MBMG Metabarcoding & Metagenomics

Halamphora taxa in Hungarian soda pans and shallow soda lakes detected via metabarcoding and microscopic analyses

Éva Ács[1,2], Tibor Bíró[1,2], Emil Boros[3], Péter Dobosy[3], Mónika Duleba[1,2], Angéla Földi[4], Keve T. Kiss[1,2], Zlatko Levkov[5], Péter Orgoványi[1], Orsolya Pálné Szén[1,2], Zsuzsa Trábert[1,2], Edit Vadkerti[1,2], István Grigorszky[6]

1 *University of Public Service, Faculty of Water Sciences, Department of Water Supply and Sewerage, Baja, Hungary*

2 *National Laboratory for Water Science and Water Security, Budapest, Hungary*

3 *Institute of Aquatic Ecology, Centre for Ecological Research, Budapest, Hungary*

4 *Eötvös Loránd University, Faculty of Science, Doctoral School of Environmental Science, Budapest, Hungary*

5 *Ss. Cyril and Methodius University in Skopje, Skopije, Republic of North Macedonia*

6 *Department of Hydrobiology, University of Debrecen, Debrecen, Hungary*

Corresponding author: Éva Ács (*acs.eva@uni-nke.hu*)

OPEN ACCESS

Abstract

The research presented investigates whether DNA-based metabarcoding can replace the morphology-based identification of diatom taxa in the ecological status assessments of aquatic habitats. When comparing data obtained with microscopy and metabarcoding, significant deviations have been noticed. One of the main reasons includes the incompleteness of the reference database used for taxonomic annotation of sequences. The database library should be complemented with species inhabiting unique habitats and having specific environmental requirements representing environmental endpoints for genetic diversification. Soda pans and soda lakes are examples of an extreme habitat with the loss of sodic character as the main threat; thus, accurate identification of species and exact information on their salinity tolerance is essential for adequate ecological status assessment. In the present study, by using microscopy and metabarcoding, we investigated taxa of the genus *Halamphora* that are common in soda pans and soda lakes. We detected six species of which *Halamphora dominici* and *H. veneta* occurred frequently and often in high abundance (it was often dominant having relative abundance higher than 5%). Analyses of DNA data confirmed the separation of the two species; as a result, the reference database library has been supplemented with sequences of *H. dominici*. Furthermore, we have confirmed that this species, which is a significant indicator of sodic character, shows a positive correlation with salinity.

Key words: diatoms, *Halamphora*, halobity, metabarcoding, soda pans/lakes

Introduction

Diatoms are reliable bioindicators for environmental conditions in aquatic habitats and, therefore, they are used in ecological status assessments (Ector et al. 2004; Lobo et al. 2016). Benthic diatoms constitute one of the five biological elements, based on which the ecological status of surface water bodies

should be assessed according to the Water Framework Directive (European Commission 2000). This process is based on diatom indices for which relative abundance data and information on pollution/trophic sensitivity and indicator values of taxa occurring in the given community are required (Berthon et al. 2011). Therefore, diatom taxa should be accurately and consistently identified. The traditional method uses microscopy to identify species, based on morphological features of the silica cell wall, the so-called frustule (Round et al. 1990). Identification of species requires deep taxonomical knowledge and expertise (Rimet and Bouchez 2012), depends on the literature used for species identification and could make identification subjective. Therefore, a more objective DNA-based method has been proposed to address taxonomic accuracy and consistency. Metabarcoding uses next generation sequencing to identify taxa in a community, based on short DNA (barcode) sequences (Hebert et al. 2003; Taberlet et al. 2012). To date, studies have compared the results of metabarcoding to morphological observations and found deviations between these two approaches (Rivera et al. 2018; Bailet et al. 2019; Kelly 2019; Mora et al. 2019; Mortágua et al. 2019; Pérez-Burillo et al. 2020; Borrego-Ramos et al. 2021; Duleba et al. 2021). One of the main reasons for the differences in the results is the incompleteness of reference databases (Rivera et al. 2018; Bailet et al. 2019; Mortágua et al. 2019; Borrego-Ramos et al. 2021; Duleba et al. 2021). It has been concluded that, currently, metabarcoding cannot replace morphology-based identification in ecological status assessments and a combination of two approaches has been proposed (Mora et al. 2019; Duleba et al. 2021).

Reference database libraries lack taxa diversity mainly from specific habitats, like soda pans (Duleba et al. 2021). Additionally, problems with misidentifications can be caused by phenotypic expression in valve morphology due to environmental stressors like changes in salinity (Håkansson and Chepurnov 1999; Trobajo et al. 2011; Leterme et al. 2013). Soda pans represent a unique habitat in the Carpathian Basin (Boros et al. 2013). These are inland alkaline soda environments providing specific living conditions within an extreme environment (Boros et al. 2017). Soda environments are characterised by high dissolved material content and frequently high trophic state and are rich in sodium and hydrogen carbonate ions (soda character) (Boros et al. 2013, 2014; Stenger-Kovács and Lengyel 2015). Therefore, soda pans are inhabited by several rare and halophilic species (Stenger-Kovács and Lengyel 2015; Ács et al. 2017, 2019). Soda pans are highly susceptible to water level changes (dilution/concentration) and poor water management when lower salinity water is channelled through or into the soda pans/lakes. Thus, monitoring the ecological status of soda pans is required. When the ecological status of soda pans is assessed, based on diatoms, accurate identification of species and information on their halobity tolerance is particularly important for correct assessment (Duleba et al. 2021).

Species within the genus *Halamphora* (Cleve) Levkov are common in soda pans (e.g. Stenger-Kovács and Lengyel (2015); Földi et al. (2018)). They were also commonly found in the study presented here. This genus was first recognised as a subgenus within the genus *Amphora* Ehrenberg ex Kützing (Cleve 1895). Levkov (2009) elevated the subgenus to genus level describing its members as having dorsoventrally linear, semi-lanceolate to semi-elliptical valves with variable, but often protracted valve ends, striae composed of round,

elliptical to transversely elongate areolae, the raphe lying on a raphe ledge near the ventral margin with dorsally curved distal ends and numerous girdle bands with one to two rows of pores. This separation from *Amphora* sensu stricto was later confirmed with phylogenetic analyses (Stepanek and Kociolek 2014).

This study aimed to present *Halamphora* taxa occurring in Hungarian soda pans and soda lakes using both microscopy and metabarcoding. We have decided to study the *Halamphora* genus because it was the fourth most common and seventh most species-rich genus amongst the 72 genera found in the soda samples we collected and its maximum relative abundance was the eighth largest. Two species from the *Halamphora* genus (*Halamphora dominici* Ács & Levkov and *Halamphora veneta* (Kützing) Levkov) were also found in almost half of the samples, they are very common species in soda pans and soda lakes and their maximum relative abundance exceeded 40 and 60%, respectively. Within this genus, species in *Halamphora veneta* group (*H. veneta*, *H. kevei* Levkov, *H. dominici* and *H. paraveneta* (Lange-Bertalot, Cavacini, Tagliaventi & Alfinito) Levkov) are not easy to distinguish. The main distinctive features include the stria density and valve shape (Levkov 2009). We have studied this group more thoroughly, testing if the combination of microscopy and metabarcoding could help delimit the species and investigated whether distinctive features in the original description of the species are appropriate to separate them. Furthermore, we intended to identify the *rbcL* DNA sequence of *H. dominici* and investigated whether it was distinct from *H. veneta*. The Diat.barcode database only included *H. veneta* sequences from this group.

Halamphora dominici is a characteristic species in Hungarian soda pans and soda lakes, as we and others have previously found (e.g. Stenger-Kovács and Lengyel (2015); Földi et al. (2018)). It is regarded as a halophilous species along with *H. kevei* (Levkov 2009; Stenger-Kovács and Lengyel 2015) showing positive correlation with salinity (Földi et al. 2018). In contrast, *H. veneta* was widespread in waters with lower salinity, from freshwater to slightly brackish waters (Levkov 2009; Stenger-Kovács and Lengyel 2015). The distribution of *H. paraveneta* was described (only from Sardinia) by Levkov (2009), further supplemented by Stenger-Kovács and Lengyel (2015) and reported to be frequent in ephemeral sodic waters. Levkov also notes that probably *H. paraveneta* has a wider distribution in Europe than reported, but is confused with *H. veneta*.

Materials and methods

Sampling

Samples were taken under the framework of a monitoring project (KEHOP-1.1.0-15-2016-00002). The objective of this programme was to assess the ecological status of Hungarian surface waters, both lentic and lotic environments with various conditions, for example, trophic state, salinity based on benthic diatom assemblages. Within this assessment, 37 soda pans and soda lakes were sampled. Samples were taken in May-June and September-October 2019, as well as May-July and August-October 2020. This resulted in 69 samples that were examined with microscopy, 26 of them also being subjected to DNA sequencing. For sequencing, samples that contained *Halamphora* taxa in considerable amounts, based on microscopy, were selected.

Due to the unfavourable weather conditions (several soda pans were dry on sampling occasions), we were unable to collect a few samples from alkaline soda pans; thus, we used soda pans samples collected in May 2021 within the framework of another project. However, we could only investigate these samples by light microscopy (LM).

In standing waters, epiphytic samples were mostly taken from green common reed (*Phragmites australis* (Cav.) Trin. ex Steud.) or, if it was unavailable, lesser bulrush (*Typha angustifolia* Linn.) or other emergent macrophytes were sampled by choosing five randomly selected stems.

To illustrate the whole distribution of *Halamphora veneta* including waters with lower salinity, we used additional samples collected in the KEHOP project including 31 samples from running waters and 51 from standing waters. Sampling from running waters was carried out in the period of March-May (and two samples from June) and September-October 2019. In rivers and streams, epilithic samples were taken from five randomly chosen cobbles; if these were absent, other available substrates, mainly emergent macrophytes, rarely dead-wood or artificial substrates (e.g. brick) were sampled.

In all cases, the five random repeats were integrated into one composite sample; the biofilm was scraped with a toothbrush into tap water. The acquired slurry was homogenised and divided into two parts. For DNA analysis, 2−3 ml was pipetted into a 15 ml sterile plastic centrifuge tube that was filled with absolute ethanol (final ethanol concentration ≥ 70%), then stored at 4 °C until processing. The rest of the slurry was preserved with buffered formaldehyde (4% final concentration) for microscopy (European Committee for Standardisation 2002).

Environmental variables

Cation (Na^+, K^+, Ca^{2+}, Mg^{2+}, NH_4^+) and anion (Cl^-, SO_4^{2-}, NO_3^-) concentrations were determined using a Dionex ICS 5000+ dual channel ion chromatograph (Thermo Scientific, Massachusetts, USA). Total nitrogen was measured with a MULTI N/C 3100 TC/TN analyser (Analytik Jena, Jena, Germany). Total phosphorus and orthophosphate concentrations were quantified spectrophotometrically, based on the methods of Murphy and Riley (1962) and Menzel and Corwin (1965).

Silicate (Si, µg l^{-1}; MSZ 1484-3:2006; Hungarian Standards Institution: HSI 2006), hydrogen carbonate and carbonate (HCO_3^- and CO_3^{2-}, mg l^{-1}, MSZ 448-11:1986, HSI 1986), chlorophyll-a concentration (CHA µg l^{-1}, MSZ ISO 10260:1993; HSI, 1993), nitrite (NO_2^-, mg l^{-1}; MSZ EN ISO 11885:2009; HSI 2009) and Secchi transparency (Secchi, cm, MSZ 260-46:1981, HSI 1981) were measured according to the national standard (see the references in Duleba et al. (2021)). The pH, electrical conductivity (µS cm^{-1}) and dissolved oxygen (DO, mg l^{-1}) were measured in situ with YSI EXO-2-S3 equipment.

Microscopy

For light microscopy, samples were treated with hydrochloric acid and hydrogen peroxide, then washed with distilled water. Cleaned diatom valves were mounted with Naphrax (CEN 2014) and investigated using a Zeiss Axio Imager Z2 microscope, equipped with differential interference contrast (DIC) optics at a 1600× magnification. At least 500 valves were identified to the species or

genus level in lake samples. For illustrating the distribution of *H. veneta*, a minimum of 400 valves were counted and identified in running water samples. The relative abundance of each taxon in the sample was calculated by dividing the counted valve number of the given taxon with the total counted valve number.

For scanning electron microscopy (SEM), part of the cleaned and washed sample was filtered through a 3 µm Isopore polycarbonate membrane filter (Merck Millipore), which was then fixed on to an aluminium stub using double-sided carbon tape and coated with gold using a rotary-pumped spatter coater, Quorum Q150R S. Ultrastructural features of diatoms were observed with a Zeiss EVO MA 10 SEM operated at 10 kV and Zeiss Sigma 300 operated at 2 kV (only the images of *Halamphora elongata* Bennett & Kociolek). The working distance varied between 10.5 and 10.7 mm.

Metabarcoding

DNA was extracted from the samples using NucleoSpin Soil Kit (Macherey-Nagel). The protocol by Vautier et al. (2020) was used; however, a mixture of two kinds of lysis buffer from the kit in a 1:1 ratio was applied.

A 312 base pair (bp) region of the *rbc*L gene was amplified and sequenced on the Illumina MiSeq platform. Primer sequences, circumstances of polymerase chain reactions (PCR), library preparation and sequencing are described in Duleba et al. (2021). The workflow included the following steps: PCR with gene-specific primers; PCR product purification, concentration, quantification and dilution to equimolar concentrations; index reaction; product purification and pooling; quality and concentration control; dilution to the same concentration (4 nM); running on the Illumina MiSeq system. The *rbc*L-specific primers used in the first PCR were developed by Vasselon et al. (2017) and supplemented with Illumina overhang P5/P7 adapters. The PCR products were purified with 1.0× AMPure XP magnetic beads (Beckman Coulter). Concentrations were quantified by using a Qubit 4 Fluorometer (Invitrogen) with the Qubit dsDNA HS Assay Kit. In the cleaned index reactions, PCR products were provided with Nextera DNA CD Indices with P5/P7 adapters and P7/P5 tags attached. Quality assessment of the samples was performed using an Agilent TapeStation System 4150 (Agilent) with Agilent High Sensitivity D1000 ScreenTape Assay Reagents.

For the run on the Illumina MiSeq system, the starting concentration of the final library pool was 4 nM. After denaturation and dilution, sequencing was performed using the Illumina MiSeq V2(500) Reagent Kit and a 2 × 250 bp read length.

Polymerase chain reactions, library preparation and sequencing were performed by Biomi Ltd. (Hungary).

One sample (Kisteleki-Müller-szék) was subjected to long read sequencing following the description in the manual of LoopSeq PCR Amplicon Kit (Loop Genomics). This sample was chosen because microscopic analyses revealed that two taxa of the *H. veneta* group, *Halamphora dominici* and *H. veneta* that have different ecological preferences in terms of salinity, were dominant.

Data analysis

Sequence data analysis followed the method of Keck et al. (2019), i.e. the official DADA2 pipeline (Callahan et al. 2016) modified and applied to diatom *rbc*L

metabarcoding (Keck et al. 2019). This included the following main steps: removing primer sequences, trimming and filtering according to read quality, dereplicating, filtering with the core sample inference algorithm of DADA2, aligning and merging paired forward and reverse reads into one contig sequence, removing chimaeras and taxonomic assignment. Diat.barcode version 9.2 (Rimet et al. 2019) was used as the reference database. Read numbers were corrected according to biovolume using the correction factor developed by Vasselon et al. (2018). For *H. dominici*, the correction factor of *H. veneta* was used. Relative abundance was calculated for each taxon in the samples dividing corrected read number of a given taxon by the total corrected read number of the sample. For more details, see Duleba et al. (2021).

All sequences that could be ranked to the genus *Halamphora* were assigned at species level with the DADA2 pipeline. *Halamphora veneta* sequences found in samples containing *H. dominici*, based on microscopy, were aligned to sequences in the National Center for Biotechnology Information (NCBI) GenBank database using the Basic Local Alignment Tool (BLAST, Altschul et al. (1990)). The search was performed in the Nucleotide collection database (nr/nt) using the Standard Nucleotide BLAST programme, megablast (highly similar sequences) algorithm with the default parameter settings. The short *rbc*L amplicon sequence variants (ASVs) were used for identifying long sequences. By using long read sequencing, we obtained sequences with different lengths (sequences identified as *H. veneta* ranged from 415–1548 nt in length, while sequences identified as *H. dominici* ranged from 379–1547 nt in length). The longest sequences per species were used in phylogeny and these sequences having been deposited in the GenBank database under accession numbers OQ588791–OQ588792.

Halamphora sequences from our samples were compared to those acquired from Diat.barcode 9.2 (Rimet et al. 2019). Alignment of sequences was performed with Clustal W (Thompson et al. 1994) implemented in MEGA7 (Kumar et al. 2016). Alignments are available as supplementary material (Suppl. materials 3, 4). The same software was used for calculating pairwise uncorrected p-distance, for selecting the best nucleotide substitution model and for performing Maximum Likelihood phylogenetic analyses; the last bootstrap test was performed in 500 replicates.

Pearson Correlation Values were calculated from data obtained by microscopy and metabarcoding using Past 4.12 (Hammer et al. 2001).

Results and discussion

Within the framework of the whole project, several environmental variables were measured and shown in Suppl. material 1. According to the international system of salinity ranges for salt content of continental waters (Hammer 1986), 22 of the studied soda pans and soda lakes were subsaline and 15 were hyposaline, 15 were natural and 22 had a disturbed status (Boros et al. 2013, 2014). Based on the dominant ions of the eight main ions, two waterbodies belonged to inland saline, 17 to alkaline soda base, two to chloride (salt) soda and two to mixed sulphate chloride soda. One lake belonged to the sulphate soda (all sample sites of Lake Fertő) and one lake to the mixed magnesium sulphate soda (all sample sites of Lake Velencei) chemical type (Boros et al. 2013, 2014). The proportion of cations and anions found during our study is presented in Fig. 1. Other physical-chemical data are provided in Suppl. material 1.
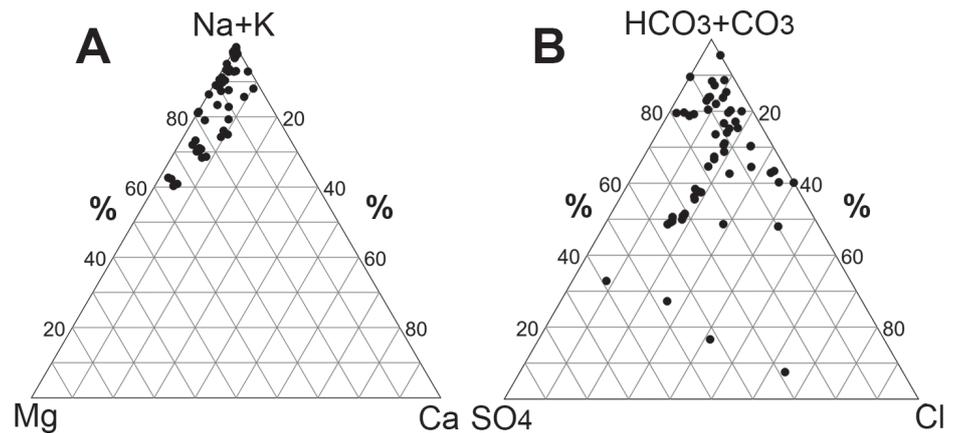
Figure 1. The ternary diagram of cation (**A**) and anion (**B**) concentrations measured in the samples.

During this study, six *Halamphora* species were identified from soda pans and soda lakes. Five of these were found by both microscopy and metabarcoding. One species (*Halamphora oligotraphenta* (Lange-Bertalot) Levkov) was only detected by microscopy because it occurred only in samples that were not sequenced. Two of the six *Halamphora* species were frequent (*H. dominici* and *H. veneta*), occurring in approximately half of the samples and they were often dominant (sometimes reaching 40−60% relative abundance and *H. dominici* was dominant in approximately half of the samples) in soda pans and lakes. The other four *Halamphora* species occurred in 5−7% of the samples and were never dominant (Suppl. material 1).

Halamphora species identified by microscopy and/or metabarcoding in the studied soda pans and soda lakes

In the following, morphological description, information from sequencing and ecological preferences in terms of salinity of the *Halamphora* taxa detected in studied soda pans and soda lakes are provided. The morphological description is based on the descriptions by Levkov (2009), as it is the most complete guide to identify *Halamphora* species by microscopy. Differences between our observations and literature are indicated. Descriptions also includes the four-letter Omnidia codes to help the search in Omnidia (Lecointe et al. 1993), a software for ecological status assessment based on diatoms; thus, Omnidia has a wide database of diatom taxa along with their ecological preferences.

Halamphora coffeaeformis (Agardh) Levkov, 1903 (HACO)
Fig. 2I−J

Short morphological description. *Length*: 23−35 (14−55) µm, *width*: (3.5)5−7.2 µm, number of dorsal striae in the middle: 19−22/10 µm (Levkov 2009). *Valves* narrow semi-lanceolate, dorsal margin convex, ventral margin weakly concave. *Apices* narrow, protracted, capitate, bent ventrally. Central area absent on both sides. Raphe slightly arcuate, very close to ventral margin. Proximal raphe end slightly deflected to dorsal valve side. Ventral striae very short, dorsal striae parallel in mid-valve becoming radiate towards apices. Puncta not visible in LM. *Ultrastructural features*: biseriate striae, this being the main characteristic of the species (Levkov 2009).
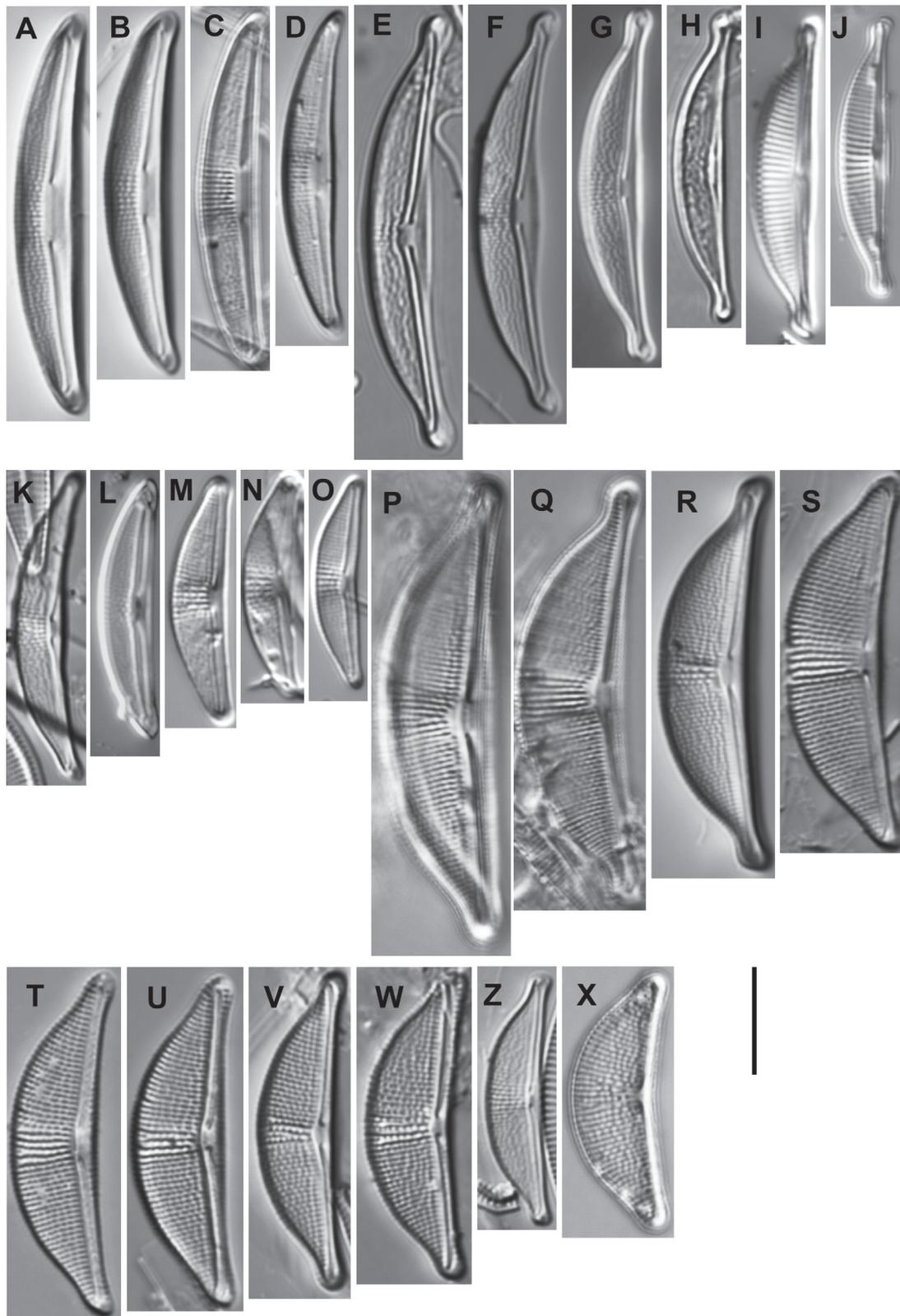
Figure 2. LM images of *Halamphora* species detected in present study. **A−D** *Halamphora paraveneta* LM (No. of sampling site: 30 in Suppl. material 1) **E−G** cf. *H. oligotraphenta* LM (No. of sampling site: 12 in Suppl. material 1) **H** *H. oligotraphenta* LM (No. of sampling site: 12 in Suppl. material 1) **I, J** *H. coffeaeformis* LM (No. of sampling site: 36 of I and 11 of J in Suppl. material 1) **K−O** *H. veneta* LM (No. of sampling site: 12 of K, No. of sampling site: 30 of L and 29 of M−O in Suppl. material 1) **P−X** *H. elongata* LM (No. of sampling site: 21 of P, U-W, X, No. of sampling site: 30 of Q, R and No. of sampling site: 19 of S, T, Z in Suppl. material 1) (scale bar 10 μm). The description of sampling sites is provided in Suppl. material 1.

Morphologically similar taxa. *Halamphora hybrida* (Grunow) Levkov has a similar valve outline, but its striae are crossed by a longitudinal line close to the dorsal margin, and the dorsal striae are formed by elongated areolae.

Detection by metabarcoding. Only one short amplicon sequence variant was detected in one sample that was assigned as *H. coffeaeformis*. It showed 0.004 − 0.072 p-distance from the *H. coffeaeformis* sequences recorded in the database. The p-distance 0.072 was with the *H. coffeaeformis* sequence (accession number FJ002103) that showed relatively high difference (0.039 − 0.04, 51 − 54 nt difference) from the other *H. coffeaeformis* sequences in the database (Suppl. material 2).

Ecology, distribution. Halophilic, cosmopolitan species found in waters with high electrolyte content and in brackish and saline inland waters. It occurred in one saline pan (Sárszentágotai-sóstó) and two shallow soda lakes (Lake Fertő and Lake Szelidi) during our survey (Fig. 8A).

Halamphora dominici Ács & Levkov, 2009 (HDOM)
Fig. 3

Short morphological description. *Length*: 10−20 µm, *width*: 3.5−4 µm, number of ventral striae on the mantle: 30−34/10 µm, number of dorsal striae: 24−28/10 µm (Levkov 2009). (The ventral striae are not be visible in LM, but in SEM, we can recognise them, for example, see Fig. 3EE). *Valves* semi-elliptical, strongly dorsiventral. Valve ends broadly rounded, not elongated. Axial area narrow, wider on ventral side. Central area absent on dorsal valve side. Raphe straight and proximal ends slightly bent dorsally. Dorsal striae punctate, radial throughout. Ventral striae on the mantle barely or are not visible under light microscopy. *Ultrastructural features*: Partial conopeum (rib adjacent to raphe) narrow, poorly developed in smaller specimens, prominent from valve surface in larger specimens. Raphe weakly to strongly curved at proximal ends. Striae uniseriate along whole length, radial and composed of elongated areolae. Areolae variable in length, generally smaller towards the valve middle and more elongated towards dorsal edge of valve and valve ends. Ventral striae not interrupted in region of central nodule. Internally, distal raphe ends terminate in weakly-developed helictoglossa. Proximal raphe ends fuse into central helictoglossa. Areolae occluded by hymens (Levkov 2009).

Morphologically similar taxa. It is similar to *Halamphora kevei*, but *H. dominici* has a more rounded (semi-elliptic) valve shape and broadly rounded, not protracted, not ventrally bent valve ends. It resembles *H. veneta*, but the latter has lower stria density (dorsal number of striae: 18−22/10 µm) with coarser punctuate striae. It is also similar to *Halamphora paraveneta* that also has coarser punctuate striae and lower stria density (dorsal number of striae: 18−21/10 µm).

Detection by metabarcoding. We found 23 short amplicon sequence variants that were assigned as *H. veneta* with DADA2, but phylogeny (Fig. 4) and BLAST search showed a closer similarity with isolate 7951-AMPH106 (accession number MG027464) that is recorded as *H. veneta* in Diat.barcode database (this explains assigning our sequences as *H. veneta* by DADA2). The corresponding long (1547 nt) *rbc*L sequence showed 99.77% similarity (two substitutions and one deletion) with this sequence according to BLAST search. In samples in which these ASVs were in high number, *H. dominici* was found
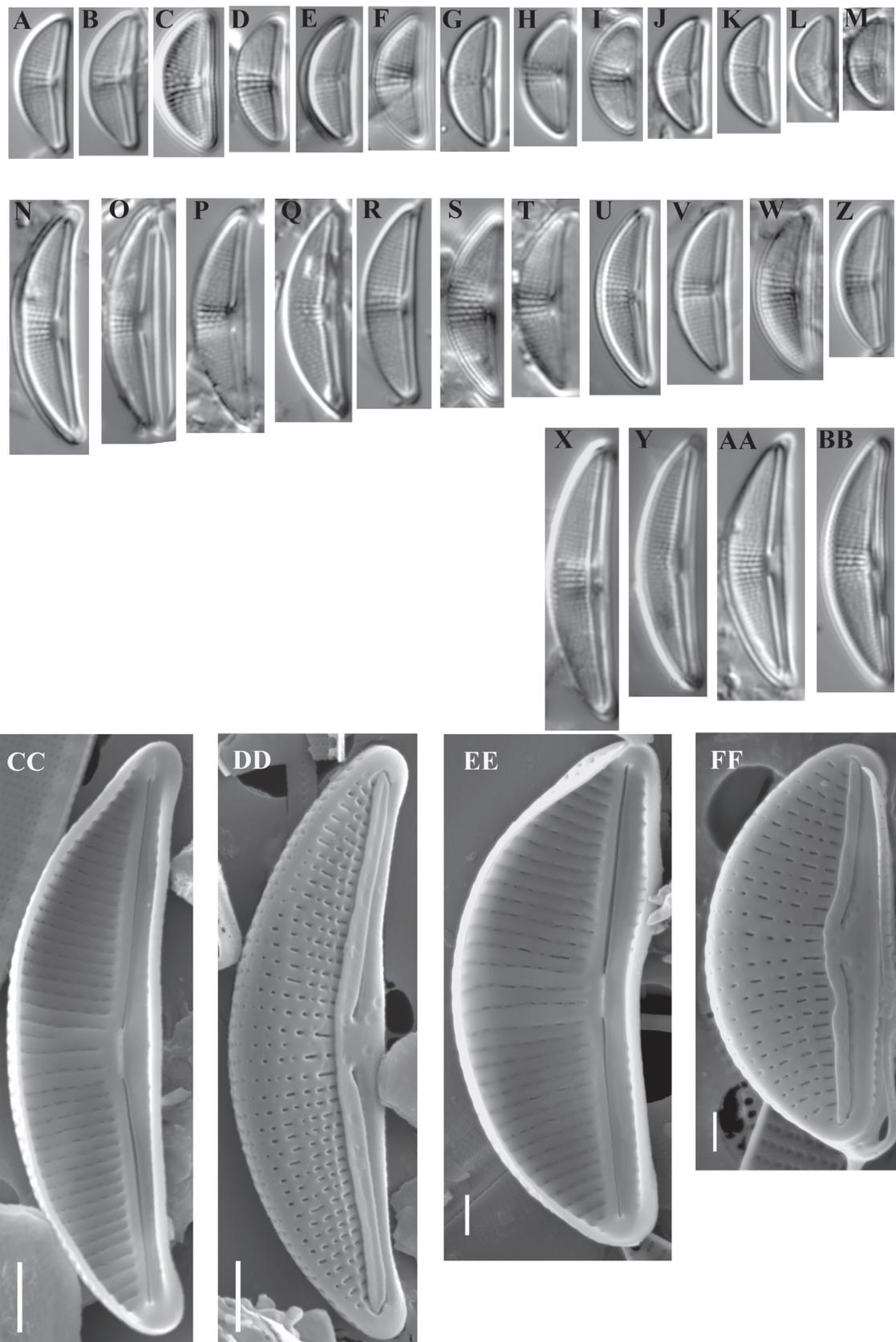
Figure 3. *Halamphora dominici* from the studied soda lakes and soda pans. **A−BB** LM. **CC, EE** SEM internal view **DD, FF** SEM external view (No. of sampling site: 25 of O, P, S, T, U, W, No. of sampling site: 27 of C, N, Q, Y, AA and No. of sampling site: 28 of others in Suppl. material 1). The description of sampling sites is provided in Suppl. material 1. Scale bars: 10 μm (**A−BB**); 2 μm (**CC, DD**); 1 μm (**EE, FF**).
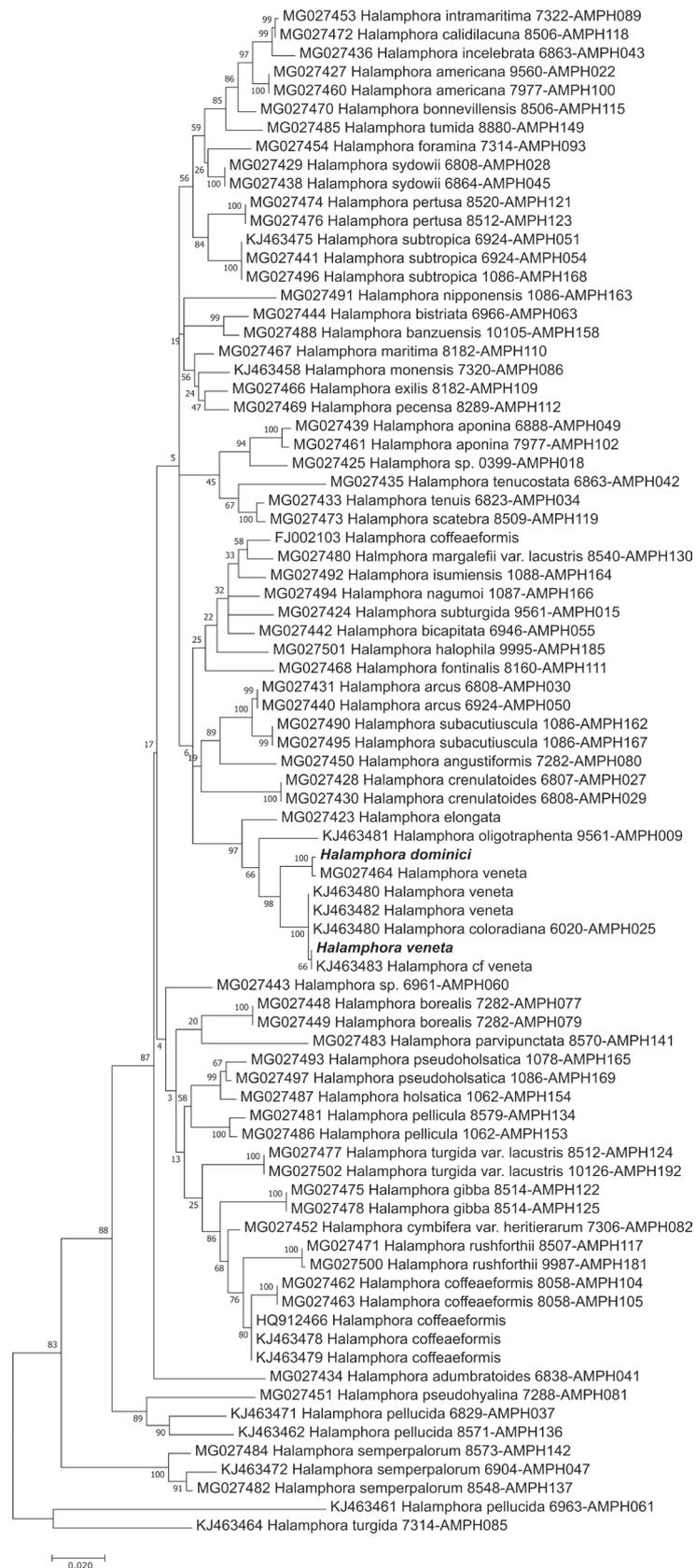
Figure 4. Maximum Likelihood tree of *Halamphora* amplicon sequence variants (ASVs) found in soda pans and soda lakes as well as *Halamphora* taxa from Diat.barcode database. *Pinnularia brebissoni*, *Fallacia pygmaea*, *Pseudofallacia monoculata*, *Caloneis lewisii*, *Stauroneis acuta* and *Craticula cuspidata* were used as outgroup. The total length of the alignment was 1530 nt. Sequences from the database are provided with NCBI GenBank accession numbers (if available) or culture ID of Thonon Culture Collection. Bootstrap values are indicated at nodes. Scale bar represents 0.01 substitutions per site.

in high abundance. This correlation indicated that the 23 ASVs belonged to *H. dominici*, suggesting that the MG027464 sequence was misidentified. The *Halamphora dominici* clade separated from *H. veneta* sequences (Fig. 4) and the long sequence showed 2% p-distance (31 nt difference, Suppl. material 2) with the long sequence of *H. veneta* confirming that *H. dominici* is a distinct species from *H. veneta*. The long sequences of *H. dominici* and *H. veneta* acquired in our study were inserted into the phylogenetic tree generated by Stepanek and Kociolek (2019) which involved several *Halamphora* taxa. The phylogeny also showed that *H. dominici* was located close to *H. veneta*, but on a distinct branch (Fig. 5). Our tree was based only on the *rbc*L gene, while Stepanek and Kociolek (2019) used four markers (SSU, LSU, *rbc*L, *psb*C) which caused some

Figure 5. Maximum Likelihood tree of *Halamphora* taxa from the work by Stepanek and Kociolek (2019) as well as *Halamphora dominici* and *Halamphora veneta* from our study. The total length of the alignment was 1560 nt. Sequences from NCBI GenBank are provided with accession numbers. Long *rbc*L sequences of *H. dominici* and *H. veneta* from our study are written in bold italics. Bootstrap values are indicated at nodes. Scale bar represents 0.02 substitutions per site.

differences in the topology of the two trees; similarly to their tree, *H. veneta* and *H. dominici* grouped with *H. oligotraphenta* and *H. elongata*.

In the case of *H. dominici*, metabarcoding helped delimit species boundaries. Morphologically, it is difficult to distinguish the larger specimens of *H. dominici* from *H. kevei*.

The difference between *H. dominici* sequence variants was below 2%, p-distance between ASVs ranged from 0 to 0.015 (0 – 4 nt difference). The *H. dominici* asv1 occurred in the most (10) soda pans and lakes (in 16 samples) and in highest abundances. Kisteleki Müller-szék and Madarász Lake showed the highest diversity of sequence variants.

Ecology, distribution. Halophilic species. Typical, characteristic, often strongly dominant species in Hungarian soda pans. Present in one soda lake (Lake Fertő) as well, but was never dominant during our survey (Fig. 8B).

Halamphora elongata Bennett & Kociolek, 2014 (HEGT)
Figs 2P–X, 6A–C

Short morphological description. *Length*: 22–57 µm, *width*: 7–11 µm, number of dorsal striae: 19–23/10 µm, number of ventral striae: 26–30/10 µm (Kociolek et al. 2014; Stepanek and Kociolek 2018). In our study: *length*: 22–40 µm (mean: 30 µm, n = 13), *width*: 5–11 µm (mean: 7.7 µm, n = 13), number of dorsal striae: 20–23/10 µm. *Valves* semi-elliptical and strongly dorsiventral. Ventral margin usually weakly concave or sometimes slightly convex. Valve ends rostrate to capitate ventrally curved and slightly protracted. Axial area narrow, often curved or with semi-circular hyaline area (it can be significantly broad in large valves) on dorsal side of the central node. Raphe located near ventral margin. Proximal raphe endings expanded and dorsally curved; distal raphe endings dorsally curved. Dorsal striae punctate, with dash-like areolae, radiating throughout and more widely spaced at mid-valve. Ventral striae short and difficult to observe in LM because of the very narrow ventral valve. **Ultrastructural features:** proximal raphe endings curved in dorsal direction, dorsal striae consisting of several elongated areolae. Area of central striae thickened and striae of this area coarser than others. In SEM, short ventral striae visible, composed of short dashes. This is a halophilic species (Kociolek et al. 2014; Stepanek and Kociolek 2018).

Morphologically similar taxa. It is very similar to *Halamphora subcapitata* (Kisselev) Levkov and their morphometric features strongly overlap, especially in LM. According to Kociolek et al. (2014), *H. elongata* differs from *H. subcapitata* in that the former is relatively narrower (7.5 – 12 µm, Levkov (2009)) with more capitate ends and its areola has external foramen areolar occlusions. We measured the width and length of 20 *H. subcapitata* individuals (nine from Stenger-Kovács and Lengyel (2015), 11 from Levkov (2009)) and 21 *H. elongata* individuals (10 from Kociolek et al. (2014) and Stepanek and Kociolek (2018); and 11 from the present study) and calculated the length/width ratio (L/W). We found no significant difference in L/W of the two species (Fig. 7); thus, it does not confirm the claim that *H. elongata* is relatively narrower. Stenger-Kovács and Lengyel (2015) presented two SEM images of *H. subcapitata*, where the external foramen areolar occlusions are clearly visible, while on our specimens, they are not (Fig. 6B, C). Several experiments correlated salinity and the structure or size of the frustule (Tuchman et al. 1984; Håkansson and Chepurnov
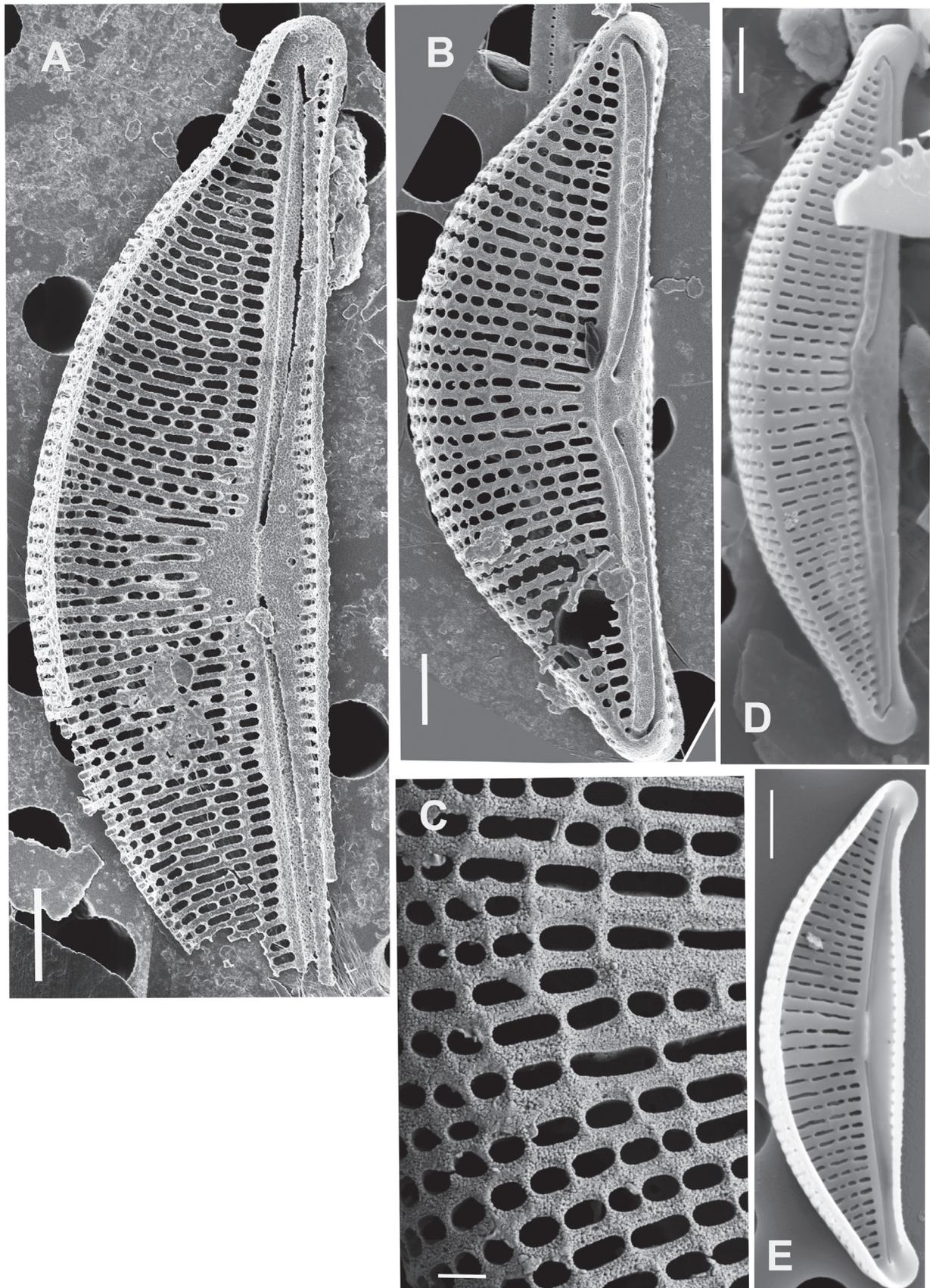
Figure 6. SEM images of two *Halamphora* species from the studied sites. **A–C** *Halamphora elongata* **D, E** *H. veneta* **A, E** internal view **B, D** external view **C** detail of **B** (No. of sampling site: 21 of A–C and No. of sampling site: 30 of D, E). The description of sampling sites is provided in Suppl. material 1. Scale bars: 3 μm (**A**); 2 μm (**B, D, E**); 500 nm (**C**).
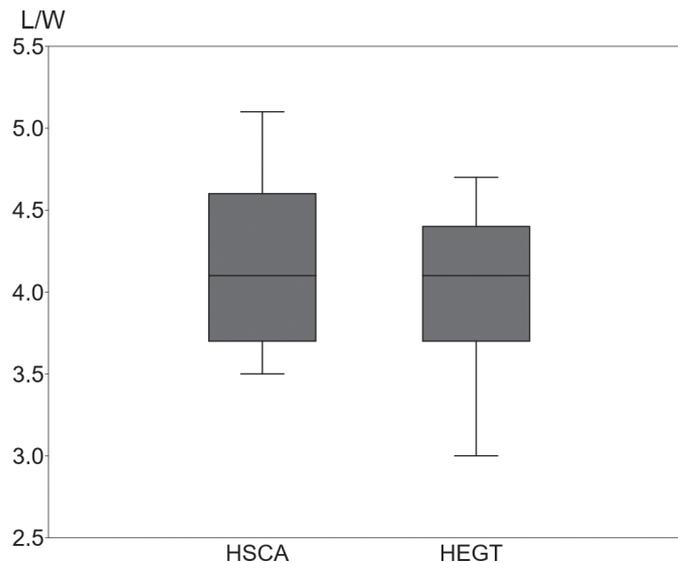
Figure 7. The length/width ratio of measured *H. subcapitata* (HSCA) and *H. elongata* (HEGT) specimens observed in our study.

1999; Trobajo et al. 2011); therefore, we think that *H. elongata* is the same species as *H. subcapitata*, but to synonymise it, a detailed analysis of the type materials is required.

Detection by metabarcoding. We detected one ASV that was assigned as *H. elongata* (in three samples) that differed only in one nucleotide (p-distance = 0.004) from the sequence in the database (the same record is present in Diat.barcode and GenBank). Unfortunately, neither Diat.barcode nor GenBank contain sequences data for *H. subcapitata*; therefore, we did not have the possibility to compare their DNA sequences, but a single nucleotide difference from the database sequence of *H. elongata* suggests that *H. elongata* was present in our samples.

Ecology, distribution. Halophilic species. This taxon occurred in two soda pans and one shallow soda lake (Lake Velencei) during our survey (Fig. 8C).

Halamphora oligotraphenta **(Lange-Bertalot) Levkov, 2009 (HOLI)**
Fig. 2E–H

Short morphological description. *Length*: 17–39 µm, *width*: 3.2–4.5 µm, number of dorsal striae in the middle: 26–30/10 µm (Levkov 2009).

Valves semi-lanceolate, dorsiventral with arched dorsal margin and straight or slightly tumid ventral margin. Valve ends shortly protracted, capitate and slightly ventrally bent. Axial area narrow and wider on the ventral side. Central area absent on dorsal side, on ventral side not differentiated from axial area. Raphe slightly arched, proximal endings distantly spaced and slightly dorsally deflected. Dorsal striae punctuate and radiating throughout. Ventral striae hard to resolve in LM. *Ultrastructural features*: distal raphe endings prolonged and dorsally curved. On ventral side, central area extends to valve margin as a fascia. Dorsal striae uniseriate and composed of areolae with various shapes. Ventral striae also uniseriate and comprised of longitudinally elongated areolae. Ventral striae interrupted in the region of central nodule (Levkov 2009).
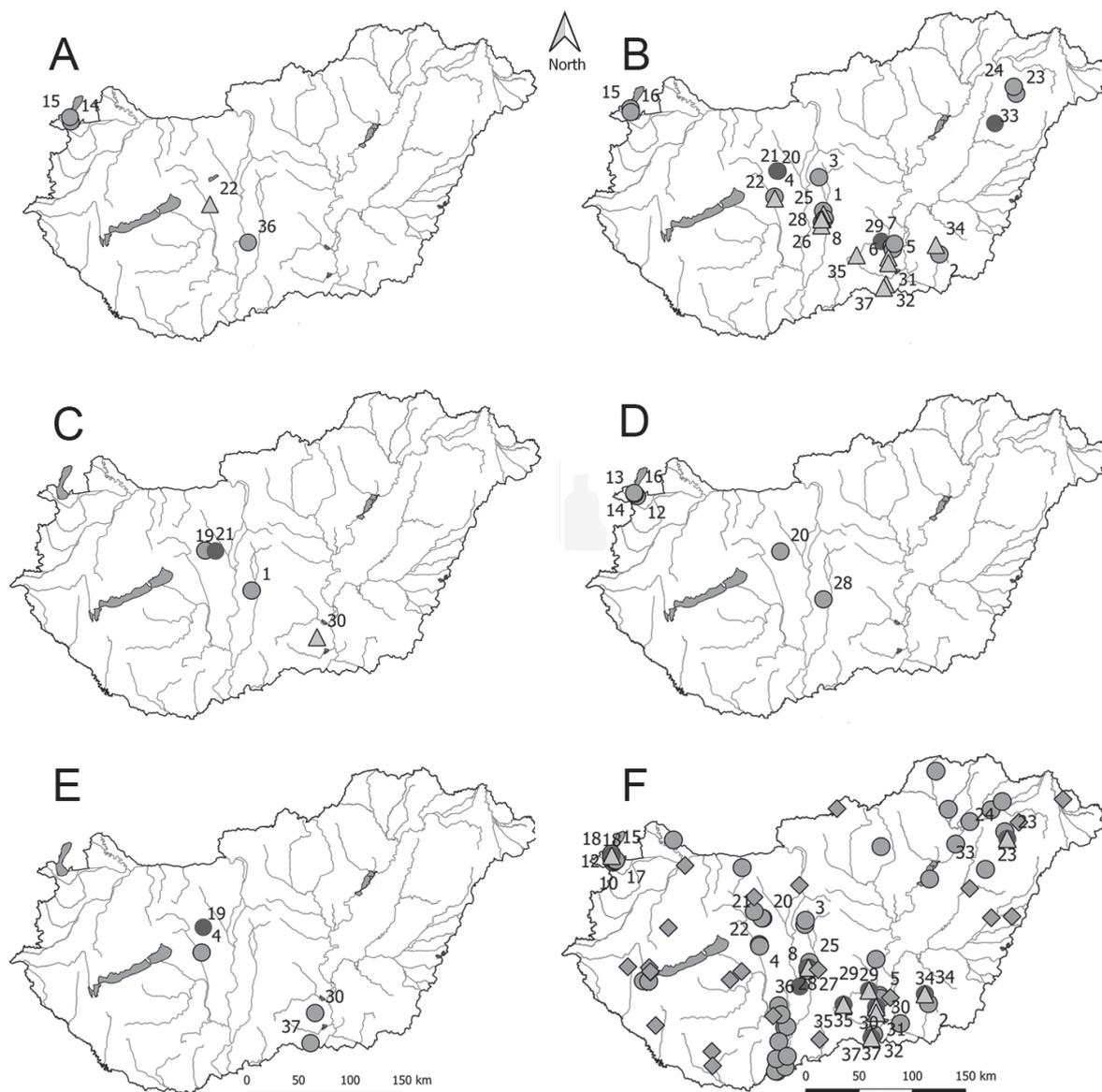
Figure 8. The occurrence of *Halamphora coffeaeformis* (**A**), *H. dominici* (**B**), *H. elongata* (**C**), *H. oligotraphenta* (**D**), *H. paraveneta* (**E**) and *H. veneta* (**F**) in Hungarian soda pans and soda lakes, based on microscopic and molecular investigations carried out in present study. In the case of *H. veneta*, we plotted all occurrences during our survey (light grey circle: based on microscopy, dark grey circle: based on metabarcoding, grey diamond: running waters based on microscopy, light grey triangular: based on microscopy and metabarcoding). This approach was applied because the distribution of *H. veneta* pointed beyond soda lakes and soda pans.

Morphologically similar taxa. This species resembles *Halamphora coraensis* Foged (Levkov), but the latter species is significantly wider (4.5−6.5 µm) and has a lower stria density (24−26 in 10 µm).

Detection by metabarcoding was not possible because this species was only observed under the microscope in samples that were not sequenced.

Ecology, distribution. Freshwater, oligotrophic species (Levkov 2009). Unfortunately, soda pan and soda lake samples in which *H. oligotraphenta* occurred were not sequenced; therefore, our DNA data could not confirm the distribution of the species in soda pans and soda lakes. This species was recorded in one soda pan and two shallow soda lakes (Lake Velencei and Lake Fertő) during our survey (Fig. 8D).

Halamphora paraveneta (Lange-Bertalot, Cavacini, Tagliaventi & Alfinito) Levkov, 2009 (HPVE)
Fig. 2A–D

Short morphological description. *Length*: 20–75 μm, *width*: 4.7–5 μm (measurement in this study), number of dorsal striae: 18–21/10 μm (Levkov 2009).

Valves semi-elliptical or semi-lanceolate and slightly dorsiventral. Valve ends rounded and slightly ventrally bent. Axial area moderately narrow, more expressed on ventral valve side. Central area absent on dorsal side and not differentiated from axial area on ventral valve side. Dorsal striae radiating throughout, coarsely punctuate. Proximal raphe ends distantly spaced. *Ultrastructural features:* Striae uniseriate and consisting of elongate areolae. Ventral striae on mantle not interrupted in region of central nodule and composed of round areolae. Proximal raphe ends slightly dorsally curved, distal ends strongly dorsally deflected (Levkov 2009).

Morphologically similar taxa. This species is very similar to *H. veneta*, differing from it in stria density (stria density of *H. veneta*: 24–30 in 10 μm) and valve end shape (more broadly rounded in *H. paraveneta*).

Detection by metabarcoding. Sequence data for *H. paraveneta* is not available either in Diat.barcode (the reference database we used for taxonomic assignment) or in NCBI GenBank database; therefore, metabarcoding could not detect it as an assigned species in our samples. However, there was one ASV (*Halamphora veneta* asv 10) that was assigned as *H. veneta*, but showed higher p-distance (Suppl. material 2) from other *H. veneta* ASVs. We assume this is the sequence of *H. paraveneta*, but it requires further investigations.

Ecology, distribution. Halophilic species. During our survey, *H. veneta* was recorded in three soda pans and one shallow soda lake (Lake Velencei) (Fig. 8E).

Halamphora veneta (Kützing) Levkov, 2009 (HVEN)
Figs 2K–O, 6D, E

Short morphological description. *Length*: (8)17–35 μm, *width*: (3.5)4–6.5 μm, number of ventral striae: 24–30/10 μm, number of dorsal striae: 18–22/10 μm (Levkov 2009; numbers in parentheses Lange-Bertalot et al. (2017)).

Valve shape semi-lanceolate with strongly convex dorsal margin and slightly concave ventral margin. Valve ends narrowly rounded, slightly protracted and ventrally bent. Axial area narrow, wider on ventral side. Central area absent on dorsal valve side. Raphe filiform, proximal ends straight or dorsally curved, distantly spaced (it is important to emphasise that this feature is expressed especially on large specimens; however, proximal endings can be more closely spaced and this character is variable). Dorsal striae can be observed with distinct punctuations in LM, radiating throughout. Ventral striae on mantle hard to resolve in LM, not interrupted at the central nodule. *Ultrastructural features:* Partial conopeum is moderately wide, gradually narrowing towards valve ends. Sometimes has ornamentation (small, round depressions) in region of the central nodule. Raphe arched, proximal ends straight and terminating in slightly expanded central pores. Distal raphe ends dorsally deflected. Striae uniseriate throughout, radiating and consisting of elongate areolae. Ventral striae not interrupted in region of central nodule and composed of round or elongate

areolae. Girdle bands open, each with large round or oval poroids located in two rows. Internally central internal costae thickened inwards and elevated from valve plane. Distal raphe ends ventrally deflected and terminating with poorly-developed helictoglossae. Proximal raphe endings fuse into a central helictoglossa. Areolae occluded by hymens (Levkov 2009).

Morphologically similar taxa. It is similar to *H. oligotraphenta*, but the latter has more capitate valve ends and live in oligotrophic freshwater habitats rich in calcium bicarbonate. It resembles *H. kevei* that has a higher stria density (number of dorsal striae: 24−30/10 μm) and more finely punctuated striae. It is most similar to *H. paraveneta*; however, the latter is usually larger (20−75 μm), although the size ranges of the smaller cells overlap. The frustule width of *H. paraveneta* (11−20 μm) and *H. veneta* (9−17 μm) also overlaps. The valve width of *H. paraveneta* is 4−9 μm also overlapping with *H. veneta* in the case of smaller specimens. Valve ends of *H. paraveneta* are bluntly and more broadly rounded. Dorsal striae are less radiate in the central region, number of dorsal striae is 18−21/10 μm, becoming more radiate and denser (22−23/10 μm) towards the valve ends. This is the case also with smaller specimens. The two species can occur together.

Detection by metabarcoding. Overall, 25 ASVs that were assigned as *H. veneta* could be accepted as *H. veneta*, showing more similarity to other *H. veneta* sequences than to the MG027464 sequence (Suppl. material 2; Fig. 4). Except for the above mentioned *Halamphora veneta* asv 10 (that could be *H. paraveneta*), they showed p-distances ranging from 0 to 0.015 (0 − 4 nt difference). The corresponding long (1548 nt) sequence showed 100% identity with *H. veneta* according to BLAST that was confirmed by phylogeny (Fig. 5).

The most widely distributed ASV was *H. veneta* asv1 occurring in 14 soda pans and lakes (19 samples). Interestingly, similarly to *H. dominici*, the highest number of *H. veneta* sequence variants was found in Kisteleki Müller-szék and Madarász Lake.

Ecology, distribution. A cosmopolitan species, common in freshwater and slightly brackish waters, often detected in waters with high nutrient content. It can tolerate drying (aerophytic species) and waters with high organic content. *H. veneta* was frequent and often a dominant species in several standing and running waters during our survey (Fig. 8F). (In the case of this species, the occurrences experienced during the entire survey are indicated on the map, showing that it was also found in many non-soda habitats, since the non-soda occurrences were also taken into account when calculating the halobity value).

Comparison of results, based on microscopy and metabarcoding

Amongst the six *Halamphora* species detected in our samples, *H. dominici* and *H. veneta* were dominant (relative abundance higher than 5%) in several samples, based on microscopy and/or metabarcoding. *H. dominici* showed corresponding occurrences, based on morphological and molecular data in 15 samples analysed by both methods; for *H. veneta* corresponding occurrences were shown in 13 samples. For *H. dominici*, a significant positive correlation was found between relative abundances, based on microscopy and metabarcoding (Pearson correlation r = 0.73, p < 0.05). *H. veneta* and *H. dominici* were only detected, based on sequencing in six and five samples, respectively.

The observation that a taxon can be detected in more samples analysed by metabarcoding than by microscopy is understandable, considering that DNA was extracted from more cells than the number of cells counted under the microscope. In two samples that were investigated by both microscopy and metabarcoding, *H. veneta* was found only under the microscope. A potential explanation includes that the observed valves belonged to dead cells.

The other four species did not reach 5% (the maximum value was 2.9% for *H. oligotraphenta*, based on microscopy) either based on morphology or DNA sequences. They occurred only in a few samples compared to the dominant two species and we could not find correspondence between microscopy- and metabarcoding-detected occurrences (except for one sample in the case of *H. elongata*). Moreover, *H. oligotraphenta* was not detected by metabarcoding at all. Reasons to explain why a taxon can only be found either by microscopy or metabarcoding may include the observation of dead cells under the microscope, incomplete DNA extraction, failure of primer annealing, detection of extracellular DNA or resting stages and unequal distribution of individuals between subsamples (Duleba et al. 2021).

Salinity demands of the two dominant *Halamphora* species

Based on the salinity system of Van Dam et al. (1994), *Halamphora dominici* belongs to category 4 (mesohalobous) (the average salinity of the waters in which the species was found was 3.1‰, while in the case of *H. veneta*, it was 1.6‰). It prefers waters with higher salinity compared to *H. veneta* as demonstrated by the total salinity violin plot (Fig. 9).

Given that the ecological status of Hungarian soda pans and shallow soda lakes depends on salinity (Földi et al. 2018), it is a major characterising environmental factor for these environments. The exact identification of species along with information on their salinity tolerance have great importance, especially for dominant species.
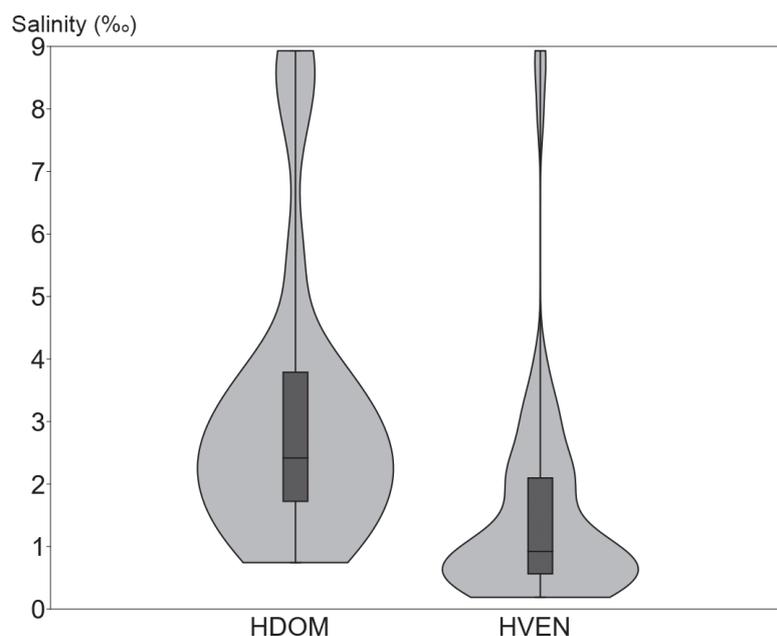

Figure 9. Violin plot of salinity (‰) in those samples where *H. dominici* (HDOM) and *H. veneta* (HVEN) occurred.

Summary

During our study, we have obtained very similar results, both by morphological and DNA analyses for prominent species. However, our results also highlight that, during sequence matching, it is of great importance that sequences are entered into the database library only after a precise morpho-taxonomic investigation of the species. This is also crucial, as there is an increasing pressure to use environmental DNA-based methods in ecological status assessments and if a dominant species is misidentified in the database, it can distort the results of ecological status assessments.

Acknowledgements

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Funding

Author contributions

Conceptualization: ÉÁ. Data curation: MD, OPS. Funding acquisition: TB. Investigation: ZT, PD, AF. Project administration: EV. Resources: EB. Visualization: PO. Writing - original draft: ÉÁ, ZL, KTK, MD. Writing - review and editing: IG.

Author ORCIDs

Tibor Bíró https://orcid.org/0000-0002-5553-9850
Orsolya Pálné Szén https://orcid.org/0000-0002-6610-3681

Data availability

All of the data that support the findings of this study are available in the main text or Supplementary Information.

References

Ács É, Földi A, Wetzel CE, Vad CF, Kiss KT, Dobosy P, Trábert Z, Grigorszky I, Engloner A, Ector L (2017) *Nitzschia austriaca* Hustedt: A characteristic diatom of Hungarian

inland saline waters including a morphological comparison with the type material. Phytotaxa 308(1): 54–65. https://doi.org/10.11646/phytotaxa.308.1.4

Ács É, Földi A, Kiss KT, Buczkó K (2019) The first Hungarian record of a rare *Gomphonema* (Bacillariophyta) species from temporary ponds. Studia Botanica Hungarica 50(1): 19–27. https://doi.org/10.17110/StudBot.2019.50.1.19

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

Bailet B, Bouchez A, Franc A, Frigerio J-M, Keck F, Karjalainen S-M, Rimet F, Schneider S, Kahlert M (2019) Molecular versus morphological data for benthic diatoms biomonitoring in Northern Europe freshwater and consequences for ecological status. Metabarcoding and Metagenomics 3: e34002. https://doi.org/10.3897/mbmg.3.34002

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

Boros E, Ecsedi Z, Oláh J (2013) Ecology and management of soda pans in the Carpathian Basin. Hortobágy Environmental Association, Balmazújváros, 551 pp.

Boros E, Horváth Z, Wolfram G, Vörös L (2014) Salinity and ionic composition of the shallow astatic soda pans in the Carpathian Basin. Annales de Limnologie - International. Annales de Limnologie 50(1): 59–69. https://doi.org/10.1051/limn/2013068

Boros E, Katalin V-B, Vörös L, Horváth Z (2017) Multiple extreme environmental conditions of intermittent soda pans in the Carpathian Basin (Central Europe). Limnologica (Online) 62: 38–46. https://doi.org/10.1016/j.limno.2016.10.003

Borrego-Ramos M, Bécares E, García P, Nistal A, Blanco S (2021) Epiphytic Diatom-Based Biomonitoring in Mediterranean Ponds: Traditional Microscopy versus Metabarcoding Approaches. Water (Basel) 13(10): 1351. https://doi.org/10.3390/w13101351

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13(7): 581–583. https://doi.org/10.1038/nmeth.3869

CEN (2014) EN 13946. Water quality – guidance for the routine sampling and preparation of benthic diatoms from rivers and lakes. European Committee for Standardization, Brussels, Belgium.

Cleve PT (1895) Synopsis of Naviculoid Diatoms. II. Kongliga Svenska-Vetenskaps Akademiens Handlingar 27: 1–219.

Duleba M, Földi A, Micsinai A, Várbíró G, Mohr A, Sipos R, Szabó G, Buczkó K, Trábert Z, Kiss KT, Bíró T, Vadkerti E, Ács É (2021) Applicability of diatom metabarcoding in the ecological status assessment of Hungarian lotic and soda pan habitats. Ecological Indicators 130: 108105. https://doi.org/10.1016/j.ecolind.2021.108105

Ector L, Charles D, Denys L, Manoylov K, Michelutti N, Rimet F, Smol J, Stevenson R (2004) Workshop report. Freshwater diatoms and their role as ecological indicators, 469–480.

European Commission (2000) Official Journal L 327 (EP, CONSIL) Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. http://data.europa.eu/eli/dir/2000/60/oj/eng

European Committee for Standardisation (2002) EN 13946 Water Quality. Guidance standard for the routine sampling and pretreatment of benthic diatoms from rivers, 18 pp.

Földi A, Ács É, Grigorszky I, Ector L, Wetzel CE, Várbíró G, Kiss KT, Dobosy P, Trábert Z, Borsodi AK, Duleba M (2018) Unexpected consequences of bombing. Community level response of epiphytic diatoms to environmental stress in a saline bomb crater pond area. PLoS ONE 13(10): e0205343. https://doi.org/10.1371/journal.pone.0205343

Håkansson H, Chepurnov V (1999) A study of variation in valve morphology of the diatom cyclo^{TEL}la meneghiniana in monoclonal cultures: Effect of auxospore formation and different salinity conditions. Diatom Research 14(2): 251–272. https://doi.org/10.1080/0269249X.1999.9705469

Hammer UT (1986) Saline Lake Ecosystems of the World. Springer Dordrecht, 616 pp.

Hammer O, Harper D, Ryan P (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis. Palaeontologia Electronica 4: 1–9.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proceedings. Biological Sciences 270(1512): 313–321. https://doi.org/10.1098/rspb.2002.2218

HSI (1981) MSZ 260-46:1981. Wastewaters analysis. Measurement of the cleaned wastewater. Determination of the volume of the sediment and the activity of the sludge. Hungarian Standards Institution, Budapest.

HSI (1986) MSZ 448-11:1986. Drinking water analysis. Determination of basicity, hydrogen-carbonate ion, carbonate ion and hydroxil ion content. Hungarian Standards Institution, Budapest.

HSI (1993) MSZ ISO 10260:1993. Water quality. Measurement of biochemical parameters. Spectrometric determination of the chlorophyll-a concentration. Hungarian Standards Institution, Budapest.

HSI (2006) MSZ 1484-3:2006. Testing of waters. Part 3: Determination of dissolved, suspended and total metals in water by AAS and ICP-OES. Hungarian Standards Institution, Budapest.

HSI (2009) MSZ EN ISO 11885:2009. Water quality. Determination of selected elements by inductively coupled plasma optical emission spectrometry (ICP-OES) (ISO 11885:2007). Hungarian Standards Institution, Budapest.

Keck F, Rimet F, Vasselon V, Bouchez A (2019) A ready-to-use database for DADA2: Diat.barcode_rbcL_312bp_DADA2. https://doi.org/10.15454/HNI1EK

Kelly M (2019) Adapting the (fast-moving) world of molecular ecology to the (slow-moving) world of environmental regulation: Lessons from the UK diatom metabarcoding exercise. Metabarcoding and Metagenomics 3: e39041. https://doi.org/10.3897/mbmg.3.39041

Kociolek JP, Laslandes B, Bennett D, Thomas E, Brady M, Graeff C (2014) Diatoms of the United States, 1 Taxonomy, ultrastructure and descriptions of new species and other rarely reported taxa from lake sediments in the western U.S.A. J. Cramer in der Gebrüder Borntraeger Verlagsbuchhandlung, Stuttgart, 188 pp.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution 33(7): 1870–1874. https://doi.org/10.1093/molbev/msw054

Lange-Bertalot H, Hofmann G, Werum M, Cantonati M (2017) Freshwater benthic diatoms of Central Europe: over 800 common species used in ecological assessment. Koeltz Botanical Books, Schmitten-Oberreifenberg, Germany, 942 pp.

Lecointe C, Coste M, Prygiel J (1993) "Omnidia": Software for taxonomy, calculation of diatom indices and inventories management. Hydrobiologia 269(1): 509–513. https://doi.org/10.1007/BF00028048

Leterme SC, Prime E, Mitchell J, Brown MH, Ellis AV (2013) Diatom adaptability to environmental change: A case study of two Cocconeis species from high-salinity areas. Diatom Research 28(1): 29–35. https://doi.org/10.1080/0269249X.2012.734530

Levkov Z (2009) *Amphora* sensu lato. A.R.G. Gantner Verlag K.G, Ruggell, 916 pp.

Lobo EA, Heinrich CG, Schuch M, Wetzel CE, Ector L (2016) Diatoms as Bioindicators in Rivers. In: Necchi JR O (Ed.) River Algae. Springer International Publishing, Cham, 245–271. https://doi.org/10.1007/978-3-319-31984-1_11

Menzel DW, Corwin N (1965) The Measurement of Total Phosphorus in Seawater Based on the Liberation of Organically Bound Fractions by Persulfate OXIDATION1. Limnology and Oceanography 10(2): 280–282. https://doi.org/10.4319/lo.1965.10.2.0280

Mora D, Abarca N, Proft S, Grau JH, Enke N, Carmona J, Skibbe O, Jahn R, Zimmermann J (2019) Morphology and metabarcoding: A test with stream diatoms from Mexico highlights the complementarity of identification methods. Freshwater Science 38(3): 448–464. https://doi.org/10.1086/704827

Mortágua A, Vasselon V, Oliveira R, Elias C, Chardon C, Bouchez A, Rimet F, João Feio M, Almeida S ((2019) Applicability of DNA metabarcoding approach in the bioassessment of Portuguese rivers using diatoms. Ecological Indicators 106: 105470. https://doi.org/10.1016/j.ecolind.2019.105470

Murphy J, Riley JP (1962) A modified single solution method for the determination of phosphate in natural waters. Analytica Chimica Acta 27: 31–36. https://doi.org/10.1016/S0003-2670(00)88444-5

Pérez-Burillo J, Trobajo R, Vasselon V, Rimet F, Bouchez A, Mann DG (2020) Evaluation and sensitivity analysis of diatom DNA metabarcoding for WFD bioassessment of Mediterranean rivers. The Science of the Total Environment 727: 138445. https://doi.org/10.1016/j.scitotenv.2020.138445

Rimet F, Bouchez A (2012) Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. Knowledge and Management of Aquatic Ecosystems 01(406): 01. https://doi.org/10.1051/kmae/2012018

Rimet F, Chonova T, Gassiole G, Gusev E, Kahlert M, Keck F, Kelly M, Kochoska H, Kulikovskiy M, Levkov Z, Maltsev Y, Mann D, Pfannkuchen M, Trobajo R, Vasselon V, Wetzel C, Zimmermann J, Bouchez A (2019) Diat.barcode, an open-access barcode library for diatoms. Scientific Reports 9(1): 15116. https://doi.org/10.1038/s41598-019-51500-6

Rivera SF, Vasselon V, Jacquet S, Bouchez A, Ariztegui D, Rimet F (2018) Metabarcoding of lake benthic diatoms: From structure assemblages to ecological assessment. Hydrobiologia 807(1): 37–51. https://doi.org/10.1007/s10750-017-3381-2

Round FE, Crawford RM, Mann DG (1990) Diatoms: Biology and Morphology of the Genera. 1st edn. Cambridge University Press, Cambridge, UK.

Stenger-Kovács C, Lengyel E (2015) Taxonomical and distribution guide of diatoms in soda pans of Central Europe. Studia Botanica Hungarica 46(Suppl.): 3–203. https://doi.org/10.17110/StudBot.2015.46.Suppl.3

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 (2018) *Amphora* and *Halamphora* from coastal and inland waters of the United States and Japan. J. Cramer, in Borntraeger Science Publishers, Stuttgart, 260 pp.

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

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

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22(22): 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Trobajo R, Rovira L, Mann DG, Cox EJ (2011) Effects of salinity on growth and on valve morphology of five estuarine diatoms. Phycological Research 59(2): 83–90. https://doi.org/10.1111/j.1440-1835.2010.00603.x

Tuchman ML, Theriot E, Stoermer EF (1984) Effects of low level salinity concentrations on the growth of *Cyclotella meneghiniana* Kütz. (Bacillariophyta). Archiv fur Protistenkunde 128(4): 319–326. https://doi.org/10.1016/S0003-9365(84)80003-2

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

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

Vasselon V, Bouchez A, Rimet F, Jacquet S, Trobajo R, Corniquel M, Tapolczai K, Domaizon I (2018) Avoiding quantification bias in metabarcoding: Application of a cell biovolume correction factor in diatom molecular biomonitoring. Methods in Ecology and Evolution 9(4): 1060–1069. https://doi.org/10.1111/2041-210X.12960

Vautier M, Vasselon V, Chardon C, Rimet F, Bouchez A, Domaizon I (2020) DNA extraction from environmental biofilm using the NucleoSpin® Soil kit (MACHEREY-NAGEL). https://doi.org/10.17504/protocols.io.bd52i88e

Supplementary material 1

Relative abundances of *Halamphora* species based on microscopy and metabarcodin

Authors: Éva Ács, Tibor Bíró, Emil Boros, Péter Dobosy, Mónika Duleba, Angéla Földi, Keve T. Kiss, Zlatko Levkov, Péter Orgoványi, Orsolya Pálné Szén, Zsuzsa Trábert, Edit Vadkerti, István Grigorszky

Data type: xlsx

Explanation note: Relative abundances of *Halamphora* species, based on microscopy and metabarcoding (this latter is indicated with "DNA" after the code of the species), as well as environmental variables measured at each sampling location and time. (See abbreviation of the species code in the text). Relative abundances indicating dominancy (> 5%) are written in bold. Abbreviations of physical-chemical variables: Cond = conductivity; DO = dissolved oxygen concentration; CHA = chlorophyll a concentration; Secchi = Secchi transparency; TN = total nitrogen concentration; TP = total phosphorus concentration. "<" stands for under detection limit. For more details, see Material and methods.

Link: https://doi.org/10.3897/mbmg.7.111679.suppl1

Supplementary material 2

Pairwise p-distance values and number of differences

Authors: Éva Ács, Tibor Bíró, Emil Boros, Péter Dobosy, Mónika Duleba, Angéla Földi, Keve T. Kiss, Zlatko Levkov, Péter Orgoványi, Orsolya Pálné Szén, Zsuzsa Trábert, Edit Vadkerti, István Grigorszky

Data type: xlsx

Explanation note: Pairwise p-distance values and number of differences between *Halamphora* amplicon sequence variants found in soda pan and soda lake samples and sequences of *Halamphora* taxa from Diat.barcode database.

Link: https://doi.org/10.3897/mbmg.7.111679.suppl2

Supplementary material 3

Supplementary Alignment 1

Authors: Éva Ács, Tibor Bíró, Emil Boros, Péter Dobosy, Mónika Duleba, Angéla Földi, Keve T. Kiss, Zlatko Levkov, Péter Orgoványi, Orsolya Pálné Szén, Zsuzsa Trábert, Edit Vadkerti, István Grigorszky

Data type: fas

Explanation note: Alignment of *Halamphora* amplicon sequence variants (ASVs) found in soda pans and soda lakes as well as *Halamphora* sequences from Diat.barcode database for Fig. 4.

Link: https://doi.org/10.3897/mbmg.7.111679.suppl3

Supplementary material 4

Supplementary Alignment 2

Authors: Éva Ács, Tibor Bíró, Emil Boros, Péter Dobosy, Mónika Duleba, Angéla Földi, Keve T. Kiss, Zlatko Levkov, Péter Orgoványi, Orsolya Pálné Szén, Zsuzsa Trábert, Edit Vadkerti, István Grigorszky

Data type: fas

Explanation note: Alignment of long *rbc*L sequences of *Halamphora dominici* and *H. veneta* acquired in this study and *Halamphora* taxa from the work by Stepanek and Kociolek (2019) for Fig. 5.

Link: https://doi.org/10.3897/mbmg.7.111679.suppl4