap values) inside this clade are high and not lower than 97, indicating a confidence in calculated evolutionary relationships.

Discussion

The newly discovered diatom species, *Halamphora witkowskii* sp. nov., has been classified within the genus *Halamphora* based on specific morphological and molecular characteristics. The classification as *Halamphora* rather than *Amphora* is primarily supported by the presence of a raphe ledge, a key characteristic of the genus *Halamphora*.

Some of the features of *Halamphora witkowskii* are generic or even familiar characteristics, such as the elongate double helictoglossa linking the internal central raphe endings. While this feature is present in both *Halamphora* and *Amphora* (Levkov 2009), it is important to note that prominent fused central helictoglossa are fairly rare within *Amphora*. This rarity further supports the classification of *H. witkowskii* within *Halamphora*. More distinctive are some aspects of the striae. Many *Halamphora* species have uniseriate striae, for example *H. coffeaeformis* (C.Agardh) Mereschkowsky, *H. mosensis* Stepanek & Kociolek, *H. isumiensis* Stepanek, Mayama & Kociolek, *H. specensa* J.G.Stepanek & Kociolek, *H. tumida* (Hustedt)

Levkov. Biseriate striae are found in several species and in some of these, the stria structure is uniform from the raphe-sternum to the distal margin, e.g. in *H. lineata* (Gregory) Levkov (Levkov 2009, pl. 244). A rather similar structure to that of *H. witkowskii* is evident in *Halamphora* "spec. 3" of Levkov (2009, pl. 240), where, towards the dorsal side, the biseriate striae are broken into sections by thicker bars (vimines), and even more in *Halamphora* "spec. 2" (Levkov 2009, pl. 239), which possesses the same row of single areolae immediately adjacent to the raphe-sternum as in *H. witkowskii*. This row is also present in for example *H. sydowii* (Levkov 2009, pl. 242, fig. 4), *H. minima* (An et. al. 2022) and *H. oceanica* (Olivares-Rubio et al. 2017).

It is particularly the distinct dorsal striae arrangement, together with morphometrics, that distinguish *H. witkowskii* from others such as *Halamphora minima* S.M.An, J.H.Kim, N.S.Kang, K.Cho, J.A.Lee & E.S.Kim (27–29 in 10 μm), *H. coffeaeformis* (C.Agardh) Mereschkowsky (19–22 in 10 μm), and *H. tumida* (Hustedt) Levkov (16–18 in 10 μm).

However, it is in conjunction with the *rbcL* data that *H. witkowskii* becomes more clearly defined compared to other *Halamphora* species.

This study included *rbcL* gene sequence analyses to understand the relationships of *H. witkowskii* within *Halamphora*. The results indicate that *H. witkowskii* belongs to clade K, as defined by Stepanek & Kociolek (2019), which also includes *H. bonnevillensis* Stepanek & Kociolek (Stepanek and Kociolek 2019), *H. sydowii* (Cholnoky) Levkov (Levkov 2009), *H. tumida* (W. Gregory) Levkov (Levkov 2009), and *H. siqueirosii* López-Fuerte, S.E.Sala and Murugan (López-Fuerte et al. 2020), *H. americana*, *H. calidilacuna*, *H. intramaritima*, *H. incelebrata* and *H. foramina*. Comparisons among these taxa show differences in morphological characteristics such as valve shape, edge structures, and features of the striae, as well as the metrics such as length and width and stria density included in Table 1. Specifically, *H. witkowskii* exhibits smaller valve sizes and smoother edges compared to *H. bonnevillensis* and *H. sydowii*. *H. bonnevillensis* has wider and generally more irregular edges, while *H. sydowii* features more pronounced projections and distinct striae patterns. Furthermore, *H. tumida* and *H. siqueirosii* have thicker and denser striae compared to *H. witkowskii*. Thus, *H. witkowskii* exhibits smaller sizes and smoother valve edges compared to *H. bonnevillensis* and *H. sydowii.* The valves of *Halamphora sydowii* have a distinctly curved dorsal margin and typically rostrate apices. Similarly, the smaller valves of *Halamphora witkowskii* also have a quite convex dorsal margin and rostrate apices.

However, the key differences between the two species are as follows: *Halamphora witkowskii* generally exhibits a less pronounced dorsal margin curvature and more acicular (needle-like) apices, while *Halamphora sydowii* shows more pronounced rostrate apices and a more curved dorsal margin. On the other hand, the valves of *Halamphora bonnevillensis* feature a narrower and dome-shaped ventral margin. In contrast, the ventral margin of *Halamphora witkowskii* is straight, while in *H. bonnevillensis*, the margin has more pronounced outward projections. Additionally, *H. bonnevillensis* is distinguished from *H. witkowskii* by its semi-lanceolate and dorsiventral valves with shorter and rostrate valve ends. These characteristics aid is helpful in separating and hence also in identifying these taxa.

Acknowledgement

Co-financed by the Minister of Science under the "Regional Excellence Initiative" Program for 2024-2027 (RID/SP/0045/2024/01).

References

An, S. M., Kim, J., Kang, N. S., Cho, K., Lee, J. A., & Kim, E. S. (2022). *Halamphora minima* (Catenulaceae, Bacillariophyta), new brackish diatom species from the mudflat in Hampyeong Bay, Korea. *Phytotaxa 572*, 115–122. https://doi.org/10.11646/phytotaxa.572.1.9

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics 25*, 1972–1973. https://doi.org/10.1093/BIOINFORMATICS/BTP348

Cleve, P.T. (1895) Synopsis of the Naviculoid Diatoms, Part II. *K. Svenska-Vetenskaps-Akad. Handl. 27*, 1–219.

Darriba, Di., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Mol. Biol. Evol. 37*, 291–294. https://doi.org/10.1093/MOLBEV/MSZ189

Degens, E. T., Wong, H. K., Kempe, S., & Kurtman, F. J. G. R. (1984). A geological study of Lake Van, eastern Turkey. *Geol. Rundschau 73*, 701-734.

Eugster, H. P., & Hardie, L. A. (1978). Saline lakes. In Lerman A (ed) *Lakes: chemistry, geology, physics* (pp. 237-293). New York: Springer.

Flouri, T., Izquierdo-Carrasco, F., Darriba, D., Aberer, A. J., Nguyen, L. T., et al. (2015). The Phylogenetic Likelihood Library. *Syst. Biol. 64*, 356–362. https://doi.org/10.1093/SYSBIO/SYU084

Guillard, R. R. L. (1975). Culture of phytoplankton for feeding marine invertebrates, in *Culture of Marine Invertebrate Animals*, edited by W. L. Smith and M. H. Chanley, pp. 26–60, Plenum Press, New York.

Guiry, M. D., & Guiry, G. M. (2022). *AlgaeBase*. World-wide electronic publication, National University of Ireland, Galway. https://www.algaebase.org; searched on 21 March 2024Jiang, H. Y., Hu, C. Q., Yang, H. P., Zhang, L. P., Peng, P. F., et al. (2015). Morphology and phylogeny of *Halamphora yongxingensis* sp. nov.(Bacillariophyta), a new marine benthic diatom isolated from Yongxing Island, South China Sea. *Phytotaxa 195*, 53–64. https://doi.org/10.11646/phytotaxa.195.1.3

Kadioğlu, M., Şen, Z., & Batur, E. (1997). The greatest soda-water lake in the world and how it is influenced by climatic change. *Ann. Geophys. 15*, 1489–1497.

Katoh, K., & Standley, D. M. (2016). A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics 32*, 1933–1942. https://doi.org/10.1093/BIOINFORMATICS/BTW108

Kazmierczak, J., & Kempe, S. (2003). Modern terrestrial analogues for the carbonate globules in Martian meteorite ALH84001. *Naturwiss. 90*, 167–172.

Kempe, S., Kazmierczak, J., & Degens, E. T. (1989). The soda ocean concept and its bearing on biotic evolutio In *Origin, Evolution, and Modern Aspects of Biomineralization in Plants and Animals* (ed. RE Crick). Plenum Press, Springer, New York, pp. 29–43.

Lange-Bertalot, H., Külbs, K., Lauser, T., Nörpel-Schempp, M., & Willmann, M. (1996). Diatom taxa introduced by Georg Krasske. Documentation and revision. Dokumnetation und Revision der von Georg Krasske beschribenen Diatomeen-Taxa. *Iconographia Diatomologica 3,* 2–358.

Legler, F., & Krasske, G. (1940). Diatomeen aus dem Vansee (Armenien). *Beiträge zur ökologie der brackwasserdiatomeen I. Beihefte zum Botanischen Centralblatt*, *60*, 335–345.

311 Lerman, A., & Stumm, W. (1989). CO2 storage and alkalinity trends in lakes. *Water Res. 23*,
312 139–146.

313 Levkov, Z. (2009) *Diatoms of Europe: Diatoms of the European inland waters and*
314 *comparable habitats. vol. 5*. Amphora sensu lato. Ruggell: A.R.G. Gantner Verlag.

315 López-Fuerte, F. O., Sala, S. E., Lora-Vilchis, M. C., & Murugan, G. (2020). *Halamphora*
316 *siqueirosii* (Bacillariophyta), a new diatom species isolated from a hypersaline evaporation
317 pond in Baja California Peninsula, Mexico. *Phytotaxa 451*, 132–144.
318 https://doi.org/10.11646/phytotaxa.451.2.3

319 Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., et al. (2020).
320 IQ-TREE 2: New models and efficient methods for phylogenetic ınference in the genomic
321 era. *Mol. Biol. Evol. 37*, 1530–1534. https://doi.org/10.1093/MOLBEV/MSAA015

322 Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast
323 and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol.*
324 *Biol. Evol. 32*, 268–274. https://doi.org/10.1093/MOLBEV/MSU300

325 Olivares-Rubio, H.F., Cabrera, L.I., Godínez-Ortega, J.L., Salazar-Coria, L., & Vega-López,
326 A. (2017) *Halamphora oceanica* (Catenulaceae, Bacillariophyta), a new species from the
327 epipelagic region of the southwestern Gulf of Mexico. *Phytotaxa 317*, 188–198.
328 https://doi.org/10.11646/phytotaxa.317.3.3

329 Özalp, S., Aydemir, B. S., Olgun, Ş., Şimşek, B., Elmaci, H., et al. (2016). Tectonic
330 deformations in the quaternary deposits of the Lake Van (Edremit Bay), Eastern Anatolia,
331 Turkey. *Bull. Miner. Res. Explor. 153*, 45–61.

332 Sala, S. E., Guerrero, J. M., & Coste, M. (2007). Valve morphology of *Amphora chilensis*
333 Hustedt (Bacillariophyceae). *Nova Hedwigia 85*, 353–364. https://doi.org/10.1127/0029-
334 5035/2007/0085-0353

335 Solak, C. N., Gastineau, R., Lemieux, C., Turmel, M., Gorecka, E., et al. (2021). *Nitzschia*
336 *anatoliensis* sp. nov., a cryptic diatom species from the highly alkaline Van Lake (Turkey).
337 *PeerJ 9*, e12220.

338 Stepanek, J. G., & Kociolek, J. P. (2013). Several new species of *Amphora* and *Halamphora*
339 fromthe western USA. *Diatom Res. 28,* 61–76.
340 https://doi.org/10.1080/0269249X.2012.735205

341 Stepanek, J. G., & Kociolek, J. P. (2018). *Amphora* and *Halamphora* from coastal and inland
342 waters of the United States and Japan, with the description of 33 new species. *Bibliotheca*
343 *Diatomol. 66*, 1–260.

344 Stepanek, J. G., & Kociolek, J. P. (2019). Molecular phylogeny of the diatom genera
345 *Amphora* and *Halamphora* (Bacillariophyta) with a focus on morphological and ecological
346 evolution. *J. Phycol. 55*, 442–456.

347 Tamura, K., Stecher, G,. & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics
348 Analysis Version 11. *Mol. Biol. Evol. 38*, 3022–3027.
349 https://doi.org/10.1093/MOLBEV/MSAB120

350 Theriot, E. C., Ashworth, M., Ruck, E., Nakov, T., & Jansen, R. K. (2010). A preliminary
351 multigene phylogeny of the diatoms (Bacillariophyta): challenges for future research. *Plant*
352 *Ecol. Evol. 143*, 278–296.

353 Van de Vijver, B., Kopalová, K., Zidarova, R., & Levkov, Z. (2014). Revision of the genus
354 *Halamphora* (Bacillariophyta) in the Antarctic Region. *Plant Ecol. Evol. 147*, 374–391.

355 Walsh, P. S., Metzger, D. A., & Higuchi, R. (1991). Chelex 100 as a medium for simple
356 extraction of DNA for PCR-based typing from forensic material. *Biotechniques 10*, 506–
357 513.

358 Wang, P., Park, B. S., Kim, J. H., Kim, J. H., Lee, H. O., & Han, M. S. (2014). Phylogenetic
359 position of eight *Amphora sensu lato* (Bacillariophyceae) species and comparative analysis
360 of morphological characteristics. *Algae 29*, 57–73.
361 https://doi.org/10.4490/algae.2014.29.2.057

362 Yılmaz, E., Mann, D. G., Gastineau, R., Trobajo, R., Solak, C. N., et al. (2024). Description
363 of *Navicula vanseea* sp. nov. (Naviculales, Naviculaceae), a new species of diatom from the
364 highly alkaline Lake Van (Republic of Türkiye) with complete characterization of its
365 organellar genomes and multigene phylogeny. *PhytoKeys 241*, 27–48.
366 https://doi.org/10.3897/phytokeys.241.118903

367 You, Q. M., Kociolek, J. P., & Wang, Q. X. (2015). Taxonomic studies of the diatom genus
368 *Halamphora* (Bacillariophyceae) from the mountainous regions of southwest of China,
369 including the description of two new species. *Phytotaxa 205*, 75–89.

370 Zhang, W., Wang, T., Levkov, Z., Jüttner, I., Ector, L., & Zhou, Q. C. (2019). *Halamphora*
371 *daochengensis* sp. nov., a new freshwater diatom species (Bacillariophyceae) from a small
372 mountain lake, Sichuan Province, China. *Phytotaxa 404*, 12–22.

373

374 **TABLE 1.** Morphometric comparation of *Halamphora witkowskii* sp. nov. with similar
375 species. ND = not documented.

Taxa	Valve length (µm)	Valve width (µm)	Dorsal striae (10 µm)	Ventral striae (10 µm)	Habitat	Reference
***Halamphora witkowskii* sp. nov.**	**14–16**	**4–5**	**20–22 biseriate**	**30-32**	**Brackish, soda lake**	**This study**
Halamphora minima	5.6–7.4	2.4–3.3	27–29 biseriate	43–45	Brackish	An et al. 2022
Halamphora eunotia	20.7-27.6	6.8-8.6	13–14 uniseriate	ND	ND	Wang et al. 2014
Halamphora coffeaeformis	23–35 (?14–55)	(3.5) 5–7.2	19–22 uninterrupted	ND	Brackish	Levkov 2009
Halamphora tumida	19-33	4-7	16–18 biseriate	22-28	Brackish	Levkov 2009
Halamphora oceanica	9-17	2.4-5.1	26-30 biseriate	48-62	Marine	Olivares-Rubio et al. 2017
Halamphora americana	23-30	4-5	19 bi- and uniseriate	28-31	Freshwater, Brackish	Stepanek & Kociolek 2018

376

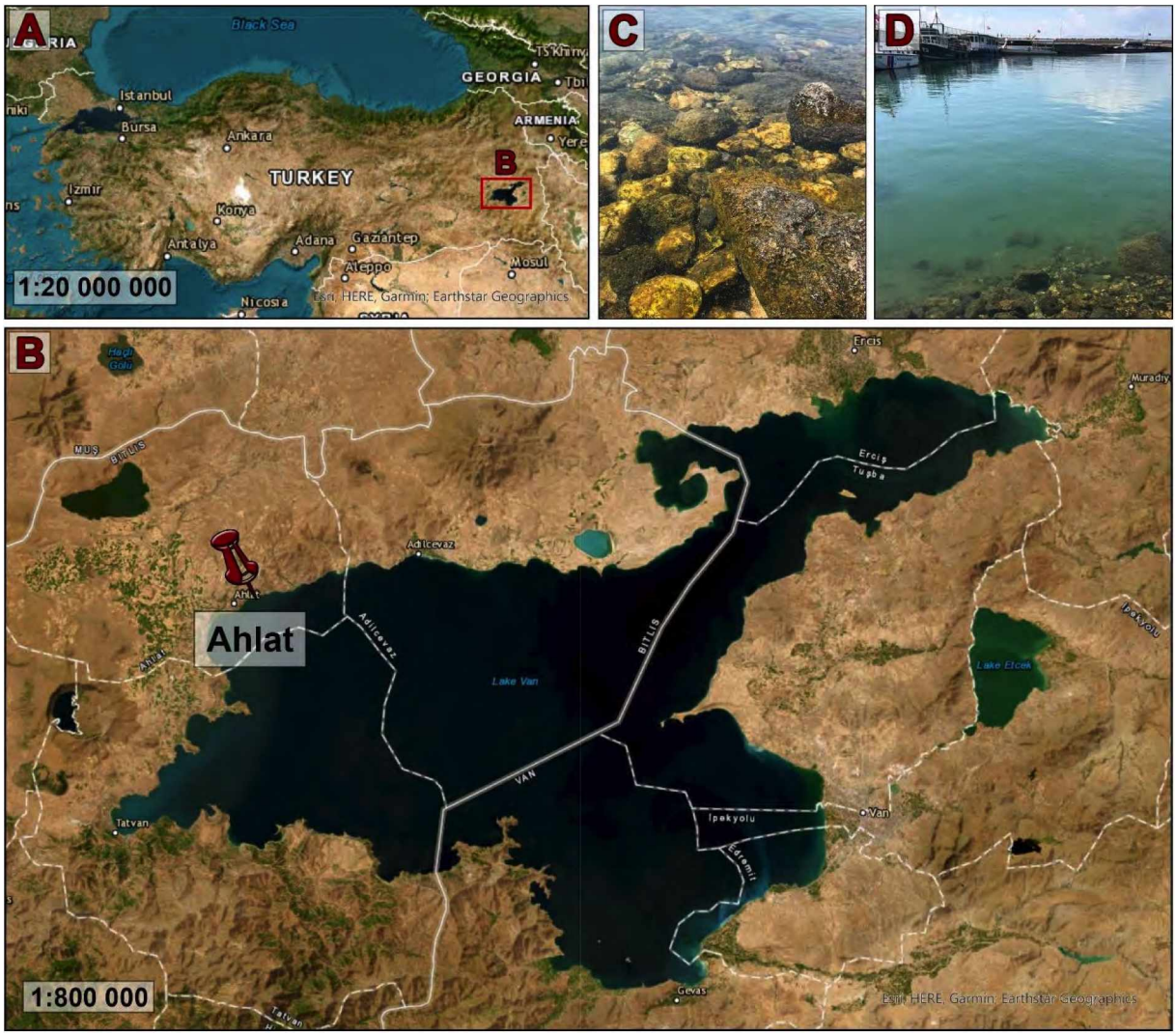


Figure 1: Map of the sampling location. A. Location of Lake Van in Turkey. The red frame indicates the position of Lake Van. B. General view of the lake. The pin indicates the position of the sampling area. C. Photo of the epilithic sampling area on the rock. D. Photo of the general sampling area. (Maps generated with Esri. (2023). ArcGIS Pro 3.1.0. Environmental Systems Research Institute).

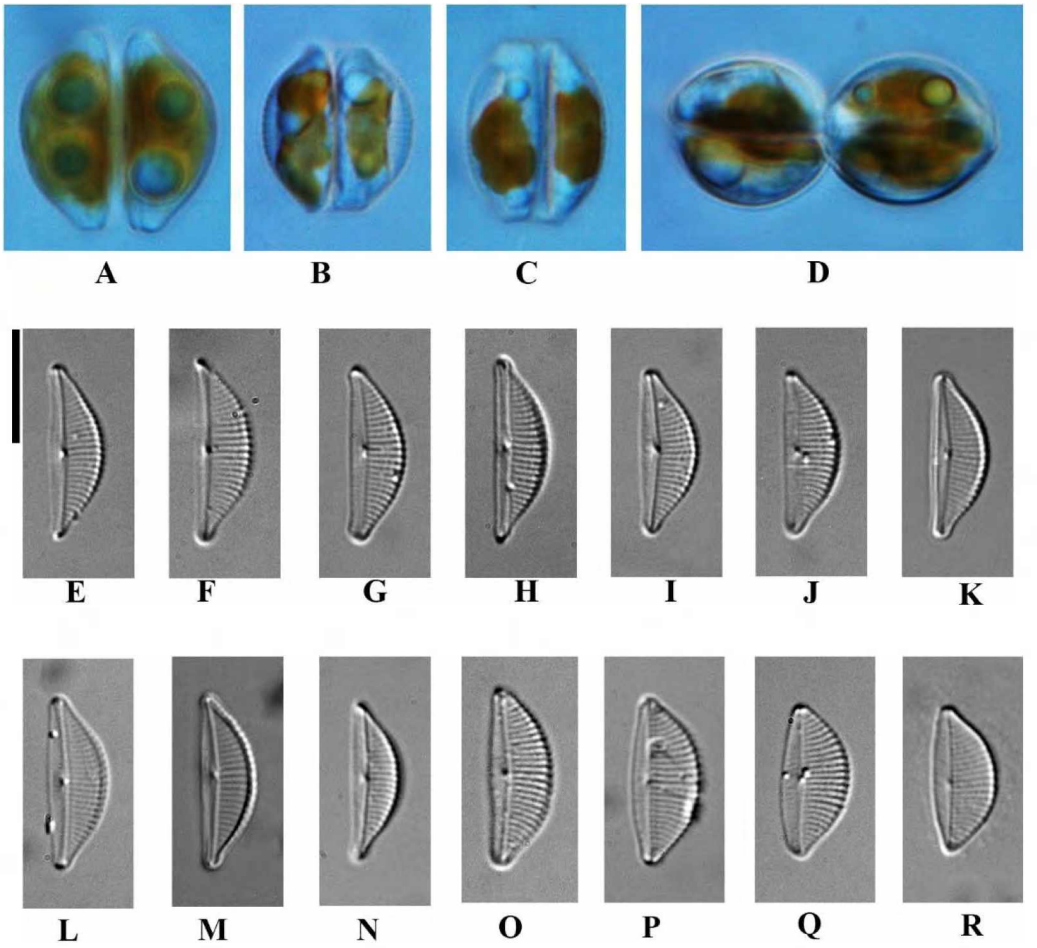


Figure 2. *Halamphora witkowskii* sp. nov. LM micrographs. (A–D): *in vivo* pictures of *Halamphora witkowskii* sp. nov. SZCZ EY2176. (E–H): LM image of a cleaned valve from environmental material. (I–N): cleaned valves of *Halamphora witkowskii* sp. nov. SZCZ EY2177. (O–R): cleaned valves of *Halamphora witkowskii* sp. nov. SZCZ EY2176. Scale bar = 10 μm.

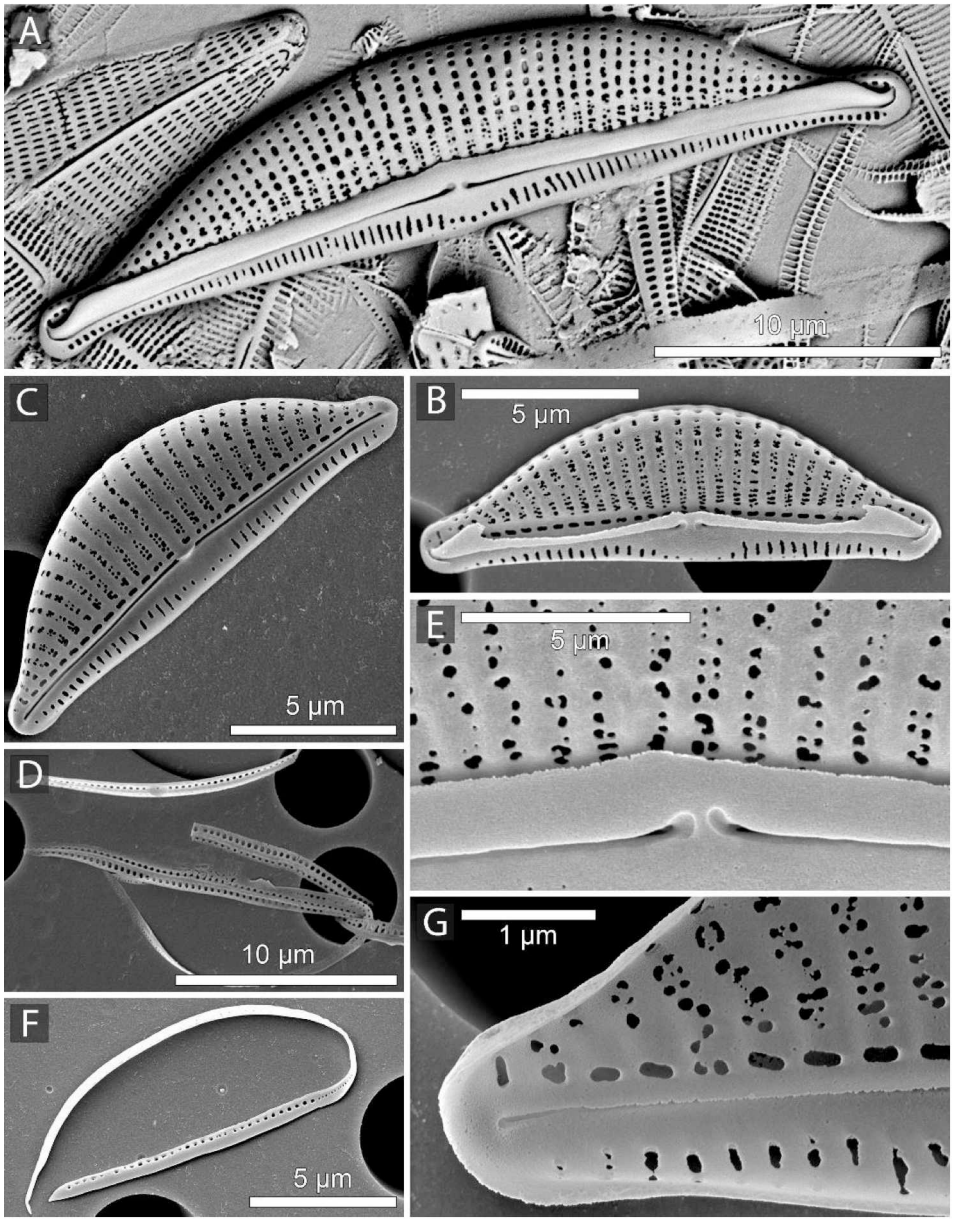
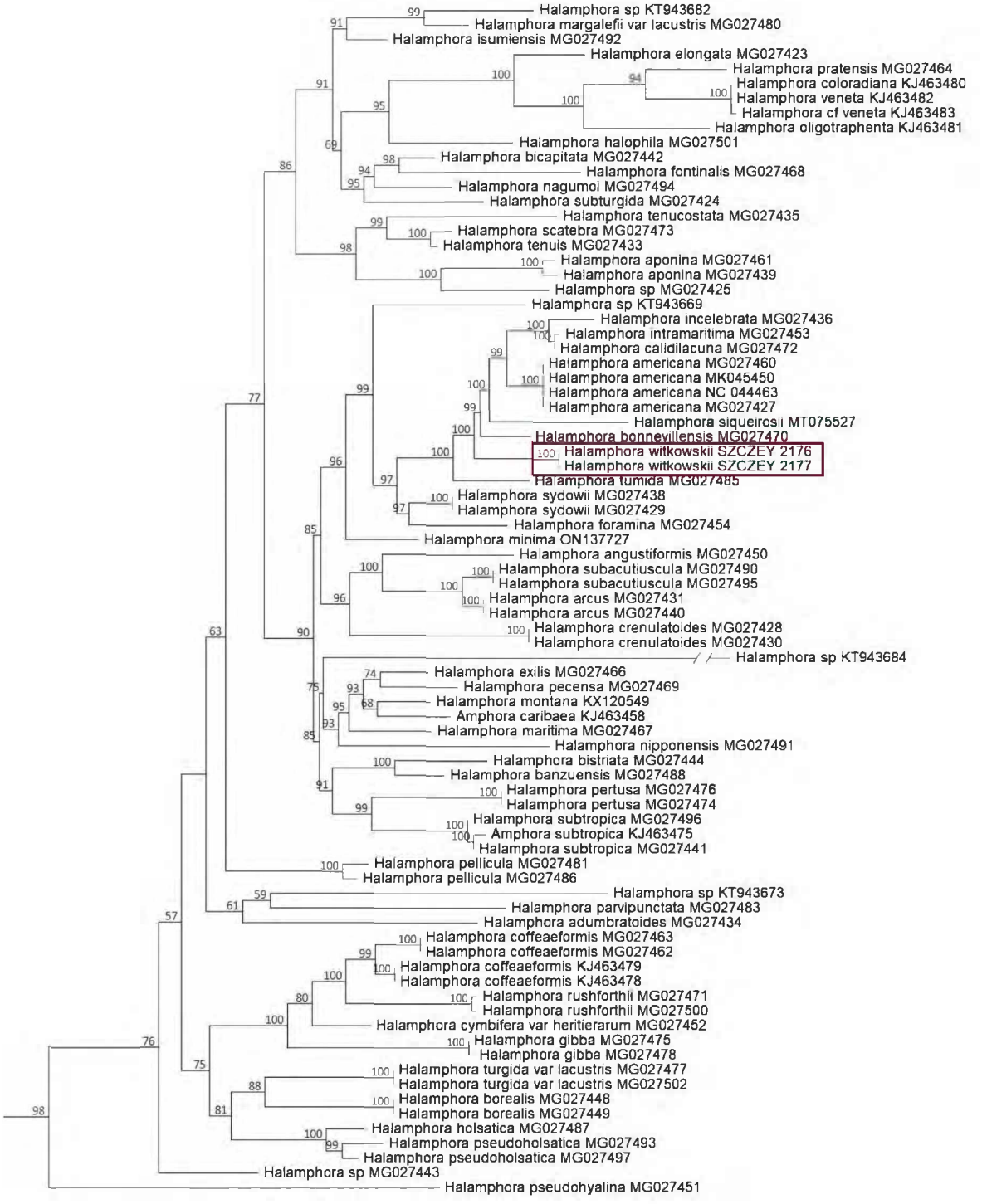


Figure 3. SEM micrographs of *Halamphora witkowskii* sp. nov. (A) External view of the entire valve from environmental material. (B) External whole valve view from culture material. (C) Internal view of the entire valve from culture material. (D) girdle bands with two rows of poroids. (E) External view of the central part of the valve, showing the raphe ledge, central raphe endings, and part of the biseriate striae. (F) girdle band with single row of poroids. (G) Detail of internal valve apex showing the proximal helictoglossa. Scale bars = (A, D) 10 µm (B, C, E, F) 5 µm, (G) 1 µm.



404

405 **Figure 4:** The clade of *Halamphora* spp. cut from Maximum Likelihood tree inferred from
406 the alignement of 260 *rbcL* sequences of diatoms and rooted with 2 *Triparma pacifica*.
407 Rectangle indicates the position of two strains of *Halamphora witkowskii*.