Phycologia (2003) Volume 42 (1), 90–97 Published 7 March 2003

Legendre shape descriptors and shape group determination of specimens in the *Cymbella cistula* species complex

JANICE L. PAPPAS[1]* AND EUGENE F. STOERMER[2]

[1]*Museum of Zoology, University of Michigan, 1109 Geddes Avenue, Ann Arbor, MI 48109-1079, USA*
[2]*Herbarium, University of Michigan, 3600 Varsity Drive, Suite 112, Ann Arbor, MI 48108-2228, USA*

J.L. PAPPAS AND E.F. STOERMER. 2003. Legendre shape descriptors and shape group determination of specimens in the *Cymbella cistula* species complex. *Phycologia* 42: 90–97.

Ten Legendre shape descriptors were calculated for 66 specimens in the *Cymbella cistula* species complex from the Great Lakes. One-hundred *x,y* co-ordinates were used to calculate the width function $W(x) = \Sigma_{n=0}^{N} b_n P_n(x)$ to obtain the Legendre coefficients (b_n) as shape descriptors. Principal components analysis (PCA) was performed on Legendre shape descriptors for the whole valve outline, to determine a shape gradient. A second PCA was performed on Legendre shape descriptors for the dorsal side of the valve outline. From this, nine shape groups were determined. From multiple discriminant analysis (DA), all but the third and eighth shape groups had near-perfect specimen assignment. Misassigned specimens from these shape groups were assigned instead to adjacent groups in PCA shape space. Overall, a curved trajectory in PCA shape space may reflect the degree of valve shape complexity. That is, there is a progression in PCA shape space from dorsally semicircular and somewhat arcuate to crescentic with a ventrally gibbous midvalve region. The distinct shape groups found indicated that several species may be present in the *C. cistula* complex.

INTRODUCTION

Legendre polynomial shape analysis has already been useful as an aid to diatom taxonomy. This quantitative technique has been applied specifically to *Gomphoneis herculeana* (Ehrenberg) Cleve (Stoermer & Ladewski 1982; Stoermer *et al.* 1984), *Tabellaria flocculosa* (Roth) Kützing (Theriot & Ladewski 1986), *Didymosphenia* M. Schmidt *in* A. Schmidt (Stoermer *et al.* 1986), *Eunotia pectinalis* (Dillwyn) Rabenhorst (Steinman & Ladewski 1987), *Surirella fastuosa* Ehrenberg (Goldman *et al.* 1990), and *Meridion* C. Agardh (Rhode *et al.* 2001). Shape has proved to be a primary, initial means to distinguish differences among individuals in a species complex.

Cymbella cistula (Ehrenberg) Kirchner contains a mixture of diverse forms (Krammer 1982). Specimens within this group grow attached to substrata by polysaccharide stalks (Hufford & Collins 1972c) and, as with other species of this genus, have dorsally directed terminal raphe fissures and ventrally directed Voigt discontinuities (Hufford & Collins 1972b; Mann 1981, 1984; Krammer 1982). Some varieties within this species complex differ from the nominate variety in the location of stigma or slight changes in valve shape, or both (Patrick & Reimer 1975). However, differences in the size–shape trajectory between populations may be indicative of species differences (Krammer 1982).

Specimens in the *C. cistula* species complex have a range of valve shapes. All have strongly dorsiventral isopolar valves [we employ descriptive terms as defined by Barber & Haworth (1981)] but subtle valve shape variations are evident. Ventrally, valve shape can range from flat to gibbous or inflated. In particular, with a semicircular dorsal valve shape, the ventral side can be flat to gibbous. Overall, valves can be crescentic or arcuate or semilanceolate with slightly produced ends. In this species complex, gradations between forms with differing degrees of inflation in the midvalve region can occur.

Historically, *C. cistula* was described generally as a half-oval that is greatly convex dorsally, almost straight ventrally, and with a central region that is slightly inflated (Hufford & Collins 1972b). Hustedt's (1930) description was that the strongly asymmetrical valves have a convex dorsal side, a concave ventral side, and a somewhat inflated middle; the ends of the valve are bluntly rounded to slightly produced and rostrate (Hustedt 1930). However, '*C. cistula*' has been used to refer to many cymbelloid taxa with isolated stigmata and there is a need for taxonomic revision (Patrick & Reimer 1975).

Other taxa are similar to *C. cistula* in valve shape and other morphological characters. *Cymbella cymbiformis* C. Agardh has papillae in the areolae (Krammer 1982), as do some specimens reported as *C. cistula.* However, other *C. cistula* do not have papillae in the areolae (Krammer 1982). *Cymbella cursiformis* Hufford & Collins was originally identified as *C. cistula* var. *gracilis* Hustedt (Hufford & Collins 1972a). However, Hufford & Collins (1972a) concede that this taxon may be a post-auxospore (or initial) cell of *C. cistula.* Alternatively, it may be a teratological form of *C. cistula* (Krammer 1982). Krammer & Lange-Bertalot (1986) consider that *C. cursiformis* is a separate taxon from both *C. cistula* and *C. simonsenii* Krammer.

Cymbella cistula var. *maculata* (Kützing) Van Heurck was separated from the nominate variety only on the basis of absence of an isolated stigma (Hufford & Collins 1972b). This taxon is similar to *C. cymbiformis* var. *nonpunctata* Fontell and *C. affinis* Kützing (Patrick & Reimer 1975). *Cymbella cistula, C. cistula* var. *gibbosa* Brun, and *C. cistula* var. *truncata* Brun have been identified in the North America Great Lakes (Stoermer & Yang 1969). However, *C. cistula* var. *gib-*

* Corresponding author (jlpappas@umich.edu).

bosa has sometimes been considered to be a separate species, *C. gibbosa* (Brun) Meister (*non* Pantocsek) (VanLandingham 1969), differing from the nominate variety only in valve shape (Patrick & Reimer 1975).

Recently, Krammer (2002) has reclassified taxa from the *C. cistula* species complex and *C. subcistula, C. neocistula, C. neocistula* var. *islandica,* and *C. neocistula* var. *lunata* have all been recognized as distinct new taxa.

For our study, Legendre shape analysis was conducted to explore variation within the *C. cistula* species complex. We chose this method of shape analysis because it is suitable for gently curving diatom outlines (Theriot & Ladewski 1986; Mou & Stoermer 1992). We were interested in determining the degree to which *C. cistula* specimens have quantitatively distinct shapes and what this implies in an ontogenetic sense.

MATERIAL AND METHODS

Background on Legendre polynomials

From Laplace's equation in the form of

$$\frac{d}{dx}\left[(1 - x^2)\frac{dP_n(x)}{dx}\right] - \frac{m^2}{1 - x^2}P_n(x) + \lambda P_n(x) = 0$$

where $x = \cos\theta$ and $\Theta(\theta) = P(\cos\theta)$ on the interval $-1 < x < 1$, for $m = 0$, we want to solve eigenfunctions in the form of

$$\frac{d}{dx}\left[(1 - x^2)\frac{dP_n(x)}{dx}\right] + \lambda P_n(x) = 0.$$

We want a solution in the neighbourhood of $x = 1$ as a power series in $(x - 1)$ given by

$$P_n(x) = (x - 1)^\alpha \sum_{k=0}^{\infty} c_k(x - 1)^k.$$

Setting coefficient c_k to zero results in the recursion formula

$$c_{k+1} = -\frac{[k(k + 1) - \lambda]}{2(k + 1)^2}c_k,$$

and $\Sigma\, c_k(x - 1)^k$ converges for $|x - 1| < 2$. Because of recursion,

$$\lambda_n = n(n + 1)$$

where $n = 0, 1, \ldots$ if and only if $P_n(x)$ is bounded by $-1 < x < 1$ (Weinberger 1995). Corresponding to λ_n with $c_0 = 1$, $P_n(x)$ is

$$P_n(x) = \sum_{k=0}^{n} \frac{(n - k)!}{(n - k)!(k!)^2 2^k}(x - 1)^k$$

which is defined as a Legendre polynomial of degree n in x (Weinberger 1995) with density function $\rho(x) = 1$ (Szokefalvi-Nagy 1965). The orthogonality condition is met because

$$\int_{-1}^{1} P_m(x)P_n(x)\,dx = 0$$

for $m \neq n$ (Farrell & Ross 1963). The first two Legendre polynomials are defined as $P_0(x) = 1$ and $P_1(x) = x$. Each successive nth degree Legendre polynomial may be calculated as

$$P_n(x) = \frac{2n - 1}{n}xP_{n-1}(x) - \frac{n - 1}{n}P_{n-2}(x).$$

Diatom shape analysis

In shape analysis, a width function, $W(x)$, is defined as

$$W(x) = \sum_{n=0}^{N} b_n P_n(x)$$

where b_n represents n-Legendre coefficients from expansion of the width function (Stoermer & Ladewski 1982). $P_n(x)$ is the non-normalized nth Legendre polynomial, where equidistant values of x range from -1 to 1. The width function is a linear combination of Legendre polynomials of degree n in x or an orthogonal polynomial regression (Stoermer & Ladewski 1982). From this, Legendre coefficients serve as shape descriptors.

Each half of a diatom is treated separately to ensure calculation of Legendre coefficients on the interval $(-1,1)$. Co-ordinates for the top half of the diatom already qualify by normalization of length to $(-1,1)$. Co-ordinates for the bottom half of the diatom are translated to their mirror image in order to qualify by normalization on $(-1,1)$. From this, each half of the diatom will have a separate set of Legendre coefficients as shape descriptors. Legendre coefficients calculated in matrix algebraic form, **b**, are

$$\mathbf{b} = [P_n(\mathbf{x})^{\dagger}P_n(\mathbf{x})]^{-1}[P_n(\mathbf{x})^{\dagger}\mathbf{y}].$$

To ensure that a sufficient number of coefficients have been calculated to represent diatom shape, reconstruction of the diatom outline is necessary. To accomplish this, x,y co-ordinates may be calculated from Legendre polynomials as

$$\mathbf{x} = \frac{n}{2n - 1}\,\frac{P_{n+2}(\mathbf{x}) + \frac{n - 1}{n}P_n(\mathbf{x})}{P_{n+1}(\mathbf{x})}$$

where $n = 0, 1, 2, \ldots,$ and

$$\mathbf{y} = \bar{b} - \sum_{k=1}^{n} \mathbf{b}_{k,1}P_{n,k}(\mathbf{x})$$

where $\bar{b}$ is the intercept or mean coefficient. From this, original and reconstructed outlines may be superimposed to check diatom shape recovery.

Imaging

Microscope slides of Great Lakes specimens mounted in Hyrax from the E.F.S. collection at the Center for Great Lakes and Aquatic Sciences (University of Michigan, Ann Arbor, MI, USA) were used in this study. Three collections were the sources of the specimens analysed; all were obtained as scrapings from substrata in the north-eastern Lake Michigan area. Collection 1102 was from samples taken in Torch Lake, MI, on June 23 1966; collections 1157 and 1163 were from samples taken near South Fox Island, MI, on June 17 1966 and June 20 1966. Of 85 specimens digitized, 66 qualified for use in shape analysis because they were flat and without distortion to the valve face. That is, the outline of a digitized specimen was highly illuminated at the same intensity in a thin, smooth line. Post-auxospore populations were present on slides from

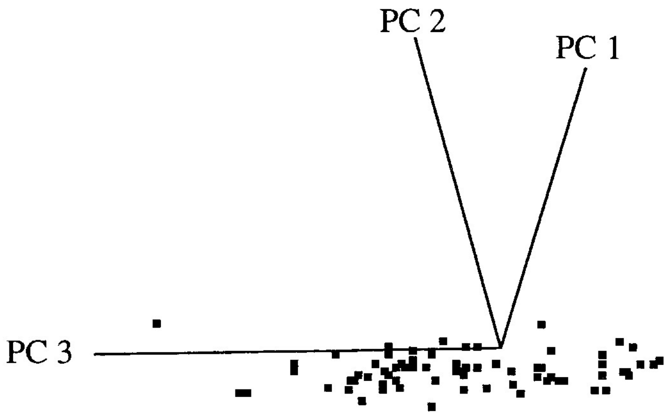


Fig. 1. First three principal components of PCA of Legendre shape coefficients for the whole valve outline of *Cymbella cistula* specimens.

near South Fox Island. The smallest specimens found, which were *c.* 30% the size of the longest specimens (cf. Edlund & Stoermer 1997), were from populations from Torch Lake.

Digitization of specimens was evaluated using an imaging station consisting of a Dialux 20 compound light microscope with a ×100 oil-immersion objective (1.32 numerical aperture), a Sony 3CCD camera attached to the microscope, and an NIH Image software (version 1.62) (Wayne Rasband at the US National Institutes of Health, http://rsb.info.nih.gov/nih-image/). Additional information on imaging is given by Stoermer (1996).

Shape group determination

The extent of shape variation was determined using principal components analysis (PCA) (Green & Carroll 1978). Principal components analysis was first performed on shape coefficients for the entire outline for each *C. cistula* valve. A second PCA was performed on shape coefficients for the dorsal side of the valve outline. The shape groups found were tested and cross-validated using multiple discriminant analysis (DA) (Green & Carroll 1978; Johnson & Wichern 1992). Discriminant analysis was performed on shape coefficients for the dorsal half of each *C. cistula* valve to test the results from the second PCA.

RESULTS

Principal components analysis

Reconstructed outlines from x,y co-ordinates calculated from the first 10 Legendre polynomials verified that the use of Legendre coefficients zero to nine as shape descriptors was sufficient to ensure diatom shape recovery.

The first PCA, using Legendre coefficients for both halves of the diatom valve, resulted in a shape gradient (Fig. 1). Eigenvalues for the first four eigenvectors were 0.948, 0.036, 0.006, and 0.003. To determine dispersion of shape variants for all specimens, a second PCA was performed on the dorsal part of the valve (Fig. 2). Eigenvalues were 0.990, 0.004, 0.003, and 0.002 for the first four eigenvectors. Ordination of the first two principal components produced nine shape groups (Fig. 3). In both PCAs, the greatest amount of shape variation was explained by the first principal component.

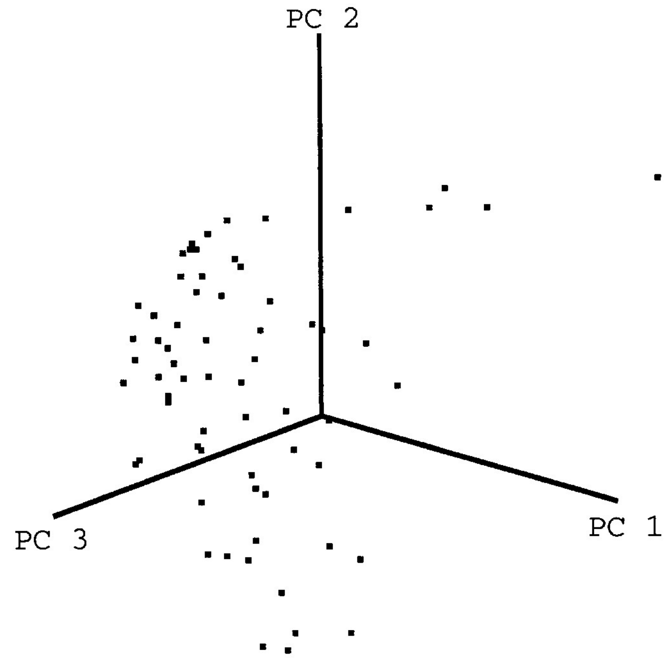


Fig. 2. First three principal components of PCA of Legendre shape coefficients for dorsiventral half of the valve outline of *Cymbella cistula* specimens.

SHAPE GROUP 1: This group had 13 members, comprising four specimens from collection 1157 and nine specimens from collection 1163. Specimens were semicircular with small ends (Fig. 4). Three specimens in this group were somewhat arcuate.

SHAPE GROUP 2: The second shape group was composed of three members from collection 1102. They were semicircular to semilanceolate, with broadly rounded ends (Fig. 5).

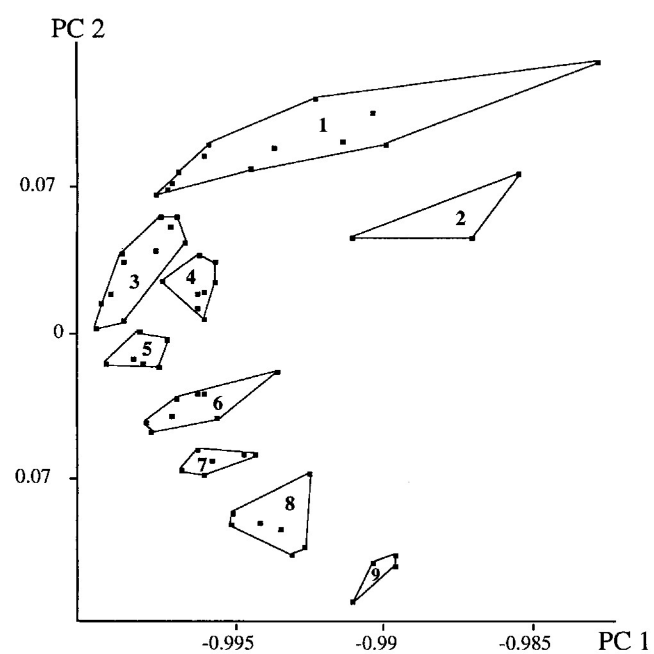
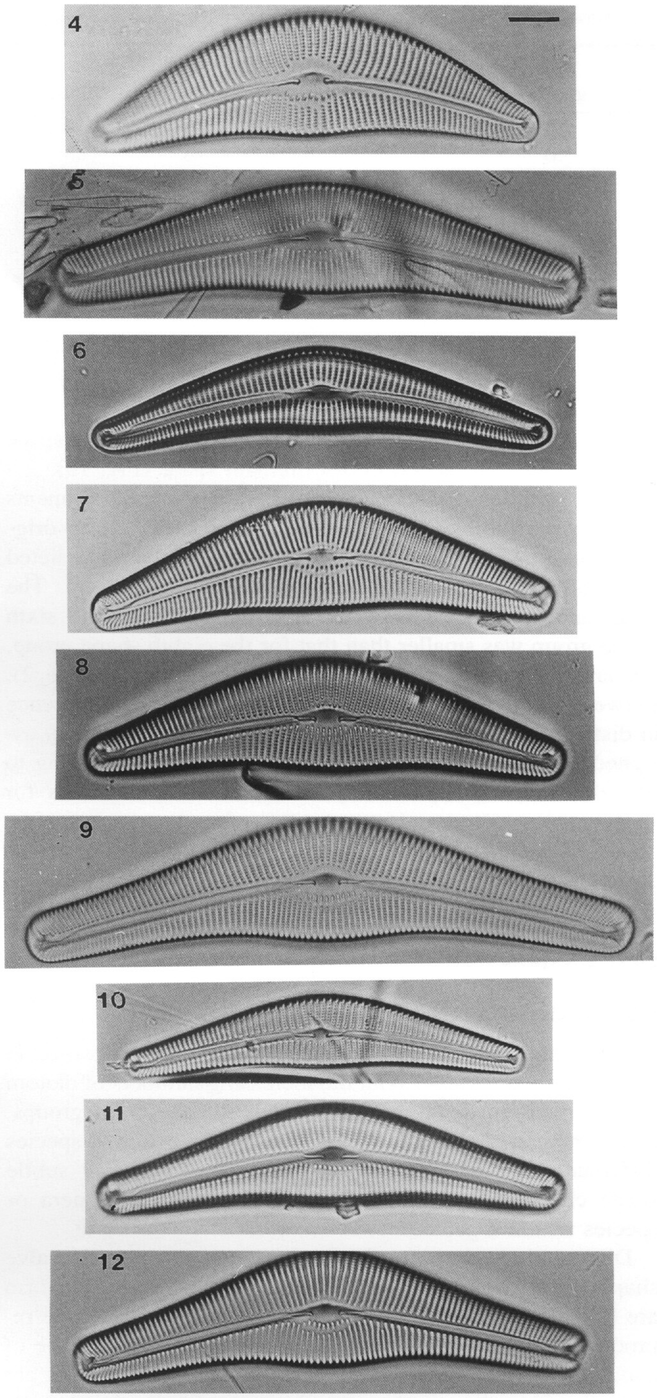


Fig. 3. First two principal components of PCA from Fig. 2. Nine shape groups are delineated.



Figs 4–12. *Cymbella* specimen examples. Scale bar = 10 µm.
Fig. 4. Shape group 1.
Fig. 5. Shape group 2.
Fig. 6. Shape group 3.
Fig. 7. Shape group 4.
Fig. 8. Shape group 5.
Fig. 9. Shape group 6.
Fig. 10. Shape group 7.
Fig. 11. Shape group 8.
Fig. 12. Shape group 9.

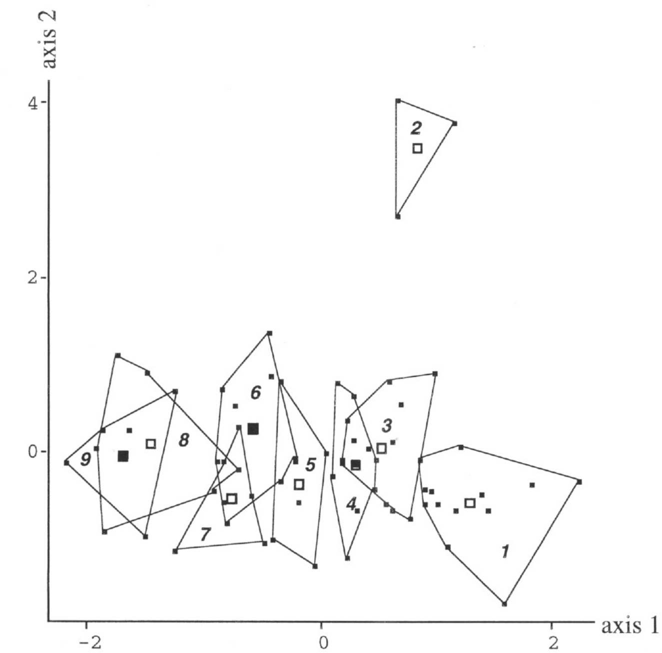


Fig. 13. First two canonical eigenvectors of multiple discriminant analysis of nine shape groups. Centroids for shape groups 1, 2, 3, 5, 7, and 8 are represented as □, and centroids for shape groups 4, 6, and 9 are represented as ■.

SHAPE GROUP 3: Most specimens were from collection 1163, with two additional specimens from collections 1157 and 1102, respectively. Specimens of this group were semilanceolate with small rounded ends (Fig. 6).

SHAPE GROUP 4: The fourth group included eight specimens, with all but one from collection 1163. The exception was from collection 1102. These specimens were similar to those in the third shape group, except that the ends were larger, but not as large as those in the second shape group (Fig. 7).

SHAPE GROUP 5: One of the six specimens was from collection 1102, with the rest from collection 1163. These specimens were similar to those in the fourth shape group, except for their very slightly produced, rostrate ends (Fig. 8).

SHAPE GROUP 6: In the sixth group, three of the eight specimens were from collection 1102 with the remainder from collection 1157. These specimens had crescentic valves with a strongly gibbous or inflated midvalve region. The valve was elongated towards produced, truncated ends (Fig. 9).

SHAPE GROUP 7: Within the seventh group were four specimens from collection 1163 and two from collection 1102. Valves were crescentic but narrow, with only a slightly inflated centre and small, rounded, produced ends (Fig. 10).

SHAPE GROUP 8: This was composed of five specimens from collection 1157 and one each from collections 1102 and 1163. These crescentic valves had slightly more gibbous regions ventrally than those in sixth shape group and had produced, truncated ends (Fig. 11).

SHAPE GROUP 9: The ninth shape group comprised one specimen from collection 1163 and three from collection 1157. Valves had moderately inflated valves centrally, compared to

Table 1. Number of correct assignments for each shape group from multiple discriminant analysis. Row totals are group membership assignments.

	Shape group 1	Shape group 2	Shape group 3	Shape group 4	Shape group 5	Shape group 6	Shape group 7	Shape group 8	Shape group 9	Total
Shape group 1	12	0	1	0	0	0	0	0	0	13
Shape group 2	0	3	0	0	0	0	0	0	0	3
Shape group 3	1	0	7	3	0	0	0	0	0	11
Shape group 4	0	0	1	7	0	0	0	0	0	8
Shape group 5	0	0	0	0	5	1	0	0	0	6
Shape group 6	0	0	0	0	0	6	2	0	0	8
Shape group 7	0	0	0	0	0	0	5	1	0	6
Shape group 8	0	0	0	0	1	0	1	4	1	7
Shape group 9	0	0	0	0	0	0	0	0	4	4

those in the sixth and seventh shape groups (Fig. 12). The valves were crescentic and had produced, truncated ends.

Multiple discriminant analysis

In DA, most of the shape group differentiation was indicated on the first canonical eigenvector, with an eigenvalue of 0.913 (Fig. 13). Shape group variation on the second canonical eigenvector, with an eigenvalue of 0.667, was defined by the separation of the second shape group from all other shape groups (Fig. 13). Eigenvalues for the third to eighth canonical eigenvectors were 0.404, 0.343, 0.187, 0.081, 0.049, and 0.020. Canonical correlation coefficients were 0.956, 0.816, 0.636, 0.585, 0.432, 0.284, 0.222, and 0.140 for the first eight DA axes extracted. Wilks' Lambda was 0.0079, and the F-statistic was 4.48.

A Monte Carlo permutation test under the full model was conducted to determine the significance of the first axis and the trace. Ninety-nine permutations were used. The first axis was significant with a P-value of 0.10 and an F-ratio of 7.09. The trace of all canonical axes was significant at 2.66 with a P-value of 0.10 and an F-ratio of 2.74.

From DA, specimen assignment to shape groups was nearly perfect with few exceptions (Table 1). Group and predicted assignments to shape groups revealed small differences in Mahalanobis distances (in double precision). That is, the difference between Mahalanobis distance for the predicted outcome and that for the original group assignment is small in all but a few cases (Table 2).

The largest misassignments occurred with respect to the third and eighth shape groups (Table 1). Of 11 specimens, 7 were assigned correctly to the third shape group (Table 1). Three misassigned specimens were allocated to the fourth shape group. Specimens in the third and fourth shape groups were proximal in shape space (Fig. 3) and overlapped in discriminant space (Fig. 13). Specimen 1163, 11, was originally assigned to the third group, but was predicted to be in the fourth shape group (Table 2). A smaller Mahalanobis distance was calculated for this specimen's assignment to the fifth shape group than the third shape group, but larger than the fourth shape group (Table 2). The difference in distances between the third and fourth shape groups for this specimen is 0.313356.

Four of seven specimens were correctly assigned to the eighth shape group (Table 1). Each of two misassigned specimens was allocated to the seventh and ninth shape groups, respectively (Table 1). Although specimens in the seventh, eighth, and ninth shape groups were separate in shape space (Fig. 3), they overlapped in discriminant space (Fig. 13).

Two specimens indicated the least favourable assignments with respect to the eighth group. Specimen 1102, 9, was originally assigned to the eighth shape group, but was predicted to be allocated to the seventh shape group (Table 2). The Mahalanobis distance between this specimen and the sixth shape group was smaller than that for the eighth shape group, but larger than that for the seventh shape group (Table 2). Between the seventh and eighth shape groups, the difference in distances was 0.805176. Specimen 1163, 26, originally assigned to the eighth shape group, was predicted to belong to the fifth shape group (Table 2). Mahalanobis distances for shape groups seven and nine are also less than that for the eighth shape group (Table 2). The maximum difference in distances is 0.328971.

DISCUSSION

Diatom shape

Shape is one of the primary morphological characters used in separating diatom taxa. With the increasing numbers of diatom genera resolved from previously large, composite groups, quantitative assessment is useful in initial resolution of species differences. Shape analysis is useful in quantifying subtle shape changes or degree of complexity between genera or species within a genus.

Diatom shape is part of valve ontogeny or overall valve shape and form development (Mann 1984). Shape and form are directly inherited in diatoms as a result of vegetative reproduction (Mann 1994); parent cells determine the shape of daughter cells because new valves and girdle bands are formed within the parent cell, which moulds them. There is therefore a high degree of constancy in valve shape, despite size diminution. Nevertheless, there are slight differences in shape between parent and daughter cells as size reduction occurs, and shape change during size diminution complicates species identification. The smallest cells of one species can be misinterpreted as a different species (Geitler 1932). In our study, Legendre coefficients were useful as quantitative shape descriptors in separating specimens in the *C. cistula* species complex.

Shape characteristics

Asymmetry in valve outline is one of the most important characters in *Cymbella* (Krammer 1982). Dorsiventral asymmetry

Table 2. Mahalanobis distances between each specimen and shape group centroid. The smallest value for each specimen, which indicates predicted group membership, is emboldened.

Collection, specimen number	Distance to group 1	Distance to group 2	Distance to group 3	Distance to group 4	Distance to group 5	Distance to group 6	Distance to group 7	Distance to group 8	Distance to group 9
1157, 1	7.683691	7.675038	5.622219	5.195555	4.904903	**2.937272**	3.592029	3.969955	5.110638
1157, 2	6.764111	8.410038	4.973529	5.054397	3.930793	**3.290821**	4.150173	5.630109	7.238462
1157, 3	**2.219559**	6.917763	2.526271	3.938414	5.034501	5.810410	5.912047	8.214039	8.914130
1157, 4	10.407508	11.495403	8.248992	7.851296	6.394437	5.796679	4.891943	4.091086	**3.574195**
1157, 7	9.278073	10.668515	7.501406	6.534846	5.611073	5.129884	4.348510	3.822675	**2.764643**
1157, 8	10.288752	10.532481	8.076159	7.489169	6.402679	4.713413	4.443165	**2.453831**	3.203353
1157, 9	**3.472694**	7.405037	4.790929	4.449484	6.266154	7.580296	7.930408	9.572264	10.394070
1157, 10	7.291810	8.388080	5.534875	4.903697	4.145583	**3.205283**	3.498447	3.224464	4.138845
1157, 11	9.451774	9.083850	7.308013	7.246223	6.703131	5.044761	4.959074	**2.942132**	3.547331
1157, 12	**3.042346**	9.074933	5.158545	5.572972	7.006158	8.048259	8.255696	10.257846	11.140059
1157, 13	**4.459355**	8.244302	6.457004	6.890608	8.552499	9.747547	10.284350	12.081084	12.725721
1157, 14	7.568575	9.416349	5.576879	5.423110	3.938913	**3.165168**	3.512142	4.494075	5.827157
1157, 16	**2.183367**	7.083578	2.289633	3.367117	4.590038	5.329480	5.361013	7.303759	8.192213
1157, 17	11.184939	11.453065	8.940610	8.488295	7.281498	6.298318	5.657423	3.486390	**2.741917**
1157, 18	7.582899	7.665348	5.153229	4.934339	3.833080	3.302301	**3.250201**	4.460325	5.507734
1157, 19	10.130447	9.361571	7.975511	7.369210	6.615584	5.157115	5.172346	**2.680294**	3.362306
1157, 20	9.918434	9.942933	7.609481	7.111712	5.978511	4.290188	3.897372	**3.354614**	4.877937
1157, 21	10.384547	10.280553	8.156116	7.263366	5.872305	5.403701	5.022908	3.282756	**2.329036**
1102, 3	3.567335	6.135106	**1.946493**	4.019812	4.642870	4.794568	4.962435	7.158587	8.169145
1102, 4	5.152131	6.996773	3.339786	3.697047	3.789185	2.719182	**2.710212**	4.411365	5.708663
1102, 6	6.534090	**3.587424**	5.396741	6.108302	6.407177	6.680711	7.930201	8.780552	9.207904
1102, 9	7.029423	8.371644	5.201919	5.298376	4.744173	3.744165	**3.060018**	3.865194	5.339092
1102, 10	4.701051	5.731397	3.694965	**3.257948**	3.853106	3.921242	5.106644	6.460507	7.223595
1102, 13	6.861384	6.630505	4.747571	4.067828	3.531632	**3.265909**	4.014734	5.225587	5.627190
1102, 16	7.078188	6.988846	5.361162	4.663868	3.61814	**3.567412**	4.846888	5.920365	6.878524
1102, 17	8.269569	**3.190891**	7.024797	7.577394	8.52404	8.191180	9.100843	9.685021	10.492104
1102, 18	7.115735	7.674615	4.923064	4.819178	4.167565	3.443334	**2.933440**	3.989564	5.045317
1102, 19	6.654362	8.497100	4.951377	4.929099	4.63786	3.773238	**3.224369**	4.543561	4.801244
1102, 20	7.550421	**3.163915**	7.276104	7.406748	8.729184	8.29753	9.685486	10.562428	11.616266
1102, 24	6.649127	5.626132	4.478295	4.205552	3.532129	**3.147237**	4.197171	4.496726	5.218577
1163, 1	8.138139	9.023262	5.668466	6.266453	4.667909	4.224283	**4.087357**	4.988051	6.136882
1163, 2	3.910699	5.135081	**2.312999**	4.016312	4.678399	4.700961	5.472878	6.975058	7.999869
1163, 3	4.235965	5.832806	2.680253	**2.505683**	2.913561	3.728187	4.535835	5.829433	6.565533
1163, 4	4.778512	6.075891	3.839925	**3.823102**	4.691787	5.349607	6.216622	8.105096	8.781781
1163, 6	5.413477	7.721264	3.775532	2.953975	**2.313159**	3.348706	3.632907	5.031911	5.772941
1163, 7	3.520603	7.230725	**3.275719**	3.390977	4.281659	5.066836	5.051107	7.274498	8.437900
1163, 8	4.822506	8.398408	4.415832	**2.926226**	4.037398	4.807726	4.787931	6.722566	7.430853
1163, 9	5.441354	8.677047	3.898309	3.762593	**2.307120**	3.862779	3.851761	6.036995	6.957560
1163, 10	**3.164859**	7.021732	3.753936	4.043556	5.149991	5.587908	6.127066	8.068213	9.163445
1163, 11	4.227585	6.445288	2.626955	**2.313599**	2.517265	3.737037	3.969836	5.791495	6.579141
1163, 12	3.882709	6.886546	3.223261	**2.487393**	3.716180	4.992636	4.968493	6.783876	7.546933
1163, 13	8.716324	8.678919	6.575015	6.782805	6.113888	4.643216	4.462692	3.347681	**3.123579**
1163, 15	4.118559	5.284579	**3.252132**	4.597178	5.307028	5.766852	6.761231	8.475625	9.488422
1163, 17	**2.303263**	5.897280	3.286057	3.033948	4.829597	5.966736	6.697324	8.740925	9.608788
1163, 18	**2.328560**	7.229634	3.333236	4.583573	5.513030	6.480853	6.843779	8.759969	9.384757
1163, 19	**2.184471**	6.888550	2.881817	3.354463	5.116878	5.689806	5.756779	7.778063	8.592389
1163, 20	**2.171150**	7.629929	3.361255	3.937239	5.307366	6.292340	6.271744	8.537027	9.406844
1163, 21	4.877305	5.450391	**3.086213**	3.120471	3.679645	3.871852	4.661375	5.893490	6.710941
1163, 22	6.444907	8.899357	4.420117	4.691697	3.590763	3.221979	**2.255862**	4.786942	6.000403
1163, 23	5.478740	6.968362	3.907054	3.003322	**2.289247**	4.038336	4.678208	6.309767	6.600136
1163, 24	**2.093676**	7.242603	3.794751	4.439860	5.871249	7.020708	7.414146	9.481838	10.280625
1163, 25	4.096842	7.331335	3.442016	**2.282540**	3.310987	4.832284	4.941986	6.474653	6.743212
1163, 26	7.811380	9.106974	6.032660	5.151929	**3.959294**	4.291691	4.255226	4.288265	4.021296
1163, 27	7.125822	8.718831	5.214695	4.228899	3.484564	**2.797678**	2.440458	3.531956	4.644177
1163, 28	8.318136	10.21528	6.738154	5.717478	4.962284	4.220718	3.523449	**3.008190**	3.742177
1163, 29	3.920529	6.427410	2.825754	**2.356181**	3.517160	4.397386	4.822137	6.864001	7.727896
1163, 30	4.329510	6.649689	**2.141798**	3.353334	3.119406	3.673845	3.858714	6.031720	6.651025
1163, 31	3.579587	7.289644	**2.666012**	4.253739	4.853593	5.385650	5.318611	7.203770	7.665933
1163, 32	**2.771720**	6.914460	2.833704	4.302359	5.239415	5.963375	6.088542	8.198240	8.832717
1163, 33	5.040092	6.775908	4.364645	**3.033767**	3.628708	5.211614	5.623293	7.212942	8.044726
1163, 35	5.300639	6.993966	3.340151	2.523518	**2.504060**	3.172351	3.105611	4.551509	5.348859
1163, 36	6.021877	8.545538	4.174704	3.522971	**2.298115**	3.318949	2.942803	4.635803	5.326748
1163, 37	4.956025	7.226747	3.984593	**2.812189**	3.237445	3.787549	4.676519	6.000351	7.072980
1163, 40	**2.941689**	7.427367	4.981729	5.707690	7.214950	8.173773	8.713518	10.653954	11.486871
1163, 41	3.878338	6.068425	**2.05434**	3.202533	3.467368	3.702271	4.056340	5.851708	6.913033
1163, 42	2.141610	6.026525	**1.986169**	3.415153	4.365610	5.012999	5.576618	7.530381	8.420609

Table 3. Weighted correlation matrix of Legendre coefficient by canonical eigenvector.

Legendre coefficient	Axis 1	Axis 2	Axis 3	Axis 4
1	−0.0593	−0.3218	0.2592	0.1498
2	0.0593	−0.3083	−0.0919	−0.0166
3	−0.1094	0.0119	−0.2946	−0.1723
4	−0.3946	0.2025	−0.0571	−0.3940
5	0.8742	−0.0780	0.1347	−0.0565
6	0.5173	−0.2161	−0.0747	0.4087
7	−0.7607	0.3209	−0.0166	0.0688
8	−0.5470	−0.0344	0.1653	−0.2324
9	−0.1125	0.5470	−0.3472	−0.0487
10	0.1815	−0.0156	0.1264	0.2870

is evident in the different heights of the dorsal and ventral valve mantles and girdle (Krammer 1981, 1982). Use of the dorsal half of the valve was critical for defining shape differences among specimens in our study and this is the first study to use partial outlines to distinguish valve shape. With Wilks' Lambda = 0.0079, centroid differences between groups were highly significant (Green & Carroll 1972).

From DA, Legendre coefficients four through seven were highly correlated with the first canonical eigenvector (Table 3). The addition of the zeroth and first Legendre coefficients indicated initial semicircularity in valve shape without expansion of the midvalve region. By the addition of the fourth Legendre coefficient, lunateness (i.e. the extent to which the valve is crescentric, arcuate, or semilanceolate) was established as the defining shape of the dorsal side in a global sense (Stoermer & Ladewski 1982). Addition of the sixth Legendre coefficient defined expansion of the midvalve region (Stoermer & Ladewski 1982), i.e. the degree of gibbousness. The eighth Legendre coefficient was highly correlated with the second canonical eigenvector (Table 3). After addition of this coefficient, the characteristic *C. cistula* shape was well defined, including differences in shape of the midvalve region ventrally. In general, shape development and complexity was defined by the addition of more and more Legendre coefficients and depicted in DA.

Shape groups and taxa names

Of the nine shape groups found in PCA, some resemble named varieties of *C. cistula.* Some specimens in the first shape group resembled *C. cistula* var. *truncata* (Stoermer & Yang 1969). Others resembled those depicted in Hufford & Collins' (1972b) treatment of *C. cistula* and somewhat like those depicted in pl. 127, fig. 8 and pl. 128, figs 1 and 2 in Krammer & Lange Bertalot (1986). Still other specimens resembled *C. cymbiformis sensu* Patrick & Reimer (1975). Three specimens, which were somewhat arcuate, do not resemble named varieties of *C. cistula.* Specimens in the first shape group also resembled *C. subcistula* in pl. 83, fig. 3 in Krammer (2002).

The second shape group, with specimens having large rounded ends, was unlike all others and was far removed from other shape groups in PC shape space (Fig. 3). They somewhat resembled *C. neocistula* var. *islandica* in pl. 93, fig. 1 in Krammer (2002).

Specimens from the third, fourth, and fifth shape groups were intermediate forms representing aspects of *C. cistula* shape variation depicted in Krammer & Lange-Bertalot's (1986) treatment of this species complex. Some specimens from the third shape group somewhat resembled *C. cursiformis.* Others resembled *C. cymbiformis* var. *nonpunctata* (Patrick & Reimer 1975; Krammer & Lange-Bertalot 1986). They also resembled *C. neocistula* var. *lunata* in pl. 89, fig. 2 in Krammer (2002). Many specimens in the fourth and fifth shape groups resembled *C. cistula* depicted in pl. 127, fig. 10 in Krammer & Lange-Bertalot (1986). They also resembled *C. neocistula* in pl. 85, fig. 3 and pl. 86, fig. 2 in Krammer (2002). It should be noted that Krammer's (2002) *C. neocistula* falls within Krammer & Lange-Bertalot's (1986) concept of *C. cistula.*

Specimens from shape groups six to nine resembled *C. cistula* var. *gibbosa* (Patrick & Reimer 1975). Specimens in the seventh shape group were narrower than those in the sixth, eighth, and ninth shape groups. In general, all these specimens were noticeably ventrally gibbous in the midvalve region. They are not represented in the treatment of *Cymbella* species by Krammer (2002).

Shape and ontogenetic inference

The curved trajectory of PC scores in shape space (Fig. 3) may represent ontogenetic projections of members in the *C. cistula* species complex. The overall shape sequence produced a gradation from a less complex valve shape that is dorsally semicircular (to arcuate) to a more complex form that is dorsally crescentic and ventrally gibbous. That is, dorsally semicircular (and arcuate) forms have less change in curvature around the periphery of the valve than crescentic forms. Because initital cells were evident in the populations sampled, this trajectory is a dynamic representation of ontogenetic changes (Williams 1994, 1996).

Valve shape is a result of many factors and interactions during valve morphogenesis (Schmid 1979, 1994). Changes in valve shape may indicate approximate successive stages within a single species life cycle. The characteristic modifications in shape that occur during post-auxospore size reduction define the developmental trajectory in diatoms (Pfitzer 1871). In our results, shape variation was discontinuous, forming discrete groups in PCA shape space (Fig. 3). The most distinct shape groups were 1 (Fig. 4), 2 (Fig. 5), 3–5 (Figs 6–8), and 6–9 (Figs 9–12), but it is possible to interpret our analysis as showing nine separate shape groups, which may represent nine separate species. Abrupt shape changes do not generally occur within a species but between species as a result of the pattern of diatom morphogenesis (Mann 1994, 1999).

The developmental sequence is a basic attribute of genomic expression embodied in shape differences (Stoermer *et al.* 1986). Developmental differences between taxa found in close proximity may be evident as shape differences (Stoermer *et al.* 1984), and these differences may be documented using quantitative shape analysis. If each shape group represents a species, then by inference, each shape group includes a developmental sequence or ontogenetic projection for that species. The shape groups within the *C. cistula* species complex will need to be tested and species differences corroborated by other types of evidence, such as cytological, molecular, or physiological findings. Because shape groups divisions were distinct, it is unlikely that the curved trajectory in PC shape space is a continuum of *C. cistula* exclusively.

Specimens used in this study comprised post-auxospore populations and populations where the smallest specimens were 30% the size of the longest specimens, which probably means that almost the whole life cycle was sampled (cf. Edlund & Stoermer 1997). Using the full size range, including sexually reproductive populations, is advantageous in relating shape change to life cycle stage (Edlund & Stoermer 1997). Results from quantitative shape analysis have potential in identifying specific valve shape changes as proxies for particular life cycle stages in ontogenetic projections. This may be indicated within each shape group found in our study.

Our study has indicated the need for careful, quantitative shape analysis, especially as the number of species identified is steadily increased (Mann 1994, 1999). Studies such as ours can be used as a springboard for, or corroboration of, further studies to verify taxonomic hypotheses based on quantitative shape analysis and species differentiation.

ACKNOWLEDGEMENTS

We thank Ted Ladewski for his novel work using Legendre coefficients as shape descriptors in diatom taxonomy. For her preliminary studies on the *Cymbella cistula* species complex, we thank Heather Carleton. This research was sponsored by NSF Grant DEP(PEET) 9521882.

REFERENCES

BARBER H.G. & HAWORTH E.Y. 1981. A guide to the morphology of the diatom frustule. *Freshwater Biological Association, Scientific Publication* 44: 1–112.

EDLUND M.B. & STOERMER E.F. 1997. Ecological, evolutionary, and systematic significance of diatom life histories. *Journal of Phycology* 33: 897–918.

FARRELL O.J. & ROSS B. 1963. *Solved problems: gamma and beta functions, Legendre polynomials, Bessel functions.* Macmillan, New York. 410 pp.

GEITLER L. 1932. Der Formwechsel der pennaten Diatomeen. *Archiv für Protistenkunde* 78: 1–226.

GOLDMAN N., PADDOCK T.B.B. & SHAW K.M. 1990. Quantitative analysis of shape variation in populations of *Surirella fastuosa. Diatom Research* 5: 25–42.

GREEN P.E. & CARROLL J.D. 1978. *Analyzing multivariate data.* Dryden Press, Hinsdale, Illinois. 519 pp.

HUFFORD T.L. & COLLINS G.B. 1972a. The freshwater diatom *Cymbella cursiformis* nom. nov. *Journal of Phycology* 8: 184–187.

HUFFORD T.L. & COLLINS G.B. 1972b. Some morphological variations in the diatom *Cymbella cistula. Journal of Phycology* 8: 192–195.

HUFFORD T.L. & COLLINS G.B. 1972c. The stalk of the diatom *Cymbella cistula*: SEM observations. *Journal of Phycology* 8: 208–210.

HUSTEDT F. 1930. Bacillariophyta (Diatomeae). In: *Die Süsswasser-Flora Mittleleuropas,* vol. 10, ed. 2 (Ed. by A. Pascher). G. Fischer, Jena. 466 pp.

JOHNSON R.A. & WICHERN D.W. 1992. *Applied multivariate statistical analysis,* ed. 3. Prentice Hall, Englewood Cliffs, New Jersey. 642 pp.

KRAMMER K. 1981. Morphologic investigations of valve and girdle of the diatom genus *Cymbella* Agardh. *Bacillaria* 4: 125–146.

KRAMMER K. 1982. Valve morphology in the genus *Cymbella* C.A. Agardh. In: *Micromorphology of diatom valves,* vol. 3 (Ed. by J.-G. Helmcke & K. Krammer), plates 1024–1148. J. Cramer, Vaduz, Germany.

KRAMMER K. 2002. *Cymbella.* In: *Diatoms of Europe. Diatoms of the European inland waters and comparable habitats,* vol. 3 (Ed. by H. Lange-Bertalot). A.R.G. Gantner Ruggell, Liechtenstein. 584 pp.

KRAMMER K. & LANGE-BERTALOT H. 1986. Bacillariophyceae. 1. Teil: Naviculaceae. In: *Süßwasserflora von Mitteleuropa,* vol. 2(3) (Ed. by H. Ettl, J. Gerloff, H. Heynig & D. Mollenhauer). G. Fischer, Stuttgart and Jena. 876 pp.

MANN D.G. 1981. A note on valve formation and homology in the diatom genus *Cymbella. Annals of Botany* 47: 267–269.

MANN D.G. 1984. An ontogenetic approach to diatom systematics. In: *Proceedings of the Seventh International Diatom Symposium* (Ed. by D.G. Mann), pp. 113–141. O. Koeltz, Koenigstein.

MANN D.G. 1994. The origins of shape and form in diatoms: the interplay between morphogenetic studies and systematics. In: *Shape and form in plants and fungi* (Ed. by D.S. Ingram & A.J. Hudson), pp. 17–38. Academic Press, London.

MANN D.G. 1999. The species concept in diatoms. *Phycologia* 38: 437–495.

MOU D. & STOERMER E.F. 1992. Separating *Tabellaria* (Bacillariophyceae) shape groups based on Fourier descriptors. *Journal of Phycology* 28: 386–395.

PATRICK R. & REIMER C.W. 1975. The diatoms of the United States exclusive of Alaska and Hawaii, vol. 2. *Monographs of the Academy of Natural Sciences, Philadelphia* 13: 42–65.

PFITZER E. 1871. Untersuchungen über Bau und Entwicklung der Bacillariaceen (Diatomeen). *Botanische Abhandlungen aus dem Gebiet der Morphologie und Physiologie* (Ed. by J. Hanstein) 1(2): 1–189.

RHODE K.M., PAPPAS J.L. & STOERMER E.F. 2001. Quantitative analysis of shape variation in type and modern populations of *Meridion* (Bacillariophyceae). *Journal of Phycology* 37: 175–185.

SCHMID A.-M.M. 1979. Influence of environmental factors on the development of the valve in diatoms. *Protoplasma* 99: 99–115.

SCHMID A.-M.M. 1994. Aspects of morphogenesis and function of diatom cell walls with implications for taxonomy. *Protoplasma* 181: 43–60.

STEINMAN A.D. & LADEWSKI T.B. 1987. Quantitative shape analysis of *Eunotia pectinalis* (Bacillariophyceae) and its application to seasonal distribution patterns. *Phycologia* 26: 467–477.

STOERMER E.F. 1996. A simple, but useful, application of image analysis. *Journal of Paleolimology* 15: 111–113.

STOERMER E.F. & LADEWSKI T.B. 1982. Quantitative analysis of shape variation in type and modern populations of *Gomphoneis herculeana. Nova Hedwigia* 73: 347–386.

STOERMER E.F. & YANG J.J. 1969. *Plankton diatom assemblages in Lake Michigan.* Great Lakes Research Division Special Report No. 47, University of Michigan, Ann Arbor. 168 pp.

STOERMER E.F., LADEWSKI T.B. & KOCIOLEK J.P. 1984. Further observations on *Gomphoneis*. In: *Proceedings of the Eighth International Diatom Symposium* (Ed. by M. Ricard), pp. 205–213. Koeltz, Koenigstein.

STOERMER E.F., QI Y.-Z. & LADEWSKI T.B. 1986. A quantitative investigation of shape variation in *Didymosphenia* (Lyngbye) M. Schmidt (Bacillariophyta). *Phycologia* 25: 494–502.

SZOKEFALVI-NAGY B. 1965. *Introduction to real functions and orthogonal expansions.* Oxford University Press, New York. 447 pp.

THERIOT E. & LADEWSKI T.B. 1986. Morphometric analysis of shape of specimens from the neotype of *Tabellaria flocculosa* (Bacillariophyceae). *American Journal of Botany* 73: 224–229.

VANLANDINGHAM S.L. 1969. *Catalogue of the fossil and recent genera and species of diatoms and their synonyms. Part III. Coscinophaena through Fibula,.* Cramer, Vaduz, Germany. pp. 1087–1756.

WEINBERGER H.F. 1995. *A first course in partial differential equations.* Dover Publications, New York. 446 pp.

WILLIAMS D.M. 1994. Ontogeny and phylogeny in the genus *Tetracyclus. Memoirs of the California Academy of Sciences* 17: 247–256.

WILLIAMS D.M. 1996. Fossil species of the diatom genus *Tetracyclus* (Bacillariophyta, "ellipticus" species group): morphology, interrelationships and the relevance of ontogeny. *Philosophical Transactions of the Royal Society of London B* 351: 1759–1782.

Accepted 15 September 2002