

RECOGNITION OF TAXONOMICALLY SIGNIFICANT CLUSTERS NEAR THE SPECIES LEVEL, USING COMPUTATIONALLY INTENSE METHODS, WITH EXAMPLES FROM THE *STEPHANODISCUS NIAGARAE* COMPLEX (BACILLARIOPHYCEAE)¹

Matthew L. Julius²

Center for Great Lakes and Aquatic Sciences, University of Michigan, 2200 Bonisteel Boulevard, Ann Arbor, Michigan 48109-2099

George F. Estabrook

Department of Biology and University Herbarium, University of Michigan, Ann Arbor, Michigan 48109-1048

Mark B. Edlund and Eugene F. Stoermer

Center for Great Lakes and Aquatic Sciences, University of Michigan, 2200 Bonisteel Boulevard, Ann Arbor, Michigan 48109-2099

ABSTRACT

Since the early 1960s, numerical techniques have produced a wide variety of methods to suggest classifications of organisms based on quantitative measurements. A long-recognized shortcoming of these methods is that they will suggest classifications for any group of organisms and any set of measurements, whether or not the clusters in the suggested classification have any natural meaning or significance. Some progress has been made in assessing the reality of clusters determined by various methods. Data simulated to reflect known cluster structure have been used to test the accuracy of different methods. Various methods have been applied to the same data sets to compare how well they realize various desirable properties. Here we define a data-based model of randomness to represent what might be meant by "no natural basis for subdivision into clusters" and use it to compare an observed measure of cluster distinctness to the distribution of this measure predicted by this model of randomness. In this way, unwarranted subdivision can be statistically avoided, and significant subdivisions can be investigated with confidence. Our methods are illustrated with some examples from the *Stephanodiscus niagarae* Ehrenb. species complex. Significant differences in morphologic expression are identified in *S. reimerii* Theriot and Stoermer in Theriot, *S. superiorenensis* Theriot and Stoermer, and *S. yellowstonensis* Theriot and Stoermer. In addition, statistically significant clusters are identified in *S. niagarae* populations from different geographic locations and in members of the same population grown in different environments. These results suggest current criteria for resolving diatom taxa may not be sufficient to discern subtle differences that occur between real species.

Key index words: computational biology; diatom(s); species recognition; statistical significance, *Stephanodiscus niagarae*; taxonomy

Scientists have been aware of the need to consider the statistical significance of clusters for many years (Tsukamura 1967, Mountford 1970, Sneath and Sokal 1973). Lenington and Flake (1975) modeled this problem with multivariate normal distributions.

Gower (1975) discussed goodness of fit criteria, one version of which, between-group distinctness, we will use to illustrate our methods. Milligan (1983) discussed external criteria to evaluate the recovery of cluster structure (where it is known to exist) by various methods, and Milligan (1996) reviewed computationally intense techniques for such evaluations. Degens (1983) construed cluster techniques as estimating an underlying structure and evaluated such estimates with maximum likelihood concepts.

In the last decade, scientists have recognized the appropriateness of methods based explicitly on evolutionary change processes for evaluating the extent to which data may evidence evolutionary change among species and higher taxa, and much effort has been invested in the development and evaluation of these methods. Relatively less attention has been given to problems of quantitative recognition of taxonomic structure near the species level. Here, we describe an approach to determine the statistical significance, under a hypothesis of no structure, of the optimal partition of a study collection of specimens representing organismal diversity near the species level. Like Milligan (1996), we use the computational power of today's ubiquitous microcomputers, but unlike earlier approaches to related problems, we hypothesize no taxonomic structure and determine the significance of apparent structure under this hypothesis. When apparent grouping is strong enough to significantly reject the hypothesis of no grouping, then there may be grounds for recognizing taxa. We define a random process to represent a hypothesis of no natural taxonomic subdivisions, and use computational power to calculate the predicted statistical distribution of a measure of the distinctness of the most distinct clusters.

To illustrate the power of this approach, we apply it to controversies related to recognition of species level taxa in some members of the *Stephanodiscus niagarae* Ehrenb. complex (Theriot 1992). Diatomists, like other biologists, have sought methods for interpreting the significance of descriptions for the morphological diversity between entities near the species level in an impartial manner. Often statements

¹ Received 24 February 1997. Accepted 24 July 1997.

² Author for reprint requests; e-mail mjulius@umich.edu.

of significant or nonsignificant differences in morphologic expression are made by an author with little or no explicit quantitative analysis. Carvalho et al. (1995) examined specimens identified as *Cyclotella caspia* Grunov from extant and fossil populations in North America and Africa. These researchers concluded that morphologic variability observed among these populations was not significant based on such authoritative comparisons of as few as two individuals from some localities. Other researchers have incorporated statistical methods to describe morphologic differences among members of a species. These methods can be straightforward bivariate analysis or more complex multivariate techniques. Kling (1992) used linear regression analysis to examine the effects of silica concentration on valve morphology in *Stephanodiscus hantzschii* Grunov. Others have used multivariate techniques incorporating three or more measurements in the analysis. Teubner (1995) used principal component analysis of eleven light microscope features to compare *Cyclotella* species.

The most comprehensive investigations of valve variability using statistical methods have been performed on the *Stephanodiscus niagarae* complex (Theriot and Stoermer 1984, Theriot 1987). Three species (*S. superiorensis* Theriot and Stoermer [Figs. 1–4], *S. reimerii* Theriot and Stoermer in Theriot [Figs. 5–8], and *S. yellowstonensis* Theriot and Stoermer [Figs. 9–11]) were distinguished from *S. niagarae* through principal component analysis of several morphologic features of the valve face. A phylogenetic study examining the evolutionary relationships of these taxa has been performed (Theriot 1992). The strength of the species classifications was not determined by a statistical significance for the groupings. We use this species complex to illustrate the methods described here and examine the statistical significance of the three species separated from *S. niagarae* (Theriot and Stoermer 1984, Theriot 1992).

We also evaluate the significance of geographically and environmentally isolated groups of specimens identified as *Stephanodiscus* cf. *niagarae*. Edlund (1992) examined natural and cultured populations of *S. cf. niagarae* and concluded that a considerable amount of morphologic variation can be attributed to variations in ambient silica concentrations. We use Edlund's (1992) data to examine groups from different geographic locations and in varying environmental conditions.

METHODS

A hypothesis of no natural taxonomic subdivisions. Suppose t specimens, representing the variation under study, have been measured for p numerical measurements. Let $X(i, m)$ be the value of measurement m for specimen i . The scientist conducting the study may have adjusted the units for the measurements so that each basis for measurement has comparable variability, e.g., rescaled so that ranges or variances are equal, or the measurements may have been chosen or weighted in some other way. These

scaling and weighting questions are very important to descriptive biology, but not essential to the methods we present here. We will suppose that $X(i, m)$ already reflects decisions about scaling and weighting.

A partition P of the t specimens into k classes (P_1, P_2, \dots, P_k) can be evaluated by this between-group distinctness criterion:

$$\sum_{i=1}^{k-1} \sum_{s=i+1}^k \sum_{j \in P_s} \sum_{m=1}^p ((X(j, m) - X(i, m)))^2 = \text{SSBGD}$$

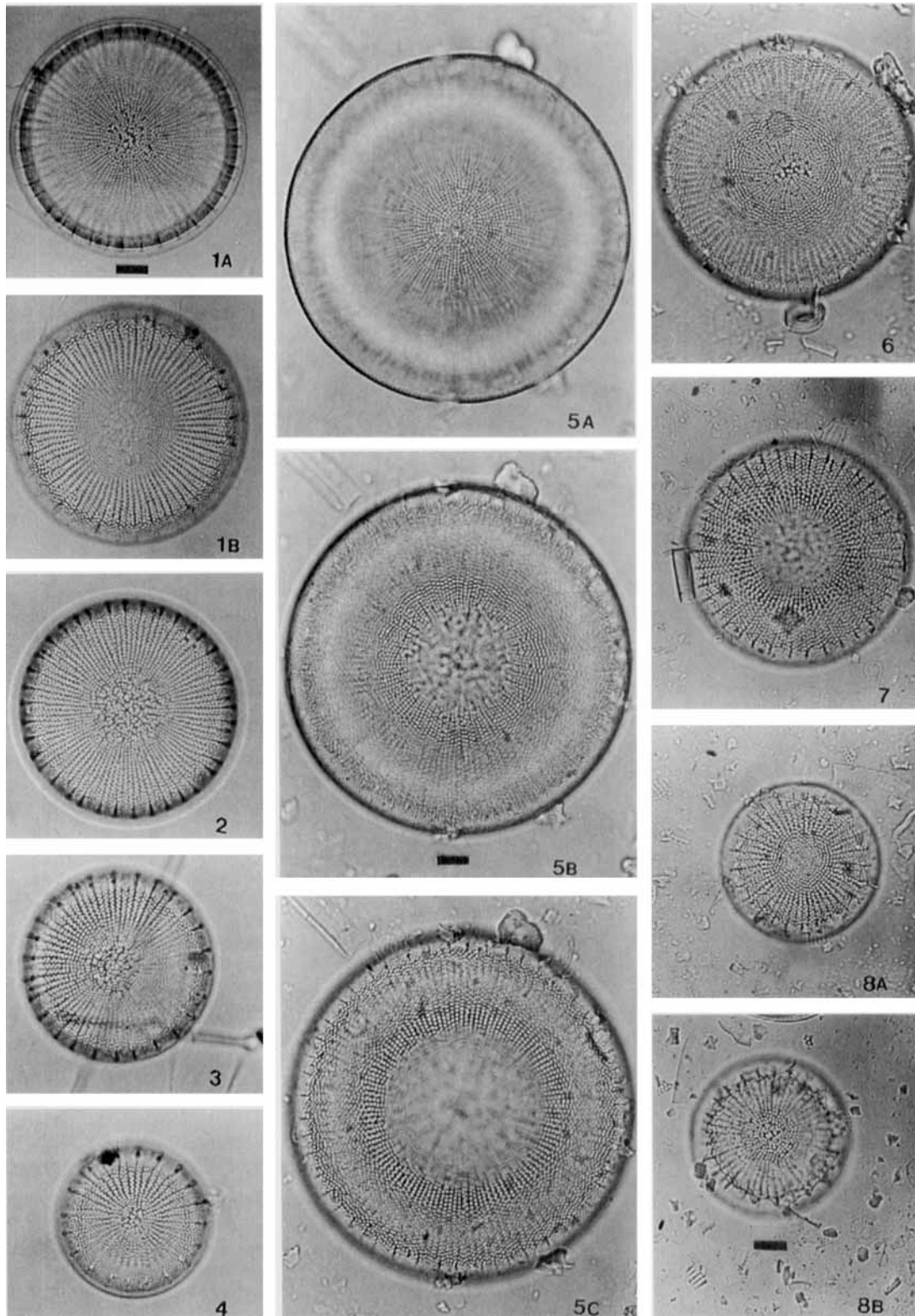
This is the sum of squares of all differences in all measurements between pairs of specimens chosen from different groups, in short, the Sum of the Squares of the Between Group Differences, SSBGD. Any well-defined criterion to evaluate the goodness of a partition could be used, but we use Gower (1975) for the sake of concreteness. Using a computer, all possible partitions are evaluated by this criterion; in this way the best ones are revealed. The question of interest is whether these partitions are substantially better than would be expected under a hypothesis that there was no natural division of these t specimens into k groups.

To answer this question we can 1) define what it might mean to have no natural division of these specimens into groups, 2) hypothesize that such a process had generated the observed data, and 3) predict, in the form of a probability distribution, the goodness of the best partition determined by the revealing procedure described above, when applied to data generated by the hypothesized process. Finally, we compare the goodness of the groups, revealed using the observed data, to the goodness predicted by the hypothesis that there are no natural groups. If the observed goodness value is typical of what we predict from the hypothesis that there are no natural groups, then this would be evidence that, although our group revealing process did dutifully reveal the best groups as it always will, the groups do not mean anything. If the observed goodness value is very much higher than that predicted by the hypothesis, then we can reject the hypothesis to argue that the groups are natural in some sense.

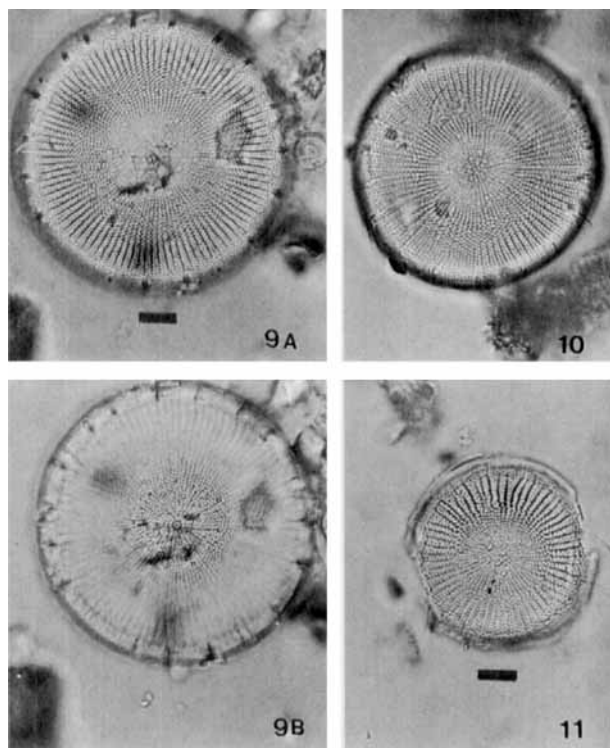
What causes data of this kind to reveal natural groups as the result of analysis by a group-revealing procedure such as that described above? It is the way the values of measurements are associated together in the various individual specimens. Suppose the values actually observed for a measurement could have been associated with any of the specimens, i.e. any assignment of the observed values to the specimens, resulting in all the same values being assigned but not necessarily to the same specimens. Suppose this were true for all measurements, and that the assignment of a value for one measurement to a specimen was completely unrelated to the assignment of a value for any other measurement to this or any other specimen. Suppose all these possible ways were equally likely. One of these possible ways is represented by the observed data, and the other possible ways all have exactly the same specimens and the same values for the measurements as the observed data. This constitutes definition 1 above. We hypothesize in 2 above that "The observed data are a 'random' (equiprobable) sample of this collection of possibilities in which there is not necessarily any natural grouping."

To predict from this hypothesis a probability distribution for the goodness of the best partitions by a group-revealing analysis (3, above), we use a computer. First we "randomly" sample (with replacement) the collection of possible data described above by equiprobably and independently choosing for each measurement a permutation of its observed values over the specimens. Then we subject this "random" dataset to the group-revealing analysis of interest and determine the goodness of the best grouping. Finally we compare the value of this goodness to the value determined from our observed data to see if it is higher. We repeat the above three steps one thousand (or more) times, each time "randomly" choosing a data set from among the possibilities hypothesized as equally likely. The fraction of times our observed goodness value was better than or equal to the "random" goodness value is a realized significance of the partition of specimens based on observed data, under a hypothesis that there is no natural grouping.

When this significance assumes values above 0.10, the interpretation is clear: there is no strong evidence that the best groups



FIGS. 1–8. Light micrographs of *Stephanodiscus superiorensis* and *S. reimerii*. FIGS. 1–4. *Stephanodiscus superiorensis*. Isotype material from E. F. Stoermer collection, slide EFS 1189. FIG. 1. A) High focus; B) low focus. FIGS. 5–8. *Stephanodiscus reimerii*. Isotype material from Edward Theriot collection, slide ECT 9. FIG. 5. A) High focus; B) midfocus; C) low focus. All scale bars = 5 μm (Figs. 1A, 5B, 8B).



FIGS. 9–11. Light micrographs of *Stephanodiscus yellowstonensis*. Isotype material from E. F. Stoermer collection, slide EFS 3614. FIG. 9. A) Low focus; B) high focus. All scale bars = 5 μm (Figs. 9A, 11).

revealed by the data are natural and such groups do not deserve recognition and do not usually require any further explanation. For values between 0.10 and ~ 0.02 the results are ambiguous; perhaps the best thing to do is to seek more measurements or sample more specimens. For values less than 0.02, it is safe to conclude that there is some natural grouping revealed by the data. However, it is not safe to conclude that the grouping is caused by anything in particular. It is important to remember that this style of argument rejects the hypothesis that “nothing is going on,” which paves the way for an interpretation, but does not in and of itself argue for any particular one. For example, highly significant clusters are not necessarily species or other higher taxa. Such recognition must include additional biologic considerations.

Experimental organisms and simulations. To illustrate these concepts, we performed four sets of simulations using measurements from specimens (populations) in the *Stephanodiscus niagarae* complex. In each of the simulations, 1000 random data sets were generated by the methods described. All simulations were performed using PHENCON 1.0 (Estabrook 1996) run on a Zenith Data Systems microcomputer equipped with an Intel 486 microprocessor. For each data set, all possible divisions into two or three groups, depending on the distinctions being tested, were evaluated for goodness. For each simulated data set, the highest goodness value revealed was compared with the goodness value observed for the real data.

Simulation 1 compared measurements from *Stephanodiscus reimerii*, *S. yellowstonensis*, and *S. superioensis* to determine if groups based on valve features observed by Theriot and Stoermer (1984) and Theriot (1992) are nonrandom. We scored 18 individuals from the following microscope slides in Dr. Theriot's and Dr. Stoermer's personal collections ECT 9, EFS 1189, and EFS 3614 for *S. superioensis* (Figs. 1–4), *S. reimerii* (Figs. 5–8), and *S. yellowstonensis* (Figs. 9–11), respectively.

Simulations 2 through 4 were performed using data from specimens identified as *Stephanodiscus cf. niagarae* (Edlund 1992). The specimens were from two geographic locations, one isolated in

culture from Lower Herring Lake, Michigan, and the other from Humboldt Lake, Saskatchewan. Individuals were sampled growing in varying silica concentrations using 20 individuals for each comparison.

Simulation 2 compared two groups from the same geographic location growing under different silica concentrations: one group growing at high silica concentrations ($6.4 \mu\text{g}\cdot\text{L}^{-1}$) and the other group at low silica concentrations ($<0.05 \mu\text{g}\cdot\text{L}^{-1}$). Two sets of comparisons were performed: Lower Herring Lake (high silica) vs. Lower Herring Lake (low silica) and Humboldt Lake (high silica) vs. Humboldt Lake (low silica).

Simulation 3 compared groups from the two different geographic locations growing under the same silica concentrations. Two sets of comparisons were performed: Lower Herring Lake (high silica) vs. Humboldt Lake (high silica) and Lower Herring Lake (low silica) vs. Humboldt Lake (low silica).

Simulation 4 compared taxa from the same geographic location and silica concentration. Four sets of comparisons were performed: Lower Herring Lake (high silica) vs. Lower Herring Lake (high silica), Lower Herring Lake (low silica) vs. Lower Herring Lake (low silica), Humboldt Lake (high silica) vs. Humboldt Lake (high silica), and Humboldt Lake (low silica) vs. Humboldt Lake (low silica).

Microscope observations. All light microscope (LM) observations were made with either a Leitz Ortholux capable of a 1.32 numerical aperture or a Leica DMRX microscope at 1200 \times using full immersion optics providing a numerical aperture of 1.40. Individual valve measurements were taken using the computer imaging system described in Stoermer (1996). Images were captured using a Sony 3 CCD color video camera using NIH Image 1.59 (written by Wayne Rasband at the U.S. National Institutes of Health and available on the Internet by anonymous ftp from zippy.nimh.nih.gov or on floppy disk from NTIS, 5285 Port Royal Road, Springfield, Virginia 22161, part number PB93-504868) run on an Apple Power Macintosh 8100/110 computer. Samples were cleaned with nitric acid or 30% hydrogen peroxide and mounted in Hyrax for LM observations.

Characters and metrics. For simulation 1 the following measurements described in Theriot and Stoermer (1984) were used for each sampled individual: 1) valve diameter, 2) mantle width, 3) areolar density in 10 μm , 4) total number of spines, 5) total number of rimoportulae, 6) total number of central fultoportulae, and 7) total number of fascicles.

For simulations 2 through 4 the following measurements, taken from Edlund (1992), were used for each sampled individual: 1) valve diameter, 2) areolar density in 10 μm , and 3) total number of fascicles.

RESULTS

Simulation 1

Significance of species designations within the *Stephanodiscus niagarae* complex. None of the 1000 simulations produced group classification values at least as high as the value from real data ($P < 0.001$). Therefore, we reject the null hypothesis at the 0.02 level of significance and conclude that statistically significant differences exist among the morphologic expression of *Stephanodiscus reimerii*, *S. yellowstonensis*, and *S. superioensis*.

Simulation 2

Populations from the same locality grown in high and low silica, compared Lower Herring Lake (high silica) vs. Lower Herring Lake (low silica). One hundred thirty-five of the 1000 simulations produced group classification values at least as high as the value from real data ($P < 0.135$). Therefore, we accept the null hypothesis at the 0.02 level of significance and con-

clude that statistically significant differences probably do not exist based on the morphologic features of specimens in Lower Herring Lake *Stephanodiscus* cf. *niagarae* populations grown in high and low silica concentrations.

Humboldt Lake (high silica) vs. Humboldt Lake (low silica). None of the 1000 simulations produced group classification values at least as high as the value from real data ($P < 0.001$). Therefore, we reject the null hypothesis at the 0.02 level of significance and conclude that statistically significant differences exist based on the morphologic features of specimens in Humboldt Lake *Stephanodiscus* cf. *niagarae* populations in high and low silica concentrations.

Simulation 3

Populations from different localities grown in similar silica concentrations, compared Lower Herring Lake (high silica) vs. Humboldt Lake (high silica). Twenty-four of the 1000 simulations produced group classification values at least as high as the value from real data ($P = 0.024$). Therefore, we accept the null hypothesis at the 0.02 level of significance and conclude that statistically significant differences do not exist between the morphologic expression of Lower Herring Lake and Humboldt Lake *Stephanodiscus* cf. *niagarae* populations in high silica concentrations.

Lower Herring Lake (low silica) vs. Humboldt Lake (low silica). None of the 1000 simulations produced group classification values at least as high as the value from real data ($P < 0.001$). Therefore, we reject the null hypothesis at the 0.02 level of significance and conclude that statistically significant differences exist between the morphologic expression of Lower Herring Lake and Humboldt Lake *Stephanodiscus* cf. *niagarae* populations in low silica concentrations.

Simulation 4

Populations from the same locality in similar silica concentrations compared Lower Herring Lake (high silica) vs. Lower Herring Lake (high silica). One hundred seventeen of the 1000 simulations produced group classification values at least as high as the value from real data ($P = 0.117$). Therefore, we accept the null hypothesis at the 0.02 level of significance and conclude that statistically significant differences do not exist between the morphologic expression of Lower Herring Lake *Stephanodiscus* cf. *niagarae* populations in high silica concentrations.

Lower Herring Lake (low silica) vs. Lower Herring Lake (low silica). One hundred sixty-one of the 1000 simulations produced group classification values at least as high as the value from real data ($P = 0.161$). Therefore, we accept the null hypothesis at the 0.02 level of significance and conclude that statistically significant differences do not exist between the morphologic expression of Lower Herring Lake *Stephanodiscus* cf. *niagarae* populations in low silica concentrations.

Humboldt Lake (high silica) vs. Humboldt Lake (high silica). Seven hundred sixty-two of the 1000 simula-

tions produced group classification values at least as high as the value from real data ($P = 0.762$). Therefore, we accept the null hypothesis at the 0.02 level of significance and conclude that statistically significant differences do not exist between the morphologic expression of Humboldt Lake *Stephanodiscus* cf. *niagarae* populations in high silica concentrations.

Humboldt Lake (low silica) vs. Humboldt Lake (low silica). Three hundred forty of the 1000 simulations produced group classification values at least as high as the value from real data ($P = 0.34$). Therefore, we accept the null hypothesis at the 0.02 level of significance and conclude that statistically significant differences do not exist between the morphologic expression of Lower Herring Lake *Stephanodiscus* cf. *niagarae* populations in low silica concentrations.

DISCUSSION

Simulation 1

Significance of species designations within the *Stephanodiscus niagarae* complex. The simulation results found a high statistical significance for the taxonomic clusterings established by Theriot and Stoermer (Theriot and Stoermer 1984, Theriot 1992), suggesting this classification system reflects nonrandom groupings. Abundant evidence also suggests these groupings are controlled by genetic differences rather than environmental variables (Theriot and Stoermer 1986, Theriot 1987). Theriot (1992) reported that cultured species of *S. yellowstonensis* and *S. superiorensis* retain distinctive morphologic expressions in various environments. Little doubt should remain about the validity of these species, and attention should be given to minor morphologic variation when applying specific epithets to members of the *S. niagarae* complex.

Simulations 2 through 4

Analysis of Edlund's (1992) data. The results of these simulations emphasize the conclusions made above. Edlund (1992) tested the effect of various silica concentrations on valve morphology. In Edlund's (1992) work, Humboldt Lake *Stephanodiscus* cf. *niagarae* had a much greater overall response to silica than Lower Herring Lake *Stephanodiscus* cf. *niagarae*. Silica levels greatly affected all three parameters measured for Humboldt Lake *Stephanodiscus* cf. *niagarae* (Pearson's $|R| > 0.3$ for all cases). Edlund's (1992) observations of Lower Herring Lake *S.* cf. *niagarae* found the effect of silica levels on valve diameter (Pearson's $R = -0.172$) and total number of fascicles (Pearson's $R = -0.161$) was greatly reduced, while areolar density in 10 μm was strongly affected by silica concentration (Pearson's $R = -0.413$). Our simulations recognized statistically significant clusters for high and low silica groupings of Humboldt Lake *Stephanodiscus* cf. *niagarae*, but significant clusters were not identified for similar groupings of Lower Herring Lake *Stephanodiscus* cf. *niagarae*. This suggests ecomorphotypes related to silica concentration may exist for the Humboldt

Lake population but not for the Lower Herring Lake population.

Our results comparing two geographic populations grown in low silica concentrations found statistically significant clusters, suggesting speciation may have occurred between the two geographically isolated populations. The results for geographic groupings in high silica concentrations, however, did not allow us to reject the null hypothesis, suggesting a speciation event has not occurred between the two geographically isolated populations. Our realized significance (0.024) for geographic groupings in high silica concentrations was very close to the critical value (0.02) and additional analysis utilizing a larger sample size and more measurements may reveal statistically significant clusters.

CONCLUSION

We support the recognition of *S. superiorensis* (Figs. 1–4), *S. reimerii* (Figs. 5–8), and *S. yellowstonensis* (Figs. 9–11). This decision is not based solely on the results of our statistical simulations. Only a combination of biological considerations and significant clusterings allow strong statements on taxonomic classification. Highly significant clusters are not necessarily species or other higher taxa, but suggest the need for additional biologic investigation.

Our results suggest, without ambiguity, that differences exist between two geographically isolated populations of *Stephanodiscus* cf. *niagarae* and that the population from Humboldt Lake has a greater sensitivity to variations in silica concentration. This has important implications on the current diatom taxonomic system. Researchers must give greater attention to details in diatom valve morphology and, in fact, strive to understand the underlying basis of more morphologic variation. For instance, although our technique found statistically significant differences in Humboldt Lake specimens growing in high and low Silica environments, we have no reason to suppose these specimens represent different species. Instead, ample evidence is available to suggest silica supply is often manifested in a diatom's morphologic variability (Paasche and Østergren 1980, Theriot 1987, Edlund 1992). Variation indicative of environmental change is currently being overlooked in paleoecologic studies (Stoermer et al. 1989), and in many cases we may not be identifying actual species using the current taxonomic system. The use of modern tools, like the technique described here, coupled with detailed biologic investigations will help establish terminal units for phylogenetic analysis. Analysis of these terminal units (species) will allow reevaluation of the current classification system, developing a taxonomy more reflective of the natural hierarchy existing in diatoms.

Contribution 588 from the Center for Great Lakes and Aquatic Sciences was supported in part by NSF Grant DEB 9521882. The comments of Dr. Edward Theriot were greatly appreciated. We

thank The University of Michigan Department of Biology for supporting BIO480, which made this work possible. PHENCON is available free from George Estabrook. It can be downloaded from his website (<http://www-personal.umich.edu/~gfred/>) or obtained on a floppy disk at 2117 Natural Science, The University of Michigan, Ann Arbor, Michigan 48109-1048.

- Carvalho, L. R., Cox, E. J., Fritz, S. C., Juggins, S., Sims, P. A., Gasse, F. & Battarbee, R. W. 1995. Standardizing the taxonomy of saline lake *Cyclotella* spp. *Diatom Res.* 10:229–40.
- Degens, P. O. 1983. Hierarchical cluster methods as maximum likelihood estimators. In Felsenstein, J. [Ed.] *Numerical Taxonomy*. NATO ASI Series, Vol. G1, Springer Verlag, Berlin, pp. 249–53.
- Estabrook, G. F. 1996. PHECON: Phenetics with confidence. Program and documentation distributed on the worldwide web at <http://www-personal.umich.edu/~gfred/>.
- Edlund, M. B. 1992. Silica-related morphological variation in natural and cultured populations of *Stephanodiscus niagarae* (Bacillariophyta). Master's thesis, University of Michigan, Ann Arbor, 102 pp.
- Gower, J. C. 1975. Goodness-of-fit criterion for classification models and other patterned structures. In Estabrook, G. F. [Ed.] *Proceedings of the Eighth International Conference on Numerical Taxonomy*. Freeman, San Francisco, pp. 38–62.
- Kling, H. J. 1992. Valve development in *Stephanodiscus hantzschii* Grunow (Bacillariophyceae) and its implications on species identification. *Diatom Res.* 7:241–57.
- Lenington, R. K. & Flake, R. H. 1975. Statistical evaluation of a family of clustering methods. In Estabrook, G. F. [Ed.] *Proceedings of the Eighth International Conference on Numerical Taxonomy*, Freeman, San Francisco, pp. 1–37.
- Milligan, G. W. 1983. Characteristics of four external criterion measures. In Felsenstein, J. [Ed.] *Numerical Taxonomy*. NATO ASI Series, Vol. G1. Springer Verlag, Berlin, pp. 167–73.
- . 1996. Clustering validation: results and implications for applied analyses. In Arabie, P., Hubert, L. J. & DeSoete, G. [Eds.] *Clustering and Classification*. World Scientific Publishing, River Edge, New Jersey, pp. 345–75.
- Mountford, M. D. 1970. A test for the difference between clusters. In Patil, G. P., Pielou, E. C. & Waters, W. E. [Eds.] *Statistical Ecology*, Vol. 3. Pennsylvania State University Press, University Park, Pennsylvania, pp. 237–57.
- Paasche, E. & Østergren, I. 1980. The annual cycle of plankton diatom growth and silica production in inner Oslofjord. *Limnol. Oceanogr.* 25:481–94.
- Sneath, P. H. A. & Sokal, R. R. 1973. *Numerical Taxonomy*. Freeman, San Francisco, 573 pp.
- Stoermer, E. F. 1996. A simple, but useful, application of image analysis. *J. Paleolimnol.* 15:111–13.
- Stoermer, E. F., Emmert, G. E. & Schelske, C. L. 1989. Morphological variation of *Stephanodiscus niagarae* Ehrenb. (Bacillariophyta) in a Lake Ontario sediment core. *J. Paleolimnol.* 2: 227–36.
- Teubner, K. 1995. A light microscopical investigation and multivariate statistical analyses of heterovalvar cells of *Cyclotella* species (Bacillariophyceae) from lakes of the Berlin-Brandenburg region. *Diatom Res.* 10:191–205.
- Theriot, E. C. 1987. Principal component analysis and taxonomic interpretation of environmentally related variation in silicification in *Stephanodiscus* (Bacillariophyceae). *Brit. Phycol. J.* 22: 359–73.
- . 1992. Clusters, species concepts, and morphological evolution of diatoms. *Syst. Biol.* 41:141–57.
- Theriot, E. C. & Stoermer, E. F. 1984. Principal component analysis of *Stephanodiscus*: Observations on two new species from the *Stephanodiscus niagarae* complex. *Bacillaria* 7:37–58.
- . 1986. Morphological and ecological evidence for two varieties of the diatom *Stephanodiscus niagarae*. In Ricard, M. [Ed.] *Proceedings of the Eighth International Diatom Symposium*. Koeltz, Koenigstein, Germany, pp. 385–94.
- Tsukamura, M. 1967. A statistical approach to the definition of bacterial species. *Jap. J. Microbiol.* 11:213–20.