scientific reports

Check for updates

OPEN

Nitzschia excavata sp. nov. (Bacillariaceae), a new diatom species from a post-mining reservoir revealed by morphology, molecular phylogeny, and metabarcoding-based biogeography

Rafał M. Olszyński[1,2]✉, David G. Mann[2,3], Piotr K. Zakrzewski[4], Łukasz Peszek[5], Éva Ács[6,7], Sára Shemesh[6,7,8,9] & Rosa Trobajo[2]

The Bogdałów post-mining reservoir (Poland) represents a slightly alkaline, moderately mineralised ecosystem formed by flooding a former lignite pit. Its anthropogenic origin and stable physicochemical conditions have enabled the development of species-rich diatom assemblages, particularly numerous *Nitzschia* (Bacillariaceae) species. To explore this diversity, an integrative approach combining microscopy and DNA-based analyses was employed. Morphological examinations were performed using light and scanning electron microscopy, as well as confocal laser scanning microscopy. Molecular phylogenetic analyses were based on the sequencing of the nuclear SSU rDNA and the chloroplast *rbc*L and *psb*C gene markers. This comprehensive study led to the discovery and formal description of *Nitzschia excavata* sp. nov., distinguishable by unique morphological features and a phylogenetically distinct lineage. Furthermore, environmental DNA metabarcoding and metagenomic database searches revealed sequences identical or closely related to the *N. excavata* sp. nov. lineage in freshwater habitats across Europe and China, indicating that this taxon has an unexpectedly broad distribution. These findings underscore the value of integrating classical morphological analysis with multi-marker molecular data in diatom taxonomy and demonstrate that anthropogenic habitats may support taxa with broader distributions than previously recognized. The study highlights the important role of metabarcoding and metagenomics in revealing cryptic diversity and clarifying the biogeographic patterns of newly described species.

Keywords *Nitzschia*, Diatoms, Integrative taxonomy, Metabarcoding, Post-mining reservoir, Freshwater biogeography

[1]Department of Algology and Mycology, Faculty of Biology and Environmental Protection, University of Lodz, ul. Banacha 12/16, 90-237 Lodz, Poland. [2]IRTA-Institute for Food and Agricultural Research and Technology, La Rápita, Catalonia, Spain. [3]Royal Botanic Garden Edinburgh, Edinburgh, Scotland, UK. [4]Department of Molecular Neurochemistry, Medical University of Lodz, ul. Mazowiecka 6/8, 92-215 Łódź, Poland. [5]Department of Agroecology and Forest Utilization, University of Rzeszów, ul. Ćwiklińskiej 1A, 35-601 Rzeszow, Poland. [6]Department of Aquatic Environmental Sciences, Faculty of Water Sciences, Ludovika University of Public Service, Ludovika tér 2, H-1083 Budapest, Hungary. [7]National Laboratory for Water Science and Water Security, Ludovika tér 2, H-1083 Budapest, Hungary. [8]Faculty of Science and Informatics, Doctoral School of Biology, University of Szeged, Közép fasor 52, H-6726 Szeged, Hungary. [9]Mutagenesis and Carcinogenesis Research Group, Institute of Genetics, HUN-REN BRC, Temesvári krt. 62, H-6726 Szeged, Hungary. ✉email: rafal.olszynski@biol.uni.lodz.pl

Nitzschia Hassall, 1845 (Bacillariophyceae) is a large and heterogeneous diatom genus comprising hundreds of described species across multiple phylogenetic lineages[1,2]. AlgaeBase currently lists over 1900 *Nitzschia* taxa (accessed 19 October 2025)[3]; however, according to the GenBank database (http://www.ncbi.nlm.nih.gov), sequences have only been submitted for 95 identified species (accessed 19 October 2025). Within *Nitzschia*, the Lanceolatae section[4] is among the most speciose. This group, characterised by a strongly eccentric keel, linear or lanceolate valves, and clearly visible fibulae[4], comprises small-celled diatoms that are notoriously difficult to identify. Recent studies based on phylogenetic analysis using *rbc*L and other markers revealed a clade of Lanceolatae that includes *N. palea* (Kützing) W. Smith, 1856, *N. lembiformis* F. Meister ex Clavero, 2009, *N. pusilla* Grunow, 1862, and *N. capitellata* Hustedt, 1930[1,2]; it is a subclade of clade 6B in the analysis of Bacillariaceae by Mann et al.[2]. These species share some morphological characteristics, including small size (widths < 5 µm), linear-lanceolate or linear shape, uniseriate striae composed of simple poroid areolae, often two (or more) longitudinal rows of areolae on the raphe canal, an absence of central raphe fissures (as opposed to the 'Obtusae' subclade of clade 6B, where they are present), and a double row of poroids on at least one girdle band[2], though none of these are unique to the subclade.

During a study of recently formed reservoirs in Poland, we discovered a *Nitzschia* that seemed to fall within clade 6B according to gene markers, but was separate from any previously sequenced species, and did not seem to agree with any described species morphologically. Reservoirs formed as a result of mining activities are highly artificial aquatic habitats frequently supporting distinctive diatom communities that may not be found in natural water bodies, as a result of their unique ecological dynamics, artificial origins, and the specific environmental conditions created by mining activities[5–8]. The mineral composition of the basin and the inflow of mine drainage water can significantly impact both the water chemistry and physical conditions. In certain cases, a single factor, such as elevated chloride levels, can dominate the environment and promote the growth of specialised taxa[6]. These anthropogenic water bodies are also known for frequently exhibiting teratological diatoms[9]. It seems that the unusual water chemistry selects for species that are rarely found elsewhere, so that, despite lying in a region (N Europe) well studied by diatomists, the reservoirs are yielding a number of species new to science e.g.[7,8,10]. Among these is *N. nandorii* Olszyński, Zakrzewski & Żelazna-Wieczorek, 2024, described from the Bogdałów reservoir by Olszyński et al.[10], who were also able to document morphologically similar species such as *N. lacuum* Lange-Bertalot, 1980, and *N. alpinobacillum* Lange-Bertalot, 1993, co-occurring in the same sample. They were also able to reveal previously undocumented ultrastructural features of these species[10].

However, a major challenge, as always with microscopic algae, is to check whether an apparently new species has been found elsewhere but identified as an existing species, or if it has been completely overlooked, perhaps because of its rarity or because it has little or no morphological signature (as is often the case in the Nitzschiae Lanceolatae). The recent development of metabarcoding provides a new tool for determining diatom distributions, using short sequences of either nuclear SSU rDNA or the plastid gene *rbc*L. SSU has been used very effectively for studying the distributions of marine planktonic diatoms, e.g., from samples collected during the TARA expedition[11], while *rbc*L, currently the favoured marker for DNA-based assessments of water quality and ecological status, has been used to advance our understanding of diatom biogeography by uncovering hidden genetic structures and extensive connectivity in freshwater assemblages. For example, Pérez-Burillo et al.[12] identified over 10,000 Amplicon Sequence Variants (ASVs) in global samples, revealing that more than half of the 1,000 most abundant ASVs occur on multiple continents, indicating broad dispersal and limited endemism. Their dataset also included environmental detections of *N. nandorii,* extending its known distribution beyond the records of Olszyński et al.[10] and demonstrating how metabarcoding can broaden the biogeographic range of newly described taxa. Previous metabarcoding research[13,14] showed that different ASVs within morphologically defined *Nitzschia* species display distinct ecological preferences, highlighting metabarcoding's ability to reveal intraspecific diversity and its possible significance. Overall, *rbc*L metabarcoding provides high genetic resolution and spatial context, making it a powerful tool for studying diatom diversity and dispersal, especially in altered environments like post-mining reservoirs.

This study aims to investigate the taxonomic and morphological diversity of Lanceolatae *Nitzschia* in a lignite post-mining reservoir. We employ an integrative approach combining morphometric and ultrastructural analyses (light and scanning electron microscopy), chloroplast imaging (differential interference contrast and confocal laser scanning microscopy), and sequence data from three genetic markers, nuclear SSU rDNA and chloroplast *rbc*L and *psb*C. We then use the gene sequence data to interrogate metabarcoding and metagenomic datasets from different continents to gain insights into the ecology and biogeography of species. Based on these multi-tiered data, we describe here *Nitzschia excavata* sp. nov. from the post-mining reservoir "Bogdałów".

Results

Environment background

The Bogdałów reservoir is a slightly alkaline freshwater environment, with a pH of around 8, moderate conductivity, and well-oxygenated water. Previous studies[5,10] indicate that the ecological status of the reservoir varies seasonally, ranging from moderate to poor in the summer and autumn, and moderate to very good in the winter and spring. According to the classification by Van Dam et al.[15] this habitat is characterised as alkaline, freshwater, and highly oxygenated, with nutrient levels ranging from mesotrophic to oligotrophic, and saprobic conditions varying from β-mesosaprobic to oligosaprobic[10].

Nitzschia excavata Olszyński, D.G. Mann and Trobajo, sp. nov.

Figures 1–2, 3a–o, u–n', 4, 5 and 6.

Holotype. Slide number: D.BOZS.191022 (environmental sample), Algae Collection Department of Algology and Mycology, University of Lodz. The holotype is illustrated in Fig. 3c (designated here).

Isotype. Slide number: SZCZ 29668, Szczecin Diatomological Collection, University of Szczecin, Poland.

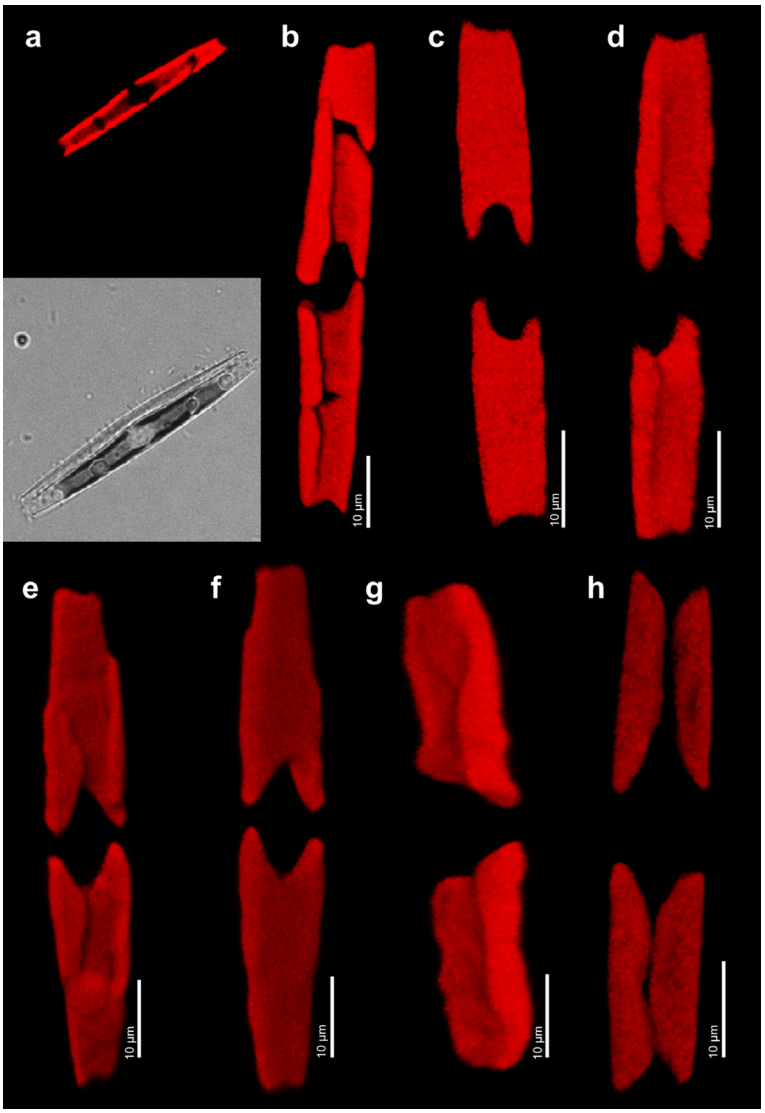

Fig. 1. Confocal Laser Scanning Microscopy photomicrographs of *Nitzschia excavata* sp. nov. chloroplasts from culture material: (**a**–**b**) strain D.LDZ35; (**c**–**d**) strain D.LDZ7; (**e**–**f**) strain D.LDZ13; (**g**–**h**) strain D.LDZ6. (**a**) General overview of a chloroplast without a frustule (upper) and with a frustule context (bottom). (**b**) The presence of four chloroplasts observed in some cells likely results from an incomplete or irregular cell division event. (**c**–**g**) Two elongated plastids are arcuate in cross-section, extending along the valve, forming a semi-tubular structure. (**h**) Four chloroplasts during the final phase of cell division.

Type locality. Poland. Greater Poland Voivodeship, Bogdałów. Post-mining reservoir Bogdałów. 52°2′53.938″N, 18°35′49.646″E.

Morphology

Morphological observations are based on both environmental material (including the type material) and cultured strains. SEM images presented in this paper were obtained exclusively from environmental material (cultured material was also studied in SEM and was morphologically consistent with the natural population, but less well preserved), whereas the LM observations and morphometric measurements presented were from both environmental and cultured material. *Nitzschia excavata* sp. nov. possesses two plastids, elongated and arcuate in cross-section, extending along the valve and onto the girdle, forming a semi-tubular structure (Fig. 1–2, see Supplementary Figs. S1–S2). Volutin granules are formed at the apices of each chloroplast and are clearly visible under light microscopy (Fig. 2f, g). These granules appear as strongly refractive, rounded inclusions and likely represent polyphosphate storage bodies (Fig. 2).

Frustules are elongate and linear-lanceolate, with protracted and slightly subcapitate apices. Valves are 29.0–50.3 µm in length (Me = 46.0, n = 83) and 3.0–4.4 µm in width (Me = 3.51, n = 83) (Fig. 3a–o, u–n'). Striae number 46–49 in 10 µm (Me = 47.0, n = 43) and are indiscernible in LM. Fibulae are visible, equidistantly spaced, and number 15–22 in 10 µm (Me = 18.0, n = 83).

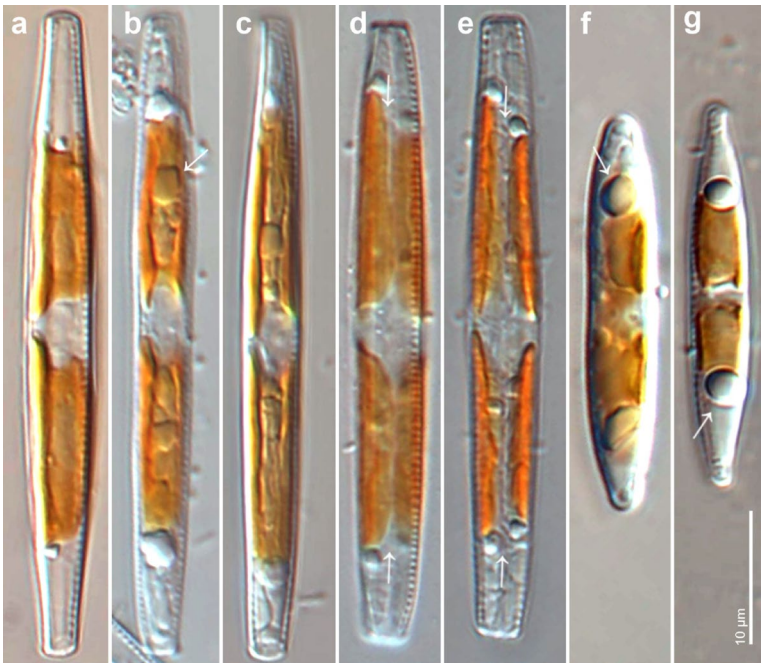

Fig. 2. Light microscope (DIC) photomicrographs of living cells of *Nitzschia excavata* sp. nov. from culture material: (**a**) strain D.LDZ37; (**b**–**e**) strain D.LDZ35; (**f**–**g**) strain D.LDZ36. (**a**) Visible semi-tubular shape of the chloroplasts, with the right and left edges appearing more clearly in a higher focal plane, while the central part is in a lower focal plane. (**b**–**c**) White arrows indicate storage material around pyrenoids. (**d**) Initial phase of cell division with the cleavage furrow (arrows). (**e**) Fully divided plastids (white arrow). (**f**–**g**) Cells with volutin granules (white arrows).

The valve structure is delicate, and chemical cleaning initially yielded valves in which the pore occlusions had been eroded away (Figs. 4b–d, f, 5a, 6c, f). Later, observation without any cleaning showed that each areola is a simple poroid closed by a hymen placed just below its outer aperture (compare Figs. 4e and f; 5b–d). Externally, the valve face is flat and lacks costae (Fig. 4a, b). Striae are uniseriate and transapically oriented (Fig. 4). Along the raphe canal, where the distinction between the raphe canal and the valve face is visible externally as an break in the areola pattern (Fig. 5c, d), two rows of single areolae can often be observed (Fig. 5b). However, the vimines between the areolae are very often reduced, resulting in the formation of transapically elongate areolae. This occurs on the wall of the raphe canal, so that in this case there is only one areola opposite each valve face stria (Figs. 4e, 5c, d), on the valve face adjacent to the raphe canal (Figs. 4c, 5c), and in the distal part of the valve face (Fig. 5d).

A marginal ridge is absent; however, there is a distinct junction between the valve face and the distal mantle, marked by a narrow strip of silica without areolae (Figs. 4b, c, 5a, d). The distal raphe endings are strongly hooked and deflected in different directions depending on the valve (Fig. 4c, f). Central raphe fissures are absent (Figs. 4b, 5a).

On the side of the raphe canal opposite the valve face, i.e., adjacent to the proximal mantle, the pore arrangement is similar to that observed on the side adjacent to the valve face, with two rows of individual areola that very often merge into a single longitudinal pore (Fig. 5e, f). Further on, the mantle begins, clearly separated from the wall of the raphe canal (Figs. 5e, 6e, f), and bears predominantly elongated pores, which likely represent fused individual areolae with reduced vimines (Figs. 4d, 6d, f).

Internally, the valve is flat and lacks any additional morphological elements (Fig. 6a, b). The fibulae are short, narrow, and equidistant (Fig. 6a–c). A wider gap between the central fibulae is absent (Fig. 6a). Each fibula is usually connected to two virgae (Fig. 6c). The distal ends of the raphe slit terminate in a helictoglossa (Fig. 6d, e).

Gene sequences. Eight monoclonal strains were established from the type material (D.BOZS.191022). Strain D.LDZ6 is designated here as the ex-holotype strain (derived from the type material). The remaining strains (D.LDZ7, D.LDZ13, D.LDZ34–38) are treated as additional reference strains and are not nomenclatural types. Sequences were deposited in the GenBank: D.LDZ6 (SSU rDNA: PX843716, *rbc*L: PX860539, and *psb*C: PX860547), D.LDZ7 (*rbc*L: PX860540, and *psb*C: PX860548), D.LDZ13 (SSU rDNA: PX843717, *rbc*L: PX860541, and *psb*C: PX860549), D.LDZ34 (*rbc*L: PX860542, and *psb*C: PX860550), D.LDZ35 (SSU rDNA: PX843718, *rbc*L: PX860543, and *psb*C: PX860551), D.LDZ36 (SSU rDNA: PX843719, *rbc*L: PX860544, and *psb*C: PX860552), D.LDZ37 (SSU rDNA: PX843720, *rbc*L: PX860545, and *psb*C: PX860553), D.LDZ38 (SSU rDNA: PX843721, *rbc*L: PX860546, and *psb*C: PX860554), PhycoBank registration. http://phycobank.org/106537.

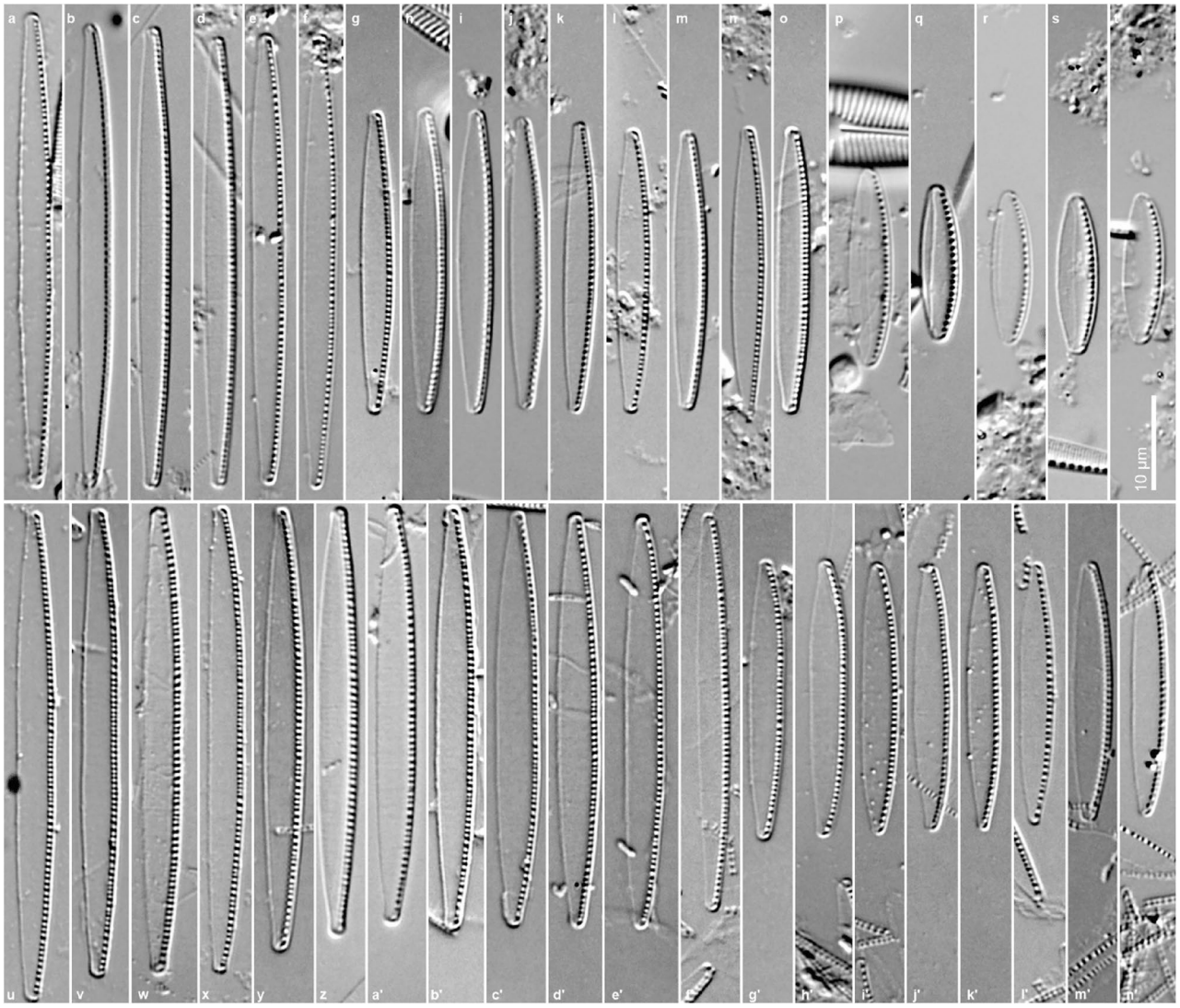

Fig. 3. Light microscopy microphotographs of *Nitzschia* spp. (**a**–**o**) *N. excavata* sp. nov. from environmental material sample D.BOZS.191022. (**c**) Holotype specimen. (**p**–**t**) *N. pusilla* (sample D.BOZS.191022). (**u**–**n'**)*N. excavata* sp. nov. from culture material: (**u**–**x**) strain D.LDZ6; (**y**) strain D.LDZ37; (**z**–**b'**) strain D.LDZ34; (**c'**–**e'**) strain D.LDZ13; (**f'**, **n'**) strain D.LDZ35; (**g'**–**h'**) strain D.LDZ36; (**i'**–**m'**) strain D.LDZ38.

Molecular phylogeny

Phylogenetic reconstructions based on the Maximum Likelihood (ML) and Bayesian Inference (BI) strategies for the concatenated SSU-*rbc*L-*psb*C matrix, as well as separate SSU, *rbc*L, and *psb*C analyses, were performed. The three-gene tree placed the newly described strains (Fig. 7) within a clade of Lanceolatae species, many of which are characterised by double rows of areolae on the raphe canal; the strains were genetically identical (Fig. 7). The *Nitzschia excavata* sp. nov. strains were sister to a clade with two strains, *Nitzschia* sp. SZCZCH1090 and *Nitzschia* cf. *ardua* L44, with maximum node support. In turn, this clade was sister to one containing two strains assigned to *Nitzschia pusilla*, TA420 and TA-45, and *N. lembiformis* strain R2, in both BI and ML phylogenetic reconstructions, with strong support (BI = 1; ML = 98). The three-gene concatenated dataset also included several additional strains identified as *Nitzschia pusilla*, namely TCC896, L1, L3, and L25, and *N. palea*. However, the resulting phylogenetic tree did not place these in the same clade as TA420 and TA-45, nor were they placed together. Instead, they occupied widely distinct positions. Strain TCC896, together with *N. palea* strains (collapsed clade), formed a sister clade to the newly described species, *N. lembiformis*, and *N. pusilla*, while strains L1, L3, and L25 clustered within a collapsed clade comprising various *Bacillariaceae* lineages and the *Durinskia kwazulunatalensis* endosymbiont. Phylogenetic trees reconstructed using separate DNA markers, i.e., *rbc*L, *psb*C, and SSU (see Supplementary Fig. S3 online) also show *N. excavata* sp. nov. on a long branch within clades composed principally of other Nitzschiae Lanceolatae, within a larger clade corresponding to clade 6B of Mann et al.[2].

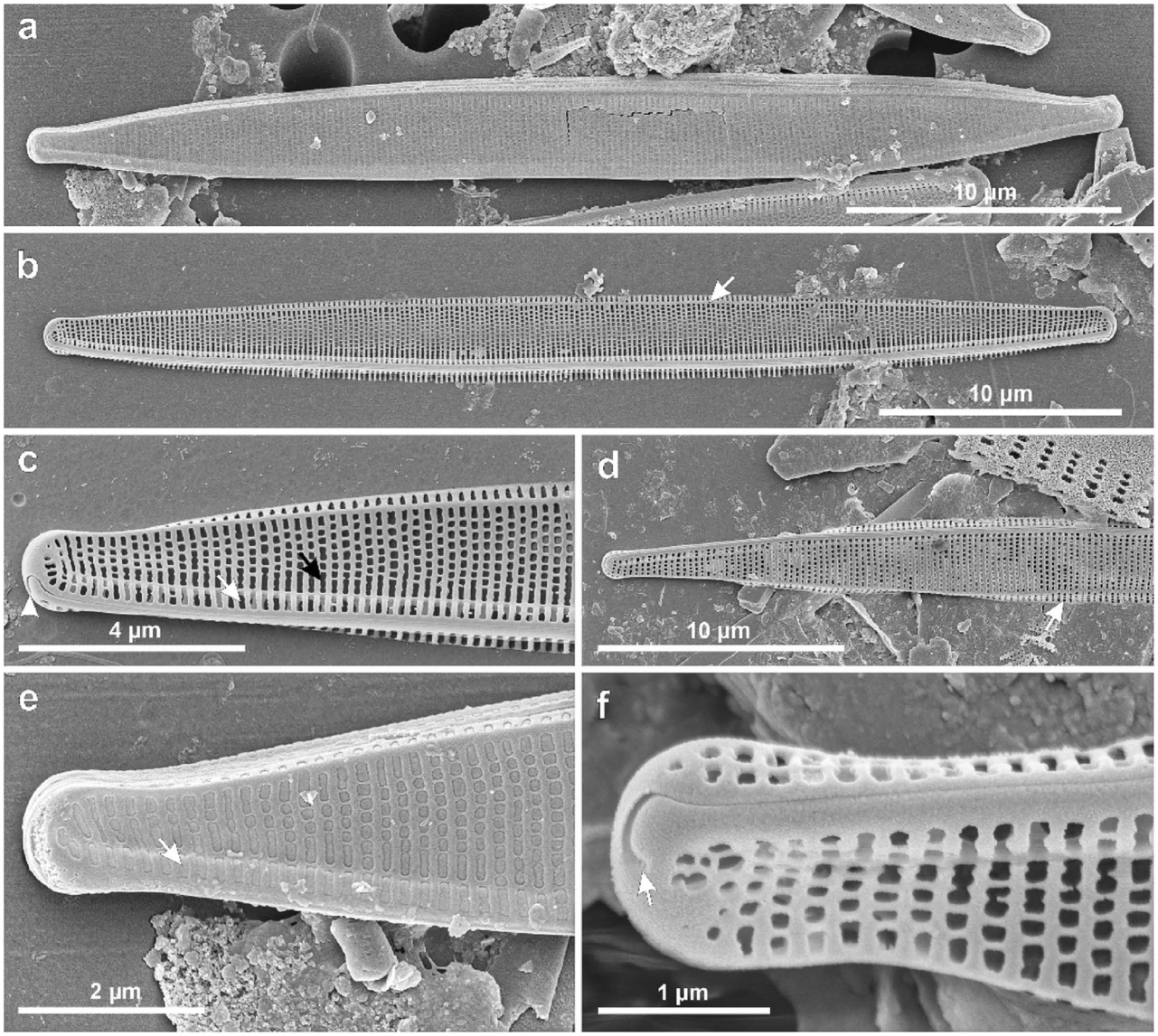

Fig. 4. Scanning electron microscopy photomicrographs of *Nitzschia excavata* sp. nov. from environmental material: (**a**, **c**, **e**, **f**) sample D.BOZM.221024; (**b**) sample D.BZ2M.221024; (**d**) D.BOZS.191022. External views. (**a**) Entire valve with visible hymenate occlusions, (**b**) Eroded valve with visible junction between valve face and distal mantle (arrow), (**c–f**) Close-ups of valve apices. (**c**) Most vimines are present, but some are reduced, forming elongate pores. The white arrow indicates pores located on the raphe canal wall; similar pore morphology is present on the valve face (black arrow). The distal raphe fissure, strongly hooked towards the proximal side, is indicated by an arrowhead. (**d**) Part of the valve with visible distal and proximal mantles; several pores are present on the distal mantle (arrow). (**e**) Close-up of the apex with visible hymenes; the raphe canal wall bears elongate pores formed by reduction of vimines (arrow). (**f**) Polar raphe fissure deflected towards the distal side.

Metabarcoding distribution analysis results

BLAST searches of 24 metabarcoding datasets (Table 1) revealed no close matches to *N. excavata* sp. nov. *rbc*L sequences. The nearest were c. 94–96% similar over the 263- or 331-bp length of the metabarcodes used. Subsequently, searches of a recent dataset from Spain (R. Trobajo et al., unpublished) revealed three sites where *rbc*L sequences identical to *N. excavata* sp. nov. lineage reference sequence were found (100% identity over the 263-bp marker region). These were the River Ebro in Miravet (Catalonia), the Ullals de Baltasar (freshwater springs at the landward margin of the Ebro Delta, Catalonia), and the Río Quípar (a small tributary of the Segura River, Murcia).

Targeted screening of *rbc*L ASVs from the regional Hungarian diatom metabarcoding study revealed the presence of sequences identical or highly similar to the *N. excavata* sp. nov. reference sequence. The fully identical ASV was detected in multiple standing-water habitats, in Lake Velencei and other waters including gravel pit lakes (Bugyi, Délegyháza, Kiskunlacháza, Lupa-sziget, Csepel, Alsózsolca and Onga), oxbow lakes (Tolnai, Bátai,

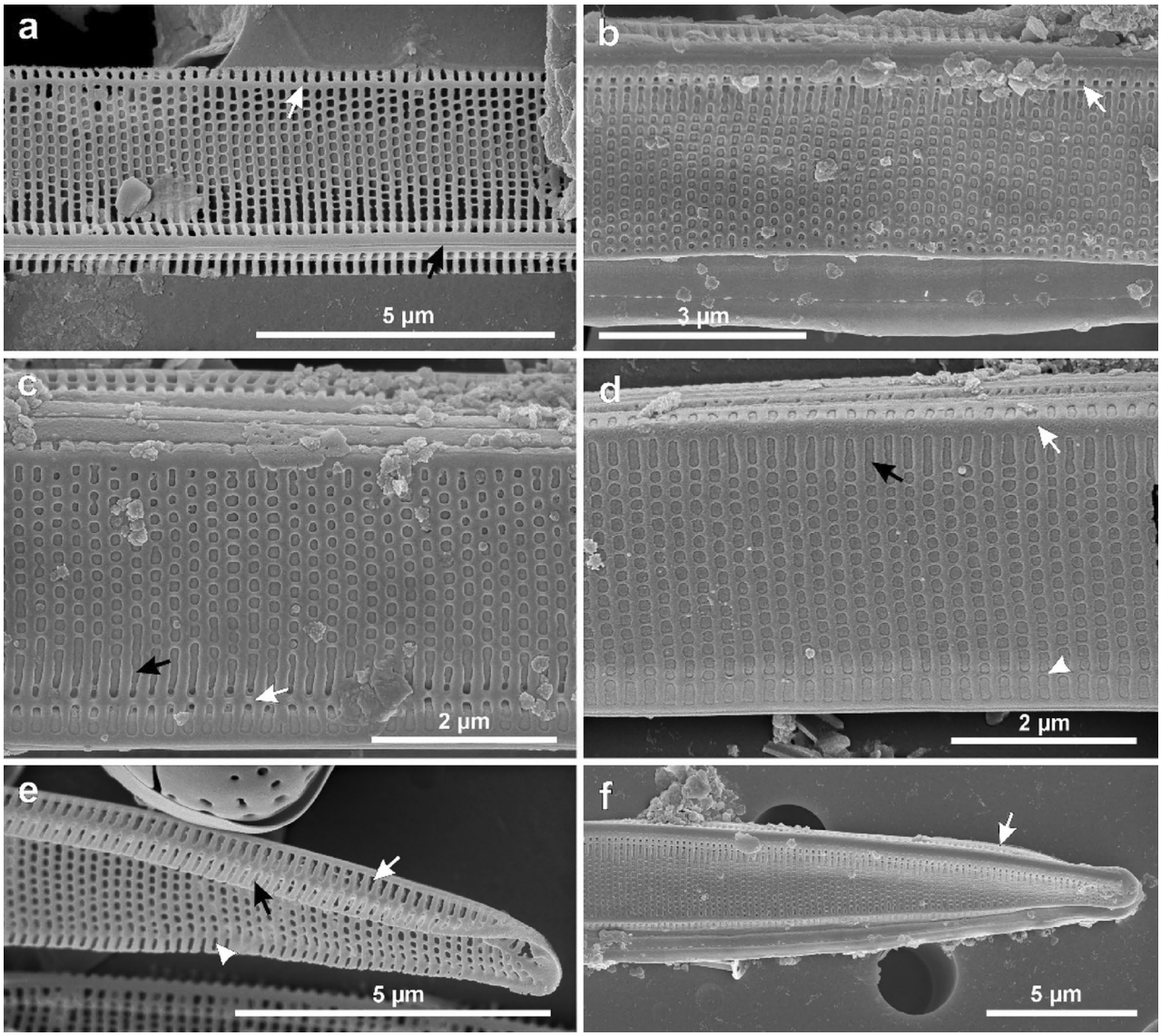

Fig. 5. Scanning electron microscopy photomicrographs of *Nitzschia excavata* sp. nov. from environmental material, sample D.BOZM.221024, close-up external view. (**a**–**d**) Central part of the valve. (**a**) Both mantles are visible, and the raphe without a central nodule (black arrow). A distinct junction between the valve face and the distal mantle is present (white arrow). (**b**) Close-up of the keel with visible poration on both sides of the raphe canal; two rows of single areolae are present on the raphe canal wall (arrow). (**c**) Elongate pores close to the raphe canal wall are indicated by a black arrow, separated by a silica bridge (white arrow). (**d**) Another specimen in which elongate pores dominate on the distal side of the valve face (black arrow); the arrowhead indicates the silica bridge separating the valve face from the raphe canal wall. On the opposite side, the junction between the valve face and the distal mantle is clearly visible (white arrow). (**e**) Oblique view of the valve apex with clearly visible raphe canal wall (white arrow) and the mantle separated from it (black arrow), both bearing elongate pores. (**f**) Elongate pores visible on the proximal part of the raphe canal wall and the mantle (arrow).

Boki and Külső Béda), and inner lakes of Lake Fertő (Herlakni, Hidegségi, Nagyhatár and Kis Herlakni) and its open water part as well. Closely related ASVs occurred at fewer sites and with lower relative abundances than the matching one (Table 2). Physicochemical data and locations are summarised in Supplementary Data S4 online.

PebbleScout searches revealed 100% matches to both *rbc*L and *psb*C reference sequences of *N. excavata* sp. nov. in SRR23621036 (Bioproject PRJNA938789) from the Ningxia Qingshui River Basin, collected at 37.19° N 105.77° E in spring 2022 (see Supplementary Data S5 and S6 online). Reads in SRR23621036 are 150 bp long. The 27 reads that exactly matched the *psb*C sequence of *N. excavata* isolate D.LDZ6 (GenBank accession PX860547) covered all but 69 bp of the 1030-bp *psb*C sequence and extended 34 bp beyond it at the 5′ end and 8 bp at the 3′ end. For *rbc*L, the 10 exactly matching reads covered most of the 1274-bp D.LDZ6 sequence, with three gaps of

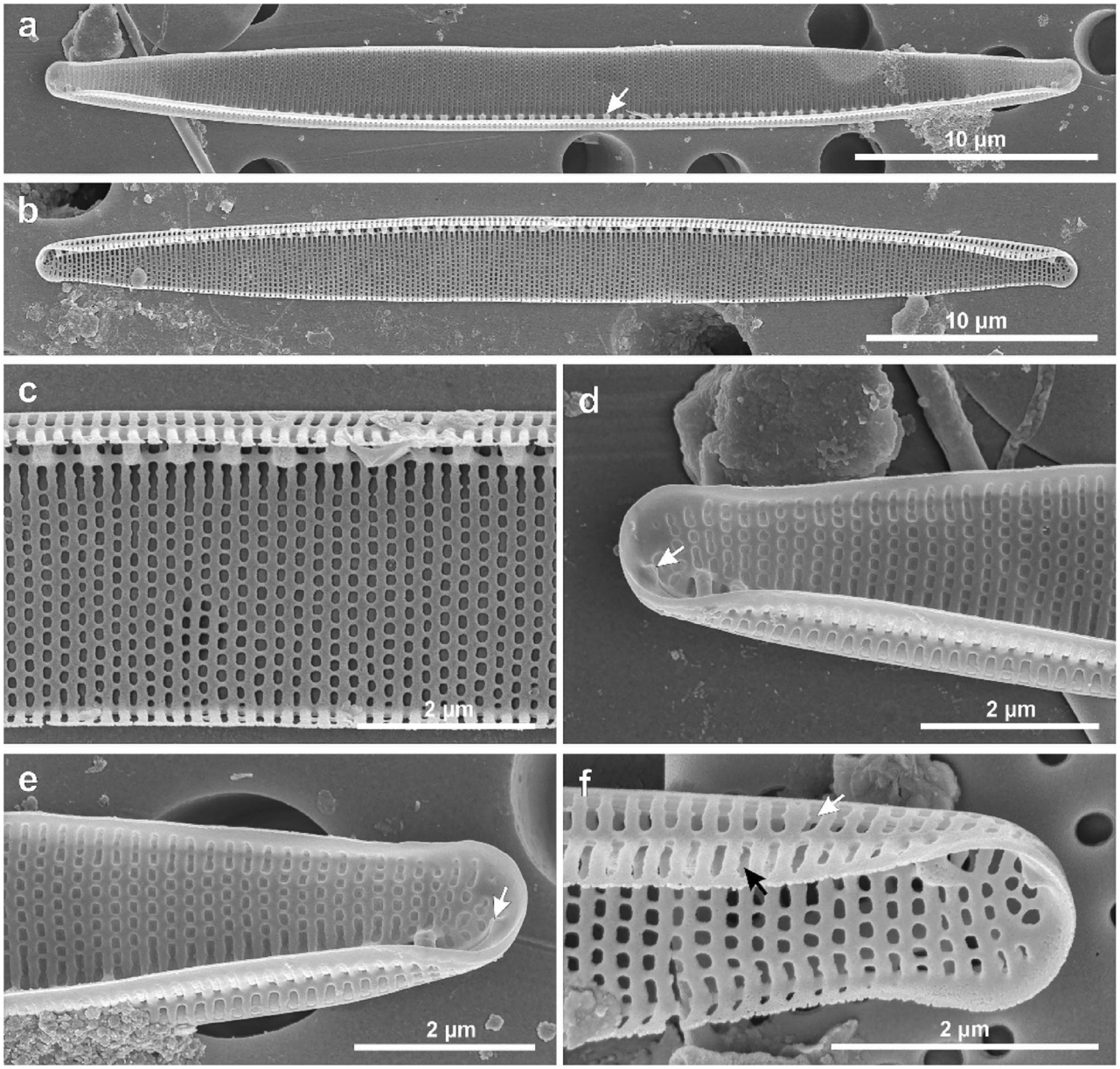

Fig. 6. Scanning electron microscopy photomicrographs of *Nitzschia excavata* sp. nov. from environmental material: (**a**, **d**–**f**) sample D.BOZM.221024; (**b**–**c**) sample D.BZ2M.221024, internal view. (**a**–**b**) Valve view. (**a**) No wider interspace is present between the two central fibulae (arrow). (**b**) Valve with equidistant fibulae. (**c**) Close-up of the central part of the valve; fibulae are narrow, and each fibula corresponds to at least two costae. (**d**–**e**) Valve apices with visible poration on the raphe canal wall and polar raphe fissures terminating with a helictoglossa (arrows). (**f**) Eroded valve with visible elongate pores on the raphe canal wall (white arrow) and the proximal mantle (black arrow).

20 bp, 107 bp, and 41 bp, respectively; the reads extended 27 bp further than the D.LDZ6 sequence at the 5′ end and 10 bp at the 3′ end (see Supplementary Data S5 and S6 online).

These records should be interpreted as evidence of the presence of the *N. excavata* lineage or closely related taxa, rather than unequivocal species-level identification based on short barcode regions.

Etymology

The species name *Nitzschia excavata* refers to the origin of the habitat where it was discovered. The specific epithet *excavata* (from Latin *excavatus*, meaning "dug out" or "hollowed") alludes to the reservoir formed in an area excavated for coal mining. The name thus highlights that the species was found in a man-made water body created after the flooding of an open-pit coal excavation site.

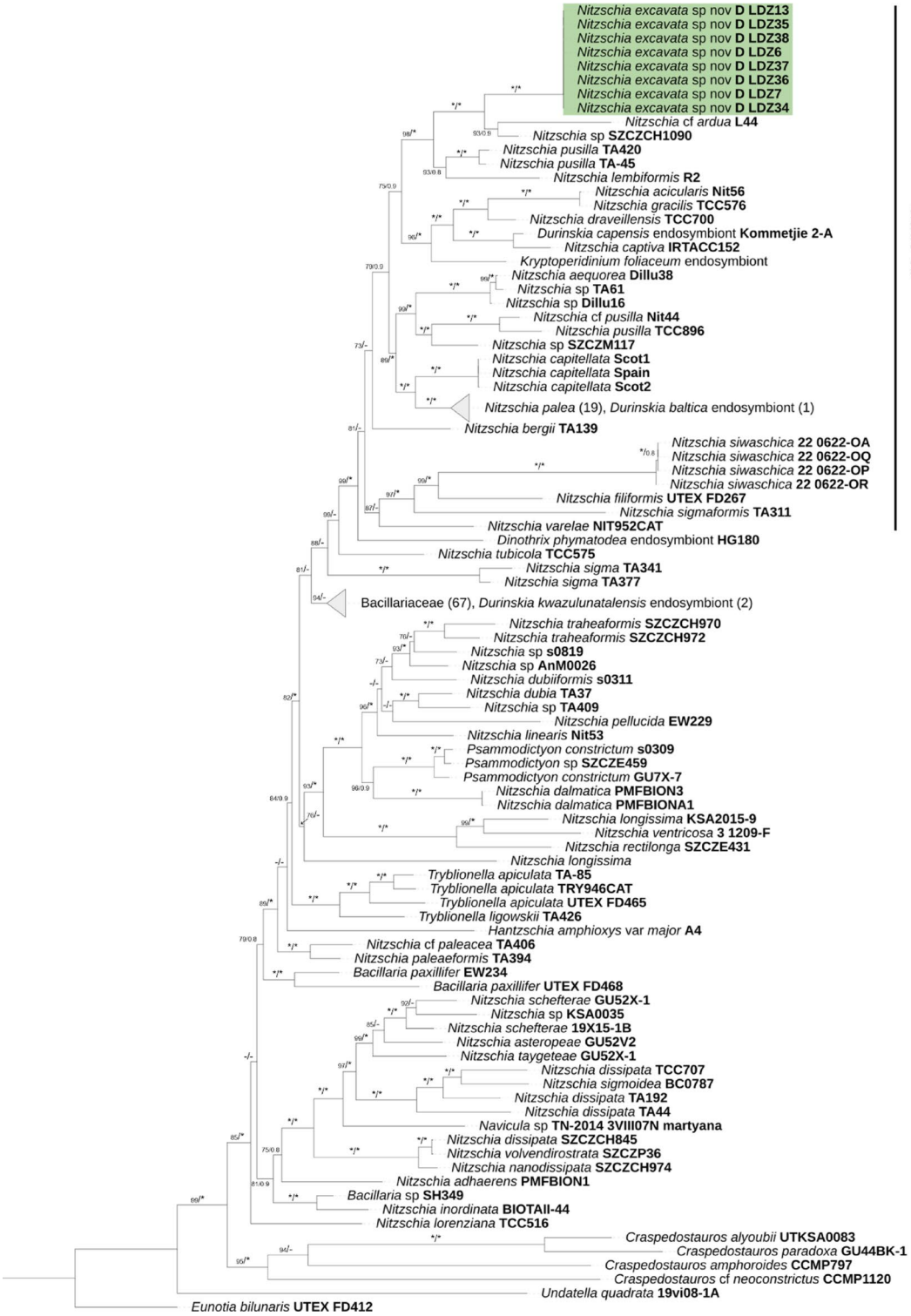

Fig. 7. The phylogenetic tree of the Bacillariaceae with *Eunotia bilunaris* as the outgroup based on the concatenated nuclear (SSU) and chloroplast (*rbc*L and *psb*C) DNA markers (total 3207 bp). The tree presents the position of newly identified *Nitzschia excavata* sp. nov. (green box). The numbers above the branches represent bootstrap values from ML analysis, followed by posterior probabilities from BI analysis. Asterisk (*) represents ML value = 100, and BI value = 1. En dash (–) represents ML value below 70, and BI value below 0.75. The topology of the tree is based on ML analysis. Clade 6b refers to Mann et al.[2].

Dataset	Repository ID where the raw data are available	DOI of the reference study	Location	No. samples	Reads
1	PRJEB41056	https://doi.org/10.1016/j.jhazmat.2021.125121	Water treatment effluents, France	48	10,645,262
2	PRJNA673880	https://doi.org/10.1371/journal.pone.0242143	Rivers, Ontario, Canada	45	14,724,682
3	PRJNA592969	https://doi.org/10.1002/eap.2205	Rivers, Ohio, USA	342	5,830,316
4	PRJNA558871	https://doi.org/10.1016/j.ecolind.2020.107070	Lake Nam Co, Tibet	23	3,511,798
5	PRJNA545290	https://doi.org/10.1002/eap.2812	Small rivers, California, USA	85	2,833,881
6	PRJNA623014	https://doi.org/10.1016/j.scitotenv.2021.147410	Small ponds, Leon, Spain	22	5,113,727
7	https://zenodo.org/record/3885810#.YIbPJqHtaUk	https://doi.org/10.1016/j.scitotenv.2020.140948	Rivers and lakes, Fennoscandia	48	8,587,305
8	https://zenodo.org/record/400160#.YK0ONaHtaUk	https://doi.org/10.1016/j.ecolind.2017.06.024	Rivers, Mayotte Island, Indian Ocean	80	3,519,207
9	https://zenodo.org/record/1157865#.Yo9c51RByUk	https://doi.org/10.1007/s13127-018-0359-5	Lakes, France	156	3,751,645
10	Own data	https://doi.org/10.1016/j.scitotenv.2020.138445	Rivers, Catalonia, Spain	307	30,620,471
11	https://data.inrae.fr/dataset.xhtml?persistentId=doi:10.15454/9EG5Z4&version=1.1	https://doi.org/10.1016/j.ecolind.2019.105775	Rivers, France	447	64,436,277
12	Data supplied by Dr Kerry Walsh (UK Environment Agency)	https://doi.org/10.1016/j.ecolind.2020.106725	Rivers, UK	1714	36,999,386
13	https://zenodo.org/records/7848220	https://eau-grandsudouest.fr/sites/default/files/2025-05/note_synthese_meta-ibd_vf.pdf	Rivers, SW France	375	17,467,126
14	PRJNA629002	https://doi.org/10.1111/mec.15646	Rivers, France and Switzerland	112	3,389,744
15	https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.57745/A7HIWL	https://doi.org/10.1016/j.scitotenv.2023.161970	Alpine lakes, France and Georgia	180	6,284,791
16	https://data.indores.fr/dataset.xhtml?persistentId=doi:10.48579/PRO/JARJYK	http://dx.doi.org/10.1016/j.scitotenv.2023.162270	Mineral springs, Massif Central, France	64	7,845,154
17	https://doi.org/10.6084/m9.figshare.11687877	https://doi.org/10.1111/mec.15696	Sediment core, Lake Constance, Germany	72	550,213
18	PRJEB48565	https://doi.org/10.1016/j.scitotenv.2022.154536	Karstic rivers, Croatia	31	717,683
19	PRJEB75497	https://doi.org/10.3390/microorganisms12081722	Rivers, Bosnia	8	326,246
20	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/MEQY4Q	https://doi.org/10.3897/mbmg.6.87497	Oxbow lake, Hungary	2	354,448
21	PRJNA1086463	https://doi.org/10.1016/j.scitotenv.2024.173502	Rivers, contiguous states of the USA	1973	31,272,891
22	PRJNA1005976	https://doi.org/10.1002/ece3.11162	low-oxygen, high-sulphur sinkholes and springs, USA	83	4,294,231
23	PRJNA957043	https://doi.org/10.1016/j.jenvman.2023.118885	Lakes, New Zealand	291	17,083,778
24	PRJNA997374	https://doi.org/10.3897/mbmg.7.110194	Marine and freshwater habitats, Antarctica	78	7,199,613

Table 1. Sources and characteristics for the 24 *rbc*L-based diatom metabarcoding datasets interrogated for the presence of *Nitzschia excavata* sp. nov. Datasets 1–12 were collected and processed by Pérez-Burillo et al.[12]. The bioproject (PRJEB and PRJNA) accessions are available at https://www.ncbi.nlm.nih.gov/nuccore/; URLs are given for other repositories. For access to the raw data for the Catalonian and UK datasets, please contact the authors. The total number of samples searched was 6586, with a total read count of 287,359,875. The number of reads of *N. excavata* sp. nov. detected was zero.

Discussion

Phylogenetic analysis

Comparisons were restricted to morphologically and phylogenetically similar taxa within the Lanceolatae group of *Nitzschia*. *Nitzschia excavata* sp. nov. belongs to the non-monophyletic group of Lanceolatae species that lack a central nodule and show no widened space between the central fibulae. This group contains species assigned by molecular data to several different clades, including clades 6B, 8A-I and 8A-II, and 8B of[2]. The newly discovered species resembles *Nitzschia palea* (also clade 6B) in outline, but there are clear differences between them. In *N. palea*, the apices are tapering and more rostrate to subcapitate, whereas in *N. excavata* sp. nov., the apices are more gradually protracted, and overall the valve outline is more linear. *N. excavata* sp. nov. has very fine striation, 46–49 striae in 10 µm, which is higher than the typical stria density in *N. palea*, usually given as 28–40 in 10 µm[16,17], though in some cases up to 49[18]. Both species have similar densities of fibulae; however, *N. excavata* sp. nov. has fibulae arranged in an equidistant pattern, which arrangement is not found in *N. palea*[16–18]. Finally, molecular analysis of the three-gene concatenated sequence indicates *N. excavata* sp. nov. is distinct from *N. palea*. Indeed, *N. excavata* sp. nov. is not sister to *N. palea* and lies on a long branch, reflecting a distant relationship to any *Nitzschia* species sequenced so far. The strains isolated are genetically identical with respect to the three markers used, which may reflect the fact that all were isolated from the same locality, but conflicts with the enormous within-species genetic variation observed in *N. palea*[1,2,19], even among isolates from the same pond (e.g. the four different LSU variants isolated from a single sample from Cala Castell, Spain[18]). This genetic separation confirms that the morphological differences are not the result of phenotypic plasticity within

Date	Site			Relative abundance in sample							
	Name	Latitude (decimal)	Longitude (decimal)	ASV_HU_001	ASV_HU_002	ASV_HU_003	ASV_HU_004	ASV_HU_005	ASV_HU_006	ASV_HU_007	ASV_HU_008
08.05.2019	Lake Velencei Brown Humic Water Area 1	47.206822	18.56955	1.483579035	0	0	0	0	0	0	0
08.05.2019	Lake Velencei Brown Humic Water Area 2	47.206822	18.56955	0.055711784	0	0	0	0	0	0	0
01.06.2019	gravel pit lake at Bugyi	47.242658	19.17845	0.215804134	0	0	0	0	0	0	0
06.09.2019	gravel pit lake at Délegyháza	47.255647	19.065003	9.824353841	0	0	0	0	0	0.117552481	0
04.09.2019	gravel pit lake at Kiskunlacháza 1	47.205167	19.094711	10.25699435	0.474918373	0.471466932	0.318222922	0	0.148411992	0	0
14.09.2019	gravel pit lake at Kiskunlacháza 2	47.199611	19.0792	0.007250265	0	0	0	0	0	0	0
27.05.2019	gravel pit lake et Lupa-sziget near Budapest	47.627253	19.076733	1.892660218	0	0	0	0	0	0	0
12.07.2020	gravel pit lake et Lupa-sziget near Budapest	47.627253	19.076733	1.86259728	0	0	0	0.192656219	0	0	0
16.09.2020	gravel pit lake et Lupa-sziget near Budapest	47.627253	19.076733	0.207022492	0	0	0	0	0	0	0.037065162
14.05.2019	Tolnai oxbow lake North	46.428489	18.824767	0.004052644	0	0	0	0	0	0	0
03.09.2019	Bátai oxbow lake	46.158006	18.802772	0.139589337	0	0	0	0	0	0	0
08.07.2020	Boki oxbow lake	45.905261	18.778064	0.016787108	0	0	0	0	0	0	0
13.05.2019	Külső Béda oxbow lake	45.936097	18.748878	0.013199261	0	0	0	0	0	0	0
17.05.2019	gravel pit lake at Alsózsolca	48.079761	20.906667	0.007564487	0	0	0	0	0	0	0
10.09.2019	gravel pit lake near Onga	48.123736	20.927036	0.007206861	0	0	0	0	0	0	0
12.07.2020	gravel pit lake (Kavicsos-tó), Csepel	47.386203	19.061256	0.016888667	0	0	0	0	0	0	0
12.09.2020	gravel pit lake (Kavicsos-tó), Csepel	47.386203	19.061256	0.133463758	0	0	0	0	0	0	0
05.09.2019	Herlakni lake - an inner lake of Fertő	47.683603	16.715772	0.004685157	0	0	0	0	0	0	0
05.09.2019	Hidegségi lake - an inner lake of Fertő	47.674411	16.732778	1.222776624	0	0	0	0	0	0	0
09.05.2019	Nagyhatár-lake - an inner lake of Fertő	47.664253	16.7155	0.00810082	0	0	0	0	0	0	0
05.09.2019	Kis herlakni lake - an inner lake of Fertő	47.69445	16.703039	0.012438233	0	0	0	0	0	0	0
09.05.2019	Fertő, open water	47.734319	16.68895	0.484542155	0	0	0	0	0	0	0
05.09.2019	Nagyhatár-lake - an inner lake of Fertő	47.664253	16.7155	0.136225185	0	0	0	0	0	0	0
02.09.2019	gravel pit lake (Kavicsos-tó), Csepel	47.386203	19.061256	0.299241383	0	0	0	0	0	0	0

Table 2. Hungarian sampling sites and ASV relative abundances. Relative abundances of amplicon sequence variants (ASVs) detected at Hungarian sampling sites, including site names, geographic coordinates (decimal degrees), and per-sample ASV abundances (ASV_HU_001–ASV_HU_009).

N. palea, but rather indicate a distinct evolutionary lineage that, in fact, is not even closely related to *N. palea*, despite the morphological similarity.

Phylogenetic analysis revealed a close relationship between the newly described species and two strains—TA45 and TA420—that were identified as *Nitzschia pusilla* and illustrated (LM) by An et al.[20]. However, as circumscribed in recent accounts[16,21], *N. pusilla* exhibits distinct morphological differences from *N. excavata* sp. nov. Its valves are more broadly rounded in outline compared to the linear-lanceolate shape of *N. excavata* sp. nov., the apices are more rounded, and the fibulae terminate before the apices (Fig. 3p–t). In contrast, in *N. excavata* sp. nov., the fibulae extend almost to the helictoglossae. Moreover, the striae density in *N. pusilla* can reach up to 55 in 10 µm[16,21], which exceeds the maximum observed in *N. excavata* sp. nov. (46–49 in 10 µm). Krammer and Lange-Bertalot[17], following Lange-Bertalot's concept of *N. pusilla*[22], illustrated specimens of *N.* cf. *pusilla* from Lake Van in Turkey (plate 79, figs 24 and 25) and from Jamaica (plate 79, fig. 21), which closely resemble *N. excavata* sp. nov. in general outline. However, these specimens lacked detailed morphological descriptions, making direct comparison difficult. Other sequenced strains identified as *N. pusilla* or *N.* cf. *pusilla* cluster in clades that are phylogenetically distinct and often distant from the two "*pusilla*" strains TA45 and TA420 from Korea and from the *N. excavata* sp. nov. lineage. For instance, SEM observations of strain TCC896 (*N. pusilla*) reveal differences in the arrangement of fibulae, which terminate earlier at the distal valve ends than in *N. excavata* sp. nov. (https://websites.rbge.org.uk/algae/research/Bacillariaceae_images.html). We conclude that *N. pusilla* urgently needs taxonomic revision to establish which, if any, of the clones that have been sequenced and identified as *N. pusilla* do actually belong to this species. *N. pusilla* was identified in the same sample as *N. excavata* sp. nov., and representative valves are presented in Fig. 3p–t.

Molecular data (Fig. 7) indicate that strain L44 (*Nitzschia* cf. *ardua*)[23] is closely related to *N. excavata* sp. nov., but its morphology differs. SEM images (https://websites.rbge.org.uk/algae/research/Bacillariaceae_images.html) show clear differences in valve outline and apices (L44 has an elliptical or linear-elliptical outline), no interruption of the striae at the valve face–mantle junction, and a smaller valve size (< 20 µm in length). Krammer and Lange-Bertalot[17], and Witkowski et al.[21] recognise *N. ardua* Cholnoky, 1961, and *N. bergii* A. Cleve, 1952 (with a question mark) as conspecific, under the name *N. bergii*, both species being characterised by a linear-elliptic valve outline and short, wedge-shaped apices with blunt ends. However, strain L44, identified as *Nitzschia* cf. *ardua*[23], is clearly different from strain TA139, identified as *N. bergii*[20]. Light microscopy of TA139 reveals a morphology distinct from that of L44. The stria density differs (L44 has more than 45 striae in 10 µm, while TA139 has 40 in 10 µm[20]) and the photo presented by An et al. fig. 3s[20] shows a valve with more pointed ends, as well as a reinforcement of the margin opposite the raphe canal that may indicate the presence of a marginal ridge, as found also by Clavero i Oms[24] in *N.* cf. *lembiformis*. In the original description of *N. ardua*, Cholnoky[25] mentions longitudinal undulation, which is also visible in the SEM microphotographs published by Clavero i Oms[24]. These images also suggest the presence of a marginal ridge, a structure occasionally visible under LM. A similar marginal ridge appears in some LM images of *N. bergii*[17,21]. Furthermore, the phylogenetic position of TA139 is unclear, since molecular analyses based on individual genes (*rbc*L, *psb*C) or concatenated sequences place TA139 *N. bergii*[20] far from *N. excavata* sp. nov. (but still in clade 6B), whereas SSU gene analysis suggests a close affinity between them, although bootstrap support is very low.

Clearly, as with *N. pusilla*, there are many unanswered questions about the identities and relationships of *N. bergii*, *N. ardua,* and *N. lembiformis*. However, regardless of whether *N.* cf. *lembiformis*, *N.* cf. *ardua,* or *N. bergii* sensu An et al.[20] are correctly named, none shows a morphological similarity to *N. excavata* sp. nov. The newly described species lacks any trace of a marginal ridge, either continuous or fragmentary, its valve face is entirely flat, and the valve shape differs from any of the species. In particular, strain L44 (*Nitzschia* cf. *ardua*)[23], although closely related to *N. excavata* sp. nov., shows clear differences in valve outline, apices, and smaller valve size (< 20 µm in length) (SEM images at https://websites.rbge.org.uk/algae/research/Bacillariaceae_images.html).

Phylogenetic analysis revealed that strain SZCZCH1090 is closely related to *Nitzschia excavata* sp. nov., forming a sister lineage with maximal bootstrap support (Fig. 7). Given this close relationship, we examined its morphology for comparison. According to metadata from the Szczecin Diatomological Collection, this strain was isolated from samples collected in Puck Bay (southern Baltic Sea). Despite genetic proximity, the frustules of SZCZCH1090 differ morphologically from *N. excavata* sp. nov. in several aspects. The valves are more elongate, with distinctly capitate apices. Additionally, although valve length in SZCZCH1090 is within the range observed for *N. excavata* sp. nov., the valves are generally slightly wider (see Supplementary Fig. S7 online), https://websites.rbge.org.uk/algae/research/Bacillariaceae_images.html). Ecological conditions also differ between the two taxa, the newly described species having been isolated from an inland anthropogenic reservoir, whereas SZCZCH1090 originates from a coastal brackish environment.

Among species for which no molecular data are available, *Nitzschia excavata* sp. nov. may be easily confused with *N. wilmotteana* Hamsher et al.[26] due to their similar linear-lanceolate shape. However, several distinct morphological features set the two species apart. *N. wilmotteana* has valves that are more distinctly linear, with narrowly rounded to rostrate, produced apices, while *N. excavata* sp. nov. has slightly broader, subcapitate apices. Additionally, *N. wilmotteana* has considerably lower densities of fibulae and striae, typically exhibiting fewer than 13 fibulae and 42 striae per 10 µm. In contrast, *N. excavata* sp. nov. is characterised by 15 to 20 fibulae and 46 to 49 striae per 10 µm. These differences point to fundamental structural variations between the two species[26]. The comparison presented in Table 3 includes representative morphologically similar species as well as the phylogenetically closest available strain.

Distribution and ecology

Concerning the distribution of *N. excavata* sp. nov., searches of 24 metabarcoding datasets, most of them publicly available (just two, from Catalonia and the UK, required permissions from the owners of the data), did not reveal a single record of the species. The datasets (Table 1) cover a very wide geographical range, from

	N. excavata sp. nov	*N. palea* agg	*N. pusilla*	*N. bergii*	*N. wilmotteana*	*Nitzschia* sp. SZCZCH1090
Length [µm]	29.0–50.3	10.8–45.4	8.0–33.0	14.0–60.0	31.0–43.5	36.3–41.0
Width [µm]	3.0–4.4	2.5–5.1	2.5–5.0	4.0–5.0	3.5–4.0	4.0–4.53
Striae in 10 µm	46–49	28.0–49.4	40.0–55.0	35.0–40.0	ca. 42	ca. 44
Fibulae in 10 µm	15–22 Equidistant	19–22 Irregularly spaced	14–20 Equidistant	14–20 Equidistant	12–14 Equidistant	15–17 Equidistant
Valve outline	Elongate and linear-lanceolate	Linear-lanceolate to linear	Linear-lanceolate, linear, elliptic	Linear-elliptic to linear	Linear to narrowly lanceolate	Linear-lanceolate with elongated ends
Apices shape	Protracted and slightly subcapitate apices	Tapering to a cuneate, acutely rounded	Obtusely to broadly rounded	Cuneate to bluntly rounded	Narrowly–rounded to prostate	Obtusely or broadly rounded
References	This study	Lange-Bertalot and Metzeltin 1996, Trobajo et al.[18]	Lange-Bertalot and Metzeltin 1996[47]	Krammer and Lange-Bertalot[17]	Hamsher et al.[26]	This study

Table 3. Comparison of *Nitzschia excavata* sp. nov. with morphologically related taxa within the Lanceolatae group.

the conterminous USA through Europe to an Indian Ocean island (Mayotte) and New Zealand, and include samples from rivers, streams, and lakes. It seems from this that *N. excavata* sp. nov. may occupy a relatively under-sampled or poorly characterised ecological niche.

The limitations of short *rbc*L sequences for identifying and recording diatom species must be borne in mind, since it is already known that a few morphologically distinct species cannot be separated using the 263-bp 'metabarcode'[14]. Hence, a positive result in a metabarcoding dataset, even a 100% match over the 263-bp region, is not an unambiguous record of *N. excavata* sp. nov. as defined here. Nevertheless, on the whole, the 263-bp metabarcode does separate species[12–14], and this property is allowing a transition from microscopical to molecular approaches in species-based river biomonitoring. At the least, therefore, a positive metabarcoding record indicates the presence of a particular species or a group of very closely related species. In the case of metagenomic datasets, like the one explored here from NW China, geographical records can be more definitive, because by assembling reads into contigs it may be possible to demonstrate correspondence with reference sequences over a much longer region than is possible with Illumina data.

In the present study positive results were obtained from metabarcoding and metagenomic datasets in Spain, NW China (in the Ningxia Hui Autonomous Region) and Hungary, suggesting that, despite its apparently exacting requirements, the *N. excavata* sp. nov. lineage is very widespread outside the type locality, Bogdałów reservoir in Poland. In addition, exact *rbc*L metabarcoding matches from multiple sites in Hungary confirm its presence in lowland standing waters and floodplain oxbows of the Carpathian Basin, further indicating that the species may be more widespread in lentic habitats than morphology-based surveys alone suggest (since these have failed to detect the species until now). More generally, our study highlights the need for great caution before claiming that a newly described species has a restricted distribution: despite 150 years of study, freshwater diatoms are still rather poorly known, especially in groups with few distinctive morphological characteristics, and we need to take every opportunity to document the variation and distributions of taxa.

The wide distribution of *N. excavata* sp. nov. may not be surprising. One possible explanation is that the species has been previously overlooked or misidentified due to its morphological similarity to members of the *N. palea* complex, which is known for its considerable morphological variability. Additionally, routine light microscopy-based surveys and current metabarcoding and metagenomic datasets are often incomplete and unevenly distributed geographically, which may lead to an underestimation of the true distribution of taxa. Therefore, it is possible that *N. excavata* sp. nov. has been present in other regions but has gone undetected or misidentified. Furthermore, the available data indicate that this species is associated with environments that have elevated concentrations of dissolved ions. Its presence in geographically distant areas, including Poland, Spain, Hungary, and northwestern China, likely reflects both a relatively wide distribution and a preference for specific ecological conditions rather than strict endemism. Regarding its occurrence in post-mining reservoirs, these habitats may not represent sites of recent speciation but rather environments that favor pre-existing taxa. The Bogdałów reservoir is supplied by a small watercourse and drainage waters from nearby mining areas. It is possible that the species exists in upstream or surrounding habitats that have not yet been documented. Alternatively, dormant stages may have been reactivated following environmental disturbances, as observed in other diatoms, such as *Skeletonema marinoi* Sarno & Zingone, 2005 from the Baltic Sea[27]. These scenarios suggest that the distribution of *N. excavata* sp. nov. is influenced by a combination of ecological filtering, historical dispersal, and limitations in detection, rather than recent evolution within anthropogenic systems.

In this context, we suggest that it is desirable that the authors of new species take advantage of the rapidly growing corpus of metabarcoding and metagenomic data to test whether new taxa do occur elsewhere, even in very distant locations, providing of course that gene sequence data are available to make a search.

Based on the available environmental data (see Supplementary Fig. S4 online) and the previous study[5,10], the species occurs predominantly at pH 8.3–8.7 and specific conductance 500–2600 µS cm^{-1}. Its relative read abundance shows moderate positive associations with Na^+, Ca^{2+}, SO_4^{2-}, and overall conductivity, suggesting a preference for ion-rich waters with elevated bicarbonate alkalinity. Negative or near-zero correlations with chlorophyll-a, total phosphorus, and chemical oxygen demand imply that the species may tolerate oligotrophic to meso-saprobic conditions (see Supplementary Fig. S8 online).

Materials and methods

Study sites

The study site is located within the Bogdałów reservoir, near the village of Bogdałów in the Greater Poland Voivodeship, to the northeast of the city of Turek[10]. This waterbody lies within the anthropogenically transformed landscape of the Konin–Turek lignite basin, an area heavily exploited for open-cast lignite and brown coal mining. The reservoir is part of the post-mining infrastructure associated with PAK KWB Konin S.A., one of Poland's largest lignite mining companies. This artificial lake was created in the mid-1990s by intentionally flooding a decommissioned lignite pit, in accordance with the regional land reclamation program. The flooding was accomplished through a combination of diverted stream inflows and drainage water from mining operations. The total surface area of the reservoir is approximately 9.5 hectares, and the volume is 600,000 m^3. While it was initially designed as a technical water storage facility for fire prevention and emergency supplies, it has since become ecologically stabilised and is now surrounded by secondary woodlands predominantly composed of pine and birch trees[28–31]. The hydrological and chemical characteristics of the Bogdałów reservoir are primarily influenced by the inflow of mine drainage and local surface waters, which together shape its overall limnological profile. The reservoir exhibits moderate mineralization and relatively stable physicochemical conditions, which have allowed for the gradual development of a diverse aquatic community. These conditions create a suitable environment for diatom assemblages that are adapted to anthropogenic or slightly altered habitats. Since 2019, the site has been the focus of systematic investigations aimed at documenting species diversity and refining ecological indicator values for post-mining aquatic ecosystems[5,10].

Sampling and culture

Sampling and laboratory procedures followed the methods outlined by Olszyński et al.[10]. Diatom samples were collected from the sandy bottom and submerged plants of Bogdałów reservoir in October 2022, with additional benthic samples with sand substrate in October 2024 (D.BOZM.221024, D.BZ2M.221024). In situ measurements were taken of environmental parameters, including temperature, pH, and conductivity. While the earlier study identified two strains, this research focused on isolating and cultivating eight new strains–D.LDZ.6, D.LDZ.7, D.LDZ.13, and D.LDZ.34–D.LDZ.38 – from samples collected in 2022 (sample number: D.BOZS.191022), under the same conditions using sterile Bacillariophycean Medium[10]. Additionally, historical site data were used to describe environmental conditions[5]. For comparison with the Bogdałów cultures, strain *Nitzschia* sp. SZCZCH1090 was analysed in LM. This strain was isolated from sediment at Chałupy, Puck Bay, Baltic Sea, and sequences for four genes (18S rRNA, partial 28S rRNA, *rbc*L, and *psb*C) were used in the study by Mann et al.[2] and are available in NCBI GenBank.

Microscopy analysis

For microscopy analyses, environmental material from three samples—D.BOZS.191022, D.BOZM.221024, and D.BZ2M.221024—were used alongside cultured strains. Cleaning of diatom frustules from environmental and cultured samples followed the oxidation protocol described by Żelazna-Wieczorek[32] and detailed in Olszyński et al.[10]. The cleaned material was mounted in Naphrax (Brunel Microscopes Ltd., UK) for light microscopy (LM) using a Nikon Eclipse 90i microscope equipped with a 100×/1.40 Plan APO VC oil-immersion objective and a Nikon DS-Ri1 digital camera. For scanning electron microscopy (SEM), aliquots of cleaned suspensions were filtered onto 5 µm Whatman Nuclepore® polycarbonate membranes (Fisher Scientific, Germany), air-dried, mounted on aluminium stubs, and sputter-coated with 20 nm gold using a Quorum Q150T ES coater (Judges Scientific plc, UK). In addition, SEM images were taken of environmental material without a cleaning protocol; the uncleaned samples were air-dried and processed using the same procedure. Observations and imaging were performed with a Hitachi SU8010 scanning electron microscope (Hitachi Ltd., Japan). For morphometric analysis, 43 specimens from culture using SEM photomicrographs and 40 specimens from environmental samples under LM were measured. The SEM photomicrographs presented in this study were obtained from the environmental material mentioned above. Morphometry of 10 frustules of *Nitzschia* sp. SZCZCH1090 was analysed to obtain valve size, as well as the density of striae and fibulae. Chloroplast structure of living cultured cells was examined by Nikon Eclipse E400 microscope equipped with a 60×/1.40 Plan APO DIC H oil-immersion objective and an OPTA-TECH HDMI series camera with a resolution of 3840 × 2040 (sensor 1/2.7″ SONY) and confocal laser scanning microscopy (CLSM) using a Leica SP8 system (Leica Microsystems, Germany), as outlined in Olszyński et al.[10].

DNA isolation, amplification, and sequencing

Genomic DNA was isolated from monoclonal diatom cultures using the Chelex® 100 protocol (Bio-Rad, USA) following Dąbek et al.[33] and as described in detail by Olszyński et al.[10]. Briefly, the cell pellets obtained from 1.5 mL of cell suspension were incubated with 10% Chelex® 100 at 95 °C for 20 min, then centrifuged, and the DNA-containing supernatants were stored at -20 °C until further analysis. Three molecular markers, nuclear SSU (18S rDNA) and the chloroplast genes *rbc*L and *psb*C, were amplified using primers listed by Olszyński et al.[10], Theriot et al.[34], and Kawai et al.[35], respectively. PCRs were performed with HOT FIREPol® DNA Polymerase (Solis BioDyne OÜ, Estonia) in 25 µL reactions under the same thermal conditions reported by Olszyński et al.[10]. All markers were amplified by PCR in a few replicates to obtain high-quality products, and were then checked on 1% agarose gels. PCR fragment purification and Sanger sequencing were performed by SEQme s.r.o. company (Czech Republic). Obtained sequences were assembled using Geneious Prime® 2023.2.1 software (Biomatters Inc., USA).

Phylogenetic analyses

The phylogenetic analyses on the studied samples and other species of the *Nitzschia* genus were performed separately for each marker, i.e., SSU, *rbc*L, and *psb*C, as well as based on a concatenated SSU-*rbc*L-*psb*C matrix (3207 bp). *Eunotia bilunaris* (Ehrenberg) Schaarschmidt, 1880, served as an outgroup. For taxa in the phylogenetic analysis with incomplete gene sampling, such regions were treated as 'missing data'. Voucher information for the specimens included in this study, with corresponding GenBank accession numbers, is presented in Supplementary Data S9 online. The performed analyses included 150, 278, and 100 sequences, respectively, for the SSU, *rbc*L, and *psb*C markers. For the concatenated analysis, the number of *rbc*L sequences was reduced to 179 to eliminate strains that only have data of *rbc*L. Sequences were aligned using the MAFFT v. 7 web server[36,37] (http://mafft.cbrc.jp/alignment/server/), where the auto strategy was applied with a scoring matrix of 200PAM, Gap opening penalty of 1.53, UniRef50 for Maft-homologs, and Plot and alignment with a threshold of 39 score were set (see Supplementary Data S10 online). The alignments were checked for poorly and ambiguously aligned regions, and small corrections were made by eye. The evolutionary models were calculated using Mega 11 (version 11.0.13) and selected according to the Akaike Information Criterion. The dataset was partitioned by gene and codon position (*rbc*L and *psb*C), whereas the SSU marker was left unpartitioned. The partitioning was performed using PartitionFinder2 software[38] (see Supplementary Data S10 online). Phylogenetic calculations were performed using ML analysis in IQ-TREE version 3[39] with the ultrafast bootstrap (UFBoot) pseudolikelihood algorithm[40], 10,000 replicates and 10,000 bootstrap alignments; and Bayesian analysis was performed in MrBayes 3.2.7[41], which provides the Markov chain Monte Carlo (MCMC) method to estimate the posterior distribution of model parameters. Trees were sampled every 1,000 generations until the average standard deviation of split frequencies reached values below 0.01 for the last 1,000 generations, and posterior probabilities were estimated from the 50% majority-rule consensus tree after elimination of the first 25% of samples as burn-in.

Metabarcoding distribution analysis

In order to explore the geographical distribution of the newly described *Nitzschia* species, existing *rbc*L metabarcoding datasets were interrogated, as listed in Table 1. Twelve of the 24 datasets had already been assembled by Pérez-Burillo et al.[12], and we used their combined FASTA file of ASV sequences to construct a database for a local BLAST search[42]. The remaining datasets were developed from the raw fastq files of forward and reverse reads obtained from the repositories listed in Table 1. Amplicon Sequence Variants (ASVs) were generated for each dataset using the R package DADA2[43], after removal of the primers using cutadapt[44]. When a dataset was formed by more than one Illumina run, each run was processed individually. The forward and reverse reads were truncated to approximately 220–240 and 160–220 nucleotides, respectively, based on their quality profiles, and reads with ambiguities or an expected error (maxEE) higher than 2 were discarded. Error rates were modelled according to the DADA2 tutorial (https://benjjneb.github.io/dada2/tutorial.html) and chimeras discarded using the function "removeBimeraDenovo". Subsequently, searches for new described *Nitzschia* were made in FASTA files of the ASVs using local BLAST.

To further evaluate the possible occurrence of the discussed *Nitzschia* species in previously investigated freshwater environments, we conducted a re-examination of *rbc*L ASVs derived from a regional diatom metabarcoding study carried out in Hungary[45]. The retrospective screening was performed on 160 lake and soda-pan samples, for which ASV lists were already available as outputs of the DADA2 pipeline. We compared the *rbc*L sequence obtained from the cultured strains with the ASVs detected in these samples to assess sequence similarity and potential identity. The relative abundances of matching and closely related ASVs across Hungarian sampling sites are summarised in Table 2. Sequence alignment was generated using ClustalW, followed by pairwise p-distance calculations. ASVs with p-distance values below 0.01 were considered to represent the same taxon or closely related ASV-level variants. Sequence information, p-distance values, and base-pair mismatches relative to *Nitzschia excavata* sp. nov. reference sequences are provided in Supplementary Data S11 online. A Spearman rank correlation analysis was applied to assess relationships between eDNA-derived relative read numbers of the target species and environmental variables measured at each sampling site. Spearman correlation was chosen due to the non-normal distribution of eDNA read numbers and several environmental parameters. Correlations were calculated separately between eDNA read numbers and each environmental variable.

In addition, we made use of the PebbleScout[46] to search metagenomic and metatranscriptomic data sets published up to the end of 2021 ("Metagenomic, Volume 1") and in 2022–23 ("Metagenomic, Volume 2"). We looked for runs with high-scoring matches to *rbc*L and *psb*C gene sequences for isolate D.LDZ6 and then queried them by BLAST to locate reads with a 100% match to part of the gene sequence. The reads were then downloaded and aligned with the query sequences.

Data availability

The datasets generated and analysed during the current study are available in the NCBI GenBank database [www.ncbi.nlm.nih.gov] under the accession numbers: PX843716–PX843721 for SSU 18S, PX860539–PX860546 for the *rbc*L, and PX860547–PX860554 for the *psb*C marker; The new species has been registered in PhycoBank [www.phycobank.org] under the number 106537 and are included in this article (and its Supplementary Information files). Further information on the metabarcoding dataset evidence is available from the authors on request.

Received: 2 February 2026; Accepted: 20 April 2026
Published online: 28 May 2026

References

1. Rimet, F. et al. Molecular phylogeny of the family Bacillariaceae based on 18S rDNA sequences: Focus on freshwater *Nitzschia* of the section Lanceolatae. *Diatom Res.* **26**, 273–291 (2011).
2. Mann, D. G. et al. Ripe for reassessment: A synthesis of available molecular data for the speciose diatom family Bacillariaceae. *Mol. Phylogenet. Evol.* **158**, 106985 (2021).
3. Guiry, M. D. & Guiry, G. M. AlgaeBase. https://www.algaebase.org/ (2023).
4. Cleve, P. T. & Grunow, A. Beiträge zur Kenntniss der arctischen Diatomeen. *K. Sven. Vetenskapsakad. Handl.* **17**, 1–121 (1880).
5. Olszyński, R. M., Szczepocka, E. & Żelazna-Wieczorek, J. Critical multi-stranded approach for determining the ecological values of diatoms in unique aquatic ecosystems of anthropogenic origin. *PeerJ* **7**, e8117 (2019).
6. Żelazna-Wieczorek, J., Olszyński, R. M. & Nowicka-Krawczyk, P. Half a century of research on diatoms in athalassic habitats in central Poland. *Oceanol. Hydrobiol. Stud.* **44**, 51–67 (2015).
7. Olszyński, R. M. & Żelazna-Wieczorek, J. *Aulacoseira pseudomuzzanensis* sp. nov. and other centric diatoms from post iron ore mining reservoirs in Poland. *Diatom Res.* **33**, 155–185 (2018).
8. Żelazna-Wieczorek, J. & Olszyński, R. M. Taxonomic revision of *Chamaepinnularia krookiformis* Lange–Bertalot et Krammer with a description of *Chamaepinnularia plinskii* sp. nov.. *Fottea* **16**, 112–121 (2016).
9. Sienkiewicz, E., Gąsiorowski, M., Sekudewicz, I., Kowalewska, U. & Matoušková, Š. Responses of diatom composition and teratological forms to environmental pollution in a post-mining lake (SW Poland). *Environ. Sci. Pollut. Res.* **30**, 110623–110638 (2023).
10. Olszyński, R. M. et al. Morphology and phylogeny of *Nitzschia nandorii* sp. Nov. (Bacillariophyceae), a new small-celled lanceolate species from a post-mining reservoir. *PhytoKeys* **241**, 1–26 (2024).
11. De Luca, D., Kooistra, W. H. C. F., Sarno, D., Gaonkar, C. C. & Piredda, R. Global distribution and diversity of *Chaetoceros* (Bacillariophyta, Mediophyceae): Integration of classical and novel strategies. *PeerJ* **2019**, e7410 (2019).
12. Pérez-Burillo, J., Mann, D. G. & Trobajo, R. Biogeography and genetic diversity of freshwater diatoms: The potential of large combined *rbc*L metabarcoding datasets. *Sci. Total Environ.* **966**, 178727 (2025).
13. Pérez-Burillo, J. et al. DNA metabarcoding reveals differences in distribution patterns and ecological preferences among genetic variants within some key freshwater diatom species. *Sci. Total Environ.* https://doi.org/10.1016/j.scitotenv.2021.149029 (2021).
14. Pérez-Burillo, J., Mann, D. G. & Trobajo, R. Evaluation of two short overlapping *rbc*L markers for diatom metabarcoding of environmental samples: Effects on biomonitoring assessment and species resolution. *Chemosphere* https://doi.org/10.1016/j.chemosphere.2022.135933 (2022).
15. Van Dam, H., Mertens, A. & Sinkeldam, J. A coded checklist and ecological indicator values of freshwater diatoms from The Netherlands. *Neth. J. Aquat. Ecol.* **28**, 117–133 (1994).
16. Lange-Bertalot, H., Hofmann, G., Werum, M., Cantonati, M. & Kelly, M. G. Freshwater benthic diatoms of Central Europe: Over 800 common species used. In *Ecological Assessment* (Koeltz Botanical Books, 2017).
17. Krammer, K. & Lange-Bertalot, H. Bacillariophyceae. 2 Teil: Bacillariaceae, epithemiaceae, surirellaceae. In *Süsswasserflora von Mitteleuropa* Vol. 2/2 (eds Ettl, H. et al.) 1–610 (Spektrum, Akad. Verlag, 1997).
18. Trobajo, R. et al. Morphological, genetic and mating diversity within the widespread bioindicator *Nitzschia palea* (Bacillariophyceae). *Phycologia* **48**, 443–459 (2009).
19. Trobajo, R. et al. The use of partial cox1, *rbc*L and LSU rDNA sequences for phylogenetics and species identification within the *Nitzschia palea* species complex (Bacillariophyceae). *Eur. J. Phycol.* **45**, 413–425 (2010).
20. An, S. M., Choi, D. H., Lee, J. H., Lee, H. & Noh, J. H. Identification of benthic diatoms isolated from the eastern tidal flats of the Yellow Sea: Comparison between morphological and molecular approaches. *PLoS ONE* **12**, e0179422 (2017).
21. Witkowski, A., Lange-Bertalot, H. & Metzeltin, D. *Diatom Flora of Marine Coasts I. Iconographia Diatomologica* Vol. 7 (A.R.G. Gantner Verlag K.G., 2000).
22. Lange-Bertalot, H. Eine Revision zur Taxonomie der *Nitzschiae lanceolatae* Grunow. Die 'klassischen' bis 1930 beschribenen Süsswasserarten Europas. *Nova Hedwigia* **28**, 253–307 (1976).
23. Rovira, L., Trobajo, R., Sato, S., Ibáñez, C. & Mann, D. G. Genetic and physiological diversity in the diatom *Nitzschia inconspicua*. *J. Eukaryot. Microbiol.* **62**, 815–832 (2015).
24. Clavero i Oms, E. Diatomees d'ambient hipersalins costaners. Taxonomia, distribució i em-premtes en el registre sedimentari. https://dialnet.unirioja.es/servlet/tesis?codigo=255151&info=resumen&idioma=SPA (2004).
25. Cholnoky, B. J. Ein beitrag zur kenntnis der diatomeenflora der venetianischen Lagunen. *Hydrobiologia* **17**, 287–325 (1961).
26. Hamsher, S., Kopalová, K., Kociolek, J. P., Zidarova, R. & van de Vijver, B. The genus *Nitzschia* on the South Shetland Islands and James Ross Island. *Fottea* **16**, 79–102. https://doi.org/10.5507/fot.2015.023 (2016).
27. Bolius, S. et al. Resurrection of a diatom after 7000 years from anoxic Baltic Sea sediment. *ISME J.* **19**, 1–9 (2025).
28. Orlikowski, D. & Szwed, L. Wodny kierunek rekultywacji w KWB "Adamów" SA – inwestycją w przyszłość regionu. *Górn. Geoinż.* **33**, 351–361 (2009).
29. Orlikowski, D. & Szwed, L. Zagospodarowanie terenów pogórniczych KWB "ADAMÓW" SA w Turku - krajobraz przed rozpoczęciem działalności górniczej i po jej zakończeniu. *Górn. Geoinż.* **35**, 225–240 (2011).
30. Stachowski, P., Kraczkowska, K., Liberacki, D. & Oliskiewicz-Krzywicka, A. Water reservoirs as an element of shaping water resources of post-mining areas. *J. Ecol. Eng.* **19**, 217–225 (2018).
31. Stachowski, P. Rekultywacja wodna na zdewastowanych obszarach pogórniczych Wschodniej Wielkopolski jako element ich zagospodarowania. *Space Soc. Econ.* **35**, 125–143 (2024).
32. Żelazna-Wieczorek, J. Diatom flora in springs of Łódź Hills (Central Poland). Biodiversity, taxonomy, and temporal changes of epipsammic diatom assemblages in springs affected by human impact. In *Diatom Monographs* Vol. 13 (ed. Witkowski, A.) 1–419 (A.R.G. Gantner K.G., 2011).
33. Dąbek, P. et al. Towards a multigene phylogeny of the Cymatosiraceae (Bacillariophyta, Mediophyceae) I: Novel taxa within the subfamily Cymatosiroideae based on molecular and morphological data. *J. Phycol.* **53**, 342–360 (2017).
34. Theriot, E. C., Ashworth, M. P., Nakov, T., Ruck, E. & Jansen, R. K. Dissecting signal and noise in diatom chloroplast protein encoding genes with phylogenetic information profiling. *Mol. Phylogenet. Evol.* **89**, 28–36 (2015).
35. Kawai, H., Hanyuda, T., Draisma, S. G. A., Wilce, R. T. & Andersen, R. A. Molecular phylogeny of two unusual brown algae, *Phaeostrophion irregulare* and *Platysiphon glacialis*, proposal of the Stschapoviales ord. nov. and Platysiphonaceae fam. nov., and a re-examination of divergence times for brown algal orders. *J. Phycol.* **51**, 918–928 (2015).
36. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2019).
37. Kuraku, S., Zmasek, C. M., Nishimura, O. & Katoh, K. ALeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* **41**, W22–W28 (2013).
38. Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T. & Calcott, B. PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol. Biol. Evol.* **34**, 772–773 (2017).
39. Wong, T. K. F. et al. IQ-TREE 3: Phylogenomic Inference Software using Complex Evolutionary Models. https://doi.org/10.32942/X2P62N (2025).
40. Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).

41. Ronquist, F. et al. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).
42. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
43. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
44. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
45. Duleba, M. et al. Applicability of diatom metabarcoding in the ecological status assessment of Hungarian lotic and soda pan habitats. *Ecol. Indic.* **130**, 108105 (2021).
46. Shiryev, S. A. & Agarwala, R. Indexing and searching petabase-scale nucleotide resources. *Nat. Methods* **21**, 994–1002 (2024).
47. Lange-Bertalot, H. & Metzeltin, D. Indicators of Oligotrophy. 800 taxa representative of three ecologically distinct laketypes Carbonate buff ered - Oligodystrophic - Weakly buff ered soft water. in Iconographia Diatomologica (ed.Lange-Bertalot, H.) **2** 1–390 (Koeltz Scientifi c Books, Königstein, (1996).

Acknowledgements

The UK metabarcoding data searched in this paper here was generated and provided by the Environment Agency of England (see also Kelly et al. 2024): we are most grateful to Dr Kerry Walsh for making the data available. We are grateful to Dr Magdalena Gapińska from the Laboratory of Microscopic Imaging and Specialized Biological Techniques, Faculty of Biology and Environmental Protection, University of Lodz, for the support with the CLSM data. We thank the Szczecin Diatomological Collection SZCZ (University of Szczecin), and in particular Dr Ewa Górecka, for providing material containing strain SZCZCH1090, which enabled detailed comparative analyses.

Author contributions

RMO, DGM and RT conceived the research. RMO and DGM wrote the original draft. RMO, DGM, RT, ÉÁ, SS, PKZ and ŁP contributed to writing, review and editing. Methodology: DNA extraction and molecular analyses were performed by RMO and PKZ; metabarcoding analyses were conducted by DGM, RT, ÉÁ and SS; SEM analyses were carried out by ŁP. RMO, DGM and RT performed the investigation. RMO, DGM, RT, ÉÁ and SS curated the data. Resources: Diatom material (environmental samples and cultures) was provided by RMO; metabarcoding datasets were provided by DGM, RT, ÉÁ and SS. RMO and ŁP prepared the visualizations.

Funding

The Hungarian part of the research presented in the article was carried out within the framework of the Széchenyi Plan Plus programme with the support of the RRF 2.3.1 21 2022 00008 project. The Royal Botanic Garden Edinburgh is supported by the Scottish Government's Rural and Environment Science and Analytical Services Division and IRTA receives support from the CERCA Programme/ Generalitat de Catalunya.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-50312-9.

Correspondence and requests for materials should be addressed to R.M.O.