Medlin, L., Hille, H. J., Stickel, S. & Sogin, M. L. 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* 71:491–9.
Melkonian, M. & Berns, B. 1983. Zoospore ultrastructure in the green alga *Friedmannia israelensis*: an absolute configuration analysis. *Protoplasma* 114:67–84.
Moestrup, Ø. 1972. Observations of the fine structure of spermatozoids and vegetative cells of the green alga *Golenkinia*. *Br. Phycol. J.* 7:169–83.
Sargent, M., Zahn, R., Walters, B., Gupta, R. & Kaine, B. 1988. Nucleotide sequence for the 18S rDNA from the microalga *Nanochlorum eucaryotum*. *Nucleic Acids Res.* 16:4156.
Smith, G. M. 1950. *The Fresh-Water Algae of the United States*. McGraw-Hill, New York, 719 pp.
Starr, R. C. & Zeikus, J. A. 1987. UTEX—the culture collection of algae at the University of Texas at Austin. *J. Phycol.* 23(Suppl.):1–47.
Swofford, D. L. 1990. PAUP—phylogenetic analysis using parsimony version 3.0L. Illinois Natural History Survey, University of Illinois, Champaign.
Takeda, H. 1991. Sugar composition of the cell wall and the taxonomy of *Chlorella* (Chlorophyceae). *J. Phycol.* 27:224–32.
Trainor, F. R. 1963. Zoospores in *Scenedesmus obliquus*. *Science (Wash. D.C.)* 142:1673–4.
Trainor, F. R. & Burg, C. A. 1965. *Scenedesmus obliquus* sexuality. *Science (Wash. D.C.)* 148:1094–5.
Trainor, F. R. & Egan, P. F. 1990. *Lagerheimia hindakii* is the unicellular stage of a *Scenedesmus*. *J. Phycol.* 26:535–9.
Watanabe, S. & Floyd, G. L. 1989a. Variation in the ultrastructure of the biflagellate motile cells of six unicellular genera of the Chlamydomonadales and Chlorococcales (Chlorophyceae), with emphasis on the flagellar apparatus. *Am. J. Bot.* 76:307–17.
——— 1989b. Comparative ultrastructure of the zoospores of nine species of *Neochloris* (Chlorophyta). *Pl. Syst. Evol.* 168: 195–219.
Watanabe, S., Floyd, G. L. & Wilcox, L. W. 1988. Ultrastructure of the zoospores and vegetative cells of *Tetraedron* and *Chlorotetraedron* (Chlorophyceae). *J. Phycol.* 24:490–5.
Wilcox, L. W. & Floyd, G. L. 1988. Ultrastructure of the gamete of *Pediastrum duplex* (Chlorophyceae). *J. Phycol.* 24:140–6.

J. Phycol. **28,** 386–395 (1992)

SEPARATING *TABELLARIA* (BACILLARIOPHYCEAE) SHAPE GROUPS BASED ON FOURIER DESCRIPTORS[1]

Daiqing Mou and Eugene F. Stoermer[2]

Center for Great Lakes and Aquatic Sciences, University of Michigan, Ann Arbor, Michigan 48109

ABSTRACT

A random sample of 490 Tabellaria *specimens was analyzed using the harmonic amplitudes of the Fourier transformations of their valve outlines as shape descriptors. Principal component analysis (PCA) was applied to the sample to reduce dimensionality. The problem of non-normal distribution of these descriptors due to cell division was solved by sub-sectioning the entire data set based on its distribution on the first three components (PC1, PC2, and PC3) of the overall PCA. Each of the subsets was then analyzed by PCA. Shape groups from subset clusters were compared with one another and then similar groups were congregated into one growth series. Eight distinct shape groups were found. The results agree with some previous classical observations on the genus and at the same time reveal many new morphological characteristics related to valve shape. These new characteristics are impossible to obtain without appropriate specimen sampling, quantitative shape description, and data analysis techniques.*

Key index words: Bacillariophyceae; Fourier descriptors; morphometric; shape analysis; Tabellaria

Taxonomy of the diatom genus *Tabellaria* is very difficult, and there is poor agreement concerning the morphological limits of specific and sub-specific taxa. Almost all morphological characters observed in the genus have been reported to be highly variable (Knudson 1952, 1953a, b, Stoermer and Yang 1969, Koppen 1973, 1975) although the range and form of variation have not been established. Traditional diatom classification has become increasingly suspect for two primary reasons. Nearly all modern studies of diatom genera have revealed greater diversity than that expressed in traditional classifications. There is some recent evidence for this in *Tabellaria* (Lange-Bertalot 1988). At the same time, modern ecological and paleoecological investigations have revealed apparently consistent morphotypes associated with particular geographical regions or ecological conditions, and investigators have increasingly resorted to arbitrary designations to convey information hidden by current formal classifications (Koppen 1978, Kingston et al. 1986). Our operational hypothesis is that unresolved order exists within morphological characteristics of *Tabellaria* and that resolution of this order will result in a more natural and informative systematics.

Valve outline shape is considered important in *Tabellaria* taxonomy for two reasons. 1) *Tabellaria* species, as currently understood, are relatively character-poor. The first and most easily available taxonomic character of this genus is the valve outline

[1] Received 15 February 1991. Accepted 15 January 1992.
[2] Address for reprint requests.

shape. 2) Although shape is variable both within and between populations, it remains relatively stable on a large scale, and it has been employed as a main diagnostic character early in the history of *Tabellaria* taxonomy (Hustedt 1931). Traditionally, *Tabellaria* shape has been dissected into characters that can be verbally described. Such descriptions are subjective and, hence, subject to errors in both description and interpretation of published works. What is a "capitate" terminal inflation to one observer may be "gradually tapering" to another. There is obvious need for a quantitative method to describe shape precisely and unambiguously.

Theriot and Ladewski (1986) applied a quantitative method that uses the Legendre polynomial expansion to quantify diatom valve outlines (Stoermer and Ladewski 1982, Goldman et al. 1990) to study *Tabellaria* shape. This method is particularly appropriate for shapes composed of relatively smooth curves and less appropriate for outlines consisting of linear sections alternating with regions of high curvature. The latter is characteristic of most *Tabellaria* species (Theriot and Ladewski 1986). To quantify *Tabellaria* shapes, a different approach is developed for this study.

Tabellaria valve outline may be satisfactorily described by a mathematically simple, closed plane curve. To be useful for taxonomic purposes, numerical descriptors applied to the shape should be able to describe all details present in a given image. The descriptors should remain constant if the change in contour they represent is only due to uniform scaling; the algebraic values of the descriptors should be directly interpretable in terms of the shape they represent.

The advantage of using numerical descriptors to represent *Tabellaria* shape is achieved not only in the accuracy in shape description but also in the variety of statistical techniques available for descriptor analysis, because the descriptors can be treated as numerical variables. The purpose of this study is to determine broad relationships within the genus as a whole using valve shape. To obtain unbiased final results, the choice of appropriate statistical techniques is as important as the choice of appropriate numerical descriptors.

An unbiased sampling technique is needed to produce unbiased statistical results. In practice, diatomists usually select valves across some size range to compare *Tabellaria* or other diatom populations. The reason for this is the need to account for the diminution–rejuvenation size series, which occurs during the reproductive cycle of most diatoms (Geitler 1932). In most populations, small cells, the result of many asexual divisions, greatly outnumber larger post-auxospore cells. To obtain an estimate of within-population morphological variance with reasonable effort, it is tempting to make arbitrary judgements in choice of specimens for analysis. In genera like *Tabellaria*, where underclassification is suspected and unquantified morphological variance associated with ecological factors may exist, *a priori* assumptions of grouping are particularly dangerous.

Due to these considerations, we developed a method of shape analysis that we feel is particularly appropriate for *Tabellaria* and applied it to a large, randomly chosen group of specimens.

MATERIALS AND METHODS

The descriptors. There are two general approaches in biological shape analysis. One treats shape variations as distortions of some initial coordinate system (Thompson 1961, Bookstein et al. 1985). This deformation model requires that there is a spatial or ontogenetic correspondence among the definable structures of organisms under study. The corresponding points are defined as landmarks. Most diatoms do not have outline landmarks. It is important to emphasize here that internal structures, such as labiate processes, are usually not qualified as landmarks for two reasons. First, their ontogeny and functional definitions are not clear. Second and most importantly, the total number of labiate processes on a valve is usually not fixed in a genus. For example, some *Tabellaria* species have one labiate process and some have two (Lange-Bertalot 1988). The deformation model represents shape variations by a fixed number of landmarks that change locations from one shape to another, and it excludes the type of variation where one point in a form becomes two or more points in another. Although the midpoints of the terminal inflation and central inflations may be defined as pseudolandmarks for *Tabellaria*, they do not convey crucial shape variations between these points. Therefore, the coordinate deformation approach cannot be applied to *Tabellaria* shape descriptions. A totally different approach from the deformation model in shape analysis is pure outline analysis, and Fourier transformation plays a major role in this method. Theoretical and experimental evidence (Granlund 1972, Zahn and Roskies 1972, Kuhl and Giardina 1982) suggests that Fourier descriptors (FD's) approximate the criteria for *Tabellaria* shape representation. FD's for closed, two-dimensional curves may be summarized into three categories.

The first type of FD is the harmonics of the Fourier expansion functions defined as the radius about an arbitrarily defined center. This type of analysis has been applied to a variety of objects ranging from sand grains (Ehrlich and Weinberg 1970) to human faces (Lu 1965). Applications in systematics include the studies of Kaesler and Waters (1972), Christopher and Waters (1974), Healy-Williams and Williams (1981), and Rohlf and Archie (1984). The main problem in applying this type of FD to *Tabellaria* shape analysis is that FD's of this type are dependent on location of the center (Full and Ehrich 1982). In applications, either a landmark (Lu 1965) or the center of mass of an image (Ehrlich and Weinberg 1970) is used as the origin of the radius. This is a severe problem since it is difficult to specify meaningful ontogenetic landmarks in diatoms. The dependency of the descriptors to the location of the center of radius is undesirable even if a landmark can be easily defined. In morphometrics, it is not unusual that landmarks change their locations while the outline remains more or less the same. The change in FD's might only reflect the drifting of a landmark in the same outline shape. Using the center of mass as a landmark only ensures identical FD's for identical outlines, and the differences among different shapes produced this way are arbitrary.

A second approach utilizes the Fourier expansion of functions that represent changes in both the x and y directions. This type of analysis was developed in the area of pattern recognition and has been used to identify handwriting (Granlund 1972). The main purpose of this type of analysis is purely grouping and discrimination. All FD's obtained by this method are sensitive to scaling, and some transformation is necessary to make the descriptor

invariant under uniform scaling, rotation, and translation. Although the transformed set of descriptors is form-invariant, it is not a complete representation of the original curve, and the original contour cannot be reconstructed from the transformed descriptors. This property essentially disqualifies the method for *Tabellaria* shape analysis.

Another approach (Giardina and Kuhl 1977) is designed for contours that are encoded as chain code (Freeman 1974). There are two sets of Fourier expansion, one for the x projection of the chain code and one for the y projection. Consequently, there are four harmonics for each order of the representation. Kuhl and Giardina (1982) showed that the four harmonics in each order form elliptical loci and may be standardized to be form-invariant. The advantage of elliptical FD's is that the curve represented may be complex, and all arbitrarily shaped, solid objects can be analyzed. The disadvantage of elliptical FD's as applied to *Tabellaria* shape analysis is the high dimensionality of the descriptors. *Tabellaria* valves form simple closed outlines and, therefore, the extra flexibility of elliptical FD's will not benefit outline representation. Since the object of this study is statistical analysis and interpretation of FD's, curve representations that have lower dimensionality are preferred.

A type of FD that is scale-invariant and of lower dimensionality than elliptical FD's was first suggested by Cosgriff (1960) and developed by Zahn and Roskies (1972). They defined the FD's as follows:

Let l be the arc length of a clockwise-oriented simple closed curve γ and $0 \leq l \leq L$. The parametric representation of the curve γ is $(x(l), y(l)) = Z(l)$. Let $\Phi(l)$ be the cumulative angular function and $\Phi(l)$ represent the net amount angular bend between the starting point $l = 0$ and point l. The domain of $\Phi(l)$ is $[0, L]$ and may be normalized to $[0, 2\pi]$, which is standard for periodic functions. The normalized variant $\Phi^*(t)$ is defined as

$$\Phi^*(t) = \Phi(Lt/2\pi) + t. \tag{1}$$

The Fourier expansion of Φ^* is

$$\Phi^*(t) = \mu_0 + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt), \tag{2}$$

where a_k and b_k are the FD's for curve γ and are known as the kth harmonics for the Fourier series.

The Fourier expansion of γ may also be written in its polar form:

$$\Phi^*(t) = \mu_0 + \sum_{k=1}^{\infty} A_k \cos(kt - \alpha_k), \tag{3}$$

where A_k is the kth harmonic amplitude and α_k is the kth harmonic phase angle. It is very easy to convert (a_k, b_k) to (A_k, α_k). If (a_k, b_k) represents a point on an X and Y coordinate system, (A_k, α_k) represents the polar coordinates of that point.

Since terms in FD's represent successively finer details on the boundary and terms of the highest orders represent only noise, FD's may be truncated to represent the curve to a specified accuracy. The truncated form of FD's $\{A_k, \alpha_k\}^{\infty}$ is $\{A_k, \alpha_k\}^{N}$. Curve reconstruction based on the truncated FD's is obtained from the following formula:

$$Z(l) = Z(0) + L/2\pi \int_0^{2\pi l/L} \exp\{i[-t + \delta_0 + \mu_0 + \sum_{k=1}^{\infty} A_k \cos(kt - \alpha_k)]\} dT, \tag{4}$$

where $Z(0)$ is the starting point of the curve and δ_0 is the initial direction.

Reconstructions of two types of *Tabellaria* valves based on the first 20 descriptors are illustrated in Figure 1. In this analysis, even-numbered descriptors contribute to symmetrical aspects of a valve, and odd-numbered descriptors contribute to its asymmetrical aspects. The higher a descriptor's order, the finer the structure on the outline it controls.

Properties of Φ^* and its FD's that are important to *Tabellaria* shape analysis are the following. 1) The terms of FD's are independent and uncorrelated. Truncation of the Fourier series does not change the values of FD's. 2) Similar shapes are usually close to each other in the FD space. 3) For a circle, $\Phi^* \equiv 0$. The FD's may be considered a measurement of how different the shape in question is from a perfect circle. 4) Φ^* is invariant under scaling, rotation, and translation. The FD's are form-invariant except that they are sensitive to the starting point of the contour.

We have developed programs to implement this analysis. Images are captured directly from a Leitz Dialux 20 microscope equipped with flat-field optics and are sent to an Apple Macintosh II computer through a ProViz digitizer. Processing begins with application of a Gaussian filter (Marr and Hildreth 1980, Canny 1986). The outlines on the smoothed image are then detected, and the contour of interest is digitized automatically.

Computational formulas for FD's obtained directly from a digitally represented curve are

$$\mu_0 = -\pi - 1/L \sum_{k=1}^{M} l_k \Delta\Phi_k \tag{5}$$

$$a_n = -1/n\pi \sum_{k=1}^{M} \Delta\Phi_k \sin(2\pi\, n\, l_k/L) \tag{6}$$

$$b_n = 1/n\pi \sum_{k=1}^{M} \Delta\Phi_k \cos(2\pi\, n\, l_k/L) \tag{7}$$

where l_k is the arc length from the starting point to the kth point, $\Delta\Phi_k$ is the change of the angular bend on the curve from the $(k - 1)$th point to the kth point, and there are $(M - 1)$ points on the curve.

Statistical analysis of the FD's. Outlines of 490 *Tabellaria* valves suitable for digitization (i.e. outlines lying parallel to the focal plane and without overlapping images) were obtained and each of the outlines was digitized from the center of a terminal inflation. Girdle bands were sometimes sampled in thickly strewn preparations where most valve outlines were overlapping. The outline shape of girdle bands in a cell is almost identical to the shape of the valve (Theriot and Ladewski 1986). We sampled only the valves whenever it was possible to avoid repeated sampling of a single specimen.

Morphological variations of *Tabellaria* valves found in our collection were well sampled. The specimens cover a wide range of North America localities. Although strict random samples are practically impossible to obtain since the availability of qualified images on a slide varies, outlines for this study were obtained in the following manner. All qualified images along certain transects on a slide were digitized until a prefixed number was reached. In other words, there were no "preferred shapes" among valves that could be digitized. Many valves that might be considered "atypical" by some taxonomists were included.

The first 30 FD's were extracted for each specimen. This number was determined by both observation of shape reconstructions and several exploratory principle component analyses (PCA's). The reconstructed curves based on the first 30 FD's showed all taxonomically important shape details on *Tabellaria* valves. All PCA's on the correlation matrix over the entire data set and subsets indicated that FD's after the first 30 coefficients are not significant. The preliminary PCA results also show that the harmonic angles are not significant. This is determined by the fact that none of the terms showed up in the first three components that explain over 80% of the variation. Consequently, the first 30 harmonic amplitudes from each of the specimens form the raw data for our shape analysis, and they are labelled A_1, A_2, A_3, etc. Besides the harmonic amplitudes, each specimen's length and width were measured, and the length-to-width ratio was calculated. Geographic origin of the specimens was also recorded. This information was used later to help determine the geometric meaning of the shape descriptors and the geographic distributions of shape groups.

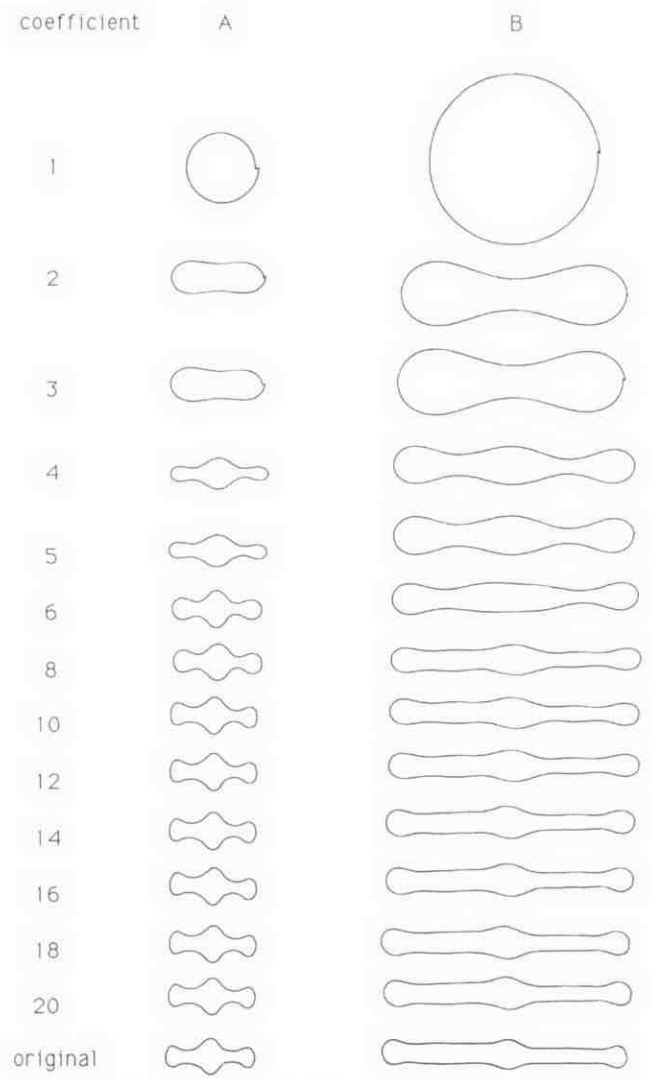

FIG. 1. Successive reconstruction of two *Tabellaria* valves based on up to 20 descriptors.

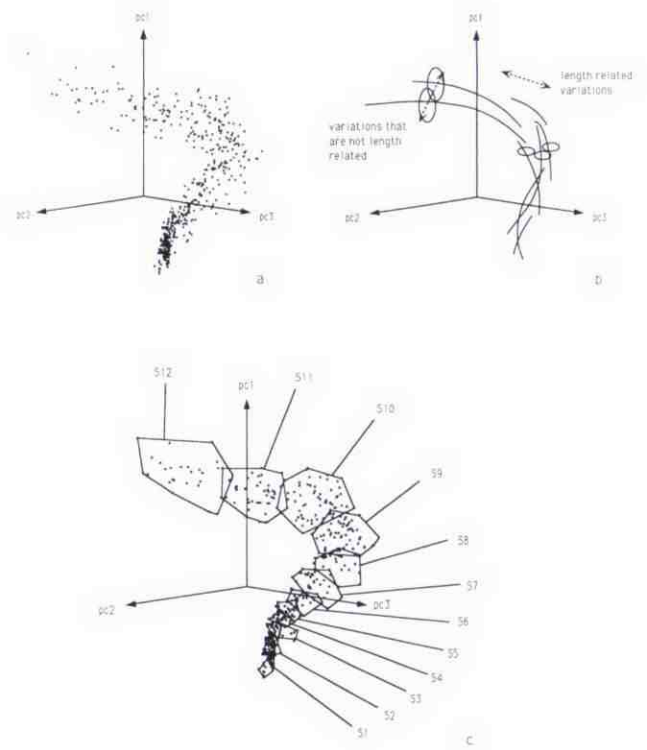

FIG. 2. The path to shape-group separation. a) Plot of the entire data set on the first three components of the global PCA. b) The conceptual model of the raw data. c) Sub-sectioning of the entire data set.

To investigate the nature and structure of the raw data, PCA on the covariance matrix was applied to them. A plot of the distribution of the 490 samples on PC1, PC2, and PC3 of this global analysis (Fig. 2a) shows a strong nonlinear functional relationship among the three variables. Nonlinear features are not unusual in biological pattern recognition. For one organism, the relationship between size change in some dimension and size change in another dimension is commonly known as allometry, and it is often exponential. In *Tabellaria,* allometry exists because cell division in a clone causes successive valve length reductions and shape changes. Although the distribution in the global PCA is not allometry, since the data are not from one clone or one species, it is certainly influenced by a combination of different allometric characteristics from many different species. Based on a rough comparison of our specimens with species described by previous studies of the genus *Tabellaria* and the information that shape variation is larger among the shorter specimens (Theriot and Ladewski 1986), and based on observations of the data distribution in the first three PC spaces from every possible angle and the fact that both PC1 and PC2 are correlated with length, a conceptual model may be built for the distribution in Figure 2a (Fig. 2b): There is more than one species in the raw data. Each species' specific allometric characteristics form a unique growth trajectory in the shape space, and this trajectory may have a different starting and finishing point from that of other species. Some species may have longer nonallometric shape variations than others, and some species might overlap with one another.

Nonlinear functional distribution of the raw data created a problem for standard multivariate techniques, since the assumption of normality is violated. For this type of distribution, PCA cannot detect the intrinsic dimensionality of the data (Fukunaga 1972). For example, the variables PC1 and PC2 from the global PCA in our analysis are not independent (Fig. 2) as they should be for normal distributions. The functional relationships among global PC's is one of the reasons why a shape analysis method called eigenshape analysis cannot be applied to *Tabellaria.* This method derives a set of empirical orthogonal shape functions by PCA of a matrix of correlations between shapes (Lohmann 1983). Essentially, it applies PCA to some points from the curve description function $\Phi^*(t)$ in equation 1. Its purpose is to find a representative shape in a group. Since the shape of each species in *Tabellaria* is a continuous growth series and a different growth stage has a different representative shape, it is wrong to try and find one representative shape knowing that there might be more than one species in the data. Eigenshape located this way would be a shape that is representative of several species at a similar growth stage, and this is not our objective. Mathematically, the problem of the empirical orthogonal functions (EOF's) obtained from our data is similar to the global PCA. Although harmonic amplitudes from truncated Fourier transformation of $\Phi^*(t)$ are not equal to points on $\Phi^*(t)$, the effect of the EOF's from our data is close to what is illustrated in Figure 2, and it would be dominated by allometric shape variations and species delineation would not be possible. Besides, eigenshape analysis requires that a fixed number of points be digitized from every specimen. For *Tabellaria,* the spatial correspondence between a fixed number of points in one outline and the same number of points in another outline is totally arbitrary, and correlations derived from these

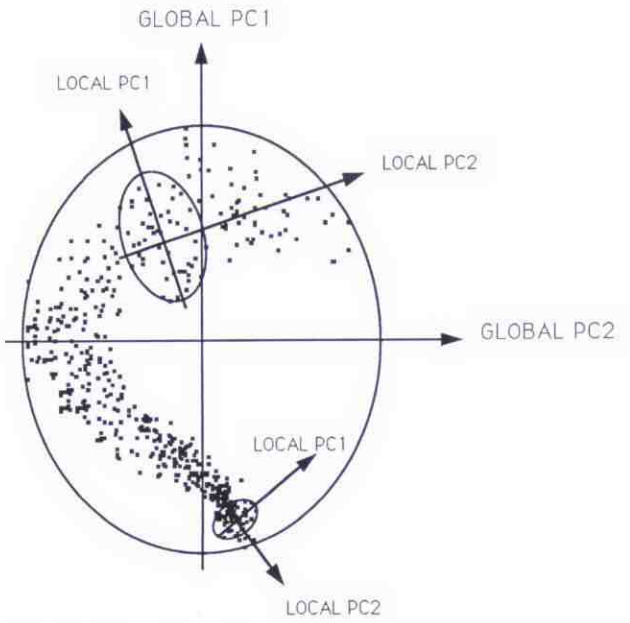

FIG. 3. A two-dimensional example of the relationships among the local PCA's and the global PCA for the raw data.

points are misleading, as are the EOF's derived from the correlations. Therefore, this method is not suitable for our study.

To delineate species in *Tabellaria*, it is crucial to understand how different sources of shape variation influence the distribution of the data. There are within-species variations and between-species variations. Within-species variations may be divided into two categories: allometric variations, which determine the shape and the location of the arcs in Figure 2b, and nonallometric variations, which determine the radius of the circles around each arc at any given point or the "thickness" of each curve, which may not be consistent throughout a growth trajectory. Between-species variations influence the "thickness" and the shape of the overall distribution. In other words, it controls how far away the growth trajectories are from one another. The overall variation may also be divided into length-related variations and non-length-related variations. It is necessary to emphasize here that although global PC1 and PC2 are allometry-related shape variations and consequently reflect valve length to certain degrees, they are not direct measures of length. Identical shapes of different magnifications would be at the same spot in the PC space even though they have different lengths. At any given point of the overall distribution, shape variations that are not length-related are composed of intraspecific nonallometric variations and interspecific variations at a similar length. Graphically, this is the cross-section of the overall distribution at that point (Fig. 2b). Shape variations at any other directions are all length-related. To delineate species, we need to magnify shape variations that are not related to length.

Our raw data consist of 30 variables. If it were possible, a plot of the data in a 30-dimensional FD space would closely resemble the distribution in Figure 2a. PCA rotated the original axes so that only three new axes are necessary to plot the raw data with only 19% loss of the original variation (Table 1). In other words, PCA is a tool to "see" the data structure in lower dimensional spaces, and it does not alter the data structure. To "see" if there is more than one growth trajectory in a certain segment of the overall distribution, PCA may be applied to this segment. A 2-dimensional analogy of the effect of this local PCA is like looking through a magnifying glass in a 30-dimensional space. We rotate the magnifying glass in the data segment until we "see" most of the variation reflected through the glass. The accuracy of PCA is usually less than 25% loss of the original variation. For any particular study, this criteria may be specified as accurately as necessary. To observe the entire distribution, we move the magnifying glass from segment to segment until every data point is observed. This moving magnifying glass effect may be achieved by dividing the overall distribution into small subsections and applying PCA to each of the subsections. A two-dimensional example of the relationships among the local PCA's and the global PCA is illustrated in Figure 3. It shows that a local PC1 might have a different orientation or have a different linear combination of the original variables from the global PC1 and other local PC1's. The same is true for PC2's and so on.

To observe species in a subset, it is easier to limit allometry-related variations in a subset so that the local PC1 is aligned toward the between-species shape variations if there is more than one species in the subset. It means that the length of the segment should be shorter than its width. The size of a subset divided this way in the data distribution where it is "thin" is smaller than where it is "thick," although data points included in the smaller subset are not necessarily less than a larger subset (Fig. 3). Using the analogy of a magnifying glass again, it is like having an oval magnifying glass, and it always adjusts the magnification so that the maximum nonallometric variation in a subset always fills up the major axis of the glass. In fact, it is not crucial to fix the minor axis of the glass; as long as the major axis of the glass is fixed, PC1's from all subsets are standardized to this distance. In practice, this means that as long as length-related shape variations are not dominant, the exact shape of the subset is not crucial to species delineation.

After several trial divisions, our raw data were divided into 12 subsets and each subset overlaps partially with its neighbors (Fig. 2c). The amount of overlapping among subsets is determined by the complexity of the raw data. The more loops and protuberances on the growth trajectories in the conceptual model, the more slowly we move the magnifying glass to observe the details. Growth trajectories in our model are assumed to be smooth, based on preliminary observations of the data and the common knowledge that diatom cell division produces gradual shape changes, and that there are no periodic repetitions of shapes in a growth series. The 12 subsets in our analysis only overlap slightly, and data points in the overlapping areas are used to check if the classifications are consistent in different subsets. The 12 subsets are labelled as S_1, S_2, S_3, etc., up to S_{12}. PCA over the covariance matrices was applied to each of the subsets. The first three components were chosen as the dominant eigenvectors because they explained at least 75% of the variance for most of the subsets (Table 1). Plots of the first three components were used to find natural shape clusters in the local regions. No strict criteria define a shape cluster in a local PCA, since the shape variance may differ from one species to another. Looking through the magnifying glass again, some clusters may be more loosely scattered than others. To make the matter even more complicated, shape variances may not be consistent within a species throughout its growth trajectory. The only way to tell one species from another is to compare the characteristics of the cluster with characteristics of clusters in other subsets. If one cluster is loosely distributed in a subset, it should be distributed in a similar pattern in the neighboring subsets. For the 12 subsets in our raw data, shape groups in each subset were compared with shape groups in its neighboring subsets. Groups that were similar were combined into one growth series. Figure 4 shows representative specimens from each group. Figure 5 shows specimen distribution in the PC space for each of the subsets.

RESULTS

Eight shape groups were found among the 490 *Tabellaria* valves we sampled. Each group is illustrated by six outlines representative of that group in Figure 4a–h. Only pure morphometric properties

of the valve outlines are discussed here. Systematic interpretation of these results will depend on further observations of nomenclatural types, observations on characters other than valve shape, and, most likely, additional sampling of populations not represented in our collection. While we will continue to pursue such research, we feel that these interim results will be of interest from the stand point of methodology, perhaps suggesting directions for research to other investigators.

Since data in this study are subsectioned based on their distribution on the first three components of the PCA of all 490 specimens (Fig. 2c), the geometric properties of each shape group may be interpreted by observing its distribution on the three axes in that analysis. For example, a shape group that has a wide range over the PC1 axis indicates a large size range in specimens sampled that are part of the group.

Group 1 (Fig. 4a) is only found in S_1 (Fig. 5a). Since it is only located in S_1, its shape is homogenous and size range is small.

Group 2 (Fig. 4b) is found in S_1, S_3, S_5, S_6, S_7, and S_8 (Fig. 5a, c, e–h). Although this group has a large size range, its shape is relatively stable.

Group 3 (Fig. 4c) is found in S_1, S_2, S_3, S_5, S_6, S_7, S_8, and S_9 (Fig. 5a–c, e–i). It has a very large size range, and the intrinsic variation of valve shape is also large.

Group 4 (Fig. 4d) is found in S_2 and S_4 (Fig. 5b, d). This group is fairly homogeneous in both size and shape.

Group 5 (Fig. 4e) is found in S_2, S_4, S_5, and S_6 (Fig. 5b, d–f). Its size range is large, and its shape is not very stable.

Group 6 (Fig. 4f) is found in S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{10}, and S_{11} (Fig. 5c–k). This group has very large variation in both size and shape.

Group 7 (Fig. 4g) is found in S_7, S_9, S_{10}, and S_{11} (Fig. 5g, i–k). The size range of specimens in this group is small, and the shape variation is relatively large.

Group 8 (Fig. 4h) is found in S_9, S_{10}, S_{11}, and S_{12} (Fig. 5i–l). Specimens in this group represent little size reduction, and the group has very large internal shape variation.

Since the subsets are designed to have smaller length-related shape variations, characteristics that are not correlated with valve length or the length-to-width ratio play a primary role in shape-group separation. Among the local PCA's of the subsets from S_2 to S_{11}, PC1's are not correlated with length or length-to-width ratio. Only two subsets, S_1 and S_{12}, have PC1's that are slightly correlated with the ratio. PC3's in all of the subset analyses are also purely shape characteristics and are not correlated with the length or length-to-width ratio.

Characteristics that are correlated with valve length or length-to-width ratio play a secondary role in shape-group delineation. PC2's in subsets from

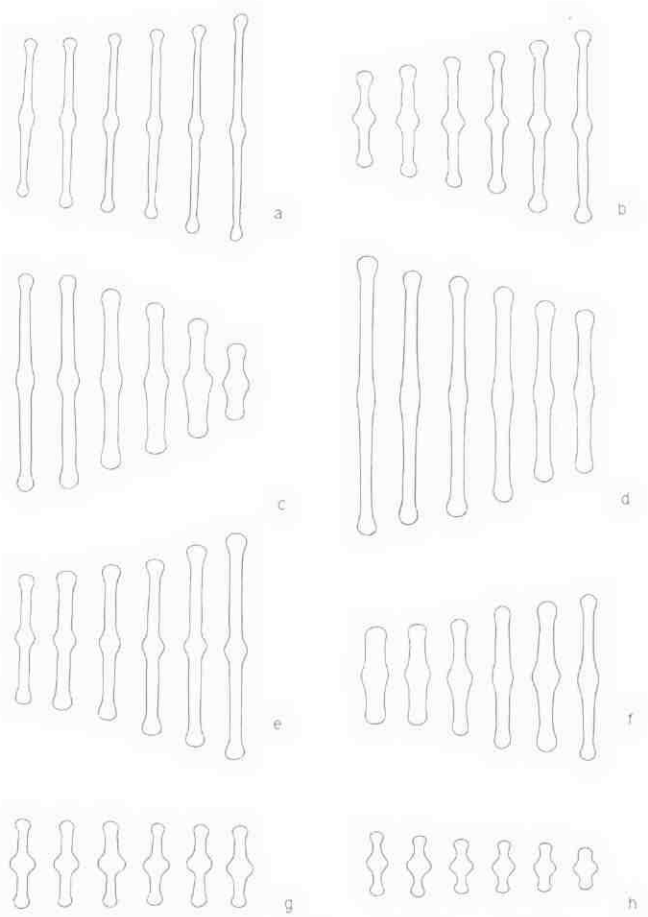

FIG. 4. *Tabellaria* valves a) Group 1. b) Group 2. c) Group 3. d) Group 4. e) Group 5. f) Group 6. g) Group 7. h) Group 8.

S_2 to S_{11} are all correlated with the ratio or length to various degrees. Among all of our sampled *Tabellaria* valves, width variation is much smaller than length variation, and width is relatively constant in some groups. Therefore, the ratio of length over width is highly correlated with the length. However, most of the PC2's have higher correlation with the ratio than with the length.

Table 1 shows that the variance explained by PC1, PC2, and PC3 together is fairly high in each of the subset analyses. This means that clusters found in the space formed by PC1, PC2, and PC3 of the local PCA's are fairly accurate. Although the geometric meaning of each of the variables or descriptors can be interpreted by reconstructions and observations, the various combinations of these descriptors such as PC1's and PC2's in this analysis are usually too complex and subtle to be verbally interpreted. This is also why verbal descriptions alone are not adequate in *Tabellaria* taxonomy.

Since S_1 to S_{12} are distributed along the global PC1, which roughly reflects valve length, it means that the valve becomes shorter as the subset order goes higher. From S_1 to S_6 (Table 1), higher-ordered descriptors such as A_{10}, A_{14}, A_{18}, and A_{22} are dom-

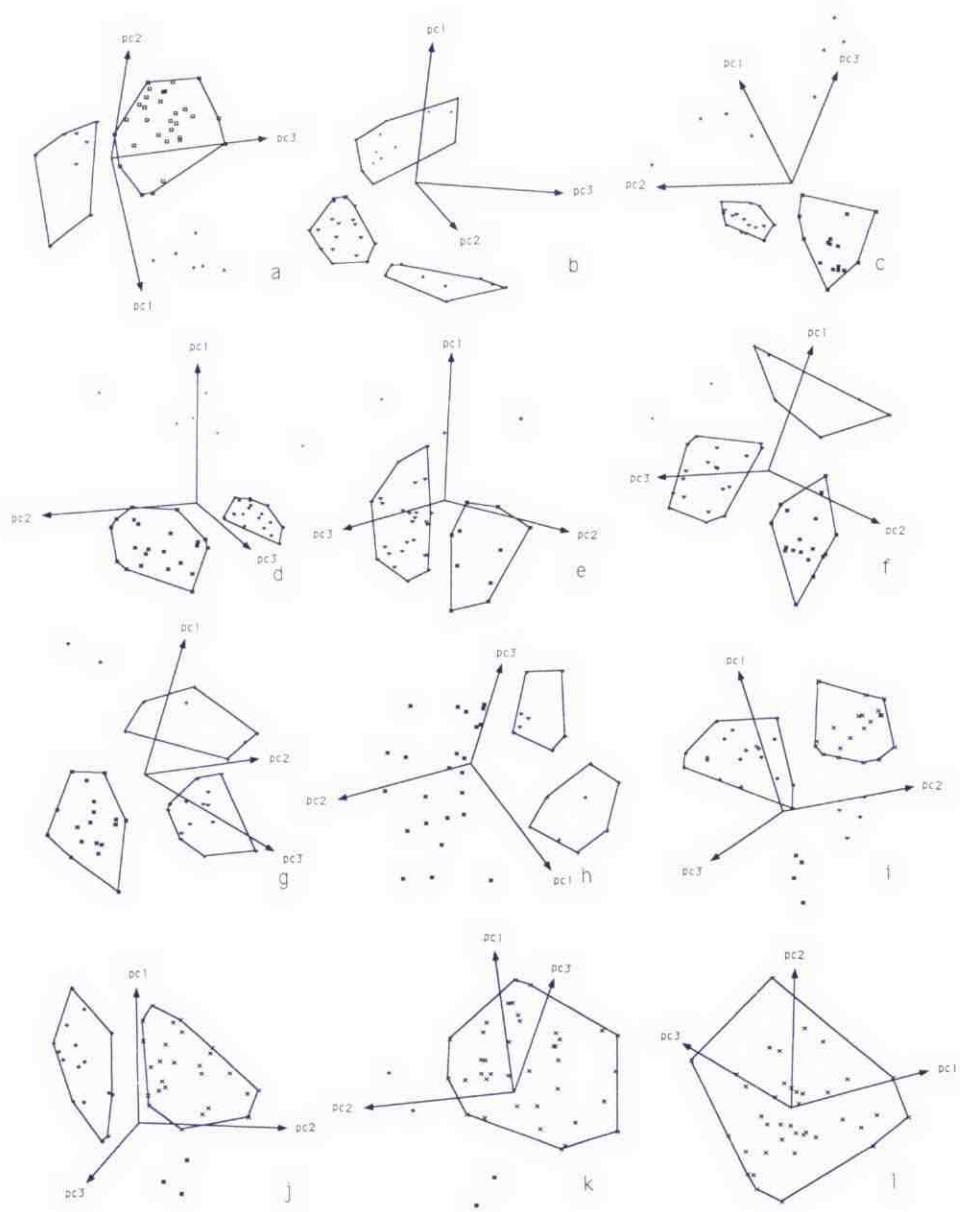

FIG. 5. Clusters of *Tabellaria* valves from S_1 to S_{12} ranging from a to l. (□) Group 1, (+) Group 2, (▽) Group 3, (○) Group 4, (■) Group 5, (× ×) Group 6, (◇) Group 7, and (×) Group 8.

inant. This indicates that shape groups with longer valves are differentiated mainly by the finer structures on the valve outline. From S_7 to S_{12} (Table 1), lower-ordered descriptors such as A_2, A_4, A_6, and A_8 are dominant. This indicates that shape groups with generally shorter valves are differentiated mainly by the more fundamental shape characters.

Asymmetrical aspects of the valve outline have proven to be important in delineating shape groups but should be interpreted with caution. Almost all of the eight shape groups have valves with some degree of asymmetry. There are more asymmetrical valves in the shorter general shape groups; however, the degree of asymmetry is greatest in the smallest specimens of all groups. A_3 is a descriptor that measures the equality of the two limb widths of a given valve. A_3 has the second highest loading on PC1 in S_{11} and PC2 in S_{12} (Table 1). A_5, which is a descriptor that measures the asymmetrical aspects of a valve along its apical axis, also has high loadings on PC3's in S_9, S_{10}, and S_{11} (Table 1).

DISCUSSION

Most of the shape groups resolved in this study have similar and overlapping length ranges and can only be separated by pure shape characteristics (e.g. Group 3 and Group 6). Each of these shape groups has its unique morphometric characteristics. The analysis shows that the size ranges from the eight groups do not completely overlap and resolve into

TABLE 1. *Variables with high loadings in global and local PCA's.*

	Main PC1		Main PC2		Main PC3		% Variance
Global	A2	(−0.100)	A6	(0.083)	A4	(0.059)	81
	A8	(−0.074)	A2	(−0.048)	A10	(0.029)	
S_1	A10	(−0.044)	A14	(−0.023)	A14	(0.013)	78
	A12	(−0.037)	A22	(0.017)	A18	(0.012)	
S_2	A10	(−0.041)	A22	(0.018)	A14	(0.016)	79
	A14	(0.029)	A18	(0.017)	A16	(0.070)	
S_3	A14	(0.041)	A4	(−0.022)	A10	(0.026)	78
	A18	(0.022)	A18	(0.020)	A8	(0.005)	
S_4	A14	(0.038)	A10	(0.010)	A16	(0.012)	72
	A18	(0.030)	A4	(0.010)	A11	(0.008)	
S_5	A10	(0.040)	A4	(0.025)	A8	(0.012)	80
	A14	(0.029)	A14	(−0.023)	A12	(−0.010)	
S_6	A10	(0.053)	A4	(0.0330)	A8	(0.018)	
	A6	(−0.024)	A14	(−0.024)	A6	(0.011)	79
S_7	A4	(0.057)	A10	(0.0380)	A10	(−0.170)	79
	A8	(0.045)	A4	(−0.026)	A6	(0.017)	
S_8	A4	(0.078)	A10	(−0.050)	A6	(−0.030)	81
	A8	(0.031)	A4	(0.014)	A8	(−0.019)	
S_9	A4	(0.096)	A8	(−0.055)	A5	(0.032)	77
	A8	(0.025)	A4	(0.023)	A3	(0.017)	
S_{10}	A4	(0.086)	A10	(−0.049)	A5	(0.026)	75
	A6	(−0.018)	A2	(−0.033)	A7	(0.017)	
S_{11}	A4	(0.076)	A2	(0.039)	A3	(−0.026)	73
	A3	(0.013)	A10	(0.034)	A5	(−0.021)	
S_{12}	A4	(−0.077)	A6	(0.038)	A2	(−0.027)	75
	A2	(−0.064)	A3	(0.026)	A3	(0.021)	

intuitively reasonable size diminution series. Once resolved, the shape series are all visually distinct and, in our estimation, approximate expectations of the shape range found in most diatom species. Since the geographic origins of the 490 specimens were recorded during the sampling process, it is easy to determine geographic distribution patterns among the shape groups. In the examples studied, we find both sympatric and allopatric shape-group distributions. In some cases only a single shape group will be found in a general (lake) or specific (collection) locality. More commonly, distinctly different shape groups coexist in the same habitat. As might be expected, our results indicate that more shape groups are found in large lakes than in small lakes. Our observations indicate that *Tabellaria* exhibits less ecophenotypic plasticity than some other diatom genera (e.g. *Stephanodiscus*, Theriot and Stoermer 1984). A formal analysis of morphological responses to ecological factors in *Tabellaria* is presently underway (Mou and Stoermer, unpubl. data).

Formal shape analysis allows distinction of groups which have different size ranges but vary only subtly in other characters. Due to the large size and shape change undergone by diatom species during asexual reproduction, it is understandably easy to confuse relationships of populations unless size-shape relationships are resolved. Once these relationships are resolved, it becomes obvious that several shape groups in this study have very different size ranges. For example, the populations resolved as Group 1 and Group 8 have completely nonoverlapping size ranges. While this might seem a trivial result, it is probable that some specimens from both populations would be considered conspecific on the basis of currently available taxonomic treatments. While our current results do not conclusively resolve relationships, they do suggest an alternate interpretation. Finally, our analysis agrees with the observation that shape variability increases as valve length decreases (Geitler 1932, Theriot and Ladewski 1986).

Our results reveal that the valve shape is a crucial character in *Tabellaria* taxonomy. With the accuracy of FD's, it is possible to separate populations with shape differences too subtle to be verbally described with adequate accuracy and completeness. With the technique of sub-sectioning of a nonlinear distribution, it is possible to address the problem of within-species size reduction and shape variation due to vegetative reproduction. Therefore, valve shape is still a reliable taxonomic character even with the interference of size and shape changes induced by cell division.

Results obtained from this study should be reasonably robust for two reasons. 1) *Tabellaria* collections in our lab are fairly extensive and representative of a wide variety of habitat types in North America. 2) Samples taken are random, and the sample size is large; thus, specimens included in this

study are unbiased representatives of the *Tabellaria* valves in our collection. Additional sampling, particularly from localities outside North America, will likely lead to discovery of shape groups not included in the present analysis; however, it is doubtful that basic trends discovered here will be significantly modified by addition of further samples.

The final results from this study supported the conceptual model illustrated in Figure 2b. Each *Tabellaria* shape group forms a curvilinear trajectory in the global shape space. These results imply that other diatom genera may have a similar growth pattern, and species may be delineated in similar ways.

The results also indicate that the large sample approach is appropriate for general numerical *Tabellaria* shape study. Large unbiased samples provide each species with representative specimens along its growth curve so that species delineation is possible. If our sample is much smaller, the data may contain a few specimens from each shape group so that the data are too sparse to show any trend. Moreover, if there are only specimens from a few segments of the global distribution, we would tend to believe that each segment is a species, and this result obviously contradicts the fact that *Tabellaria* species go through length reduction. Actually, clusters obtained in this way are collections of different species at similar length. Although it is likely that we did not have the complete range for every shape group in our sample, the data are large enough to show some basic distributional properties of the shape groups.

Since most of the shape groups resolved range over several subsets, it is obvious that they have highly non-normal distributions. Consequently, it is impossible to apply many commonly employed multivariate tests of discrimination to these groups because the assumption of normality is not valid. Especially, the mean group differences cannot be obtained. To test the validity of the group separation, it will ultimately be necessary to analyze taxonomic characters other than outline to see if they segregate accordingly. These matters will be discussed in a future paper.

We thank the Natural History Museum, London, particularly Ms. P. A. Sims; the Academy of Natural Sciences, Philadelphia, particularly Dr. C. W. Reimer; and the California Academy of Sciences, particularly Dr. J. P. Kociolek, for loan of material. We would also like specifically to recognize the contributions of Mr. T. B. Ladewski to studies of diatom shape analysis, including *Tabellaria*, in our lab. Contribution number 545 of the Center for Great Lakes and Aquatic Sciences, University of Michigan, was funded by NSF grant BSR-9006727.

Bookstein, F. L., Chernoff, B., Elder, R. L., Humphries, J. M., Smith, G. R. & Strauss, R. E. 1985. Morphometrics in evolutionary biology. *Acad. Nat. Sci. Pa., Spec. Publ.* 15:1–277.

Canny, J. 1986. A computational approach to edge detection. *IEEE Trans., PAMI* 8:679–98.

Christopher, R. A. & Waters, J. A. 1974. Fourier series as a quantitative descriptor of miosphere shape. *J. Paleontol.* 48: 697–709.

Cosgriff, R. L. 1960. Identification of shape. Ohio State Univ. Res. Foundation, Columbus, Rep. 820-11, ASTIA AD, pp. 254–792.

Ehrlich, R. & Weinberg, B. 1970. An exact method for characterization of grain shape. *J. Sed. Petrol.* 40:205–12.

Freeman, H. 1974. Computer processing of line drawing images. *Comp. Surv.* 6:57–97.

Fukunaga, K. 1972. *Introduction to Statistical Pattern Recognition.* Academic Press, New York, 369 pp.

Full, W. E. & Ehrlich, R. 1982. Some approaches for location of centroids of quartz grain outline to increase homology between Fourier amplitude spectra. *Math. Geol.* 14:43–55.

Geitler, L. 1932. Der Formwechsel der pennaten Diatomeen. *Arch. Protistenk.* 78:1–226.

Giardina, C. R. & Kuhl, F. P. 1977. Accuracy of curve approximation by harmonically related vectors with elliptic loci. *Comp. Graph. Image Process.* 6:277–85.

Goldman, N., Paddock, T. B. B. & Shaw, K. M. 1990. Quantitative analysis of shape variation in populations of *Surirella fastuosa. Diatom Res.* 5:25–42.

Granlund, G. H. 1972. Fourier preprocessing for handprint character recognition. *IEEE Trans.* C-22:195–201.

Healy-Williams, N. & Williams, D. F. 1981. Fourier analysis of test shape of planktonic foraminifera. *Nature (Lond.)* 289: 485–7.

Hustedt, F. 1931. Die Kieselalgen Deutschlands, Österreichs und der Schweiz unter Berücksichtingung der übringen Länder Europas sowie der angrenzenden Meersgebiete. Band 7, Teil 2, Lieferung 1. *In* Rabenhorst, L. [Ed.] *Kryptogamenflora von Deutschland, Österreichs und der Schweiz.* Akademische Verlagsgesellschaft Geest und Portig K-G., Leipzig, Deutschland, pp. 25–31.

Kaesler, R. L. & Waters, J. A. 1972. Fourier analysis of ostracod margin. *Geol. Soc. Am. Bull.* 83:1169–78.

Kingston, J. C. & others. 1986. *PIRLA Diatom Iconograph.* A privately published reference guide for the project "Paleo-Limnological Investigation of Recent Lake Acidification." PIRLA Report Series No. 3, Indiana University, Bloomington, 53 plates, 1059 figures.

Knudson, B. M. 1952. The diatom genus *Tabellaria.* I. Taxonomy and morphology. *Ann. Bot., N.S.* 15:421–40.

——— 1953a. The diatom genus *Tabellaria.* II. Taxonomy and morphology of the plankton varieties. *Ann. Bot., N.S.* 17:131–55.

——— 1953b. The diatom genus *Tabellaria.* III. Problems of intra-specific taxonomy and evolution. *Ann. Bot., N.S.* 17: 598–609.

Koppen, J. D. 1973. Distribution of the species of the diatom genus *Tabellaria* in a portion of the north-central United States. Ph.D. thesis, Department of Botany, Iowa State University, Ames, 270 pp.

——— 1975. A morphological and taxonomic consideration of *Tabellaria* (Bacillariophyceae) from the north-central United States. *J. Phycol.* 11:236–44.

——— 1978. Distribution and aspects of the ecology of the genus *Tabellaria* Ehr. (Bacillariophyceae) in the north-central United States. *Am. Midland Nat.* 99:383–97.

Kuhl, F. P. & Giardina, C. R. 1982. Elliptic Fourier features of a closed contour. *Comp. Graph. Image Process.* 18:236–58.

Lange-Bertalot, H. 1988. Die Gattung *Tabellaria* unter besonderer Berücksichtigung an *Tabellaria ventricosa* Kützing (Bacillariophyceae). *Nova Hedwigia* 46:413–31.

Lohmann, G. P. 1983. Eigenshape analysis of microfossils: a general morphometric procedure for describing changes in shape. *Math. Geol.* 15:659–72.

Lu, K. J. 1965. Harmonic analysis of the human face. *Biometrics* 21:491–505.

Marr, D. & Hildreth, E. 1980. Theory of edge detection. *Proc. R. Soc. Lond. B* 207:187–217.

Rohlf, F. J. & Archie, J. W. 1984. A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Syst. Zool.* 33:302–17.

Stoermer, E. F. & Ladewski, T. B. 1982. Quantitative analysis

of shape variation in type and modern populations of *Gomphoneis herculeana*. *Nova Hedwigia, Beih.* 73:347–86.
Stoermer, E. F. & Yang, J. J. 1969. Plankton diatom assemblages in Lake Michigan. Special Report No. 47, Great Lakes Research Division, University of Michigan, Ann Arbor, 268 pp.
Theriot, E. C. & Ladewski, T. B. 1986. Morphometric analysis of shape of specimens from the neotype of *Tabellaria flocculosa* (Bacillariophyceae). *Am. J. Bot.* 73:224–9.
Theriot, E. C. & Stoermer, E. F. 1984. Principal components analysis of character variation in *Stephanodiscus niagarae* Ehrenb.: morphological variation related to lake trophic status. *Proc. Int. Diatom Symp.* 7:97–111.
Thompson, D. W. 1961 [1917, 1942]. *On Growth and Form*, abridged ed., Bonner, J. T. [Ed.]. The University Press, Cambridge, 346 pp.
Zahn, C. T. & Roskies, R. Z. 1972. Fourier descriptors for closed plane curves. *IEEE Trans.* C-22:269–81.

J. Phycol. **28,** 395–399 (1992)

ALEXANDRIUM SATOANUM SP. NOV. (DINOPHYCEAE) FROM MATOYA BAY, CENTRAL JAPAN[1]

Katsuhisa Yuki[2]

Matoya Oyster Research Laboratory, Matoya, Isobe-cho, Mie, 517-02 Japan

and

Yasuwo Fukuyo

Faculty of Agriculture, University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo, 113 Japan

ABSTRACT

The gonyaulacoid dinoflagellate Alexandrium satoanum *Yuki et Fukuyo sp. nov. is described from Matoya Bay, Pacific coast of central Japan. The species is distinctive in its conical epitheca with almost straight sides and dorsal concavity of the hypotheca. The plate formula is Po, Pc, 4′, 6″, 6c, 10s, 5‴, and 2⁗, including two accessory plates inside the sulcus. The apical pore plate is triangular and possesses an anterior attachment pore at the right margin. The first apical plate does not make contact with the apical pore plate and lacks a ventral pore. A posterior attachment pore lies at the center of the posterior sulcal plate. In Matoya Bay, vegetative cells occur as solitary cells or sometimes in pairs during late spring and early summer in low concentrations. In connection with this study, the following new combination is proposed:* Alexandrium pseudogonyaulax *(Biecheler) Horiguchi ex Yuki et Fukuyo comb. nov.*

Key index words: **Alexandrium satoanum** *sp. nov.; Dinophyceae; Matoya Bay; seasonal occurrence; thecal morphology*

The transfer of species belonging to the so-called *tamarensis/catenella* group of gonyaulacoid dinoflagellates to the genus *Alexandrium* Halim has been argued by Balech (1985, 1989) and has found general acceptance (Steidinger 1990). This transfer increased the number of species of the genus *Alexandrium* to over 20. Most recently, Balech (1990b) advocated subdividing the genus *Alexandrium* into two subgenera, *Alexandrium* and *Gessnerium* Halim. Of these, the former consists of the species having direct or indirect contact between the apical pore plate (Po) and the first apical plate (1′) and accordingly includes the species formerly assigned to the genus *Protogonyaulax* Taylor by Taylor (1979) and Fukuyo et al. (1985, 1988). The latter consists of the species lacking the Po and 1′ contact as in *A. monilatum* (Howell) Taylor and *A. balechii* (Steidinger) Taylor. In Japan, 11 species of *Alexandrium* have previously been recorded from coastal waters and rockpools (Fukuyo et al. 1985, Kita and Fukuyo 1988, Yuki and Yoshimatsu 1990, and others).

During microscopical observations on dinoflagellates in samples obtained from Matoya Bay, Pacific coast of central Japan, our attention was focused on an *Alexandrium* species having a distinctive appearance. We ascertained that some taxonomically important characters of this species differed distinctly from those of the species previously described in the genus. In this paper, we propose a new taxon for this dinoflagellate and give a detailed description of its thecal plate arrangement.

In Balech's (1990b) revised diagnosis of *Alexandrium*, our new species belongs to the subgenus *Gessnerium*, in which the most ventral epithecal plate is disconnected from the apical pore plate. Although this plate should be regarded as the first precingular (1″) in the conventional Kofoid system, we describe it as the first apical (1′) in order to consider the plate homology and to avoid unnecessary confusion with

[1] Received 8 July 1991. Accepted 15 January 1992.
[2] Address for reprint requests.