This manuscript is contextually identical with the following published paper:

Mónika Duleba , Keve T. Kiss , Angéla Földi , József Kovács , Koraljka Kralj Borojević , Levente F. Molnár , Andjelka Plenković-Moraj , Zsuzsanna Pohner , Cüneyt Nadir Solak , Bence Tóth , Éva Ács (2015) Morphological and genetic variability of assemblages of Cyclotella ocellata Pantocsek/C. comensis Grunow complex (Bacillariophyta, Thalassiosirales). Diatom Research
Vol. 30, Iss. 4, pp. 283-306. **DOI:** 10.1080/0269249X.2015.1101402

The original published pdf available in this website:
http://www.tandfonline.com/doi/full/10.1080/0269249X.2015.1101402

Morphological and genetic variability of assemblages of *Cyclotella ocellata* Pantocsek/*C. comensis* Grunow complex (Bacillariophyta, Thalassiosirales)

MÓNIKA DULEBA[1], KEVE T. KISS[1], ANGÉLA FÖLDI[2], JÓZSEF KOVÁCS[3], KORALJKA KRALJ BOROJEVIĆ[4], LEVENTE F. MOLNÁR[2], ANDJELKA PLENKOVIĆ-MORAJ[4], ZSUZSANNA POHNER[5], CÜNEYT NADIR SOLAK[6,7], BENCE TÓTH[1] & ÉVA ÁCS[1*]

[1]MTA *Centre for Ecological Research, Danube Research Institute, H-1113 Budapest, Karolina út 29, Hungary*

[2]*Doctoral School of Environmental Sciences, Eötvös Loránd University, H-1117 Budapest Pázmány Péter sétány 1/A, Hungary*

[3]*Institute of Geography and Earth Sciences, Eötvös Loránd University, H-1117 Budapest Pázmány Péter sétány 1/C, Hungary*

[4]*Division of Biology, Faculty of Science, University of Zagreb, Rooseveltov trg 6, HR-10000 Zagreb, Croatia*

[5]*Department of Microbiology, Institute of Biology, Eötvös Loránd University, H-1117 Budapest Pázmány Péter sétány 1/C, Hungary*

[6]*Department of Biology, Arts and Science Faculty, Dumlupınar University, Kütahya, 43100, Turkey*

[7]*Department of Palaeoceanology, Institute of Marine Sciences, University of Szczecin, Mickiewicza 18, Szczecin, 70-383, Poland*

[*]Corresponding author.Email: acs.eva@okologia.mta.hu

Abstract

Centric diatom taxa in the *Cyclotella ocellata* and *C. comensis* complexes show high morphological variability and often apparently continuous morphological transitions. In this study we investigated natural assemblages of the *C. ocellata/C. comensis* complex from Hungarian and Croatian lakes and from Turkish streams using morphological and molecular methods. The studied assemblages contained cells with morphologies resembling *C. ocellata* as well as other, closely related, species: *C. comensis*, *C. pseudocomensis*, *C. costei*, and *C. trichnoidea*. The goal of our paper was to assess whether the observed morphological differences were due to intraspecific variability or suggest the existence of several, putatively distinct species.

Ten morphometric characters were measured, which, either individually, or in pairs, did not differentiate the nominal taxa in our assemblages. However, multivariate discriminant analysis has revealed a group including *C. ocellata* and *C. trichonidea* morphologies could be separated from another containing *C. comensis*, *C. pseudocomensis* and *C. costei*.

A nuclear (18S rDNA) and a chloroplast (*rbc*L) gene were amplified and partially sequenced from environmental DNA or from isolated cells. The sequences showed little variability among the assemblages and nominal species. Although general congruence of molecular and morphometric separation supports the species level separation of *C. ocellata/trichonidea* from the probably conspecific *C. comensis*/*pseudocomensis*/*costei*, sequence divergences between the groups are in the same range as within them, so that a onspecificity of all four taxa cannot be unequivocally excluded.

Key words: Cyclotella ocellata*,* Cyclotella comensis, *European and Turkish assemblages, morphometric analysis, 18S rDNA,* rbc*L*

Running title: Variability of *Cyclotella ocellata/C. comensis* complex

Introduction

Centric diatom species constitute a remarkable proportion of phytoplankton communities (e.g. in Lake Geneva, Rimet 2013). Resolution of micromorphological characteristics by electron microscopy has been the basis on which a large number of new taxa in this group have been described. High taxonomic resolution is needed for precise environmental bioindication (Hausmann & Lotter 2001, Rimet & Bouchez 2012), especially in the area of water quality assessment (for example in the implementation of the EU Water Framework Directive), and for reconstructing past environmental conditions (Birks 1994). However, the often questionable taxonomic value of minor morphological variations used in the description of new taxa has created great uncertainty in the delimitation of some groups of centric diatom taxa.

Valve morphology in centric diatoms is often more variable than stated in the original description of the species. Teubner (1995), Kiss et al. (1996, 1999) and Hegewald &Hindákova (1997) presented continuous transitions between forms recognizable as separate taxa in the genus *Cyclotella* (Kützing) Brébisson. The presence of heterovalvate cells with valves that could be identified as separate species (Teubner 1995, Hegewald & Hindákova 1997) represents another type of morphological variability spanning boundaries between hypothesized taxa.

The centric diatom *Cyclotella ocellata* was first described by Pantocsek (1901) from Lake Balaton. Since then it has been recorded at several sites worldwide, mainly in lakes (e.g. Lake Las Madres, Spain, Kiss et al. 1996) but also in rivers (e.g. River Danube, Hungary, Kiss et al. 2012), and springs and small streams (e.g.Türkmen Mountain, Turkey, Solak & Kulikovskiy 2013). Moreover, it has been detected as a fossil in sediments (e.g. Pleistocene-Holocene sediments of Lake El'gygytgyn, Russia, Cherepanova et al. 2010).

Cyclotella ocellata shows great morphological variability, as demonstrated by a reinvestigation of its type material (Kiss et al. 1999), but also by observations from other localities including the eutrophic Lake Dagow in Germany (Schlegel & Scheffler 1999), Lake Las Madres in Spain (Kiss et al. 1996), Rapel reservoir in Chile (Rivera et al. 2003), Lake Khubsulug in Mongolia (Genkal & Popovskaya 2008), Lake El'gygytgyn in Northeast Siberia (Cremer et al. 2005), and Gallberg-pons in NW Germany (Hegewald & Hindáková 1997). Due to its great morphological variability and the occurrence of intermediate forms with other species, *C. ocellata* is considered part of a species-complex with its close congeners, *C. krammeri* Håkansson, *C. rossii* Håkansson, *C. tripartita* Håkansson, *C. kuetzingiana* Thwaites*, C. polymorpha* B. Meyer & Håkansson and *C. comensis* Grunow (Edlund et al. 2003, Cherepanova et al. 2010). The delimitation of other taxa within the *C. ocellata* group was also found to be problematic (e.g. Wunsam et al. 1995, Knie & Hübener 2007, Genkal & Popovskaya 2008). Teubner (1995) showed the lack of clear distinction between *C. krammeri*, *C. kuetzingiana* var. *planetophora* Fricke, *C. kuetzingiana* var. *radiosa* Fricke, *C. ocellata* and *C. comensis*. *Cyclotella trichonidea* Economou-Amilli also resembles *C. ocellata,* the most conspicuous distinguishing feature being the radial undulation of its valves (Economou-Amilli 1982) creating an angular outline in valve view. This species was first described from Lake Trichonis in Greece (Economou-Amilli 979), but, beside the type locality, it has only been recorded from a few other sites (e.g. Lake Amvrakia in Greece, Danielidis et al. 1996).

Cyclotella comensis was first described from Lake Como (Hustedt 1930), has a wide geographic distribution (e.g in Europe, Wunsam et al. 2005, in North America, Werner & Smol 2006, in Asia, Alfasane et al. 2013). It is considered an indicator of oligotrophic to mesotrophic conditions (Wunsam et al. 1995) and is also significant for paleolimnological investigations (e.g. Hausmann & Lotter 2001, Wolin & Stoermer 2005). Wunsam et al. (1995) distinguished four morphs in the surface sediment of Alpine lakes, while Hausmann & Lotter

(2001) classified six morphs differing mainly in the length of the striae, based on principal component analysis of sediment surface samples from the Swiss Alps. The latter also found a relationship between the morphs and mean summer air temperature. Scheffler & Morabito (2003) described only one morph, *C. comensis* morphotype *minima* in addition to the typical form.

Cyclotella pseudocomensis Scheffler was described from the Großer Boberowsee (Scheffler 1994) and later the co-occurrence of two distinct morphotypes of *C. pseudocomensis* were observed, even in unialgal cultures, accompanied by a lack of variation in ITS-2 sequences, as well as a morphological transition between both morphs during auxosporulation (Scheffler et al. 2003). Based on morphological features of specimens from Lake Como, Scheffler & Morabito (2003) considered *C. pseudocomensis* a synonym of *C. comensis* and its morph (*C. pseudocomensis* morphotype *minima* Scheffler et al. (2003) asynonym of *C. comensis* morphotype *minima*. Later, Scheffler et al. (2005) concluded that *C. comensis* was a dimorphic species, comprising the highly variable morph *comensis* and the morph *minima,* which shows slight variability in the shape and structure of the central area. Werner & Smol (2006) distinguished three additional morphs ('*rossi*', '*socialis*', 'fine'). *Cyclotella costei* was described by Druart & Straub (1988) from Lake Paladru, although Scheffler & Morabito (2003) treated it as a distinct taxon, they considered it as a morph of *C. comensis*.

Using molecular data, Alverson et al. (2007) produced a phylogenetic analysis of order Thalassiosirales, but only included *C. ocellata* from the *C. ocellata/C. comensis* complex. In their analysis, *Cyclotella* was found to be non-monophyletic, *C. ocellata* and *C. bodanica* Grunow diverged from other *Cyclotella* species (*C. meneghiniana* Kützing and related species). Kistenich et al. (2014) investigated the *C. comensis* group using both morphological and molecular methods. Their results indicated that *C. comensis* was indistinguishable from

C. pseudocomensis and *C. costei,* supporting the results of Scheffler & Morabito (2003). Moreover, they used *C. ocellata* as a reference taxon, which seemed to be closely related to the *C. comensis* group. Scheffler & Morabito (2003) also found similarities between the valves of *C. comensis* and *C. ocellata*.

In the present study, natural assemblages of the *C. ocellata/C. comensis* complex from different Hungarian, Croatian and Turkish freshwater sites were studied using morphological and molecular methods. We investigated whether morphological / morphometric differences between assemblages and morphs were greater than differences within these, and whether such differences, if present, were accompanied by molecular differences. The goal of our paper is to contribute to the clarification of the morphologically and genetically distinct species in the *C. ocellata/C. comensis* complex.

Materials and methods

Sampling

Phytoplankton samples were collected from below the water surface at sampling sites shown on the map, prepared using the ESRI ArcInfo 9.3 GIS program (Fig.1). Detailed site descriptions are given in Table 1.

Conductivity (Cond.), pH, dissolved oxygen (DO), and water temperature (T) were measured *in situ* with a WTW multiline portable device. Turbidity was also measured *in situ* with Lovibond PC Checkit®. Ammonium (NH_4^+) was measured according to ISO 7150–1:1984. Nitrate (NO_3^-) was measured by the sodium salicylate method (Vijayasarathy 2011). Orthophosphate (SRP) and total phosphorus (TP) were measured according to Eaton et al. (2005), by the ammonium molybdate method, and the manual digestion method, respectively. Carbonate (CO_3^{2-}) and hydrocarbonate (HCO_3^-) ions were determined by volumetric titration, using a 0.1 M HCl solution. Calcium (Ca^{2+}) and magnesium (Mg^{2+}) ions were determined by

complexometric titration, using a 0.01 M EDTA solution. Chlorophyll-*a* concentration (chl-*a*) was determined spectrophotometrically, after extraction with hot methanol (Iwamura et al. 1970). The measured water chemical parameters are presented in Table 2.

Morphological analysis

Samples were allowed to settle to concentrate the frustules, which were then cleaned with hydrochloric acid and hydrogen peroxide, and subsequently washed in distilled water, filtered onto a 3 μm-mesh polycarbonate membrane, fixed on SEM stubs, and coated with gold-palladium. The prepared samples were investigated with a Hitachi S-2600N scanning electron microscope. For detailed morphological analyses of a given assemblage, SEM micrographs were taken of the outside and inside views of 15–30 randomly selected valves.

The following morphological features of the valves were analysed and measured on the micrographs (Fig. 2): valve diameter (d), diameter of the central area (dCA), number of orbiculi depressi (OD), number of puncta (puncta), number of striae on the edge of the valve face (stria), number of divided striae (div stria), number of marginal fultoportulae (MFP), number of radially symmetric valve sectors (ns), number of central fultoportulae (CFP), number of rimoportulae (RP), orientation of the rimoportula labium (RP ang). The diameter of the central area was divided by the cell diameter (dCA/d). Assuming valves and central area to be circular the number of striae, number of divided striae and number of MFP were referred to 10 μm valve perimeter (stria/10, div stria/10, MFP/10) , the numbers of OD and puncta were calculated to 10 μm^2 central area (OD/10CA, puncta/10CA).

The structural elements of the valves were measured and analysed according to the methods of Genkal (1977, 1984), taking into account the work of Theriot (1987). The terminology follows that of Anonymous (1975), Ross et al. (1979) and Theriot & Serieyssol (1994).

Statistical analysis

For statistical analysis, d, ns and derived variables (dCA/d, stria/10, div stria/10, MFP/10, OD/10CA, puncta/10CA) were used. To reveal bivariate relationships between morphotypes a scatter plot matrix was prepared in R. To find out if morphs were separate groups based on the combination of variables, canonical variates analysis (CVA) was performed using PAST 2.16 (Hammer et al. 2001).

Molecular methods

Samples from gravel-pit lakes at Dunaharaszti, Nyékládháza, Szany, Hegyeshalom, Himód, Kunsziget and Lake Visovac were included in the molecular analysis. The samples were either used fresh or preserved in absolute ethanol (Reanal) for DNA analysis.

Environmental DNA was extracted from samples in which the presence of only one of the investigated species was determined by scanning electron microscopy. The sample was mixed with a lysis buffer (10 mM Tris, 1 mM Na_2-EDTA, 200 mM NaCl, 0.02 % SDS, pH=8), frozen at –20°C for 30 min, heated at 95°C for 15 min, then treated with glass beads using a cell mill. Proteins were digested using Proteinase K (recombinant, Fermentas) at 56°C for 3 h. DNA was then isolated using the DNeasy® Plant Mini Kit (Qiagen).

From samples in which more than one species of interest was observed in SEM (the gravel-pit lake at Nyékládháza and Visovac lake), individual cells were isolated using a micropipette under an inverted microscope (Olympus IX70) and transferred into 100 µL sterile TE buffer (10mM Tris, 1 mM Na_2-EDTA, pH=8) or 2X Phire Plant PCR Buffer (Phire Plant PCR Kit, Thermo Scientific).Cell isolation from the Nyékládháza sample resulted in two samples: sample I included cells with visible ODs, while sample II consisted of small

cells without visible ODs. DNA was extracted using heat denaturation (95°C 15 min), protein digestion with Proteinase K (recombinant, Fermentas), and centrifugation (20000 g 5 min).

Polymerase chain reactions were performed on the18S rRNA and *rbc*L genes using the primers listed in Table 3. The primers had been designed for an earlier study (Duleba et al. 2014), and their specificity was checked with Primer-BLAST (Ye et al. 2012) and in in PCRs on isolated cells and community DNA from the River Danube. Primers which proved to be specific mainly to the Thalassiosirales were used in present study. Mismatch between the primer and a given sequence on the 3′ end region (especially at the third site) of the primer were considered to prevent its annealing. Additionally, Medlin-A was applied which had been designed for eukaryotes by Medlin et al. (1988). Because of the low number of cells and the low concentration of extracted DNA, the polymerase chain reaction (PCR) was performed in semi-nested design. The PCR mixture contained the following components in the total volume of 25 µl: 1.25 U DreamTaq™ DNA Polymerase (Thermo Scientific), 200 mM of each deoxinukleoside-triphosphate (Fermentas), 1X DreamTaq Buffer (Thermo Scientific), 0.325 µM of each primer, and 1–3 µl of template. In the cases of Nyékládháza II and Visovac, a Phire® Plant Direct PCR Kit (Thermo Scientific) was used according to the manufacturer's instructions: a reaction mix containing 0.4 µl Phire® Hot Start II DNA Polymerase, 1X Phire® Plant PCR Buffer, and 0.5 µM of each primer was added to the isolated cells.

The PCR amplifications with DreamTaq DNA Polymerase used the following heat protocol: initial denaturation at 98°C for 5 min, 32 cycles at 94°C for 1 min, 52–60°C for 30 sec, 72°C for 1–1.5 min, and a final extension at 72°C for 10 min. The Phire Kit PCR heat protocol was: initial denaturation at 98°C for 5 min, 40 cycles at 98°C for 5 sec, 60°C for 5 sec, 72°C for 30 sec, and a final extension at 72°C for 1 min. The exact annealing temperature for each primer pair was determined on community DNA (from River Danube) by gradient PCR: it was 60°C for the MedlinA-Sk-1550R and Sk-900f-Sk-1550R primer pairs, 52°C for

the Sk-600F-Sk-1550R primer pair, 55°C for the rbcL66F-Sk-rbcL975R and Sk-rbcL400F-Sk-rbcL975R primer pairs. The duration of the extension step in the cycles was chosen according to the expected length of the PCR product. The sequencing reactions and capillary electrophoreses were performed by Biomi Ltd. For 18S rDNA Sk-600F, Sk-900F and Sk-1550R, for rbcL Sk-rbcL400F and Sk-rbcL975R were used as sequencing primers (Table 3). All nucleotide sequences are available from the DDBJ/EMBL/GenBank databases under the accession numbers KJ755337- KJ755348 (listed in Table 6).

Phylogenetic analysis

The final 18S rDNA sequences were assembled from overlapping sequence fragments. Assembled 18S rDNA and *rbc*L sequences were aligned to sequences in NCBI GenBank using BLAST (Altschul et al. 1990) to find sequences with the highest similarity. Sequences were compared to those made available by Kistenich et al. (2014): *C. comensis* from Baggersee (Austria), *C. pseudocomensis* from Haussee (Germany), *C. ocellata* from Kiesgrube-Krugsdorf (Germany), *C. costei* from Gültzsee (Germany) and Fernsteinsee (Austria). Parismony networks to illustrate the distance between sequence types were drawn manually. Uncorrected p-distance values were calculated manually and with MEGA 6 (Tamura et al. 2013). The sequences were aligned to sequences downloaded from GenBank using Clustal W implemented in MEGA 6 (Tamura et al. 2013). The aligned sequences were trimmed at both ends to the same length. The 'Find best DNA models' option in MEGA 6 was used to determinate the most appropriate substitution model for the DNA sequence evolution of each gene. Maximum likelihood phylogenetic trees were constructed in MEGA 6 (Tamura et al. 2013) using the suggested substitution models: the Hasegawa-Kishino-Yanomodel (Hasegawa et al. 1985) with gamma distribution for 18S rDNA and the Tamura-Nei model (Tamura and Nei 1993) with gamma distribution for *rbc*L. Phylogenies were tested

with the boostrap method in 500 replications. Bayesian analyses were run on datasets in combination. Posterior probability of distribution was estimated using Metropolis-coupled Markov Chain Monte Carlo (MCMC) as implemented in MrBayes 3.2 (Ronquist et al. 2012). Default priors were used, 400000 generations were used and the first 25 % of samples were discarded as burn-in. The convergence diagnostic used was the average standard deviation of split frequencies across independent analyses.

Results

Morphological and morphometrical investigation

Based on qualitative and quantitative characteristics five morphs ('*ocellata*', '*trichonidea*', '*comensis*', '*pseudocomensis*' and '*costei*') could be distinguished (for quantitative characteristics of the morphotypes, see Table 4). Transitional forms between them constituted a sixth group.

The morphs could not be distinguished in internal view. Each valve had a single (sometimes 2-3) rimoportula, internally sessile with a round external aperture, situated on the valve face on an interstria; the orientation of the lip varied. The inner apertures of the alveoli were round or elongated. Costae were usually equal in length, but those bearing a fultoportula were often shorter.As the classical form of the species, frustules of the '*ocellata*' morph (e.g. Figs 3–12) are disc-shaped, mainly solitary, although short chains were occasionally found. The valve face has up to six orbiculi depressi (OD), complementary to up to six large papillae. Papillae are often broken off, leaving only the papilla postament visible (Fig. 15). The central part of the valve has a colliculate surface and its diameter varies between 20 - 80 % of the valve diameter, Figs 3–11). Besides the OD, there can be relatively small puncta on the external valve face. Among them the external opening of the single (occasionally 2-4) central fultoportula (CFP) is difficult to observe. The CFP is usually surrounded internally by two (1–3) satellite

pores. Most valves are radially undulate, but a slight tangential undulation of the central area can be seen on quadradiate and triradiate valves in a tilted position (Figs 6, 8, 10), resulting in 3–5 axes of symmetry. The marginal part of the valve has 14 to 24 striae in 10 μm. Intervening interstriae differ in length and a few can be bifurcated. Small granules are observed on the interstriae near the margin on most valves, and also sporadically on the striae. Usually, every third to fifth (range: each to every sixth) interstria bears a marginal fultoportula (MFP) with two satellite pores.

The '*trichonidea*' morph is identical to '*ocellata*' except that its frustules are quadrangular (Figs 101, 104). The central area of the 'comensis' morphotype (e.g. Figs 87–89, 92), always has a slight tangential undulation. The valve face has no OD, but a strongly colliculate central area with small puncta. The external opening of the single CFP is usually visible among the puncta. The '*pseudocomensis*' morph (Fig. 53) is identical to the '*comensis*' morph except in the following: the central area shows slight tangential undulation (less distinct than the '*comensis*' morph): the colliculate pattern of the central area is less pronounced than in the '*comensis*' morph: the valve face has no OD, but several smaller or larger puncta (much numerous than the '*comensis*' morph). The 'costei' morph (e.g Figs 44-48) differs from the latter two by having small granules on the valve face. The central area is never tangentially undulate and the marginal part of the valve has 16 to 24 striae in 10 μm.

We observed substantial frustule morphological variability in the *C. ocellata/C. comensis* assemblages investigated, with a mixture of different morphs occurring at a single location in several cases. In certain lakes a 'morph shift' was observed, e.g. in 2012 the Dunaharaszti gravel-pit lake contained mainly the '*costei*' morph, while the '*ocellata*' mophotype was dominant in 2013. The change from 'more adorned' to smooth valve surfaces are shown (Figs 3–11, 15–22, 27–35, 39–48, 51–59, 63–72, 75–85, 87–96) and a summary of the range of morphometric featuers can be found in Supplementary material Table 1.

The Balaton, Dunaharaszti-2013, Yalova and Inli assemblages contained mainly the '*ocellata*' morph (Figs 4–15, 27–34, 103–111, Supplementary Material Table 2). This morph was also dominant in Himód (Figs 15–16), although this sample included a few transitional (Fig. 17) and some 'costei' (Figs 18–22) forms. The '*comensis*' morph was the most abundant in Kunsziget and Hegyeshalom (75 and 57% of the studied valves, respectively, Figs 63–68, 75–82), while the '*costei*' morph constituted 75% of the observed specimens in Dunaharaszti-2012 (Figs 44–48). The '*trichonidea*' form occurred only in Visovac (Figs 101, 104), the '*pseudocomensis*' form was recorded in Nyékládháza (Fig. 53) and Szany. Nyékládháza had the most diverse assemblage, containing five of the six morphs (Figs 51–62).

Of the studied morphometric features, only the number of orbiculi depressi (OD) displayed any differences between the studied assemblages (Supplementary Table 1). An OD was lacking in all members of the Dunaharaszti -2012, Kunsziget, Szany and Hegyeshalom assemblages. Although the presence of an OD is considered a characteristic feature of *C. ocellata*, we also found numerous individuals lacking ODs in the Himód, Nyékládháza and Visovac assemblages (Figs17–22, 53–59,102–103).

Pairwise combinations of variables did not show any separation of the morphs (see scatter plot matrix in Supplementary Material Fig. 1). However, a CVA performed to test whether a combination of all measured variables can distinguish the five morphs revealed two main groups (Fig. 111.). One of them included predominantly '*comensis*', '*pseudocomensis*', '*costei*' and transitional forms (*C*. cf *comensis* group), the other was mainly constituted of '*ocellata*' and '*trichonidea*' morphs (*C*. cf *ocellata* group). Within the first group, most valves of the '*costei*' morph showed slight differentiation from the other forms. Density of puncta (Puncta/10CA) was the most determinant parameter, with striae (div stria/10) and OD density (OD/10 CA) also contributing to the separation of the *C*. cf *ocellata* and *C*. cf *comensis* groups. Diameter of the central area relative to that of the valve (dCA/d) and density of MFPs

(MFP/10) contributed little to the separation. The specimens from Balaton, Dunaharaszti-2013, Yalova and Inli mainly included the '*ocellata*' morph based on our initial morphological classification, all placed in the *C*. cf *ocellata* group in the CVA. There were two exceptions in the Dunaharaszti-2013 assemblage: a '*costei*' and and an '*ocellata*' form. Valves of Dunaharaszti-2012, Kunsziget, Szany and Hegyeshalom belonged predominantly to the *C*. cf *comensis* group (except one '*ocellata*' form from Dunaharaszti 2012, a single '*comensis*' from both Kunsziget and Hegyeshalom, and a '*costei*' and a '*pseudocomensis*' from Szany). Himód, Visovac and Nyékládháza specimens were split between the two groups. The '*ocellata*' and '*trichonidea*' forms of these assemblages belonged to the *C*. cf *ocellat*a group (with two exceptions from Himód which fell into the other group), the other forms fell mainly into the *C*. cf *comensis* group (with one exception from Visovac).

Based on the morphological investigation we therefore refer to the Nyékládháza I, Himód, Visovac, Dunaharaszti-2013 samples as *C.* cf *ocellata*, and to the Nyékládháza II, Kunsziget, Szany, Dunaharaszti-2012, Hegyeshalom samples as *C*. cf *comensis*.

Molecular investigation

Sequencing success and primer specificity

We successfully amplified and sequenced both markers (18S rDNA and *rbc*L) from three assemblages; only 18S rDNA from four, and only *rbc*L from two assemblages (for details, see Table 5). The sample from Visovac contained isolated *C. trichonidea*-like cells and cells of the classic form of *C. ocellata*; but we nevertheless obtained a 876 nt long clean partial 18S rDNA sequence from this sample.

Mixed *rbc*L sequences were acquired from the Dunaharaszti-2013 and Hegyeshalom assemblages, from which the 18S rDNA was also sequenced. This might have been a result of the applied *rbc*L primers not being strictly Thalassiosirales–specific. According to Primer-

BLAST, Sk-rbcL400F and Sk-rbcL975R primers perfectly match a sequence from *Craticula cuspidata* (Kützing) D. G. Mann, and several other non-Thalassiosiralean diatoms, e.g. *Asterionella formosa* Hassall, *Diatoma tenue* C. Agardh with one or just a few non-3′-terminal mismatches. These mixed sequences were discarded from further analyses.Only amplification of 18S rDNA was attempted from Visovac and Nyékládháza II.

The Nyékládháza I sample provided a clean 18S rDNA sequence (Supplementary Material Fig. 5), but its *rbc*L electropherogram contained some small peaks under the dominant peaks read by the sequencer (Supplementary Material Fig. 6). The 18S rDNA sequence from the Nyékládháza II sample (Supplementary Material Fig. 7) also showed small peaks that suggest the amplification of this gene from cells of other morphs/species. The read sequences of these samples were used to discriminate the sequence types; the Nyékládháza I *rbc*L sequence and the Nyékládháza II 18S sequence expressions cover these read sequences.

Genetic divergence

We obtained three partial 18S rDNA sequence groups (Fig. 112) which differ from each other at one or two nucleotide positions (corresponding to up to 0.25 %), and from the previously determined sequence from *C. ocellata* strain LB8 (Alverson et al. 2007) (referred to as type D in Fig. 112) at one to three positions (corresponding to up to 0.34 %, Tables 5, 6). We also compared our *C*. cf *ocellata* and *C*. cf *comensis* sequences with those from Kistenich et al. (2014), matching them to our sequences (Fig. 112). The latter sequence lengths were half the length of ours (403 nt). They differed from our sequences at zero to one positions (0–0.25 %) and at one to two positions (0.25–0.5 %) from *C. ocellata* LB8. Over the entire >800 nt region, there were three variable nucleotide sites of which two were located in the region studied by Kistenich et al. (2014). The most commonly observed sequences combining all our and Kistenich et al. (2014) sequences is group B in Fig. 112, which was obtained from

numerous strains / assemblages with mostly the *comensis*/*pseudocomensis*/*costei* morphology. The only exception is the Visovac sequence, from an assemblage containing only '*ocellata*' and '*trichonidea*' morphologies. When compared to this sequence group, one substitution (transversion) occurred only in the Dunaharaszti-2013, Nyékládháza I, and Kiesgrube-Krugsdorf samples (group A). These sequences were assigned to *C.* cf. *ocellata* based on their morphology (Nyékládháza I contained cells with visible ODs containing all except the '*trichonidea*' morph, corresponding to *C. ocellata*). One substitution (transversion) was unique to group C (Nyékládháza II, cells without ODs isolated from the same assemblage as Nyékládháza I). It should be noted that the divergent site separating groups B and C was outside the region studied by Kistenich et al (2014). Thus their sequences (group B in Fig. 112) could also represent type C. One substitution (transition) was observed in *C. ocellata* LB8 only (group D).

In the case of partial *rbc*L sequences, we obtained three distinct groups of sequences (four with the addition of the sequences from Kistenich et al., 2014) (Fig. 113), differing from each other at one to three positions (i.e., up to 0.66 %). One of the sequence groups (A) was identical to that available from *C. ocellata* strain LB8, as well as the single *C. ocellata rbc*L sequence obtained by Kistenich et al. (2014) (Tables 5, 7), whose sequences were longer than ours (507 nt). Those representing the *C. comensis* group differed at up to two positions from our sequences for this morphological group. The *rbc*L sequences corresponding to the ocellata vs. comensis/pseudocomensis/costei morphologies differ at two (one transition and one transversion were detected only in group A) or three positions, whereas up to two nucleotide differences occur among the sequences within the latter group (one transition was unique to group C; a transition at another site was observed only in group D). All the differences occurred in third codon positions.

Phylogeny

In the phylogenetic tree compiled from the combined data (Fig.114), the *C.* cf. *ocellata* sequences fell into the same group as the Kiesgrube-Krugsdorf (Kistenich et al. 2014) and the LB8 (Alverson et al. 2007) *C. ocellata* sequences. This group appeared as sister group to the *C*. cf *comensis/comensis/pseudocomensis/costei* clade, and with the latter clade as the sister group to *C. bodanica*. The closest relatives of the *C. bodanica/ocellata/comensis* clade were species of *Stephanodiscus* Ehrenberg.

Based on the partial 18S rDNA sequences (Supplementary Material Fig.2), sequences from our assemblages and *C. costei*, *C. comensis* and *C. pseudocomensis* sequences determined by Kistenich et al. (2014) formed a hardly subdivided group together with *C. ocellata* LB8. Within this group, the LB8 strain and our 18S rDNA group A (Dunaharaszti-2013) were slightly separated from the other assemblages (groups B and C, represented by *C.* cf *comensis* Szany and *C.* cf *comensis* Nyékládháza II, respectively). The C. *ocellata*/cf. *ocellata*/cf *comensis* cluster was part of a larger cluster containing *C. bodanica* and all the *Discostella* V. Houk & R. Klee, *Cyclostephanos* Round and *Stephanodiscus* species.

The phylogenetic tree of partial *rbc*L sequences (Supplementary Material Fig.3) similarly showed a clade including all rbcL groups from our assemblages and *C. costei*, *C. comensis* and *C. pseudocomensis* sequences determined by Kistenich et al. (2014), and *C. ocellata* LB8. Within this cluster, groups B, C, D involving *C.* cf *comensis* (Kunsziget, Szany and Dunaharaszti-2012 assemblages) andthe *C. comensis, C. pseudocomensis*, and *C. costei* sequences were separated from group A (*C. ocellata* LB8 determined by Alverson et al., 2006).

Morphometric versus genetic data

Based on morphometric parameters the distinguished morphs constituted two groups, suggesting that the morphs could represent two morphospecies (*C*. cf *ocellata* and *C*. cf *comensis*). In general, sequence data corresponded to this separation. Assemblages representing group A 18S rDNA (Dunaharaszti-2013, Nyékládháza I = cells from Nyékládháza with obvious ODs) and *rbc*L sequence (Nyékládháza I) fell predominantly into the *C*. cf *ocellata* morphospecies. Valves of assemblages with group B and C 18S rDNA (Dunaharaszti-2012, Hegyeshalom, Szany, Nyékládháza = cells from Nyékládháza without ODs) and group B and D *rbc*L (Dunaharaszti-2012, Szany, Kunsziget) mainly belonged to *C*. cf *comensis* morphospecies. Nevertheless, Nyékládháza I and II were problematic because the small peaks that could result from inadequate isolation. There were two exceptions from this general pattern of correspondence between sequence divergence and the *ocellata*/*comensis* morphological split. The Himód assemblage displayed both '*costei*' and '*ocellata*' (along with transitional) forms, but nevertheless had a single, group A, *rbc*L sequence (Fig. 113). The Visovac assemblage consisted of '*ocellata*', '*trichonidea*' and transitional forms, of which the first two had ODs. Cells with ODs (presumably '*ocellata*' and '*trichonidea*' forms) comprised the sample used for DNA analysis. In spite of this, their 18S rDNA sequence was identical with a '*comensis*' sequence (type B; Fig. 112).

Discussion

Morphological and morphometric investigation

The classical form of *C. ocellata* with 3–5 ODs and papillae was typical of only five of the eleven assemblages. For the purposes of comparison, the classical form is based on previous publications (Pantocsek 1901, Håkansson 1993, Kiss et al 1996, 1999). In the original description of *C. ocellata* (Pantocsek, 1901: fig. 318), this is shown with 3 ODs and a relatively narrow striated part. In the same publication, Pantocsek also described another

species as *Cyclotella crucigera* (later synonymized with *C. ocellata* by Fricke in Schmidt (1906). Hustedt (1928), Håkansson (in Krammer & Lange-Bertalot 1991a) and Kiss et al. (1999) also considered *C. crucigera* a synonym of *C. ocellata.*

A detailed reinvestigation of the type material (Kiss et al. 1999) revealed a higher degree of morphological variability: the number of ODs varied up to ten, and could differ between the epi- and hypovalves of the same frustule, as could the arrangement of the marginal fultoportulae (Kiss et al. 1999: figs 16, 17).The arrangement of the marginal fultoportulae can be irregular; the central part can be colliculate and there can be up to a dozen puncta. A similarly high degree of morphological variability was observed in the Las Madres population in Spain (Kiss et al. 1996), in the eutrophic Lake Dagow in Germany (Schlegel & Scheffler 1999), in the Rapel reservoir in Chile (Rivera et al. 2003), and a natural population and clones from Gallberg-pons in NW Germany (Hegewald & Hindáková 1997). Sometimes ODs were missing, only valves with a colliculate central area and/or puncta in central area being observed at the last site. Cremer et al. (2005) also illustrated several valves without ODs, but with colliculate or completely flat valves. Edlund et al. (2003) found slightly quadrangular valves in a Lake Hovsgol (Mongolia) population of *C. ocellata*. Genkal & Popovskaya (2008) found several frustules with flat valve surfaces and a slight tangential undulation of the central part alongside the classical form of *C. ocellata* in Lake Khubsulug (Mongolia). Therefore, the presence of valves with flat surfaces in the Hungarian and Croatian assemblages is not surprising.

SEM micrographs show that a wide range of *C. ocellata,* '*C. comensis* Grunow–*C. pseudocomensis* Scheffler–*C. costei* Druart et Straub'–like and transitional forms can be found (Figs 4–111). The morphological distinction of the individual nominal species is, however, very difficult. Even the OD characteristic of *C. ocellata* cannot be found in all

individuals, and because ODs do not break through the valve surface, valves with and without ODs are usually indistinguishable from interior views using SEM.

Scheffler & Morabito (2003) published many SEM micrographs of *C. comensis* from Lake Como. Based on their morphological characterisation and these micrographs, the following important traits that potentially distinguish *C. comensis* from *C. ocellata* can be noted: the central area of *C. comensis* is more or less tangentially undulate, sometimes radially or irregularly corrugated, rarely, covered by radial or irregular depressions (resembling ODs and puncta) widely differing in size, and in the number of irregularly arranged depressions and protrusions. Striae are flat to slightly elevated, straight and of different lengths. Scheffler & Morabito (2003) stated: 'The single valves of this species are similar to the structures of *C. ocellata* and can not be clearly classified' (Scheffler & Morabito 2003, fig. 11). Houk et al. (2010) published many SEM micrographs of *C. comensis* with slight tangential undulation and radial patterns over the central area (most valves have radially arranged deep holes, which are identical to ODs). Scheffler & Morabito (2003) found many small frustules with flat valve faces and called these *C. comensis* morphotype *minima*. This form is identical with the small valves of *C. ocellata* from Lake Khubsugul (Genkal & Popovskaya 2008), which becomes particularly clear if we compare Fig. 24 in Scheffler & Morabito (2003) with Fig.2 in Genkal & Popovskaya (2008).

These transitional forms from 'classical' (radial symmetry of the central part) *C. ocellata* through a flat valve centre to 'classical' (tangential undulation of the central part) *C. comensis* were found in several assemblages. In Lake Himód, Lake Dunaharaszti-2012) and Lake Nyékládháza, mixtures of classical *C. ocellata* forms ('*ocellata*' morph), flat intermediate forms, and valves with very slight tangential undulation ('*comensis*' morph) were found. In Lakes Hegyeshalom, Kunsziget, and Szany, the valves were flat or displayed a characteristically tangential undulation at the valve centre ('*comensis*' type valves,

resemblingthose shown in Houk et al. 2010, tab. 215, figs 4, 5). In Lake Visovac, some valves showed the classical radially segmented *C. ocellata* pattern (sometimes with slightly quadrangular contours that was characteristic of '*trichonidea*' morph) whereas others had a flat valve face.

Many authors have also drawn attention to the taxonomic uncertainties of the *C. ocellata* 'group'. Based on light microscopy, Teubner (1995) showed that the taxonomic delimitation of *C. krammeri*, *C. kuetzingiana* var. *planetophora* Fricke, *C. kuetzingiana* var. *radiosa* Fricke, *C. ocellata*, *C. comensis* was problematic because of the polymorphism of the valve. She demonstrated the variation between both valves of the same frustule and between different frustules of *C. ocellata* and raised the question of whether morphological variants might represent different stages in cell wall development or different life cycle stages, or might reflect growth under differing environmental conditions. Genkal & Popovskaya (2008) stated that many representatives of the genus *Cyclotella* had a very similar internal valve view: *C. ocellata*, *C. tripartita*, *C. rossii*, *C. polymorpha*, *C. kuetzingiana*, *C. hispanica* Kiss, Hegewald et Acs (Genkal & Popovskaya 2008).

The morphometric parameters studied by us did not show separation of the '*ocellata*' and the '*trichonidea*' morph, although Ecomou-Amilli (1982) did find differences between *C. ocellata* and *C. trichonidea*. However, she did not discuss transitional forms. Since certain features considered characteristic of *C. trichonidea* (Ecomou-Amilli 1982) were not observed in our '*trichonidea*' morph (e.g. no remarkable difference in silicification between MFP-bearing and non-MFP bearing costae), these specimens may represent transitional forms between *C. ocellata* and *C. trichonidea*. It is therefore possible that these taxa are not the identical, but closely related species.

Although Scheffler (1994) established *C. pseudocomensis* as a new species distinct from *C. comensis*, he later (Scheffler & Morabito 2003, Scheffler et al. 2005) transferred it to *C.*

comensis. In spite of this Houk et al. (2010), who also observed a nearly complete transition between the typical valve morphologies of the two species, nevertheless treated them as separate species because of their different ecological preferences. According to their results, *C. comensis* prefers oligotrophic conditions, while *C. pseudocomensis* occurs in oligotrophic/mesotrophic to moderate eutrophic lakes. However Genkal et al. (2015) found *C. comensis* in mesotrophic and eutrophic water bodies (e.g. eutrophic Lake Rapsudozero, Republic of Karelia). Kistenich et al. (2014) compared *C. comensis*, *C. pseudocomensis*, and *C. costei* strains from eight German and Austrian lakes. Their morphological analysis indicated that the strains formed a morphological continuum rather than three distinctly separate groups. Therefore, they concluded that the three morphospecies comprised a single species complex of *C. comensis*, including *C. pseudocomensis* and *C. costei*. The values of morphological features (numbers of CF and RP, position of MFP and RP) observed in our assemblages fit within that found by Kistenich et al. (2014). Our CVA supported the conspecificity of *C. comensis*, *C. pseudocomensis* and *C. costei*.

Although the morphometric analysis based on the combination of several parameters separated two morphospecies, alone or in pairs none of the studied variables clearly distinguishes *C. ocellata* and *C. comensis.*Visual distinction of the morphospecies was difficults and, in several assemblages, valves with intermediate, flat central area were numerous and could be hardly classified. In addition phenotypic plasticity of single species has been demonstrated by Kiss et al. 2002. One valve of a cell of *C. hispanica* had a plain valve face with fine striation while the other valve had remarkable depressions and elevations, and well-developed striation. The effects of the environment on valve morphology have also been shown for *C. meneghiniana* (Håkansson & Chepurnov 1999). Overall, our results suggested that these species were closely related, but their separation is possible based on combination of characteristics (mainly OD, striae and punctum density.

Wunsam et al. (1995) found that total phosphorus and conductivity were the two most important environmental variables explaining the distribution of *Cyclotella* taxa. In their study the ranges of conductivity and total phosphorus for *C. ocellata* were 137–465 μS cm^{-1} and 4–144 μg L^{-1} respectively. In our study in lakes where *C.* cf *ocellata* occurred the range of conductivity was higher (240–866 μS cm^{-1}) and total phosphorus was a little lower (9–114 μg L^{-1}). In lakes where *C.* cf *comensis* was present the range of conductivity spanned 425 to 872 μS cm^{-1} and total phosphorus ranged between 42–102 μg L^{-1}. Our results support the findings of Kistenich et al. (2014), who stated that the determined genodemes were not linked to particular trophic conditions. Moreover Fritz et al. (1993) and Stoermer &Yang (1969) observed *C. ocellata* in ultra-oligotrophic lakes. It would seem, therefore that *C. ocellata* can occur over a wide trophic range, and thus cannot be taken to be an indicator of trophic conditions.

Molecular investigation:

Genetic divergence and the use of molecular barcodes

The sequenced part of the 18S rDNA involved the V4 subregion, which was found to contain enough variation for the identification of diatom species by Zimmermann et al. (2011). The latter proposed that an approximately 400 bp-long segment of the 18S rRNA gene including this subregion is sufficient as a molecular barcode for diatom species. We found a 0–0.34% (0–3 nt) difference between the studied assemblages (Nyékládháza I, II, Szany, Hegyeshalom, Visovac, Dunaharaszti-2012, Dunaharaszti-2013), and compared to the *C. ocellata* strain LB8 (Alverson et al. 2007). Two of the three divergent sites were in the V4 subregion. Zimmermann et al. (2011) found that the average within-species p–distance varied between 0–0.53 % (a 0–2 nt difference in a 375 nt region; 123 taxa, mainly pennates). Their analyses showed that intrageneric (i.e. interspecific) variation was significantly higher than

intraspecific, within-strain variation. However, this does not apply to *Stephanodiscus* (to which *C. ocellata* proved to be a closer than other *Cyclotella* species, Alverson et al. 2007), because its interspecific variation is lower (Zimmermann et al. 2011, Ki 2009). Luddington et al. (2012) supported the use of the V4 subregion as an effective diatom barcode. They suggested a threshold of p=0.02 (2 %in a 333nt region) for interspecific variation because it separated 96.9 % of the Thalassiosirales, Cymatosirales and Lithodesmiales species studied by them. A p-distance of 0.01 (1 %) increased the efficacy of species separation, but overlapped the intraspecific variation of several species, including within *Cyclostephanos*, sister to *Stephanodiscus* (Alverson et al. 2007). Neither *Stephanodiscus*, nor *C. ocellata* formed part of Luddington et al.'s (2012) study.

In contrast to *Stephanodiscus*, high interspecific divergences were detected in *Cyclotella* (p–distance 1.8 %, excluding *C. ocellata* and *C. bodanica*) and *Discostella* (p–distance 1.3 %) by Jung et al. (2009), although they were using the complete 18S rDNA. In their study the lowest difference was 0.5 %, between *C. atomus* Hustedt and *C. striata* (Kützing) Grunow, while they judged the variation among *C. meneghiniana* isolates (0.3 %) to be almost negligible, due to the effect of concerted evolution on ribosomal genes.

In their investigation of *C. comensis, C. pseudocomensis* and *C. costei* Kistenich et al. (2014) used partial 18SrDNA and *rbc*L as molecular markers, including the 18S rDNA region suggested by Zimmermann et al. (2011), but applied the thresholds (interspecific 2 % in a 333 nt region; intraspecific <0.7 %) suggested by Luddington et al. (2012). Based on this and 0–0.25% variation (in a 403 nt region) Kistenich et al. (2014) considered that their strains belonged to the same species. Furthermore, their *C. ocellata* strain (0–0.5 % different from the former strains) could not be separated from them. Although other markers showed greater distances between the *C. comensis/pseudocomensis/costei* group and *C. ocellata*, a close relationship between the two groups was indicated.

Applying a 2 % species threshold in the region suggested by Zimmermann et al. (2011) to our sequences would indicate that all belong to the same species. The Visovac sample containing *C*. cf *ocellata* and a sequence identical to the '*comensis*'-'*pseudocomensis*'-'*costei*' one seemed to support this. In addition, the minimum difference between the sequences corresponding to *C*. cf *ocellata* (groups A and D) the sequences belonging to *C*. cf *comensis* (groups B and C) was 1 nt (0.11%). However, 1 nt difference could occur within the morphospecies (between groups A and D, as well as groups B and C). Moreover, high intraspecific variability of 18S rDNA has been observed in other Thalassiosiralean genera (e.g. *Skeletonema* Greville, Alverson & Kolnick 2005). However, this region could not separate *Stephanodiscus* and *Cyclostephanos* species (see Supplementary Material Fig. 1.). Therefore it is more probable that the region proposed by Zimmermann et al. (2011) does not have enough variability for species separation, not only in *Stephanodiscus* (as the authors found), but also in *Cyclostephanos* and the *C. ocellata* group.

In addition to 18S rDNA we used a partial sequence of the *rbc*L gene, which has also been proposed as a barcoding marker (e.g. Hamsher et al. 2011, Lee et al. 2013). This gene is more variable than 18S rDNA (Evans et al. 2007, Lee et al. 2013). The separation of intra- and interspecific variability with *rbc*L is made difficult by several factors (listed in Zimmermann et al. 2011, Hamsher et al. 2011, Stoof–Leichenring et al. 2012). Furthermore, there is no consensus as to which region should be used and what thresholds should be adopted for different diatom groups. We sequenced 628 nt region downstream from the start codon and found a 0 to 0.66 % (0–3 nt) difference in this 454–499 nt region between the Hungarian assemblages (Nyékládháza I, Kunsziget, Himód, Dunaharaszti-2012, Szany) and *C. ocellata* strain LB8. It is difficult to comparenthis to previous studies using regions of various length and location, and different groups of diatoms. According to Evans et al. (2007) 0.4 % divergence (5 nt) in 1400 bp *rbc*L between 'elliptical' *Sellaphora pupula* (Kützing)

Mereschkovsky isolates and 2.2 % (30 nt) divergence in the same region between *Sellaphora laevissima* (Kützing) D. G. Mann isolates indicated cryptic species diversity. The divergence in 18S rDNA between the two *S. laevissima* isolates was 0.4 % (8 nt). In the *rbc*L–3P (a 748 bp region at the 3′ end that partly overlapped with our sequences in the 5′ direction) proposed as a barcode by Hamsher et al. (2011) the lowest difference between *Sellaphora* species was 1–2 bp. In contrast, other authors found high intraspecific variability. Rimet et al. (2014) found 2% variability in 1270 nt of *rbc*L of *Nitzschia palea* (Kützing) W. Smith strains and Stoof-Leichenring et al. (2012) found 1.3–8 % in 76 nt region of *S. laevissima* and *S. pupula*. MacGillivary & Kaczmarska (2011) determined an intra/interspecific threshold of 1 % for the class Bacillariophyceae and 2 % for the Mediophyceae in a 540 nt region of *rbc*L, which overlapped with our sequences by 219–316 nt in 5′ direction.

Kistenich et al. (2014) demonstrated a<0.4 % variation between *C. comensis, C. pseudocomensis, C. costei* strains in the 507 nt region. This difference was considered as the intraspecific variability of a single species that was separated from *C. ocellata* by –differences of 0.4–0.6 %. The 0.50–0.75 % (2–3 nt in 402 nt) distance found between A *rbc*L group from assemblages displaying *C*. cf *ocellata* morphospecies and other three sequence groups belonging to *C*. cf *comensis* can be enough to separate two species. However, sequence groups C and D corresponding to the same morphological group also differ from each other at 2 nucleotide positions.).

The genetic divergence was low between the morphospecies and variable within morphospecies. This was especially true for the 18S rDNA of which probably the complete region should be sequenced to decide on the conspecificity of *C. ocellata* and *C. comensis*. The partial *rbc*L sequences performed better, their divergence was higher than Kistenich et al. (2014) found between *C. comensis, C. pseudocomensis* and *C. costei*. Moreover, the

variability in *rbc*L was more corresponding to the morphospecies. Our results did not provide enough evidence in the favour of the conspecificity of *C. ocellata* and *C. comensis.*

Phylogeny

In previous studies, the only representative of the *C. ocellata* group included in phylogenetic analyses was *C. ocellata* LB8, and these showed that this species, along with *C. bodanica*, was separated from other *Cyclotella* species (Alverson et al. 2007, Jung et al. 2009, Lee et al. 2013). This was supported by our results. Kistenich et al. (2014) who compared *C. comensis/pseudocomensis/costei* complex to *C. ocellata* based on molecular markers did not provide phylogenetic tree. In our trees a slight separation of the studied assemblages was observed, indicating two clusters (*C.* cf *ocellata* and *C.* cf *comensis*) that may represent two closely related species. Although Kistenich et al. (2014) emphasized the importance of morphological and molecular investigation of clonal cultures in relation to this kind of taxonomic issue, our results showed that communities with one Thalassiosirales species along with primers specific mainly to Thalassiosirales also could be used. The primers we used could amplify the wide range of Thalassiosirales, and we therefore expected mixed sequences from samples containing cryptic species within *C. ocellata/comensis* (see electropherograms in Supplementary Material Figs 4–9).

Conclusion

From the literature it can be clearly seen that there appears to be a consensus that the *C. ocellata* complex should be considered a combination of morphologically diverse but related species: *C. ocellata, C. krammeri, C. rossii, C. tripartita, C. kuetzingiana,* and *C. polymorpha*. However, comparing the morphological characteristics of the assemblages, it is

clear that further analyses are needed to decide the extent of the morphological variability of this taxon.

Partial sequences of 18S rDNA and *rbc*L showed little (or in the case of Visovac no) variation in samples with different morphs, and the difference was also low when compared to *C. ocellata, C. comensis, C. pseudocomensis,* and *C. costei.* The latter three taxa are probably conspecific, closely related to *C. ocellata* (Kistenich et al. 2014), as our results seem to support. The distinction of *C. ocellata* and *C. comensis* based on the combination of morphometric parameters and their slight but evident phylogenetic separation contradict the hypothesis of their conspecificity. Our results suggest that *C. ocellata* and *C. comensis* are two very closely related species that have only recently diverged from one another.

Acknowledgements

This research was conducted at the MTA Centre for Ecological Research, Danube Research Institute, H-1113 Budapest, Karolina út 29, Hungary and in the Department of Microbiology, Institute of Biology, Eötvös Loránd University, H-1117 Budapest Pázmány Péter sétány 1/C, Hungary.

The authors would like to thank Sonja Kistenich, Mirko Dreßler and Dr. Thomas Hübener for the *C. costei, C. comensis* and *C. pseudocomensis* sequences, Neela Enke, Regine Jahn and Bánk Beszteri and two anonymous reviewers for their valuable comments on the manuscript, Paul Thatcher, Zsuzsa Szilágyi and Virág Pozderka for the language correction. This study was supported by Hungarian National Science Foundation (KTIA-OTKA 80140) and also received support from the Turkish National Science Foundation (TUBITAK-116Z004).

References

AGARDH C.A. 1824. *Systema algarum*. Literis Berlingianis, Lunde. 312 pp.

ALFASANE M. A., ULLAH M. U. & KHONDKER M. 2013. Limnology of Lake Rainkhyongkain with a new record of *Marchantia polymorpha* L. var *aquatica* Nees. *Bangladesh Journal of Botany* 42: 223–229.

ALTSCHUL S. F., GISH W., MILLER W., MYERS E.W. & LIPMAN D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

ALVERSON A.J. 2014. Timing marine–freshwater transitions in the diatom order Thalassiosirales. *Paleobiology* 40: 91–101.

ALVERSON A. J., JANSEN R.K .& THERIOT E.C. 2007. Bridging the Rubicon: Phylogenetic analysis reveals repeated colonizations of marine and fresh waters by thalassiosiroid diatoms. *Molecular Phylogenetics and Evolution* 45: 193–210.

ALVERSON A. J. & KOLNICK L. 2005. Intragenomic nucleotide polymorphism among small subunit rDNA paralogs in the diatom genus *Skeletonema* (Bacillariophyta). *Journal of Phycology* 41: 1248–1257.

ANONYMOUS. 1975. Proposals for standardization on diatom terminology and diagnoses. *Nova Hedwigia, Beiheft* 53: 323–354.

BIRKS H.J.B. 1994. The importance of pollen and diatom taxonomic precision in quantitative palaeoenvironmental reconstructions. *Review of Palaeobotany and Palynology* 83: 107–117.

CHEREPANOVA M.V., USOL'STEVA M.V., PUSHKAR V.S. & DUBROVINA Y.F. 2010. Morphogenesis in *Cyclotella ocellata*–Complex from Lake El'gygytgyn (Chukchi Peninsula) during the Pleistocene–Holocene. *Paleontological Journal* 44: 1–10.

CREMER H., WAGNER B., JUSCHUS O. & MELLES M. 2005. A microscopical study of diatom phytoplankton in deep crater Lake El' gygytgyn, Northeast Siberia. *Archiv für Hydrobiologie Supplement Algological Studies* 116: 147–169.

DANIELIDIS D. B., SPARTINOU M. & ECONOMOU-AMILLI 1996. Limnological survey of Lake Amvrakia, western Greece. *Hydrobiologia* 318: 207–218.

DAVIES D.L. & BOULDIN D.W. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis* 1: 224–227.

DRUART J.C. & STRAUB F. 1988. Description de deux nouvelles Cyclotelles (Bacillariophyceae) de milieux alcalins et eutrophes: *Cyclotella costei* nov. sp. et *Cyclotella wuetrichiana* nov. sp. *Schweizerische Zeitschrift für Hydrobiologie* 50: 182–188.

DULEBA M., ECTOR L., HORVÁTH ZS., KISS K.T., MOLNÁR L.F., POHNER ZS., SZILÁGYI ZS., TÓTH B., VAD CS.F., VÁRBÍRÓ G. & ÁCS É. 2014. Biogeography and phylogenetic position of a warm-stenotherm centric diatom, *Skeletonema potamos* (C.I. Weber) Hasle and its long-term dynamics in the River Danube. *Protist* 165: 715–729.

DUNN J.C. 1973. A fuzzy relative of the ISODATA process and its use in detecting compact well–separated clusters. *Journal of Cybernetics* 3: 32–57.

EATON A.D., CLESCERI L.S., RICE E.W. & GREENBERG A.E. (Ed. by) 2005. *Standard Methods for the Examination of Water and Wastewater* (21st ed.), APHA, AWW, WEF, Washington, DC, USA. 1296 pp.

ECONOMOU-AMILLI A. 1979. Two new taxa of *Cyclotella* Kützing from Lake Trichonis, Greece. *Nova Hedwigia* 31:467–477.

ECONOMOU-AMILLI A. 1982. SEM-studies on *Cyclotella trichonidea* (Bacillariophyceae). *Archiv für Hydrobiologie Supplement Algological Studies* 30: 25–34.

EDLUND M.B., WILLIAMS R.M. & SONINKHISHIG N. 2003. The planktonic diatom diversity of ancient Lake Hovsgol, Mongolia. *Phycologia* 42: 232–260.

EVANS K.M., WORTLEY A.H. & MANN D.G. 2007. An assessment of potential diatom ‘‘barcode’’ genes (*cox*1, *rbc*L, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist* 158: 349–364.

FRITZ S.C., KINGSTON J.C. & ENGSTROM D.R. 1993. Quantitative trophic reconstruction from sedimentary diatom assemblages: a cautionary tale. *Freshwater Biology* 30: 1–23.

GENKAL S.I. 1977. K metodike podscheta nekotorykh taksonomicheski znachimykh strukturnykh ehlementov stvorki u diatomovykh vodoroslej sem. Thalassiosiraceae Lebour emend. Hasle (Bacillariophyta). [On counting of some taxonomically significant structural elements of valves in the diatom algae of the family Thalassiosiraceae Lebour emend. Hasle (Bacillariophyta).] *Botanicheskii Zhurnal* 62: 848–851. [In Russian.]

GENKAL S.I. 1984. O morfologicheskoj izmenchivosti osnovnykh elementov stvorki u vidov roda *Stephanodiscus* (Bacillariophyta). [On morphological variability of the main structural valve elements in the species of the genus *Stephanodiscus* (Bacillariophyta).] *Botanicheskii Zhurnal* 69: 403–408. [In Russian.]

GENKAL, S. I., CHEKRYZHEVA, T. A. & KOMULAYNEN, S. F. 2015. To the taxonomy of *Cyclotella comensis* (Bacillariophyta). *Botanicheskii Zhurnal* 100: 389–394. [In Russian.]

GENKAL S.I. & POPOVSKAYA G.I. 2008. Morphological variability of *Cyclotella ocellata* from lake Khubsugul (Mongolia). *Diatom Research* 23: 75–91.

GRUNOW A. 1882. *Cyclotella comensis*. In: *Synopsis des Diatomées de Belgique: Atlas* (Ed. by H. van Heurck), pp. 78–103. Ducaju et Cie, Anvers

HÅKANSSON H. 1993. Morphological and taxonomic problems in four *Cyclotella* species (Bacillariophyceae). *Diatom Research* 8: 309–316.

HÅKANSSON H. & CHEPURNOV V. 1999. A study of variation in valve morphology of the diatom *Cyclotella meneghiniana* in monoclonal cultures: the effect of auxospore formation and different salinity conditions. *Diatom Research* 14(2): 251–272.

HAMMER Ø., HARPER D.A.T. & RYAN P.D. 2001. PAST: Paleontological Statistics software package for education and data analysis. *Palaeontologica Electronica* 4: 9 http://palaeo-electronica.org/2001_1/past/issue1_01.htm [accessed 29. 04. 2014.]

HAMSHER S.E., EVANS K.M., MANN D.G., POULÍČKOVÁ A. & SAUNDERS G.W. 2011. Barcoding diatoms: Exploring alternatives to COI–5P. *Protist* 162: 405–422.

HASEGAWA M., KISHINO H. &YANO T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evoltuion* 22:160–174.

HAUSMANN S. & LOTTER F.A. 2001. *Cyclotella comensis* and its importance for quantitative temperature reconstructions. *Freshwater Biology* 46: 1323–1333.

HEGEWALD E. & HINDÁKOVA A. 1997. Variabilität von einer natürlichen Population und von Klonen des *Cyclotella ocellata*–Komplexes (Bacillariophaceae) aus dem Gallbergweiher, Nordwestdeutschland. *Archiv für Hydrobiologie Supplement Algological Studies* 86: 17–37.

HOUK V., KLEE R. & TANAKA H. 2010. Atlas of freshwater centric diatoms with a brief key and descriptions. Part III. Stephanodiscaceae A. *Cyclotella, Tertiarius, Discostella*. *Fottea* 10 (Suppl): 1–498.

HUSTEDT F. 1930. Die Kieselalgen Deutschlands, Österreichs und der Schweiz unter Berücksichtigung der übrigen Länder Europas sowie der angrenzenden Meeresgebiete. In: *Dr. L. Rabenhorst 's Kryptogamenflora von Deutschland, Österreich und Schweiz* (Ed. by L. Rabenhorst). Akademische Verlagsgesellschaft, Leipzig: 7.

IWAMURA T., NAGAI H. & ICHIMURA S. 1970. Improved methods for determining contents of chlorophyll, protein, ribonucleic acid and deoxyribonucleic acid in planktonic

populations. *Internationale Revue der gesamten Hydrobiologie und Hydrographie*. 55: 131-147.

JUNG S.W., HAN M.-S. & KI J.-S. 2010. Molecular genetic divergence of the centric diatom *Cyclotella* and *Discostella* (Bacillariophyceae) revealed by nuclear ribosomal DNA comparisons. *Journal of Applied Phycology* 22: 319–329.

KI J.-S. 2009. Comparative molecular analysis of freshwater centric diatoms with particular emphasis on the nuclear ribosomal DNA of *Stephanodiscus* (Bacillariophyceae). *Algae* 24: 129–138.

KISTENICH S., DREßLER M., ZIMMERMANN J., HÜBENER T., BASTROP R. & JAHN R. 2014. An investigation of the morphology and genetics of *Cyclotella comensis* and closely related taxa. *Diatom Research* 29: 423–440.

KISS K.T., KLEE R., ECTOR L. & ÁCS É. 2012. Centric diatoms of large rivers and tributaries in Hungary: morphology and biogeographic distribution. *Acta Botanica Croatia* 71: 311–363.

KISS K.T., KLEE R. & HEGEWALD E. 1999. Reinvestigation of the original material of *Cyclotella ocellata* Pantocsek (Bacillariophyceae). *Archiv für Hydrobiologie Supplement Algological Studies* 93: 39–53.

KISS K.T., ROJO C. & ÁLVAREZ-COBELAS M. 1996. Morphological variability of a *Cyclotella ocellata* (Bacillariophyceae) population in the lake Las Madres (Spain). *Archiv für Hydrobiologie Supplement Algological Studies* 82: 37–55.

KISS K.T., HEGEWALD E. & ÁCS É. 2002. *Cyclotella hispanica* a new dimorphic centric diatom species. *Algological Studies* 106: 1–6.

KNIE M. & HÜBENER T. 2007. Morphological variability of the *Cyclotella ocellata–krammeri–rossii* complex in field samples and cultures. In: *Proceedings of the 1st Central European Diatom Meeting 2007* (Ed. by Kusber W.-H. & Jahn R.), pp. 83–86.

KOVÁCS J., KOVÁCS S., MAGYAR N., TANOS P., HATVANI I.G. & ANDA A. 2014. Classification into homogeneous groups using combined cluster and discriminant analysis. *Environmental Modelling & Software* 57: 52–59.

LEE M.-A., FARIA D.G., HAN M.-S., LEE J. & KI J.-S. 2013. Evaluation of nuclear ribosomal RNA and chloroplast gene markers for the DNA taxonomy of centric diatoms. *Biochemical Systematics and Ecology* 50: 163–174.

LUDDINGTON I.A., KACZMARSKA I. & LOVEJOY C. 2012. Distance and character-based evaluation of the V4 region of the 18S rRNA gene for the identification of diatoms (Bacillariophyceae). *PLoS ONE* 7: e45664.

MACGILLIVARY M.L. & KACZMARSKA I. 2011. Survey of the efficacy of a short fragment of the *rbc*L gene as a supplemental DNA barcode for diatoms. *Journal of Eukaryotic Microbiology* 58: 529–536.

MEDLIN L., ELWOOD H.J., STICKEL S. & SOGIN M.L. 1988. The characterization of enzymatically amplified eukaryotic 16S–like rRNA–coding regions. *Gene* 71: 491–499.

PANTOCSEK J. 1901. Die Kieselalgen oder Bacillarien des Balaton. Im Auftrage des ungarischen geographischen Gesellschaft auf Basis eigener Aufsammlungen. In: *Resultate der wissenschaftlichen Erforschung des Balatonsees*. II. Band. Anhang zur II. Section des 2. Theiles, K. und K. Hofbuchdruckerei des Victor Hornyánszky, Budapest. 112 pp.

RIMET F., BOUCHEZ A. (2012). Biomonitoring river diatoms: Implications of taxonomic resolution. *Ecological Indicators* 15: 92–99.

RIMET F. (2013). *Phytoplancton du Léman (The phytoplankton of the Lake Geneva)*. Rapport Commission internationale pour la protection des eaux du Léman Campagne 2012, 2013: 82–93. [in French]

RIMET F., TROBAJO R., MANN D. G., KERMARREC L., FRANC A., DOMAIZON I. & BOUCHEZ A. 2014. When is sampling complete? The effects of the geographical range and marker choice on perceived diversity *Nitzschia palea* (Bacillariophyta). *Protist* 165: 245–259.

RIVERA P., CRUCES F. & VILA I. 2003. *Cyclotella ocellata* Pantocsek (Bacillariophyceae): Primera cita en Chile y comentarios sobre su variabilidad morfologica. *Gayana Botánica* 60: 123–131.

RONQUIST F., TESLENKO M., VAN DER MARK P., AYRES D.L., DARLING A., HÖHNA S., LARGET B., LIU L., SUCHARD M.A. & HUELSENBECK J.P. 2012. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.

ROSS R., COX E.J., KARAYEVA N.I., MANN D.G., PADDOCK T.B.B., SIMONSEN R. & SIMS P.A. 1979. An amended terminology for the siliceous components of the diatom cell. *Nova Hedwigia, Beiheft* 64: 513–533.

SCHEFFLER W. 1994. *Cyclotella pseudocomensis* nov. spec. (Bacillariophyceae) aus norddeutschen Seen. *Diatom Research* 9: 355–369.

SCHEFFLER W. & MORABITO G. 2003. Topical observations on centric diatoms (Bacillariophyceae, Centrales) of Lake Como (N. Italy). *Journal of Limnology* 62: 47–60.

SCHEFFLER W., NICKLISCH A. & HEPPERLE D. 2003. Dimorphism in *Cyclotella pseudocomensis* (Heterokontophyta, Bacillariophyceae) as revealed by morphological, ecological and molecular methods. *Advances in Limnology* 58: 157–173.

SCHEFFLER W., NICKLISCH A. & SCHÖNFELDER I. 2005. Beiträge zur Morphologie, Ökologie und Ontogenie der planktischen Diatomee *Cyclotella comensis* Grunow: Untersuchungen an historischem und rezentem Material. *Diatom Research* 20: 171–200.

SCHLEGEL I. & SCHEFFLER W. 1999. Seasonal Development and Morphological Variability of *Cyclotella ocellata* (Bacillariophyceae) in the Eutrophic Lake Dagow (Germany). *International Review of Hydrobiology* 84: 469–478.

SOLAK C.N. & KULIKOVSKIY M. 2013. Species composition and distribution of centric diatoms from Türkmen Mountain (Sakarya River Basin/Turkey). *Turkish Journal of Botany* 37: 589–596.

STOERMER E.E. & YANG J.J. 1969. *Plankton diatom assemblages in Lake Michigan*. Great Lakes Research Division, Institute of Science and Technology, University of Michigan, Ann Arbor, Michigan, GLRD, Special Report No. 47. 268 pp.

STOOF-LEICHENRING K.R., EPP L.S., TRAUTH M.H. & TIEDEMANN R. 2012. Hidden diversity in diatoms of Kenyan Lake Naivasha: a genetic approach detects temporal variation. *Molecular Ecology* 21: 1918–1930.

TAMURA K. & NEI M.1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. *Molecular Biology and Evolution* 10: 512–526.

TAMURA K. STECHER G., PETERSON D., FILIPSKI A. & KUMAR S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30: 2725–2729.

TEUBNER K. 1995. A light microscopical investigation and multivariate statistical analyses of heterovalvar cells of *Cyclotella*–species (Bacillariophyceae) from lakes of the Berlin–Brandenburg region. *Diatom Research* 10: 191–205.

THERIOT E. 1987. Principal component analysis and taxonomic interpretation of environmentally related variation in silicification in *Stephanodiscus* (Bacillariophyceae). *British Phycological Journal* 22: 359–373.

THERIOT E. & SERIEYSSOL K. 1994. Phylogenetic systematics as a guide to understanding features and potential characters of the centric diatom family Thalassiosiraceae. *Diatom Research* 9: 429–450.

VIJAYASARATHY P.R. 2011. *Engineering Chemistry* (2nd ed.). PHI Learning Private Ltd., New Delhi, pp. 310.

WERNER P. & SMOL J.P. 2006. The distribution of the diatom *Cyclotella comensis* in Ontario (Canada) lakes. *Nova Hedwigia* 130: 373–391.

WOLIN J. A. & STOERMER E. F. 2005. Response of a Lake Michigan coastal lake to anthropogenic catchment disturbance. *Journal of Paleolimnology* 33: 75–94.

WUNSAM S., SCHMIDT R. & KLEE R. 1995. *Cyclotella*–taxa (Bacillariophyceae) in lakes of the Alpine region and their relationship to environmental variables. *Aquatic Sciences* 57: 360–386.

YE J., COULOURIS G., ZARETSKAYA I., CUTCUTACHE I., ROZEN S. & MADDEN T.L. 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13: 134.

ZIMMERMANN J., JAHN R. & GEMEINHOLZER B. 2011. Barcoding diatoms: evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. *Organisms, Diversity and Evolution* 11: 173–192.

Supplemental material

Supplemental data for this article can be accessed ….

Table 1. Description of sampling sites. gp=gravel pit lake, Hu=Hungary, Cr=Croatia, Tr=Turkey, ad=average depth, md=maximal depth

Name of water/ closest town/type of water body	**GPS/country**	**Area and depth**	**Trophic level**	**Anthropogenic influence**	**Special remarks**
Lake Balaton/ Alsóörs/ natural	N 46º58′58.52″ E 17º58′37.57″ Hu	Area: 59300 ha, ad: 3.2 m, md: 12.2 m	mesotrophic	recreation	biogenic chalk precipitation
Lake Himód/Himód/gp	N 47º31′49.7″ E 17º01′29.8″ Hu	Area: 6 ha md: about 20 m	eutrophic	fishing lake rich in fish	
Lake Dunaharaszti/Dunaharaszti /gp	N 47º20′30.88″ E 19º07′03.07″ Hu	Area: 10.4 ha, md: 4-5 m	mesotrophic	jet-ski track	
Lake Nyékládháza/ Nyékládháza/ gp	N 47º58′55.26″E 20º53′26.09″ Hu	Area: 361 ha, md: 30 m	oligo-meso-trophic	group of currently mined lakes	
Lake Hegyeshalom/ Hegyeshalom/gp	N 47º54′06.1″ E 17º08′53.8″ Hu	Area: 70 ha, md: 50 m	meso-eu-trophic	diving, fishing, surfing	
Lake Csiszlói/ Kunsziget/ gp	N 47º43′58.7″ E 17º30′05.6″ Hu	Area: 3.7 ha md: about 20 m	eutrophic	fishing lake rich in fish	
Lake Szany/Szany/gp	N 47º26′53.7″ E 17º18′19.8″ Hu	Area: 5 ha md: about 20 m	eutrophic	fishing lake	

Table 1 continued. Description of sampling sites. gp=gravel pit lake, Hu=Hungary, Cr=Croatia, Tr=Turkey, ad=average depth, md=maximal depth

Name of water/ closest town/type of water body	**GPS/country**	**Area and depth**	**Trophic level**	**Anthropogenic influence**	**Special remarks**
Lake Visovac/ Losovac/natural, warm, monomictic.	N 43º48′05.35″ E 15º58′34.79″ Cr	Area: 790 ha, maximal depth: 55 m	oligo-meso-trophic	national park	situated on Krka River bounded by two travertine barriers
İnli Plateau/Spring	N 39°27'48.42″ E 30°20'02.28″ Tr		oligotrophic		typical holocrene mountain spring
Yalova/ Stream	N 40°36′40.01″ E 29°13′00.14″ Tr		eutrophic		small stream of Marmarean river basin

Table 2. Water chemical parameters measured at the sampling sites (except at İnli spring). Cond.=conductivity, Turb.=turbidity, DO=dissolved oxygen

water	Cond. ($\mu S\ cm^{-1}$)	pH	DO ($mg\ L^{-1}$)	Turb. (NTU)	Cl^- ($mg\ L^{-1}$)	CO_3^{2-} ($mg\ L^{-1}$)	HCO_3^{2-} ($mg\ L^{-1}$)	Ca^{2+} ($mg\ L^{-1}$)	Mg^{2+} ($mg\ L^{-1}$)	TP ($\mu g\ L^{-1}$)	PO_4–P ($\mu g\ L^{-1}$)	NO_3–N ($mg\ L^{-1}$)	NH_4–N ($mg\ L^{-1}$)	Chl *a* ($\mu g\ L^{-1}$)
Lake Balaton	750.0	8.6	11.3	4.0	88.0	27.0	274.0	37.0	66.0	26.8	13.5	0.01	0.040	6.30
Lake Himód	504.8	8.5	8.7	4.7	52.4	16.9	85.9	53.7	15.3	113.6	20.6	0.1	0.002	15.43
Lake Dunaharaszti	871.9	7.2	8.0	4.3	143.3	8.4	180.4	53.1	52.5	42.3	9.7	0.21	0.051	–
Lake Nyékládháza	866.0	8.0	9.4	7.7	47.5	0.0	106.7	115.3	–	–	–	–	–	–
Lake Hegyeshalom	611.4	8.0		0.1	59.0	16.9	107.4	58.1	18.3	64.6	19.4	2.4	0.000	2.38
Lake Csiszlói at Kunsziget	835.9	8.1	8.0	3.0	64.9	16.9	128.8	76.4	40.6	101.8	27.8	0.2	0.006	2.24
Lake Szany	424.8	6.5	8.2	10.2	60.2	8.4	60.1	50.0	11.6	90.6	20.6	0.1	0.002	4.59
Lake Visovac	452.0	7.9	9.7	-	-	-	-	–	–	9.0	6.0	0.1	0.005	6.78
Yalova stream	240.0	7.6	8.9	-	10.0	-	-	–	–	–	40.0	–	–	–

Table 3. Primers used in the molecular study

Name	Marker	Sequence (5' to 3')	Reference
Medlin-A[1]	18S rDNA	AACCTGGTTGATCCTGCCAGT	Medlin et al. (1988)
Sk-600F[2,s]	18S rDNA	AAATCCCTTATCGAGTATCA	Duleba et al. (2014)
Sk-900F[s]	18S rDNA	TTGGTTTGCGAGTCAAAGTA	Duleba et al. (2014)
Sk-1550R[1,2,s]	18S rDNA	TCTCGGCCAAGGTTATAT	Duleba et al. (2014)
rbcL66F[1]	*rbc*L	TTAAGGAGAAATAAATGTCTCAATCTG	Alverson et al. (2007)
Sk-rbcL400F[2,s]	*rbc*L	ATTAACTCTCAACCATTCATGC	Duleba et al. (2014)
Sk-rbcL975R[1,2,s]	*rbc*L	CAACATCATCACCTAAATAGTG	Duleba et al. (2014)

[1]primers for the first PCRs, [2]primers for seminested PCRs, [s]sequencing primers.

Table 4. Morphometric data of the morphs.

	valve diameter	diameter of the central area	number of striae	number of striae in 10 µm	number of marginal fultoportulae	number of puncta	number of orbiculi depressi	number of divided striae	number of radially symmetric valve sectors
'*ocellata*'	4.0–15.5	1.7–9.3	25–87	14.2–24.3	6–29	0–19	0–6	0–19	0–5
'*trichonidea*'	8–15.5	3.2–10.3	47–87	14.8–19.5	13–21	0–3	3	0–6	3
'*costei*'	4.3–9.2	1.4–4.1	28–57	15.9–24.4	4–15	0–19	0	0–14	0–2
'*comensis*'	4.8–9.3	2.1–4.4	29–54	16.8–23.4	6–13	4–24	0	0–9	0–7
'*pseudocomensis*'	6.3–10.8	3.8–5.9	38–63	17.0–19.2	7–11	11–16	0	0–3	0–2
'*transitional*'	3.0–10.2	1.5–5.3	26–60	16.4–26.0	5–17	0–18	0–3	0–14	0–3

Table 5. Accession number of sequences acquired from *Cyclotella* cf *ocellata* and *C.* cf *comensis* assemblages and their identity with *C. ocellata* strain LB8 sequences from GenBank.

Sample	**Type of DNA sample**	**18S rDNA sequence**	**Identity with *C. ocellata* (nt/nt)**	***rbc*L sequence**	**Identity with *C. ocellata* (nt/nt)**
Himód	community DNA	–	–	KJ755338.1	467/467 (100 %)
Dunaharaszti-2013	community DNA	KJ755342.1	794/795 (99.87 %)	–	–
Dunaharaszti-2012	community DNA	KJ755348.1	909/911 (99.78 %)	KJ755339.1	451/454 (99.34 %)
Nyékládháza I.	isolated cells	KJ755343.1	874/875 (99.89 %)	KJ755337.1	499/499 (100 %)
Nyékládháza II.	isolated cells	KJ755345.1	873/876 (99.66 %)	–	–
Hegyeshalom	community DNA	KJ755346.1	807/809 (99.75 %)	–	–
Kunsziget	community DNA	–	–	KJ755340.1	451/454 (99.34 %)
Szanyi	community DNA	KJ755347.1	855/857 (99.77 %)	KJ755341.1	452/454 (99.56 %)
Visovac	isolated cells	KJ755344.1	874/876 (99.77 %)	–	–

Table 6. Pairwise differences (uncorrected p-distance values expressed in percentage) in 18S rDNA among the *Cyclotella* cf. *ocellata* and *C*. cf *comensis* assemblages and *C. ocellata* strain LB8 (accession number DQ514904). Numbers of the different nucleotides are in parentheses.

	Dunaharaszti-2013	Dunaharaszti-2012	Nyékládháza I.	Nyékládháza II.	Hegyeshalom	Szany	Visovac	*C. ocellata* LB8
Dunaharaszti-2013	–							
Dunaharaszti-2012	0.13 (1)	–						
Nyékládháza I.	0 (0)	0.11 (1)	–					
Nyékládháza II.	0.25 (2)	0.11 (1)	0.23 (2)	–				
Hegyeshalom	0.13 (1)	0 (0)	0.12 (1)	0.12 (1)	–			
Szany	0.13 (1)	0 (0)	0.12 (1)	0.12 (1)	0 (0)	–		
Visovac	0.13 (1)	0 (0)	0.11 (1)	0.11 (1)	0 (0)	0 (0)	–	
C. ocellata LB8	0.13 (1)	0.22 (2)	0.11 (1)	0.34 (3)	0.25 (2)	0.23 (2)	0.22 (2)	–

Table 7. Pairwise differences (uncorrected p-distance values expressed in percentage) in *rbc*L among the *Cyclotella* cf *ocellata* and *C*. cf *comensis* assemblages and *C. ocellata* strain LB8 (accession number DQ514832). Numbers of the different nucleotides are in parentheses.

	Himód	Duna-haraszti 2012	Nyéklád-háza I.	Kunsziget	Szany	*C. ocellata* LB8
Himód	–					
Duna-haraszti-2012	0.66 (3)	–				
Nyéklád-háza I.	0 (0)	0.66 (3)	–			
Kunsziget	0.66 (3)	0 (0)	0.66 (3)	–		
Szany	0.44 (2)	0.22 (1)	0.44 (2)	0.22 (0)	–	
C. ocellata LB8	0 (0)	0.66 (3)	0 (0)	0.66 (3)	0.44 (2)	–

Figure captions

Fig. 1. Sampling sites. Stars: samples only used for morphometric analysis, black points: samples used for DNA and morphometric analysis. 1 = Lake Balaton, 2 = Lake Himód, 3 = Lake Dunaharaszti, 4 = Lake Nyékládháza, 5 = Lake Hegyeshalom, 6 = Lake Csiszlói at Kunsziget, 7 = Lake Szany, 8 = Lake Visovac, 9 = İnli spring, 10=Yalova stream.

Fig. 2. Morphometric parameters measured, and morphological features counted, on the valves from outside (a) and inside (b) views.
Abbreviations: d = valve diameter, dCA = diameter of the central area, OD = orbiculi depressi, MFP = marginal fultoportula, CFP = central fultoportula, RP = rimoportula, RP ang = orientation of the rimoportula labium

Figs 3–14. *Cyclotella* cf *ocellata* assemblage from Lake Balaton (Figs 3–11. classical '*ocellata*' form). Scale bar = 5 μm.

Figs 15–26. *Cyclotella* cf *ocellata* assemblage from Lake Himód (Figs 15–16. Classical '*ocellata*' form, Fig. 17. Transitional form, Figs 18–22. '*costei*' form). Scale bar =5 μm (Figs 15–18, 20, 22, 24–26); 2.5 μm (Figs 19, 21, 23).

Figs 27-38. *Cyclotella* cf *ocellata* assemblage from Lake Dunaharaszti in 2013 (Figs 27–34. Classical '*ocellata*' form, Fig.35. '*costei*' form).Scale bar = 5 μm (Figs 27, 31–33, 35–38); 2.5 μm (Figs 28–29); 2 μm (Figs 30, 34).

Figs 39–50. *Cyclotella* cf *comensis* assemblage from Lake Dunaharaszti in 2012 (Fig. 39. classical '*ocellata*' form, Figs 40–42. '*comensis*' form, Fig. 43. transitional form, Figs 44–48. '*costei*' form). Scale bar =5 μm (Figs 39, 41–44, 47, 49); 3 μm (Figs 45–46); 2 μm (Figs 40, 48, 50).

Figs 51–62. *Cyclotella* cf *ocellata*/*comensis* assemblage from Lake Nyékládháza (Figs 51–52. classical '*ocellata*' form, Fig. 53. '*pseudocomensis*' form, Figs 54–55. '*comensis*' form, Figs 56–58. aff. '*comensis*' form, Fig. 59. '*costei*' form). Scale bar = 10 μm (Figs 51, 52, 61, 62); 5 μm (Figs 53, 55, 57, 59, 60); 2.5 μm (Figs 54, 56, 58).

Figs 63–74. *Cyclotella* cf *comensis* assemblage from Lake Hegyeshalom (Figs 63–68. '*comensis*' form**,** Figs 69–72. '*costei*' form). Scale bar =5 μm (Figs 66, 68, 69, 73); 2.5 μm (Figs 63, 64, 72, 74); 2 μm (Figs 65, 67, 70, 71).

Figs 75–86. *Cyclotella* cf *comensis* assemblage from Lake Csiszlói at Kunsziget; Figs 75–82. '*comensis*' form, Figs 83–85. '*costei*' form). Scale bar = 5 μm (Figs 75–77, 80–82, 84); 2.5 μm (Figs 85–86); 2 μm (Figs 78, 79, 83).

Figs 87–98. *Cyclotella* cf *comensis* assemblage from Lake Szany (Figs 87–89, 92. '*comensis*' form, Figs 90, 91, 93. aff. '*costei*' form, Figs 94–95. '*costei*' form, Fig. 96. initial cell). Scale bar = 10 μm (Fig. 96); 5 μm (Figs 87, 89, 92, 93, 95, 97, 98); 2.5 μm (Figs 88, 90, 91, 94).

Figs 99–110. *Cyclotella* cf *ocellata* assemblage from Lake Visovac (Figs 99–104; Figs 99–100. classical '*ocellata*' form, Figs 101, 104. aff. '*trichonidea*' form, Figs 102–103. aff. '*comensis*' form); Inli (Figs 105–106. aff '*ocellata*' form) and Yalova (Figs 107–110; Fig. 107. classical '*ocellata*' form). Scale bar = 10 μm (Fig. 104); 5 μm (Figs 99–101, 105, 106, 109, 110); 2.5 μm (Figs 102–103); 2 μm (Fig. 107); 1 μm (Fig. 108).

Fig. 111. Scatter plot of canonical variates analysis. The circle represents the '*ocellata*', the star the '*trichonidea*', the square the '*costei*', the triangle the '*comensis*', the diamond the '*pseudocomensis*', the cross the transitional forms. See abbreviations of variables in the Morphological analysis section of the Materials and methods.

Fig. 112. Parsimony network of the partial 18S rDNA sequence types.

Fig. 113. Parsimony network of the partial *rbc*L sequence types.

Fig.114. Bayesian inferred phylogenetic tree of combined data constructed using Hasegawa-Kishino-Yanomodel with gamma distribution for 18S rDNA and Tamura-Nei model with gamma distribution for *rbc*L. *Cyclotella meneghiniana* LS03–01 and Tl1 strains and *C. striata* were used as outgroup. Scale bar represents 0.009 substitutions per site. Posterior probabilities (express in percent) are indicated at the nodes. Sequences acquired in this study are in bold. Sorting of sequences into groups is explained in the text.

Supplementary Material Fig. 1. Scatter plot matrix. See abbreviations of variables in the Morphological analysis section of the Materials and methods. Symbols represent the morphs: circle='*ocellata*', star='trichonidea', triangle='*comensis*', diamond='*pseudocomensis*', square='*costei*', cross=transitional form.

Supplementary Material Fig. 2. Maximum likelihood tree for 18S rDNA sequences constructed using Hasegawa-Kishino-Yano model with gamma distribution. *Cyclotella meneghiniana* LS03–01 and Tl1 strains and *C. striata* were used as outgroup. Scale bar represents 0.005 substitutions per site. Available GenBank accession number of the used sequences is provided in the name of the taxa. Bootstrap values are indicated at the nodes. Sequences acquired in this study are in bold. *C*. cf *ocellata* Dunaharaszti 2013 represents the sequence type A, *C*. cf *comensis* Szany the type B, *C*. cf *comensis* Nyékládháza II the type C. *Cyclotella ocellata* LB8 was considered type D.

Supplementary Material Fig. 3. Maximum likelihood tree for *rbc*L sequences constructed using Tamura-Nei model with gamma distribution. *Cyclotella meneghiniana* LS03–01 and Tl1 strains and *C. striata* were used as outgroup. Scale bar represents 0.01 substitutions per site. Available GenBank accession number of the used sequences is provided with the name of taxa. Bootstrap values are indicated at the nodes. . Sequences acquired in this study are in bold. *C*. *ocellata* LB8 represents the sequences type A, *C*. cf *comensis* Szany the type B, *C*. *costei* Gültzsee the type C, *C*. cf *comensis* Kunsziget the type D.

Supplementary Material Fig. 4. Selected parts of electropherogram of *C.* cf *ocellata* Himód acquired with Sk-rbcL-400F primer. Arrows indicate the sites diverging between the sequence groups.

Supplementary Material Fig. 5. Selected parts of electropherograms of *C.* cf *ocellata* Nyékládháza I acquired with Sk-600F and Sk-1550R primers. Arrows indicate the sites diverging between the sequence groups.

Supplementary Material Fig. 6. Selected parts of electropherogram of *C.* cf *ocellata* Nyékládháza I acquired with Sk-rbcL-975R primer. Arrows indicate the sites diverging between the sequence groups.

Supplementary Material Fig. 7. Selected parts of electropherograms of *C.* cf *comensis* Nyékládháza II acquired with Sk-600F and Sk-1550R primers. Arrows indicate the sites diverging between the sequence groups.

Supplementary Material Fig. 8. Selected parts of electropherogram of *C.* cf *comensis* Dunaharaszti 2012 acquired with Sk-600F primer. Arrows indicate the sites diverging between the sequence groups.

Supplementary Material Fig. 9. Selected parts of electropherogram of *C.* cf *comensis* Dunaharaszti 2012 acquired with Sk-rbcL-400F primer. Arrows indicate the sites diverging between the sequence groups.

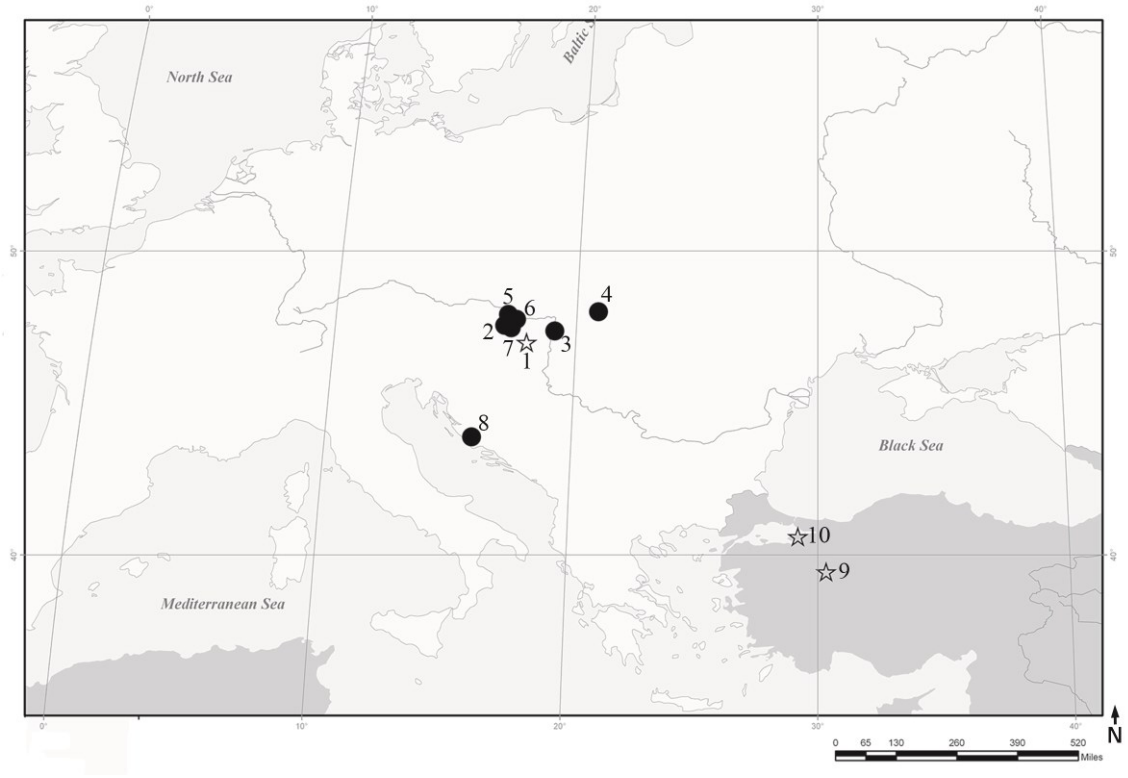
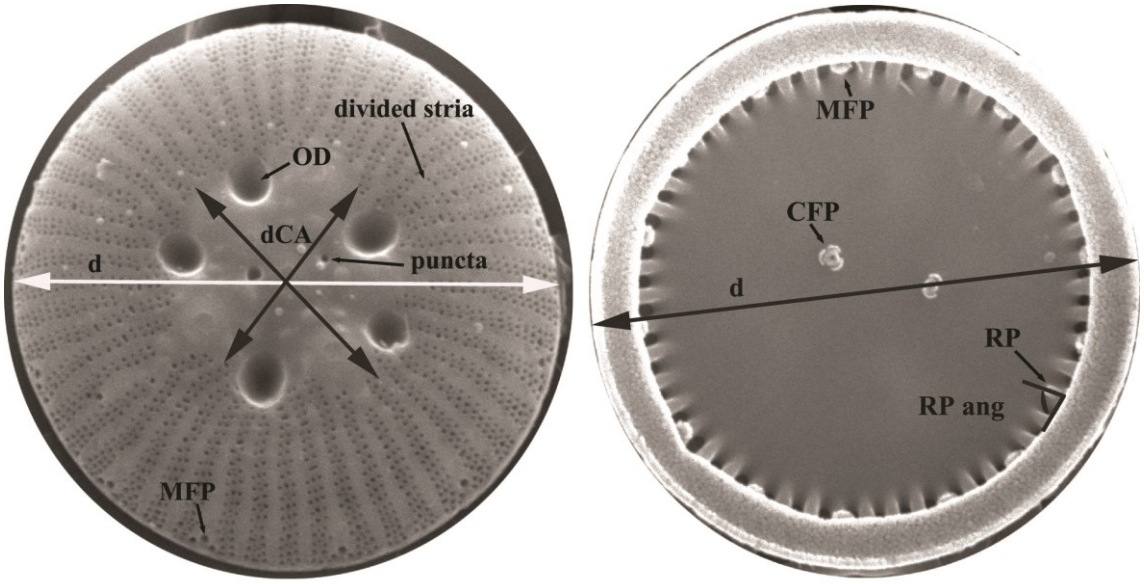
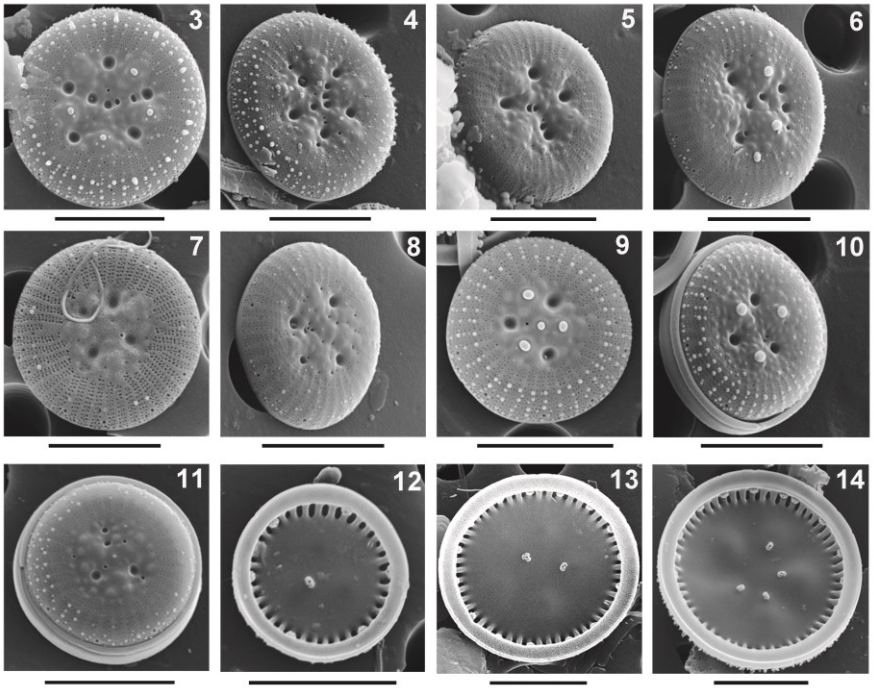

North Sea
Baltic Sea
Black Sea
Mediterranean Sea
1
2
3
4
5
6
7
8
☆9
☆10
N
0 65 130 260 390 520
Miles

Fig. 1

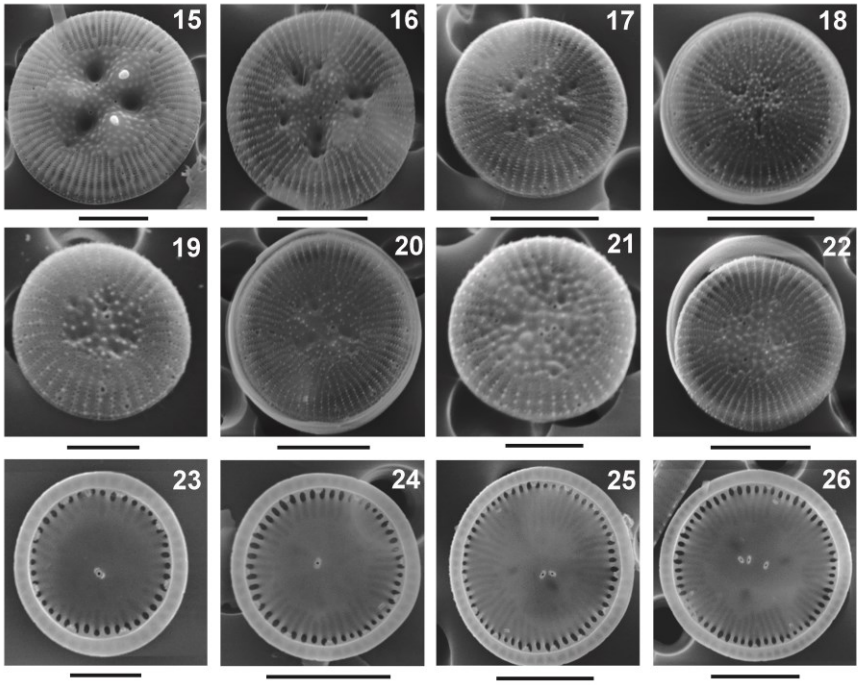


(a) (b)

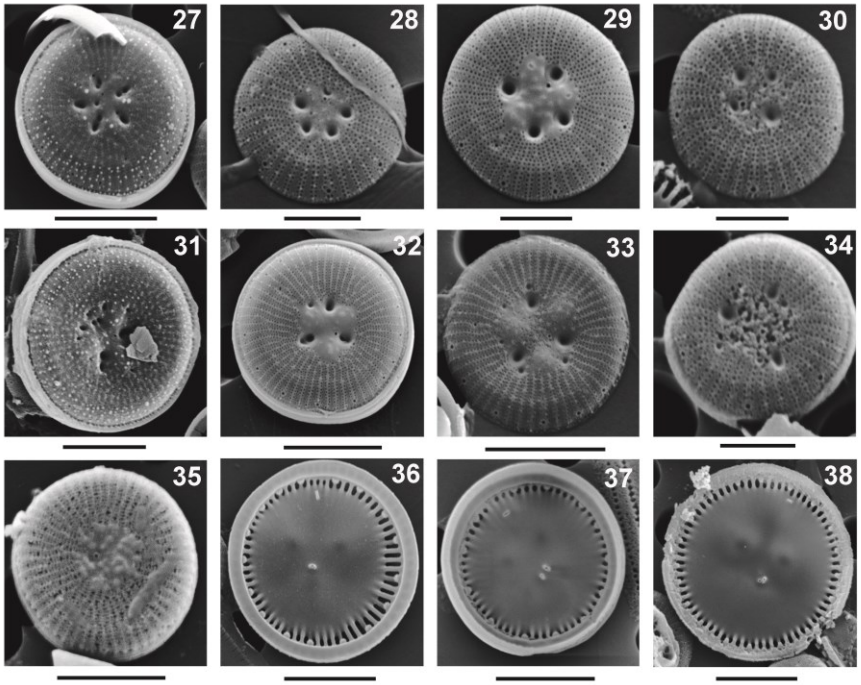
Fig. 2



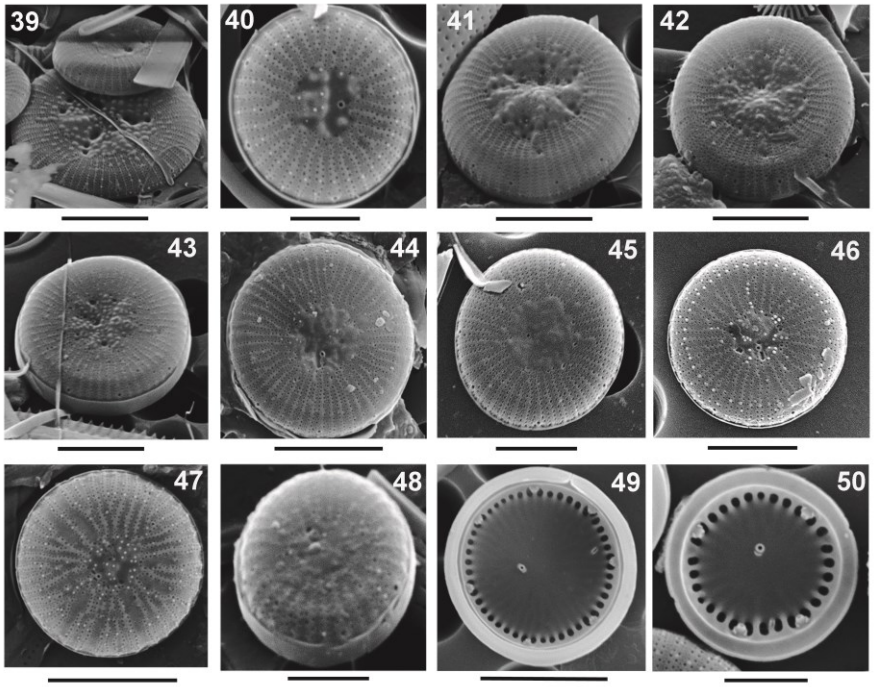
Figs 3–14



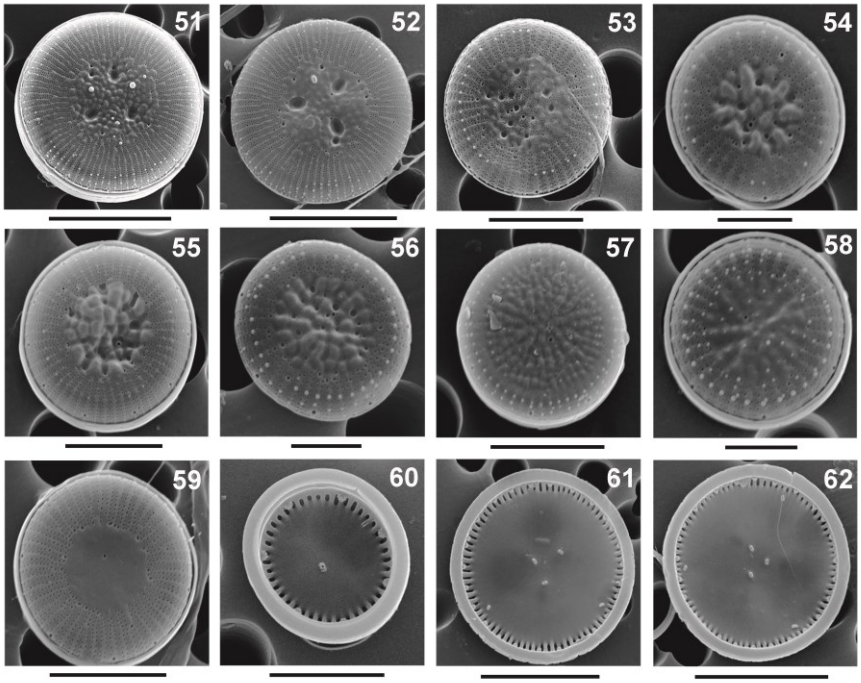
Figs 15–26



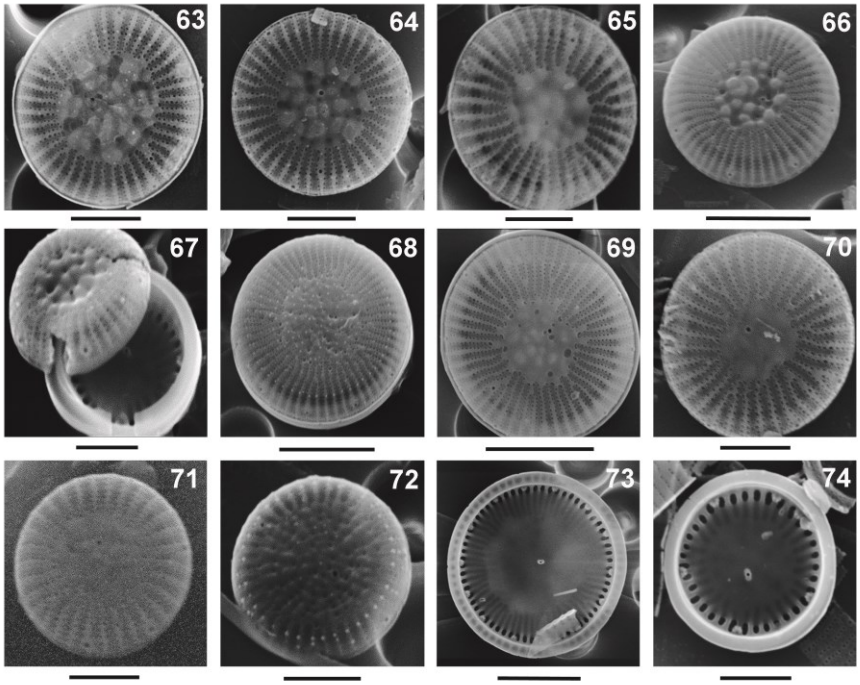
Figs 27–38



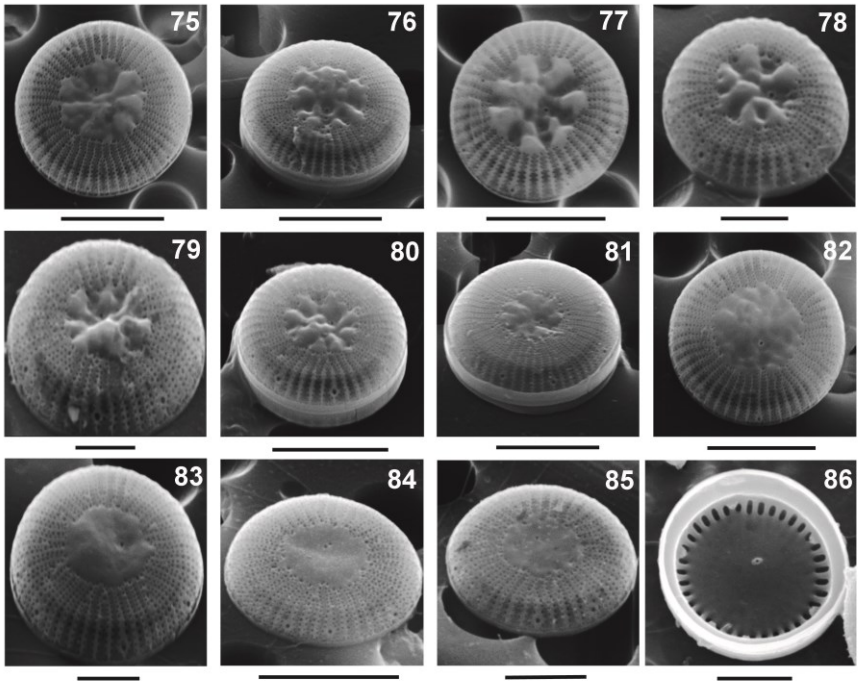
Figs 39–50



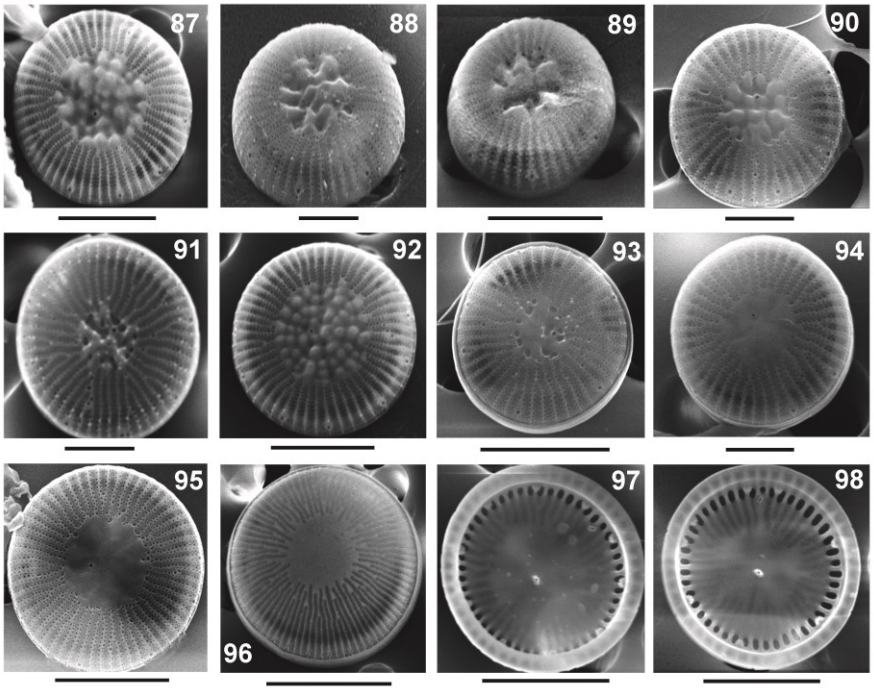
Figs 51–62



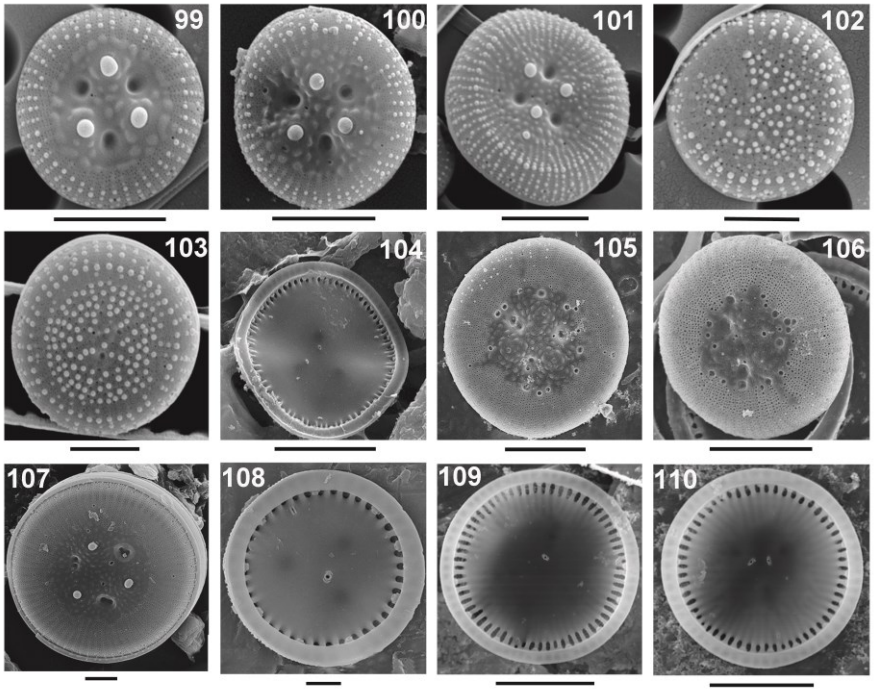
Figs 63–74



Figs 75–86



Figs 87–98



Figs 99–110

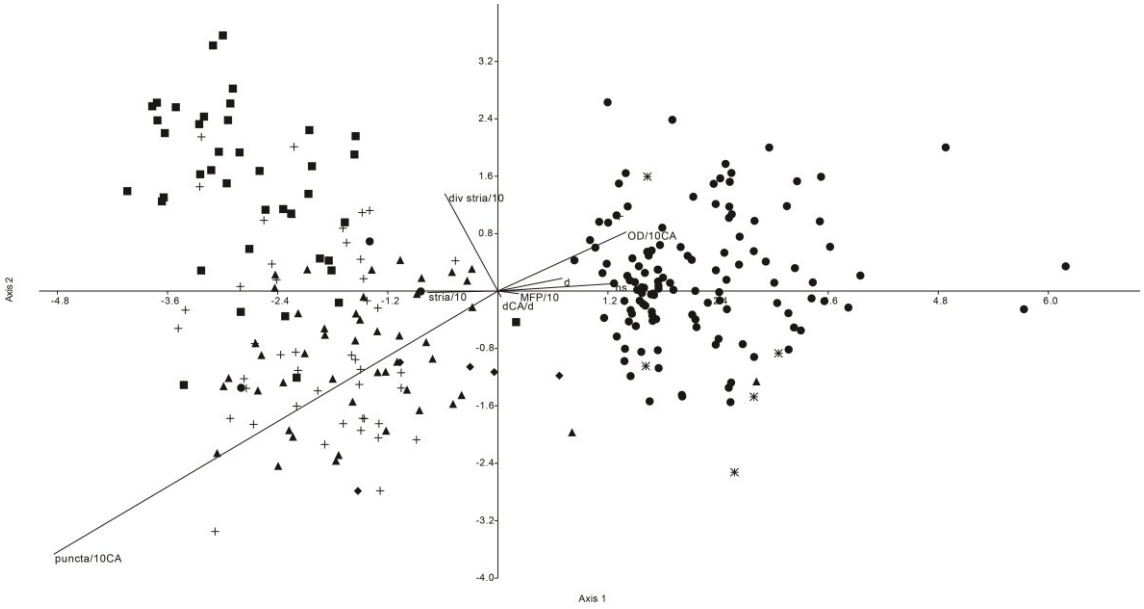

div stria/10
OD/10CA
d
ns
stria/10
MFP/10
dCA/d
puncta/10CA
Axis 1
Axis 2

Fig. 111

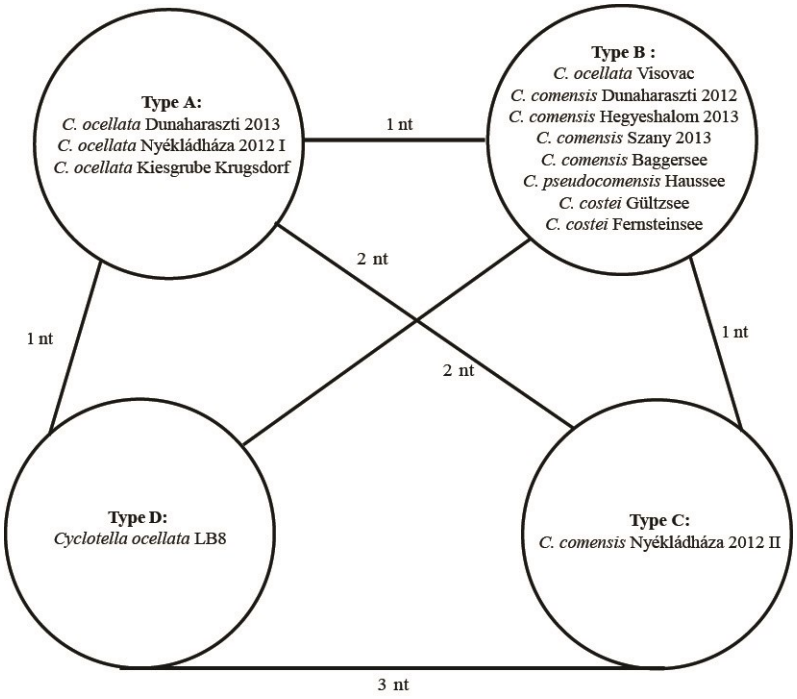

Type A:
C. ocellata Dunaharaszti 2013
C. ocellata Nyékládháza 2012 I
C. ocellata Kiesgrube Krugsdorf
Type B :
C. ocellata Visovac
C. comensis Dunaharaszti 2012
C. comensis Hegyeshalom 2013
C. comensis Szany 2013
C. comensis Baggersee
C. pseudocomensis Haussee
C. costei Gültzsee
C. costei Fernsteinsee
Type D:
Cyclotella ocellata LB8
Type C:
C. comensis Nyékládháza 2012 II
1 nt
2 nt
1 nt
2 nt
1 nt
3 nt

Fig. 112

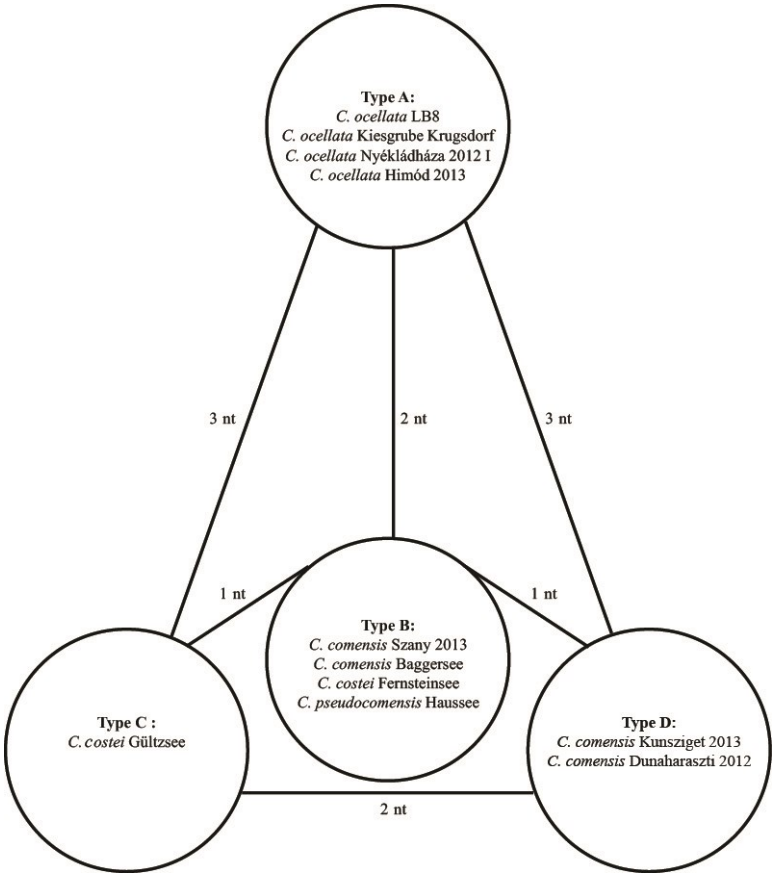

Type A:
C. ocellata LB8
C. ocellata Kiesgrube Krugsdorf
C. ocellata Nyékládháza 2012 I
C. ocellata Himód 2013
3 nt
2 nt
3 nt
1 nt
1 nt
Type B:
C. comensis Szany 2013
C. comensis Baggersee
C. costei Fernsteinsee
C. pseudocomensis Haussee
Type C :
C. costei Gültzsee
Type D:
C. comensis Kunsziget 2013
C. comensis Dunaharaszti 2012
2 nt

Fig. 113

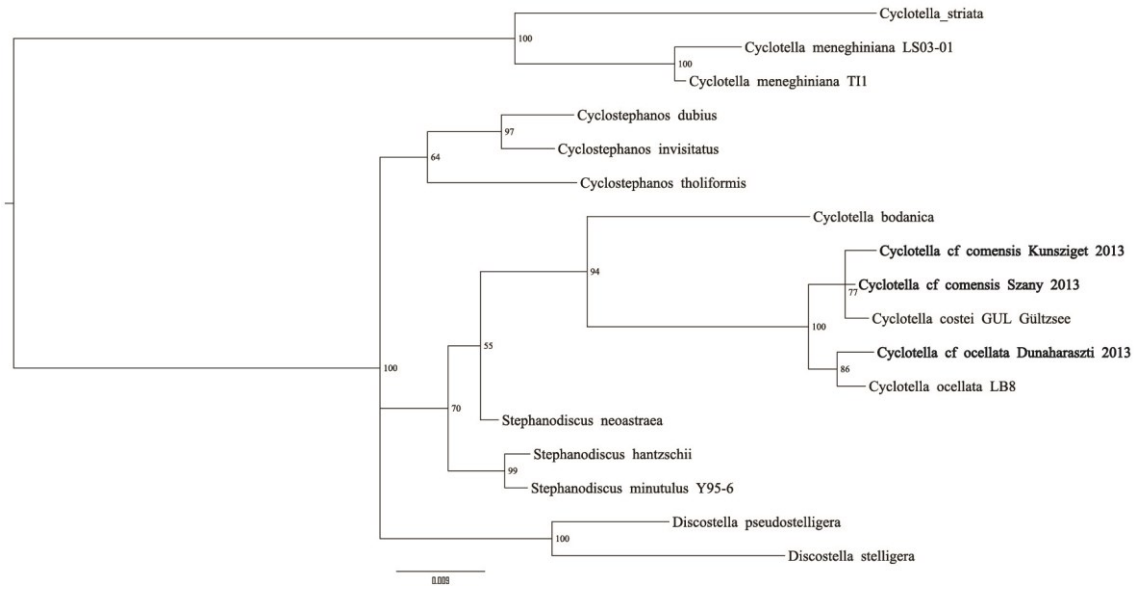

Cyclotella_striata
100
Cyclotella meneghiniana LS03-01
100
Cyclotella meneghiniana TI1
Cyclostephanos dubius
97
Cyclostephanos invisitatus
64
Cyclostephanos tholiformis
Cyclotella bodanica
Cyclotella cf comensis Kunsziget 2013
94
Cyclotella cf comensis Szany 2013
77
Cyclotella costei GUL Gültzsee
100
55
Cyclotella cf ocellata Dunaharaszti 2013
100
86
Cyclotella ocellata LB8
70
Stephanodiscus neoastraea
Stephanodiscus hantzschii
99
Stephanodiscus minutulus Y95-6
Discostella pseudostelligera
100
Discostella stelligera
0.009

Fig. 114