Phycological Research

Phycological Research 2014; **62:** 282–293

Morphometric variation of *Seminavis pusilla* (Bacillariophyceae) and its relationship to salinity in inter-dune lakes of the Badain Jaran Desert, Inner Mongolia, China

Patrick Rioual,* Yanbin Lu, Guoqiang Chu, Bingqi Zhu and Xiaoping Yang
Key Laboratory of Cenozoic Geology and Environment, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China

SUMMARY

We used light and scanning electron microscope analyses to quantify morphometric features (valve length, width, stria density, lineola density and valve curvature) from the observation of valves representing *Seminavis pusilla*. Cluster analysis based on Gaussian mixture models and the expectation-maximization algorithm was used for delineating two species, *Seminavis pusilla sensu stricto* and *Seminavis lata* (Krammer) Rioual comb. et stat. nov. By comparison with *S. pusilla*, *S. lata* is characterized by wider valves and lower stria density. The two species have also markedly different ecology. *S. pusilla* is most abundant in the most saline lakes of the dataset, while *S. lata* is most abundant in the less saline lakes. Our results indicate that combining the two species into *S. pusilla sensu lato* would lead to a loss of ecological information and a decrease of the performance of transfer functions developed for quantitative reconstruction of past salinity from fossil diatom assemblages in sediment cores.

Key words: Badain Jaran Desert, Expectation-Maximization algorithm, morphometric analysis, model-based clustering, *Navicymbula*, salinity, *Seminavis lata* comb. et stat. nov., *Seminavis pusilla*.

INTRODUCTION

The genus *Seminavis* was erected by D.G. Mann (in Round *et al*. 1990) to accommodate species previously classified in *Amphora* that possess uniseriate striae with slit-like areolae, as in *Navicula* Bory *sensu stricto* (*s.s.*) (Danielidis & Mann 2002). However, the plastid arrangement and the structures of striae, raphe and girdle indicate that *Seminavis* is not closely related to *Amphora* but instead has been derived from diatoms allied to *Navicula s.s.* (Round *et al*. 1990). Most species of *Seminavis* described so far are marine epipelic and epiphytic diatoms (Round *et al*. 1990; Witkowski *et al*. 2000; Danielidis & Mann 2002, 2003; Garcia 2007; Wachnicka & Gaiser 2007). However, *Seminavis* can also occur in non-marine (albeit brackish) habitats, as evidenced by the inclusion in *Seminavis* of *Cymbella pusilla* Grunow (=*Seminavis pusilla* (Grunow) Cox & Reid 2004) and *Amphora strigosa* Hust. (=*Seminavis strigosa* (Hust.) Danielidis & Economou-Amilli).

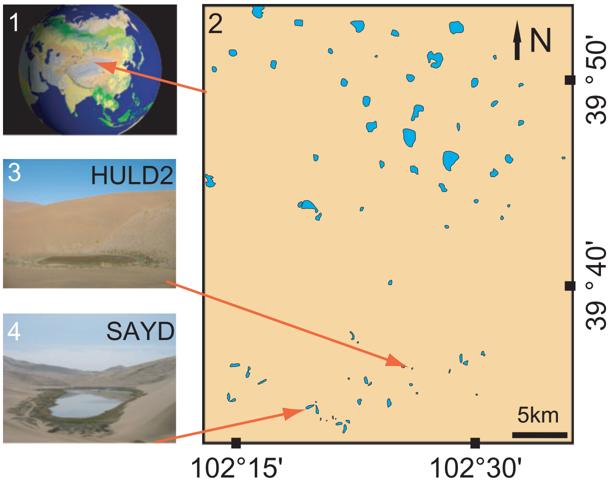
Seminavis pusilla is common in brackish waters in both continental and coastal areas, as well as in calcium-rich springs (Snoeijs 1993; Krammer 2003). Unfortunately, its transfer has been overlooked by most authors working on non-marine diatoms, perhaps in part because Krammer (2003) had erected the genus *Navicymbula* based on the same species, and the combination *Navicymbula pusilla* (Grunow) Krammer has been regularly used in the literature (e.g. Metzeltin *et al*. 2009; Nakov *et al*. 2009; Stavreva-Veselinovska & Todorovska 2010; Hofmann *et al*. 2011; Potapova 2011; Rioual *et al*. 2013). Besides the nominate variety, Krammer (2003) distinguished two other varieties that differ in size and shape: *Navicymbula pusilla* var. *lata* Krammer and *N. pusilla* var. *rotunda* Krammer. However, no morphological and/or morphometric analysis was proposed to support the separation of these varieties. Moreover, most authors do not distinguish different varieties for this taxon (e.g. Gasse 1986; Cumming *et al*. 1995; Rumrich *et al*. 2000).

During a study of the diatom flora of surface sediments of the inter-dunal lakes of the Badain Jaran desert (Inner Mongolia, China), *S. pusilla sensu lato* (*s.l.*) was found in all samples investigated and was often the dominant diatom. The purpose of this paper is to illustrate morphological variation of *S. pusilla*, to decide whether or not the distinction of morphotypes is

---

**To whom correspondence should be addressed.*
Email: prioual@mail.igcas.ac.cn

Communicating editor: K. Toyoda
Received 7 June 2013; accepted 7 April 2014.
doi: 10.1111/pre.12063



Figs 1–4. Geographical location of the study region. 1. Satellite image of Asia (image credit: Visible Earth, NASA). 2. Close-up on the south-east corner of the Badain Jaran Desert showing the location of the lakes sampled in this study. 3. View of the subsaline lake Hulusitai Dong Hu (HULD2). 4. View of the mesosaline lake Sayinwusu Dong Hu (SAYD).

justified and to investigate its ecological distribution in the lakes of the Badain Jaran desert.

The Badain Jaran sandy desert, located in the western part of Inner Mongolia (China) is characterized by its unique landscape in which coexist megadunes, the tallest ones in the world, and over one hundred permanent spring-fed lakes (Yang 2000; Dong *et al.* 2004) (Figs 1,2). These lakes show a wide range in salinity, from subsaline to hypersaline (Hofmann 1996; Yang & Williams 2003) (Figs 3,4).

For this study we focused on the southeastern corner of the Badain Jaran desert, between latitudes 39°33′ to 39°44′ N and the longitudes 102°14′ to 102°31′ E. Surface morphology consists of sand dunes with a maximum height of 460 m. The regional climate is strongly continental and hyperarid. At the Ekenhuduge weather station located at the southern edge of the desert, the mean annual temperature is 7.7°C, mean monthly temperatures ranges from –10°C in January to +25°C in July. The average annual precipitation is 118 mm but inter-annual variation is high. Diurnal temperatures in summer months range from 0 to >40°C (Yang & Williams 2003; Yang *et al.* 2003b; Ma *et al.* 2009).

MATERIAL AND METHODS

Field and laboratory methods

Using an Uwitec gravity corer, surface-sediment samples (the uppermost 1 cm) were taken from the deepest point of 26 lakes between June 2007 and October 2009 (Table 1). These sediment samples were prepared using the water-bath method (Renberg 1990). Sub-samples of the cleaned, homogenized suspension were diluted with distilled water and left to settle onto glass coverslips and dry overnight in ambient conditions. Coverslips were mounted on glass slides with Naphrax (a commonly used mountant with a refraction index of 1.73). Counting was made with a Leica DM LB2 light microscope (Leica Microsystems, Wetzlar, Germany) using bright-field and phase contrast oil immersion optics at 1000 × magnification. At least 500 diatom valves were enumerated to establish abundance values expressed as relative percentages.

Morphometric analysis was done using both light and scanning electron microscopy.

For light microscope (LM) analysis, measurements of valve length, width and dorsal stria density were made on digitized images calibrated against a slide micrometer using the public domain ImageJ software (Rasband 2013). Dorsal stria density was measured in the region of the valve that extends from the apical margin of the central area towards the poles and immediately alongside the raphe-sternum (Mann *et al.* 2004). Practically, the distance spanned by ten striae located between the margin of the central area up to the Voigt discontinuity (Fig. 5) was measured and algebraically converted to a number of striae per 10 μm. In rare instances where it was not possible to measure ten striae, at least eight striae were measured (Cortese & Gersonde 2007; Kaczmarska *et al.* 2009). Thirty valves were investigated from each lake apart from two lakes in which *Seminavis* was not abundant (Table 1). In total 750 valves were investigated under the LM.

For scanning electron microscope (SEM) analyses, aliquots of cleaned samples from two lakes were air-dried onto glass coverslips of 12 mm diameter. These were mounted onto aluminum stubs using Ted Pella double-coated carbon conductive adhesive tape (Ted Pella, Inc., Redding, CA, USA) and coated with gold using a Polaron SC7640 sputter coater (Quorum Technologies, East Grinstead, UK) for viewing on a LEO 1530 VP (LEO Electronenmikroskopie GmbH, Oberkochen, Germany) at the state key laboratory of paleobiology and stratigraphy (Chinese Academy of Sciences, Nanjing, China). Taxonomic metrics (length, width at mid-valve and width near the apices taken at 95% of the length, dorsal stria density and lineola density) were taken from digitized images using ImageJ. For lineola density, the length of five lineolae was measured on the stria immediately adjacent to the Voigt discontinuity (Fig. 5). These values were converted to number of lineolae per 10 μm (Cortese & Gersonde 2007). Dorsal and

Table 1. Location of the 26 lakes sampled, relative abundances of *Seminavis pusilla s.l.* in the surface sediment samples and number of specimens included in the light microscopy (LM) study

Lake name	Code	Latitude	Longitude	Salinity (g L^{-1})	EC (µS cm^{-1})	*S. pusilla s.l.* %	*n*
Hulusitai Dong Hu 2	HULD2	39.36.032N	102.26.037E	1.1	1874	42.1	30
Mudan Jilin	MUDA	39.35.744N	102.14.591E	1.3	1371	22.1	30
Wosigetu Nuoer	WOSI	39.36.872N	102.30.486E	1.3	1443	2.5	10
Halareritu	HALA	39.37.353N	102.22.200E	1.3	1951	16.2	30
Baoritelegai Nan Hu	BAON	39.36.030N	102.28.434E	1.4	2161	37.8	30
Baoritelegai Dong Hu	BAOD	39.36.391N	102.29.295E	1.5	2488	16.1	30
Zhunaogeqi	ZHUN	39.35.255N	102.26.033E	1.6	2303	38.4	30
Badan Dong Hu	BADD	39.33.120N	102.21.859E	1.6	2478	16.6	30
Baoritelegai Xi Hu	BAOX	39.36.452N	102.29.199E	1.8	2803	11.0	30
Aer Jilin	AERJ	39.34.319N	102.14.485E	2.0	2441	16.6	30
Women Jilin	WOME	39.34.974N	102.16.569E	2.1	2860	30.5	30
Xiao Haizi	XIAH	39.33.515N	102.29.217E	2.1	3096	41.8	30
Dugui Jilin Dong Hu	DUGD	39.35.167N	102.23.430E	2.3	3257	10.7	30
Hulusitai Dong Hu 1	HULD1	39.36.092N	102.25.484E	2.4	2357	7.2	30
Shao Bai Jilin	SHAO	39.34.037N	102.14.778E	2.4	3019	9.0	30
Halareritu Dong Hu	HALD	39.37.182N	102.22.508E	3.4	4775	37.6	30
Haisen Jilin	HAIS	39.34.842N	102.26.409E	3.9	5721	29.9	30
Hulusitai Xi Hu	HULX	39.36.124N	102.25.435E	4.0	6066	29.4	30
Dunde Jilin	DUND	39.35.714N	102.15.729E	4.9	6019	32.1	30
Dugui Jilin Xi Hu	DUGX	39.35.252N	102.23.174E	5.4	3506	6.9	30
Shuang Haizi Xi Hu	SHHX	39.43.674N	102.18.463E	6.0	8246	1.7	20
Badan Bei Hu 2	BADB2	39.33.399N	102.21.180E	9.8	11552	20.7	30
Huhewuzhu'er Xi Hu	HHWX	39.33.398N	102.20.397E	16.6	19696	36.8	30
Sayinwusu Xi Hu	SAYX	39.34.014N	102.19.867E	20.3	21547	11.5	30
Sayinwusu Dong Hu	SAYD	39.33.804N	102.20.176E	25.0	29219	46.2	30
Badan Bei Hu 3	BADB3	39.33.387N	102.21.312E	30.3	38149	33.3	30

The two lakes chosen for scanning electron microscope (SEM) analysis are shaded in grey. The lakes are listed in increasing order of salinity. Further details on the lakes and samples can be found in Rioual *et al.* (2013). EC, electrical conductivity.

ventral curvatures were measured following the method defined by Bixby and Zeek (2010) (Fig. 5). In total, 69 valves were investigated under the SEM.

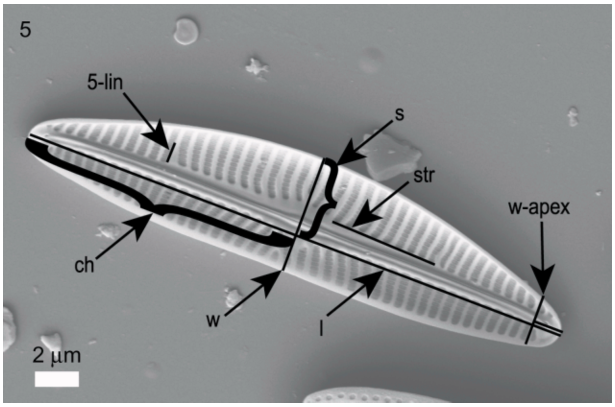


Fig. 5. Scanning electron microscope (SEM) image of *Seminavis pusilla s.l.* Interior view of the whole valve and sketch of measurements: l, valve length; w, width; w-apex, width at the apex (=95% of valve length); 5-lin, length across 5 lineolae; str, length across dorsal striae measured along the raphe (used for the conversion to stria density measurement); ch, chord; s, sagitta. Chord and sagitta were used for curvature measurements.

Environmental data

Surface water samples (~0.3 m depth) were collected by hand from the deepest part of each lake as determined either with a hand-held acoustic depth meter in deep lakes or with a ruler for shallow lakes. Lakes were sampled in May 2007, October 2008 and October 2009. In the original dataset, 26 environmental variables were measured. The complete list and the methods used to measure these various environmental variables are given in Rioual *et al.* (2013). In this study, however, we will focus only on the salinity gradient. This is because significant relationships between the distribution of *S. pusilla s.l.* could be established only with salinity and with environmental variables highly correlated with salinity (correlation established using Student's *t*-test, $\alpha = 0.001$). Such variables include, electrical conductivity (EC), alkalinity and concentrations in Cl^-, SO_4^{2-}, Mg^{2+}, K^+, Na$^+$, total nitrogen, and dissolved organic carbon. Salinity, expressed in g L^{-1} was calculated by summing the concentrations of all the major cations (Ca^{2+}, Mg^{2+}, K$^+$, Na$^+$) and anions (Cl$^-$, SO$_4^{2-}$, CO$_3^{2-}$ and HCO$_3^-$).

Data analysis

Preliminary analyses on LM observations

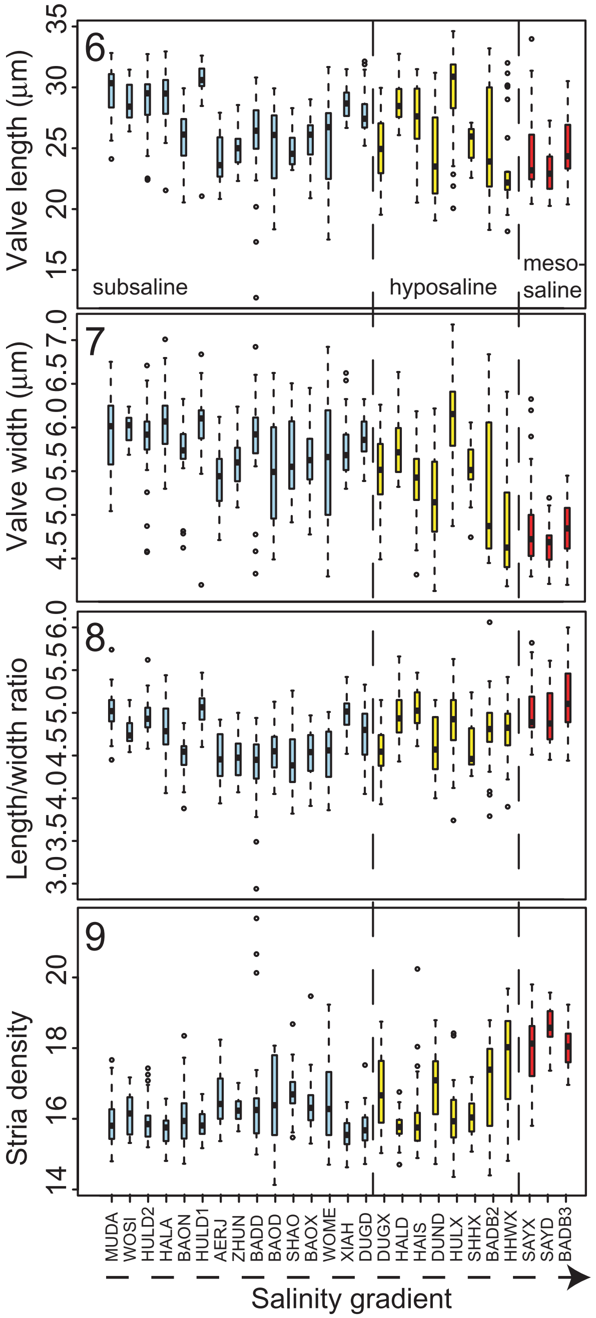
Initially, populations of *S. pusilla s.l.* were not split into morphotypes. Boxplots depicting the quantile distributions of the valve length, width, length/width ratio and density of dorsal striae were used to examine how the morphometric characteristics of the populations of *S. pusilla s.l.* varied in the 26 lakes investigated (Figs 6–9). These LM observations were also used for a preliminary assessment of the relationship between the changes in valve morphometry and the salinity gradient.

Cluster analysis of LM data

A model-based cluster analysis of the LM dataset, that includes measurements for valve length, width and stria density for 750 valves of *S. pusilla s.l.*, was used to discern and describe subgroups of individuals. With this method, Gaussian finite mixture models are fitted through the data by Expectation-Maximization (EM) algorithm and the Bayesian Information Criteria (BIC) was used to chose the best model, including its number of components (or clusters). In addition to the classification presented as a scatterplot matrix (Fig. 10), the uncertainty of the classification and the density estimates, plotted as contour surface, are also provided for all pairs of variables (see Appendix S1 in Supporting Information). Cluster analysis was performed in R using the package mclust version 4 (Fraley *et al.* 2012; R Development Core Team 2013).

Analyses on SEM observations

SEM was used to study in more detail the populations of *S. pusilla s.l.* from two lakes of contrasting salinity. In particular, SEM analysis was required in order to investigate ultrastructural characters such as lineola density. Valves of *S. pusilla* were abundant and well preserved in the surface sediment of the subsaline lake Hulusitai Dong Hu (HULD2, Fig. 3) and the mesosaline lake Sayinwusu Dong Hu (SAYD, Fig. 4). To establish which morphological characters best discriminate between these two populations, single character analyses and bivariate analyses were performed. For single character analyses, non-overlapping notches on boxplots were used to infer statistically significant dissimilarity at the 95% confidence interval (Chambers *et al.* 1983). Notched boxplots were also



Figs 6–9. Box plots showing the quantile distributions of valves characters for populations of *Seminavis pusilla s.l.* observed under the light microscope (LM) in the 26 lakes of the dataset. 6. Valve length. 7. Valve width. 8. Length/width ratio. 9. stria density (= number of striae in 10 µm). The 25–75 percent quartiles (excluding outliers) are drawn using a box. The median is shown with a horizontal line inside the box. The whiskers represent the upper and lower 'inner fence', that is, are drawn from the edge of the box up to the largest/lowest data point less than 1.5 times the box height. Outliers, i.e. values outside the inner fences, are shown as circles if they lie further from the edge of the box than three times the box height. Lakes were ordered according to their salinity in increasing order. Lake classification follows Hammer *et al.* (1983).

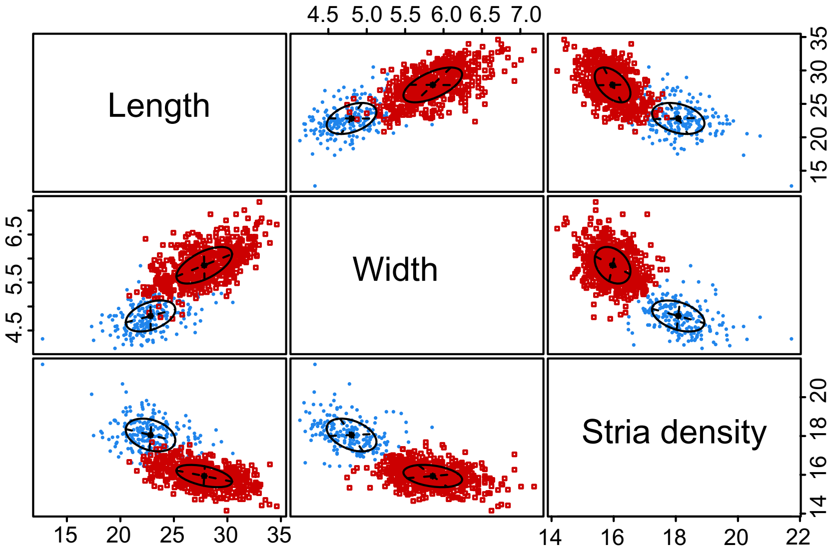
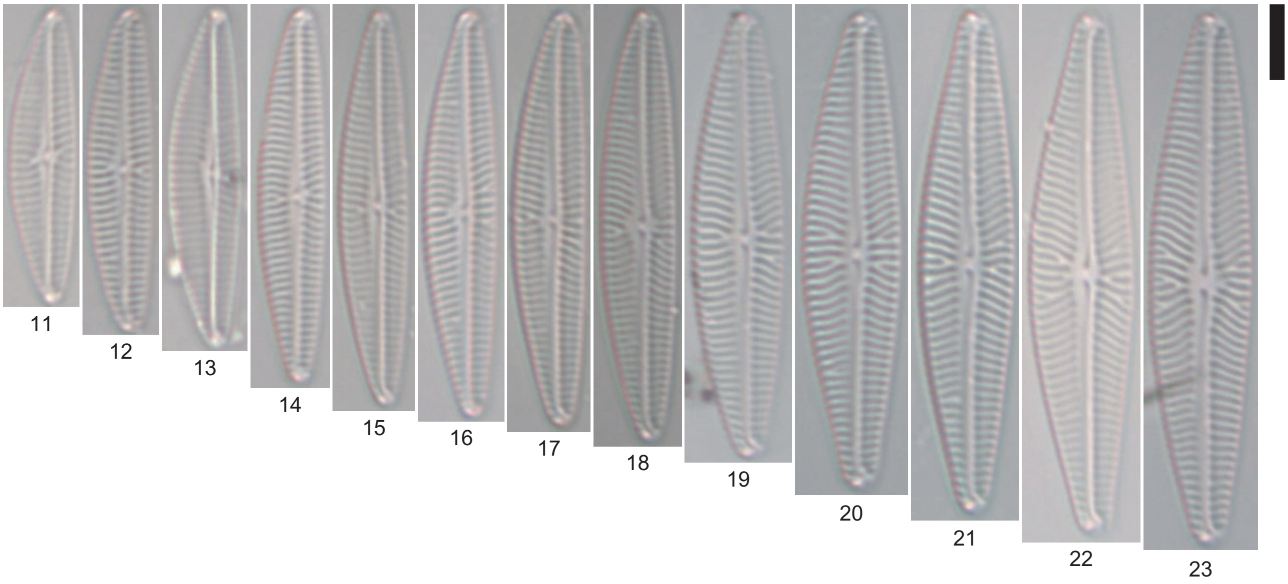


Fig. 10. Matrix of biplots showing the classification obtained by fitting Gaussian mixture models on the LM dataset of *Seminavis pusilla s.l.* (*n* = 750). The model-based clustering was performed on three morphological variables (valve length in µm, valve width in µm and stria density in number of striae in 10 µm). One cluster corresponds to *Seminavis pusilla* (blue dots), the other corresponds to *Seminavis lata* (red squares). The ellipses shown are the multivariate analogs of the standard deviations of the components (=clusters). Plots for the uncertainty of the classification and density estimates are given in Appendix S1 in Supporting Information.



Figs 11–23. LM images of *Seminavis*. Scale bar = 5 µm. 11–18. Populations of *Seminavis pusilla* from Sayinwusu Dong Hu. 19–23. Populations of *Seminavis lata* from Hulusitai Xi Hu (19,20) and Badan Bei Hu 2 (21–23).

used to identify outlier specimens from each population (Figs 32–39). Bivariate analyses based on scatterplots were then used to confirm separation between the two populations (Fig. 40). Boxplots and scatterplots were drawn using R (R Development Core Team 2013).

Relationship with salinity

Generalized additive models (GAMs), as available in the program CanoDraw 4.1 (ter Braak & Šmilauer 2002), were fitted to appraise the influence of salinity (= the predictor variable) on the distribution of *S. pusilla s.l.* in the LM dataset. To avoid the influence of uneven sampling intensity along the gradient, salinity data were log-transformed. The model complexity of GAMs was chosen by the stepwise selection procedure using the Akaike information criterion (AIC). AIC considers not only the goodness of fit but also parsimony, penalizing more complex models (Lepš & Šmilauer 2003). Poisson distribution and log link function were used in the models.

We also tested the effects of separating or merging morphotypes of *S. pusilla s.l.* on the performance of diatom-based transfer functions for salinity and EC that were built from the complete diatom dataset from the Badain Jaran desert lakes (Rioual *et al.* 2013). The transfer functions were constructed by

Table 2. Mean and standard deviation of valve characters of *Seminavis pusilla* and *Seminavis lata* observed under the LM and SEM

Morphometric feature	Mean and range	SD	Mean and range	SD
LM observations	*S. pusilla* (*n* = 212)		*S.lata* (*n* = 538)	
Valve length (μm)	22.7 (12.7–30.5)	2.3	27.8 (20.8–34.6)	2.6
Valve width (μm)	4.8 (4.1–5.8)	0.3	5.9 (4.7–7.2)	0.4
Stria density/10 μm	18.1 (16.4–21.7)	0.8	15.9 (14.1–17.7)	0.6
SEM observations	SAYD (n = 36)		HULD2 (n = 30)	
Valve length (μm)	24.3 (19.2–30.4)	2.4	29.7 (27.7–31.5)	1.0
Valve width (μm)	4.6 (4.1–5.3)	0.3	5.9 (5.4–6.3)	0.2
Stria density/10 μm	18.2 (17.0–19.2)	0.5	17.4 (15.2–16.3)	0.5

SD, standard deviation.

Weighted-Averaging Partial Least Square (WA-PLS, ter Braak & Juggins 1993) using the program C2 v. 1.6.5 (Juggins 2010).

RESULTS

Distribution of *Seminavis pusilla s.l.*

Visual inspection under the LM of valves of *S. pusilla s.l.* indicated the presence of two forms which co-exist in most lakes, as shown by the wide range in length, width and stria density measures in many, but not all, lakes of the dataset (Figs 6–9). However, there are obvious trends across the salinity gradient as specimens in subsaline (0.5–3 g L^{-1}) and hyposaline (3–20 g L^{-1}) lakes tend to have wider valves (Fig. 7) with lower stria density (Fig. 9) than specimens in mesosaline lakes (20–50 g L^{-1}). There are no such obvious trends in valve length (Fig. 6) and length/width ratio (Fig. 8).

Results of cluster analysis of the LM dataset

The best model, that is, with the highest BIC value, was the VVV model with ellipsoidal distribution, varying volume, shape, and orientation and with two components (Fig. 10). With this model, the 750 valves of the LM dataset are split into two clusters. Summary statistics for the two clusters are given in Table 2. One cluster, that includes 212 valves, is characterized by narrow valves with high stria density. Valves from this cluster (Figs 11–18) match with illustrations given in Krammer (2003) for *Navicymbula pusilla* var. *pusilla*. We will refer to this cluster as *Seminavis pusilla*. The second cluster, with 538 valves, is characterized by wider valves with lower stria density (Figs 19–23) that match the illustrations given in Krammer (2003), for *Navicymbula pusilla* var. *lata*.

Taxonomic transfer

The fact that valves from the two clusters co-occur in the same samples preclude the use of the variety level as varieties should be geographically separated from each other. Therefore, we propose a new combination for the formal transfer of *Navicymbula pusilla* var. *lata* to the genus *Seminavis* and to elevate this variety to the species level.

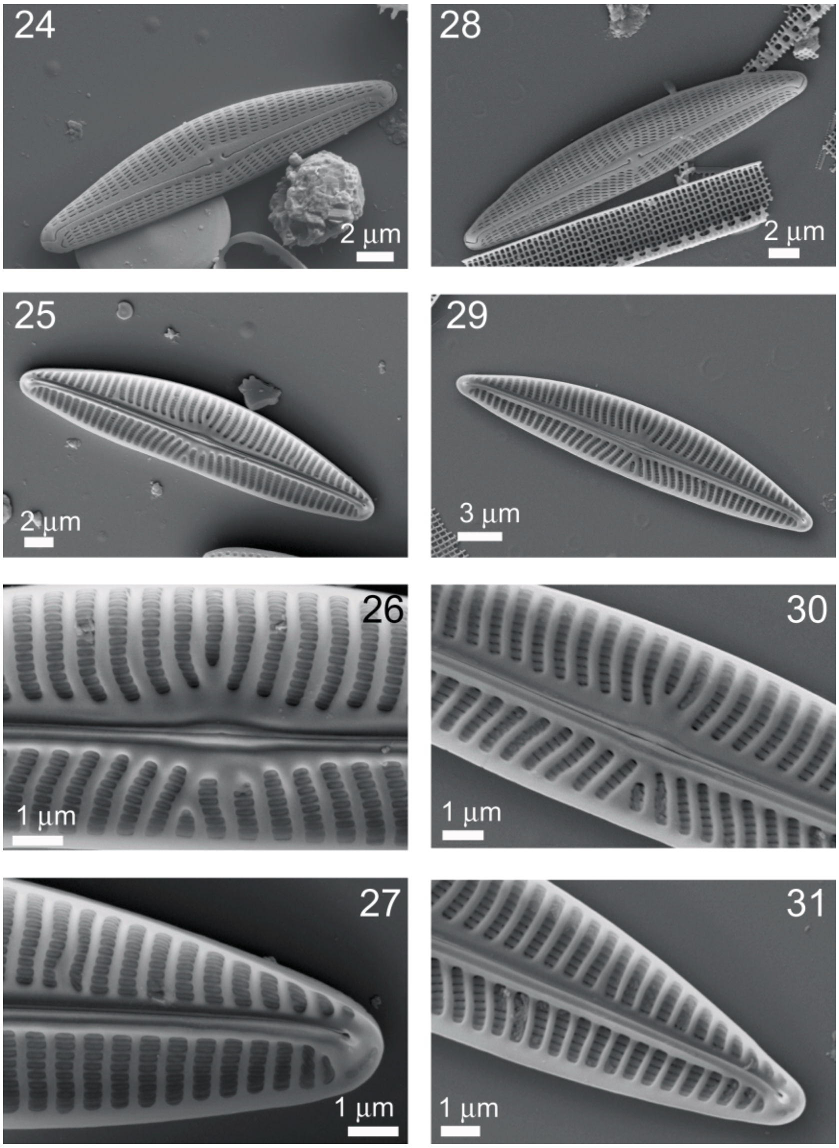
Seminavis lata (Krammer) Rioual comb. et stat. nov.

Basionym: *Navicymbula pusilla* var. *lata* Krammer 2003 in Krammer 2003: Diatoms of Europe volume 4 *Cymbopleura, Delicata, Navicymbula, Gomphocymellopsis, Afrocymbella*, Plate 138, Figs 14–17.

Krammer (2003) distinguished var. *lata* by its broadly rounded apices. Our results indicate that var. *lata* also differs from the nominate variety by its width (4.7–7.2 instead of 4.1–5.8 μm) and stria density (14.1–17.7 instead of 16.4–21.7 striae/10 μm).

Complementary SEM observations of qualitative characters

In the SEM, specimens from HULD2 that match *S. lata* (Figs 24–27) and those from SAYD that match *S. pusilla* (Figs 28–31) show no difference in ultrastructural features. The areolae are formed of transapically elongated lineolae. The exterior fissure of the raphe is straight, and is characterized by proximal ends dorsally deflected and distal ends dorsally recurved following a subterminal ventral deflection (Figs 24,28). The interior raphe fissure is located on a narrow costa with its proximal ends slightly dorsally curved (Figs 26,30) and well-developed helictoglossae at the distal ends (Figs 27,31). The Voigt fault is well-developed internally and externally. These observations fully conform to the description given by Krammer (2003).



Figs 24–31. SEM images of *Seminavis pusilla s.l.* Populations of *Seminavis lata* from lake HULD2 (24–27) and *Seminavis pusilla* from SAYD (28–31). 24 and 28. Exterior view of the whole valve showing transapically elongated lineolae, straight raphe with dorsally deflected proximal ends and dorsally recurved distal ends. 25 and 29. Interior view of the whole valve. 26 and 30. Interior view of the centre showing the central nodule with a narrow costa bearing the raphe fissure and uncorroded hymens occluding the lineolae. 27 and 31. Interior view of the distal end showing the helictoglossae.

Separation between species of *Seminavis* based on SEM observations

Analyses for single characters

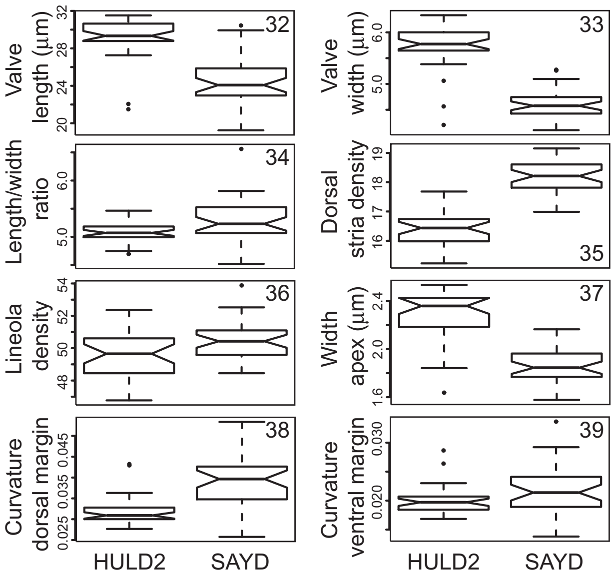
Notched boxplots show that valves from the subsaline lake HULD2 were dissimilar to those of the mesosaline lake SAYD in length (Fig. 32), width (Fig. 33), dorsal stria density (Fig. 35), width of apex (Fig. 37) and curvature of the dorsal margin (Fig. 38). Conversely, length/width ratio (Fig. 34), lineola density (Fig. 36) and curvature of the ventral margin (Fig. 39) were not significantly different between the two populations. Two or three valves from HULD2 appear as outliers in this population.

Bivariate analysis

Bivariate analysis confirms the clear separation between the two populations as well as the presence of three outliers in the population from HULD2 (Fig. 40). In particular, the scatterplot length vs width reveals that the two populations are clearly separated and occupy different size trajectories (cf. Mann *et al*. 2004). In summary, SEM observations confirmed the results based on LM observations (Table 2).

Species response models

Results of the Generalized Additive Modeling for the two species on the salinity and EC gradients are summarized in Table 3. Note that the regression results are almost identical for the two gradients. The response curves for the two species along the salinity gradient are given in Fig. 41 (response curves along the EC gradient, not shown, are almost identical). The two species have clearly different distributions along the salinity gradient with *S. pusilla* showing an increasing trend with



Figs 32–39. Box plots showing the quantile distributions of valves characters for populations of *Seminavis pusilla* and *S. lata* observed under the SEM in two lakes of the dataset HULD2 and SAYD. The 25–75 percent quartiles (excluding outliers) are drawn using a box. The median is shown with a horizontal line inside the box. The whiskers represent the upper and lower 'inner fence', i.e. are drawn from the edge of the box up to the largest/lowest data point less than 1.5 times the box height. Outliers, i.e. values outside the inner fences, are shown as circles if they lie further from the edge of the box than three times the box height. 32. Valve length (µm). 33. Valve width (µm). 34. Length/width ratio. 35. Dorsal stria density (number of striae in 10 µm). 36. Lineola density (number of lineolae in 10 µm). 37. Width of apex (=width at 95% length, in µm). 38. Curvature of dorsal margin (μm^{-1}). 39. Curvature of ventral margin (μm^{-1}). The SAYD population matches with the description of *S. pusilla*, while the HULD2 population is dominated by *S. lata* (see text for details).

increasing salinity, while highest abundances of *S. lata* were observed at low salinity. In most lakes, both species co-occur although *S. lata* was not found in the two most saline lakes in the dataset and *S. pusilla* was absent from several lakes of intermediate salinity or at the low end of the salinity gradient. It is interesting to note that, if we do not segregate the two species, but instead combined their abundances as *S. pusilla s.l.*, there is no significant relationship between this 'combined' taxon and the salinity gradient as shown by GAM (Table 3).

WA-PLS transfer functions

Separating the two species of *Seminavis* clearly improved the performances of the WA-PLS salinity and EC transfer functions, in terms of r^2, root-mean-square error of prediction and more noticeably in terms of maximum bias (Table 4). The salinity weighted-averaging optima and tolerance range (calculated by adding and subtracting the standard error from the optimum) of the two species are very distinct. *S. lata* has its optimum and range in subsaline waters while *S. pusilla* has a mesosaline optimum and a distribution that extends from hyposaline to mesosaline waters.

DISCUSSION

Our results show that specimens of *S. pusilla s.l.* collected from the lakes of the Badain Jaran Desert can be classified into two species based on morphometric differences, especially valve width and stria density. For *S. lata* our measurements of length and width slightly extend the range of taxonomic metrics given in the original description by Krammer (2003) (length 25–37 µm, width 5.5–7.2 µm). Stria density was not given in the original description of this taxon.

Apart from Krammer (2003), no distinctions have been made between morphotypes of *S. pusilla s.l.* in previous studies. We could speculate on which varieties were observed in such by examining LM figures, as well as the estimated salinity and/or EC optima. For example, light micrographs given in Cumming *et al.* (1995) suggest that both *S. pusilla* and *S. lata* were present in their samples (plate 45, Figs 19–24). Correspondingly, the optimum for salinity given by Wilson *et al.* (1996), a study that includes the results from Cumming *et al.* (1995), is very close to the one we obtained from the Badain Jaran dataset when we combined both species (Table 5). Similarly, the optima given for *Cymbella pusilla* by Fritz *et al.* (1993), Gell (1997) and Yang *et al.* (2003a) suggest that the populations they studied were a mixture of both species of *Seminavis*, although no illustrations were given to confirm this. The specimen illustrated by Gasse (1986), taken from a hot spring with high conductivity in Ethiopia, belongs to *S. pusilla*, although the optimum given in Gasse *et al.* (1995) could also correspond to a mixture of both species. By contrast, the high optimum and tolerance range given in Reed (1998) for the Spanish dataset, most likely indicate that the populations investigated in that case are only of *S. pusilla*.

A few other illustrations of *S. pusilla s.l.* can be found in the literature, but they seem to confirm that both species occur worldwide. For example, the unique specimen shown in Rumrich *et al.* (2000) from Laguna Santa Maria (Chile) belongs to *S. lata*, while those from Uruguay, illustrated by Metzeltin *et al.* (2005), are close to *S. pusilla*, although their stria density is lower than in the populations from the Badain Jaran lakes. It is also likely that other species of *Seminavis* are still to be formally described, as shown by Metzeltin *et al.*

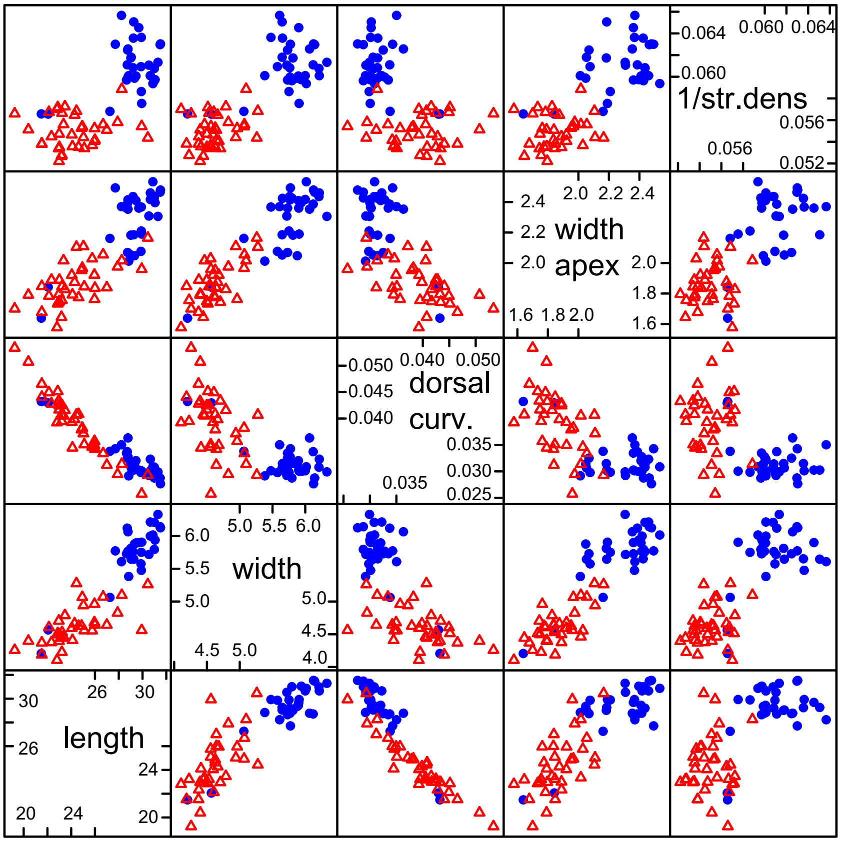


Fig. 40. Matrix of biplots of morphometric characters comparing populations from HULD2, dominated by *Seminavis lata* (filled circles), and SAYD composed of *Seminavis pusilla* (open triangle). Measurements, made on SEM images, are in μm for valve length, width, width of apex and μm^{-1} for dorsal curvature. Stria density is in number of striae per 10 μm.

Table 3. Generalized additive models regression results on the salinity and electrical conductivity gradients for the two *Seminavis* species and their combined abundances

Taxon	Predictor	Sm.df	Distribution	Link function	F	*P*	
Seminavis lata	Salinity	2	Poisson	Log	6.75	0.0050	
	EC	2	Poisson	Log	5.83	0.0091	
Seminavis pusilla	Salinity	1	Poisson	Log	40.0	<0.0000	
	EC	1	Poisson	Log	42.7	<0.0000	
Combined spp.	Salinity	2	Binomial	Logit	0.92	0.411	ns
	EC	2	Binomial	Logit	0.99	0.386	ns

EC, electrical conductivity; ns, not significant; Sm.df, degrees of freedom of the smoother term.

(2009), who found specimens that they called *Navicymbula cf. pusilla* in rivers of northeastern Mongolia. However, light micrographs of this taxon (Metzeltin *et al.* 2009; plate 156, figs 48–53) show that it does not belong to either species, discussed here, being too narrow for *S. lata* and having much lower stria density than *S. pusilla*.

Leterme *et al.* (2010) showed that areola size in *Cocconeis placentula* Ehrenberg varies with fluctuating salinity levels, and in particular, that areola size decreases with decreasing salinity. We did not measure lineola size in our study or find any trends in the lineola density between the two species investigated or across the salinity gradient. However, the lower stria density of *S. lata* compared with *S. pusilla* suggests that the percentage of 'open area' on the surface of the frustule of *S. pusilla s.l.* also decreased with decreasing salinity. Leterme *et al.* (2010) suggested that an increase in 'open area' could compensate for the decrease in diffuse layer thickness at higher salinity. Interestingly, the same phenomenon can be seen in *Navicula gregaria* Donkin as Cox (1995) observed that specimens from freshwater have narrower pore openings and smaller relative proportions of pore to solid silica than those from brackish waters.

Trobajo *et al.* (2011) also found that salinity was the main factor affecting valve length, width, stria and fibula density of several freshwater and brackish pennate diatom species. However, there was no consistent pattern in the effect of salinity on valve width as some species showed an increase in valve width with salinity, some showed a decrease and some were narrowest at intermediate salinities. By contrast stria density was found to be less affected by salinity and, among the various morphological characters investigated, can be considered to be more constant and informative in terms of taxonomy (Trobajo *et al*. 2004, 2011). If this can be generalized, the clear difference in stria density between the species of *Seminavis* would suggest that their separation is taxonomically valid and not a function of phenotypic plasticity associated with salinity.

In our dataset, salinity is the overwhelming environmental gradient and most of the variable that could have an influence on diatom species distribution are highly correlated with it. Based on our data, it is therefore difficult to assess how *S. pusilla s.l.* responds to other variables. However, culture experiments have shown that nutrient concentrations and brine composition (sulfate versus bicarbonate) are also important in explaining the distribution of *S. pusilla* (Saros & Fritz 2000, 2002).

Finally, because separating varieties clearly improves the performance of transfer functions, our results illustrate the importance of using finer taxonomic reconstruction for paleo-environmental reconstructions.

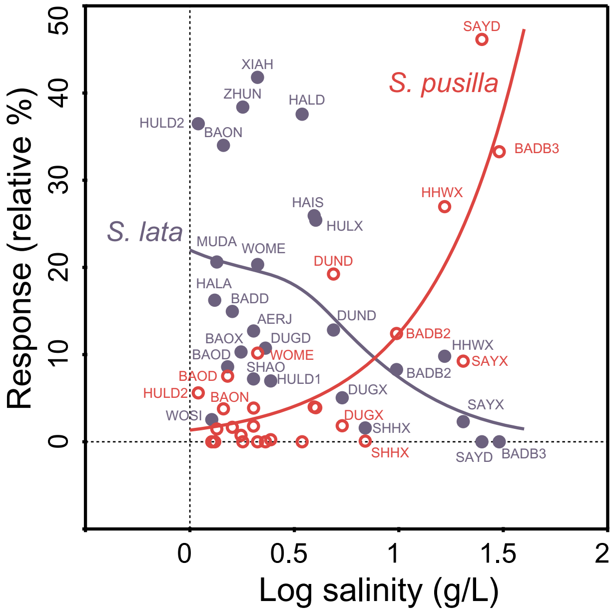


Fig. 41. Distribution along the salinity gradient of *Seminavis lata* (filled circles) and *Seminavis pusilla* (empty circles) in the 26 surface sediment samples included in the training-set and fitted generalized additive models of the taxa response. Data based on LM observations. The split between the two species was made in each sample according to the model-based cluster analysis.

Table 4. Performance of the weighted-averaging partial least square (WA-PLS) (component 2) transfer functions developed with and without distinguishing between the two species of *Seminavis*

Model	Variable	r^2	Max bias	boot r^2	boot Max bias	RMSEP
Seminavis pusilla sensu lato	Salinity (log g L^{-1})	0.96	0.159	0.85	0.337	0.183
Two species separated	Salinity (log g L^{-1})	0.97	0.090	0.87	0.238	0.171
Seminavis pusilla sensu lato	EC (log mS cm^{-1})	0.96	0.161	0.87	0.336	0.162
Two species separated	EC (log mS cm^{-1})	0.97	0.105	0.90	0.276	0.154

Boot, bootstrapping method used for internal validation; RMSEP, root mean square error of prediction.

Table 5. WA-PLS optima and tolerance (optima ± standard error) for salinity and electrical conductivity (EC) for the two species of *Seminavis* (taken separately and as a combined taxon) and comparison with the literature

Variable	Reference	Region/country	Taxon	Occur.	Max%	N2	Opt	Opt – SE	Opt + SE
Salinity (g L^{-1})	1	N Great Plains	*Seminavis pusilla* (agg)	22	48.7	n/a	8.5	3.2	23.1
Salinity (g L^{-1})	2	South-east Australia	*S. pusilla* (agg)	83	80.4	28.1	7.2	4.4	19.1
Salinity (g L^{-1})	3	British Columbia	*S. pusilla* (agg)	66	99.0	24.8	5.5	0.3†	89.5†
Salinity (g L^{-1})	this study	Badain Jaran	*S. pusilla* (agg)	26	46.2	23.3	4.8	3.4	6.8
Salinity (g L^{-1})	this study	Badain Jaran	*Seminavis lata*	24	41.8	20.7	2.0	1.6	2.5
Salinity (g L^{-1})	this study	Badain Jaran	S. *pusilla*	20	46.2	13.8	16.9	11.0	26.0
EC (mS cm^{-1})	4	Spain	*S. pusilla* (agg)	27	61.3	9.13	26.3	12.2	56.6
EC (mS cm^{-1})	5	North and East Africa	*S. pusilla* (agg)	52	64.6	14.8	9.1	3.4	24.5
EC (mS cm^{-1})	6	Qinghai/Tibet	*S. pusilla* (agg)	10	7.3	4.9	10.5	n/a	n/a
EC (mS cm^{-1})	this study	Badain Jaran	*S. pusilla* (agg)	26	46.2	23.3	6.7	4.9	9.1
EC (mS cm^{-1})	this study	Badain Jaran	S. *lata*	24	41.8	20.7	2.9	2.4	3.4
EC (mS cm^{-1})	this study	Badain Jaran	S. *pusilla*	20	46.2	13.8	19.7	13.0	30.0

References: [1]Fritz *et al.* (1993); [2]Gell (1997); [3]Wilson *et al.* (1996); [4]Reed (1998); [5]Gasse *et al.* (1995); [6]Yang *et al.* (2003a). †For Wilson *et al.* (1996), we show the range of salinity in which *S. pusilla* occurred instead of the standard error (SE) on the optima.

ACKNOWLEDGMENTS

This work was funded by the Chinese Academy of Sciences (grant number: kzcx2-yw-119). With thank Dr Yan Fang, State key laboratory of paleobiology and stratigraphy, Chinese Academy of Sciences, Nanjing, for his assistance with the scanning electron microscopy and Dr Luca Scrucca, Università degli studi di Perugia, Italy, for his help with the 'mclust' R package.

REFERENCES

Bixby, R. J. and Zeek, E. C. 2010. A simple method for calculating valve curvature. *Proc. Acad. Nat. Sci. Philadelphia* **160**: 73–81.

Chambers, J. M., Cleveland, W. S., Kleiner, B. and Tukey, P. A. 1983. *Graphical methods for data analysis*. Wadsworth & Brooks/Cole Publishing Company, Pacific Grove, USA.

Cortese, G. and Gersonde, R. 2007. Morphometric variability in the diatom *Fragilariopsis kerguelensis*: implications for Southern Ocean paleoceanography. *Earth Planet. Sci. Lett.* **257**: 526–44.

Cox, E. J. 1995. Morphological variation in widely distributed diatom taxa: taxonomic and ecological implications. *In* Marino, D. and Montresor, M. (Eds) *Proceedings of the Thirteenth International Diatom Symposium*. Biopress Limited, Bristol, pp. 335–45.

Cox, E. J. and Reid, G. 2004. Generic relationships within the Naviculineae: a preliminary cladistic analysis. *In* Poulin, M. (Ed.) *Proceedings of the Seventeeth International Diatom Symposium*. Biopress Limited, Bristol, pp. 49–62.

Cumming, B. F., Wilson, S. E., Hall, R. I. and Smol, J. P. 1995. Diatoms from British Columbia (Canada) lakes and their relationship to salinity, nutrients and other limnological variables. *In* Lange-Bertalot, H. (Ed.) *Bibliotheca Diatomologica*, *Vol. 31*. Cramer, Berlin, pp. 1–207.

Danielidis, D. B. and Mann, D. G. 2002. The systematics of *Seminavis* (Bacillariophyta): the lost identities of *Amphora angusta, A. ventricosa* and *A. macilenta*. *Eur. J. Phycol.* **37**: 429–48.

Danielidis, D. B. and Mann, D. G. 2003. New species and new combinations in the genus *Seminavis* (Bacillariophyta). *Diatom Res.* **18**: 21–39.

Dong, Z., Wang, T. and Wang, X. 2004. Geomorphology of the megadunes in the Badain Jaran Desert. *Geomorphology* **60**: 191–203.

Fraley, C., Taftery, A. E., Murphy, T. B. and Scrucca, L. 2012. MCLUST Version 4 for R: normal mixture modeling for model-based clustering, classification, and density estimation, Technical Report no. 597, Department of Statistics, University of Washington.

Fritz, S. C., Juggins, S. and Battarbee, R. W. 1993. Diatom assemblages and ionic characterization of lakes of the Northern Great Plains, North America: a tool for reconstructing past salinity and climate fluctuations. *Can. J. Fish. Aquat. Sci.* **50**: 1844–56.

Garcia, M. 2007. *Seminavis atlantica* Garcia, a new psammic diatom (Bacillariophyceae) from southern Brazilian sandy beaches. *Braz. J. Biol.* **67**: 765–9.

Gasse, F. 1986. East African diatoms, taxonomy ecological distribution. *In Bibliotheca Diatomologica*, *Vol. 11*. J. Cramer, Berlin, pp. 1–201.

Gasse, F., Juggins, S. and Ben Khelifa, L. 1995. Diatom-based transfer functions for inferring past hydrochemical characteristics of African lakes. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **117**: 31–54.

Gell, P. A. 1997. The development of a diatom database for inferring lake salinity, Western Victoria, Australia: towards a quantitative approach for reconstructing past climates. *Aust. J. Bot.* **45**: 389–423.

Hammer, U. T., Shamess, J. and Haynes, R. C. 1983. The distribution of algae in saline lakes of Saskatchewan, Canada. *Hydrobiologia* **105**: 1–26.

Hofmann, G., Werum, M. and Lange-Bertalot, H. 2011. *Diatomeen im Süßwasser-Benthos von Mitteleuropa*. ARG Gantner Verlag K.G., Ruggell.

Hofmann, J. 1996. The lakes in the SE part of Badain Jaran Shamo, their limnology and geochemistry. *Geowissenschaften* **14**: 275–8.

Juggins, S. 2010. C2 data analysis. Version 1.6.5. University of Newcastle, Newcastle.

Kaczmarska, I., Ehrman, J. M., Moniz, M. B. J. and Davidovich, N. 2009. Phenotypic and genetic structure of interbreeding populations of the diatom *Tabularia fasciculata* (Bacillariophyta). *Phycologia* **48**: 391–403.

Krammer, K. 2003. *Cymbopleura, Delicata, Navicymbula, Gomphocymbellopsis, Afrocymbella*. *In* Lange-Bertalot, H. (Ed.) *Diatoms of Europe*, *Vol. 4*. ARG Gantner Verlag K.G., Ruggell, pp. 1–530.

Lepš, S. and Šmilauer, P. 2003. *Multivariate Analysis of Ecological Data Using CANOCO*. Cambridge University Press, Cambridge.

Leterme, S. C., Ellis, A. V., Mitchell, J. G. *et al.* 2010. Morphological flexibility of *Cocconeis placentula* (Bacillariophyceae) nanostructure to changing salinity levels. *J. Phycol.* **46**: 715–19.

Ma, J., Edmunds, W. M., He, J. and Jia, B. 2009. A 2000 year geochemical record of palaeoclimate and hydrology derived from dune sand moisture. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **276**: 38–46.

Mann, D. G., McDonald, S. M., Bayer, M. M. *et al.* 2004. The *Sellaphora pupula* species complex (Bacillariophyceae): morphometric analysis, ultrastructure and mating data provide evidence for five new species. *Phycologia* **43**: 459–82.

Metzeltin, D., Lange-Bertalot, H. and García-Rodríguez, F. 2005. Diatoms of Uruguay. *In* Lange-Bertalot, H. (Ed.) *Iconographia Diatomologica*, *Vol. 15*. ARG Gantner Verlag K.G., Ruggell, pp. 1–736.

Metzeltin, D., Lange-Bertalot, H. and Soninkhishig, N. 2009. Diatoms in Mongolia. *In* Lange-Bertalot, H. (Ed.) *Iconographia Diatomologica*, *Vol. 20*. ARG Gantner Verlag K.G., Ruggell, pp. 1–703.

Nakov, T., Spaulding, S. A. and Theriot, E. C. 2009. A preliminary phylogeny of the Cymbelloid diatoms using molecular data with emphasis on the Genera *Didymosphenia* M. Schmidt and *Navicymbula* Krammer. *J. Phycol*. **45**: 28.

Potapova, M. 2011. *Navicymbula pusilla*. Diatoms of the United States. [cited 28 April 2014]. Available from: http://westerndiatoms.colorado.edu/taxa/species/navicymbula_pusilla

R Core Team. 2013. R: a language and environment for statistical computing, R Foundation for Statistical Computing. [cited 28 April 2014]. Available from: http://www.R-project.org

Rasband, W. 2013. Image J 1.46r, National Institute of Health, USA. [cited 28 April 2014]. Available from: http://imagej.nih.gov/ij

Reed, J. M. 1998. A diatom-conductivity transfer function for Spanish salt lakes. *J. Paleolimnol.* **19**: 399–416.

Renberg, I. 1990. A procedure for preparing large sets of diatom slides from sediment cores. *J. Paleolimnol.* **4**: 87–90.

Rioual, P., Lu, Y., Yang, H. *et al*. 2013. Diatom-environment relationships and a transfer function for conductivity in lakes of the Badain Jaran Desert, Inner Mongolia, China. *J. Paleolimnol.* **50**: 207–29.

Round, F. E., Crawford, R. M. and Mann, D. G. 1990. *The Diatoms. Biology & Morphology of the Genera*. Cambridge University Press, Cambridge.

Rumrich, U., Lange-Bertalot, H. and Rumrich, M. 2000. Diatomeen der Anden. Von Venezuela bis Patagonien/Tierra del Fuego. *In* Lange-Bertalot, H. (Ed.) *Iconographia Diatomologica*, *Vol. 9*. ARG Gantner Verlag K.G., Ruggell, pp. 1–673.

Saros, J. E. and Fritz, S. C. 2000. Changes in the growth rates of saline-lake diatoms in response to variation in salinity, brine type and nitrogen form. *J. Plankton Res.* **22**: 1071–83.

Saros, J. E. and Fritz, S. C. 2002. Resource competition among saline-lake diatoms under varying N/P ratio, salinity and anion composition. *Freshw. Biol.* **47**: 87–95.

Snoeijs, P. 1993. *Intercalibration and Distribution of Diatom Species in the Baltic Sea*, *Vol. 1*. Opulus Press, Uppsala.

Stavreva-Veselinovska, S. and Todorovska, A. 2010. Ecology of the diatomic flora in thermo-mineral springs of Katlamovska Banja in the Republic of Macedonia. *Ecol. Balkanica* **2**: 1–6.

ter Braak, C. J. F. and Juggins, S. 1993. Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. *Hydrobiologia* **269/270**: 485–502.

ter Braak, C. J. F. and Šmilauer, P. 2002. CANOCO for Windows version 4.5. Wageningen, Biometris, Plant Research International.

Trobajo, R., Cox, E. J. and Quintana, X. D. 2004. The effects of some environmental variables on the morphology of *Nitzschia frustulum* (Bacillariophyta), in relation its use as a bioindicator. *Nova Hedwigia* **79**: 433–45.

Trobajo, R., Rovira, L., Mann, D. G. and Cox, E. J. 2011. Effects of salinity on growth and on valve morphology of five estuarine diatoms. *Phycol. Res.* **59**: 83–90.

Wachnicka, A. H. and Gaiser, E. E. 2007. Characterization of *Amphora* and *Seminavis* from South Florida, USA. *Diatom Res.* **22**: 387–455.

Wilson, S. E., Cumming, B. F. and Smol, J. P. 1996. Assessing the reliability of salinity inference models from diatom assemblages: an examination of a 219-lake data set from western North America. *Can. J. Fish. Aquat. Sci.* **53**: 1580–94.

Witkowski, A., Lange-Bertalot, H. and Metzeltin, D. 2000. Diatom flora of marine coasts, volume 1. *In* Lange-Bertalot, H. (Ed.) *Iconographia Diatomologica*, *Vol. 8*. ARG Gantner Verlag K.G., Ruggell, pp. 1–925.

Yang, X. D., Kamenik, C., Schmidt, R. and Wang, S. 2003a. Diatom-based conductivity and water-level inference models from eastern Tibetan (Qinghai-Xizang) Plateau lakes. *J. Paleolimnol.* **30**: 1–19.

Yang, X. P. 2000. Landscape evolution and precipitation changes in the Badain Jaran Desert during the last 30 000 years. *Chin. Sci. Bull.* **45**: 1042–7.

Yang, X. P. and Williams, M. A. J. 2003. The ion chemistry of lakes and late Holocene desiccation in the Badain Jaran desert, Inner Mongolia, China. *Catena* **51**: 45–60.

Yang, X. P., Liu, T. S. and Xiao, H. 2003b. Evolution of megadunes and lakes in the Badain Jaran Desert, Inner Mongolia, China during the last 31,000 years. *Quat. Int.* **104**: 99–112.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Uncertainties of classification and density contour for *Seminavis* LM dataset.