Multivariate complexity analysis of 3D surface form and function of centric diatoms at the Eocene–Oligocene transition

Janice L. Pappas

Museum of Paleontology, University of Michigan, 1109 Geddes Avenue, Ann Arbor, MI 48109-1079, USA

ARTICLE INFO

ABSTRACT

Complexity is important in the course of evolution, but consensus of what complexity analysis entails is elusive. In this study, multivariate complexity is measured and analyzed in terms of magnitude of change and the trends behind those changes. Multivariate morphological and functional complexities were analyzed using 3D surface models of centric diatom genera from the Eocene to Oligocene, a transitional time when temperatures cooled, sea levels rose, and glaciation increased. Diatoms were chosen for study because of their importance in biostratigraphy, biogeochemical cycling, productivity, food web dynamics, and sensitivity to environmental conditions. Probabilistic analysis using a Markov chain indicated an increase in total complexity across the Eocene–Oligocene transition (EOT). Causal inference via structural equation modeling indicated weakly driven functional and morphological complexity trends over the EOT. Morphological and functional complexity trends differed with respect to predation resistance as responses to ecological complexity as environmental and climate change occurred across the EOT. Macroevolutionary patterns of morphological and functional complexity with respect to ecological complexity did not necessarily coincide over time.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

That organisms become more complex over time has been espoused as the dominant viewpoint in biology (e.g., Lamarck, 1809; Darwin, 1859; Miconi, 2008; McShea and Brandon, 2010; Auerbach and Bongard, 2014) because evolution has occurred from unicellular to multicellular entities. The nominal basis for this claim is that by counting cells, complexity can be determined (e.g., Wicken, 1979; McShea, 1991; McShea and Brandon, 2010). Aside from this obvious pronouncement, no consensus exists on the definition of complexity (especially over multiple disciplines), analyses have entailed discordant approaches, and equivocal results have been produced (McNamara, 2006).

Complexity of organisms has been based on various attributes such as size (e.g., McShea and Brandon, 2010; McShea and Hordijk, 2013), body plan (Valentine, 2000), heterogeneity or degree of differentiation (Darwin, 1859), and pattern arrangement of number of parts (e.g., McShea, 1991, 1992; McShea and Venit, 2001). Complexity is usually determined at a one- (1D) or two-dimensional scale (2D) or level (e.g., McShea, 1991; McShea and Brandon, 2010) where, e.g., bits of information and their sequential structure are assessed (e.g., McShea, 1992; Adami, 2002). Whenever questions arise about the definition of complexity or parts, qualifiers (e.g., McShea and Venit, 2001), nuanced descriptions (e.g., McShea, 1992; McShea and Brandon, 2010), or negation terms (e.g., Wicken, 1979; McShea, 1991) are invoked in order to legitimize the methods used in complexity analysis. Definitions of parts and whole entities with respect to complexity are described, which are dependent on boundary determinations, which in turn are context and scale dependent and must be defined a priori for complexity analysis to begin. Applying such constraints across the board presents difficulties in generalizing complexity measurement and analysis and arriving at a consensus on what is meant by complexity.

At the other end of the spectrum, complex systems are analyzed by treating different organisms as "particles" and applying a particular algorithm such as an artificial neural network via agent-based modeling (e.g., Yaeger et al., 2008) to determine large scale complexity trends. While results have been intriguing (e.g., Auerbach and Bongard, 2014), computational complexity sometimes substitutes for biological complexity (e.g., Adami, 2002), and meaningful attributes of actual organisms and their significance are not necessarily recovered by such methods. In spite of this, many 1D and 2D systems-based assessments express complexity change over time as probabilistic (e.g., Yaeger et al., 2008; McShea and Brandon, 2010). By using a probability-based framework, the likelihood of a future event of complexity change and its trend over time can be inferred as a result rather than merely counting and cataloging the frequency of past events of complexity change. To start, measuring complexity as a deterministic quantity aids in circumventing the need for elaborate, specific descriptions and conceptualizations in order to constrain analyses to fit the purposes of isolated studies.

Complexity as a deterministic quantity is useful in studying a single event or determining a long term macroevolutionary trend where

E-mail address: jlpappas@umich.edu.

probabilities of outcomes enable assessment at the *n*-systems level. For biological complexity, an *n*-complex system relating functionality to form is important in the evolution and coevolution of organisms. Different types of complexity as deterministic quantities can be aggregated as *n*-dimensional measures. Complexity as a multidimensional quantity enables its assessment as *n*-macroevolutionary trends, with potentially different magnitudes and/or drivers for each trend. Understanding the differences may enable discernment of long term patterns of complexity changes in evolution.

1.1. Purposes of the study

Multicellular taxa are typically chosen for study in complexity analysis (e.g., McShea, 1992), since it is taken for granted that only multicellular organisms qualify. However, complexity assessment is a comparative exercise, and among the unicells, diatoms possess complex morphology (Harwood et al., 2007), representing a monophyletic group at the class or phylum level (Williams and Kociolek, 2007). Diatoms account for at least 40% of total productivity, 50% of organic carbon sequestration in the sediment (Falkowski et al., 2004), and are important in the global carbon (Hopkinson et al., 2011) and silica (Knoll, 2003) cycles. Diatoms are biogeochemically and stratigraphically important in characterizing paleoenvironments and paleoclimates as a result of their ubiquitous distribution and diversity (e.g., Harwood et al., 2007; Jordan and Stickley, 2010). Upwelling affected silica cycling and productivity in surface waters throughout the Cenozoic, and since the Oligocene, diatoms are the dominant contributors to the global silica cycle as evidenced by their high sedimentation rate (Harwood et al., 2007; Witkowski et al., 2014). Limiting conditions for silica availability (Conley, 2002; Harwood et al., 2007) and other nutrients in the oceans control not only diatom productivity, evolution and diversification (Jordan and Stickley, 2010), but also diatom morphological features and the degree to which they are silicified (Barron and Baldauf, 1989).

The Eocene–Oligocene transition (EOT) is an important period of changing composition of diatom taxa during the Cenozoic (Jordan and Stickley, 2010) because of changes from increasing levels of atmospheric CO_2 with greenhouse temperatures and high sea levels during the Eocene to low levels of atmospheric CO_2, with icehouse temperatures and low sea levels with ice sheet expansion in the Oligocene (Zachos et al., 2008; Jordan and Stickley, 2010). Centric diatoms contributed to a major shift in global carbon cycling and increasing oxygenation of surface waters during the change from greenhouse to icehouse conditions (Armbrust, 2009). Since the Cretaceous, centric diatoms have endured, survived and proliferated, exhibiting a wide variety of morphologies (Harwood et al., 2007). Although centric diatoms were prevalent during the Cenozoic, morphological and functional complexity are neither easily discernible upon cursory inspection of the taxa nor is the relation between complexity and environmental and climate conditions evident during this time.

To measure complexity in centric diatoms from the Eocene and Oligocene, three-dimensional (3D) surface models will be devised using parametric 3D equations (Pappas, 2005a, 2005b, 2008; Pappas and Miller, 2013), and their numerical solution will be used in measurement of morphological complexity. Functional complexity will be measured as a quantity with respect to surface area extracted from measurement of morphological complexity. Morphological and functional complexity will be aggregated as multivariate quantities, and from this, rank ordering of taxa from the Eocene to the Oligocene will serve as a vector of complexity change over time. Magnitude of short and long term complexity change will be probabilistically analyzed and measured with regard to steady-state conditions. Causal inference will be used to determine mode of complexity as an evolutionary trend using multivariate modeling and probabilistic analysis. The context of morphological and functional complexity will be examined in terms of ecological changes that occurred during the EOT.

2. Methods

Measuring complexity is not the same as measuring diversity. Complexity is not about counting the number of all the different taxa that existed in the Eocene and Oligocene. To attempt to consider every last possible taxon, especially rare or accidental occurrences, is neither attainable nor necessary. Rather, complexity is about the number of different kinds of surface morphological patterns common among taxa. Surface morphological patterns are redundant over a number of different taxa (e.g., Barber and Haworth, 1981), and for a circumscribed time period from the Eocene to the Oligocene, a finite number of patterns occur. Such patterns fall into succinct categories based on surface geometry.

Geometrically, whole diatom frustules are capped cylinders. On their exterior valve faces, layers of pores or foramina are arranged in various patterns (Round et al., 1990). Areolae or puncta can be spaced regularly or irregularly, change in size or shape, be radiating from one spot, arranged concentrically from the center, or lined up in a parallel fashion (Barber and Haworth, 1981). Rows of areolae or puncta may form a discontinuous, random, sparse, or dense pattern (Barber and Haworth, 1981). In spite of the vast array of possible combinations of features on their frustule surfaces, redundant centric diatom structural geometric features are commonly recognized over many taxa.

These surface morphological patterns are discernable from taxa in the Eocene and Oligocene, where diatoms are indicative of the prevailing environmental and climate conditions and are accumulated in the sediments via rapid burial or from lack of exposure to high alkalinity and high temperature pore waters (e.g., Barron and Baldauf, 1989, 1995). For selection purposes, taxa may be grouped by traditional morphological classification categories since a monophyletic tree of diatom taxa has yet to be realized (e.g., Williams and Kociolek, 2007; Theriot et al., 2010, 2011). Classification groups specified in Fenner (1985) and Round et al. (1990) are used as general non-phylogenetic taxonomic bins although these descriptors have been used in phylogenetic assessments (e.g., Williams and Kociolek, 2007; Theriot et al., 2010, 2011).

Taxon groups in the Coscinodiscophyceae (e.g., Round et al., 1990; Williams and Kociolek, 2007) characterize many of the centric taxa from the Eocene and Oligocene consisting of the Coscinodiscales, Stictodiscales, Arachnoidiscales, and Asterolamprales (e.g., Fenner, 1985; Round et al., 1990). Of these taxon orders, particular gross geometries on the valve face can be described in terms of different sizes and degrees of pores and distinct arrangements of surface features in a radiating pattern (e.g., Barber and Haworth, 1981; Round et al., 1990). From actual centric taxa obtained from Eocene and Oligocene sediments, published accounts from the Deep Sea Drilling Project (DSDP), Ocean Drilling Program (ODP), Integrated Ocean Drilling Program (IODP), and the United States Geological Survey (USGS) are used to determine the pool of taxa from which to select exemplars of surface geometry having the specific arrangement of pores or radiating features mentioned above. As a result, diatom biostratigraphic markers qualify as the best means to select exemplars of surface morphological patterns for complexity analysis.

The basis of diatom biostratigraphy rests on evolution and extinction of taxa documented by their first and last appearances (Scherer et al., 2007). Multiple taxa representing a given time period make the stratigraphic picture more robust and stable, and transitions between diatom assemblages are an indication of evolutionary change (Scherer et al., 2007). Taxa used in stratigraphic zonations have widespread distributions and are usually larger and more well-preserved than their smaller, rarer counterparts. Diatom assemblages incurred major turnover events at times of rapid high latitude cooling and changes in oceanic circulation of surface waters and are an indicator of this chronostratigraphic evolutionary record (Barron and Baldauf, 1989; Scherer et al., 2007). Zonations based on first and last abundant appearances reiterate evidence of diatom evolutionary changes (Scherer et al., 2007). Diatom evolution is exemplified by unique taxa found in biostratigraphic zones.

The particular centric diatom genera that were chosen from which to make 3D surface models have been restricted to those in the Coscinodiscophyceae with an approximate circular valve and cross-sectional shape and cylindrical 3D shape. This will enable the elimination of overall shape as a confounding factor and make modeling of the taxa for comparative purposes simpler. Taxon selection representing surface morphology was based on specific criteria: (1) the importance of the taxon in diatom biostratigraphy; (2) relative abundance for a given time period; and/or (3) whether a species from a genus was a marker at first or last occurrence. These taxa may have originated at an earlier time or occur in both the Eocene and Oligocene, but many published reports were used to obtain the best consensus of information from which to choose centric diatom genera. Diatom zonation is an ongoing refinement process governed by first and last taxon appearances, reconciling different zonation schemes, incomplete stratigraphic sequences, and matching diatom to calcareous nanofossil zones (e.g., Barron and Baldauf, 1989) or magnetostratigraphy (e.g., Barron and Baldauf, 1995). Opal deposition at different latitudes, regions and oceans, and non-continuous cores of diatom-containing sediments are considerations for matching diatom zones, and dissolution, bioturbation, or diagenesis introducing biases in discerning the fossil diatom record (e.g., Harwood et al., 2007) further complicate the matching of different zonation schemes. However, there is much agreement in diatom zone correlations so that particular centric diatom genera can be construed to be representative or important in either the Eocene or Oligocene for the purposes of this study. For the entire formal probability selection process, see the Appendix A.

The Eocene–Oligocene centric diatom genera are assembled from published accounts in Hajós (1976), Gombos (1983a, 1983b, 1987), Gombos and Ciesielski (1983), Fenner (1984a, 1984b, 1985), Barron (1985), Baldauf and Barron (1987), Saito et al. (1988), Baldauf and Monjanel (1989), Fenner and Mikkelsen (1990), Fourtanier (1991), Harwood and Maruyama (1992), Barron and Mahood (1993), Barron and Baldauf (1995), Arney et al. (2003), Barron et al. (2004), Bohaty et al. (2011), Oreshkina (2012), and Gladenkov (2012).

Representative genera chosen as the basis for 3D surfaces models are: *Actinoptychus* C. G. Ehrenberg, typically found in the Early to Middle Eocene (Fenner, 1985); *Arachnoidiscus* Deane ex Pritchard, which is considered to be robust to preservation issues (Fenner and Mikkelsen, 1990) and a representative of the Early Eocene (Fourtanier, 1991); *Asterolampra* C. G. Ehrenberg, where *Asterolampra marylandica* Ehrenberg is representative of the Middle to earliest Late Eocene (Fenner, 1984a, 1984b, 1985; Fenner and Mikkelsen, 1990; Barron and Baldauf, 1995) and the genus is prevalent (Fenner, 1985); *Brightwellia* J. Ralfs in A. Pritchard, where *Brightwellia hyperborea* Grunow is representative of the Middle Eocene (Gombos, 1983b; Fenner, 1985; Barron and Baldauf, 1995; Oreshkina, 2012); *Craspedodiscus* C. G. Ehrenberg, where *Craspedodiscus oblongus* (Greville) Grunow in A. Schmidt is representative of the Early Eocene (Fenner, 1985; Fourtanier, 1991; Barron and Baldauf, 1995).

Genera chosen for study representing the Oligocene are: *Cestodiscus* Greville, where *Cestodiscus reticulatus* Fenner is important in the Early Oligocene (Fenner, 1985; Barron and Baldauf, 1995; Barron et al., 2004), with 50 to 80% of diatom assemblages being composed of *Cestodiscus* species; *Coscinodiscus* C. G. Ehrenberg, where *Coscinodiscus excavatus* Greville is important in the Early Oligocene (Fenner, 1985; Barron and Baldauf, 1995; Barron et al., 2004) and many other species are present throughout the Oligocene (Barron and Mahood, 1993); *Rocella* Hanna, where *Rocella vigilans* Fenner (Oligocene) and *Rocella gelida* (Mann) Bukry (Oligocene into the Miocene) (Gombos and Ciesielski, 1983; Fenner, 1984a, 1984b, 1985; Baldauf and Barron, 1987; Baldauf and Monjanel, 1989; Barron and Baldauf, 1995) are important. Two additional genera are chosen to be used as Oligocene representatives based on relative abundances even though these taxa may have originated in the Upper Cretaceous (Harwood and Nikolaev, 1995). The genera are: *Stellarima* Hasle and Sims, the species of which

occur frequently (Baldauf and Barron, 1987; Saito et al., 1988), with the genus containing long-ranging taxa (Scherer et al., 2000; Witkowski and Harwood, 2011); *Stictodiscus* R. K. Greville, where species are common, *Stictodiscus kittonianus* Greville is a key marker species (Baldauf and Barron, 1987; Barron and Mahood, 1993; Arney et al., 2003), with the genus containing long-ranging taxa (Scherer et al., 2000). *Actinoptychus*, *Arachnoidiscus*, and *Stictodiscus* are benthic, *Actinoptychus* also has some planktonic species, and all other taxa are planktonic (Baldauf and Barron, 1987). All taxon names were checked using the following: G Dallas Hanna Database (Catalog of Diatom Names, California Academy of Sciences, On-line Version updated 19 Sep. 2011. Compiled by Elisabeth Fourtanier and J. Patrick Kociolek. Available online at http://research.calacademy.org/research/diatoms/names/index.asp); ITIS (http://www.itis.gov); WoRMS (http://www.marinespecies.org).

Diatom specimens are not well preserved over long periods of geologic time (Barron and Baldauf, 1995; Harwood et al., 2007), because opaline silica will dissolve in alkaline waters and warm temperatures where the longer the diatom is exposed to seawater, the more likely it is to dissolve. Opal deposition is subject to diagenesis (Barron and Baldauf, 1995) and will accumulate in upwelling regions where there has been high productivity (Fenner, 1985). Using a 3D surface modeling technique enables control of the quantity (large numbers of specimens can be made), quality (no taphonomic alterations), and fidelity (models capturing essential geometric characteristic features of morphology) of model specimens. Time and expense of acquiring the appropriate specimens for study is also minimized. Published micrographs on diatom taxa are readily available via the Internet on which to base 3D surface morphology models. Diatom biomechanical data is scant, and using 3D surface models to extrapolate values in concert with empirical data enables functional complexity analysis.

The 3D surface morphologies for each diatom genus are constructed as hybrid surface patterns based on light and/or scanning electron micrographs of taxa as pictured in the following: Hajós (1976) (*Arachnoidiscus Coscinodiscus*), Gombos (1983a) (*Asterolampra, Brightwellia, Stellarima*), Gombos and Ciesielski (1983) (*Asterolampra, Brightwellia, Coscinodiscus, Rocella, Stictodiscus*), Gombos (1987) (*Brightwellia, Craspedodiscus*), Fenner (1984a, 1984b) (*Cestodisdus, Coscinodiscus*), Fenner (1985) (*Asterolampra, Brightwellia, Cestodiscus*), Barron (1985) (*Rocella*), Baldauf and Barron (1987) (*Actinoptychus, Asterolampra, Cestodiscus, Coscinodiscus, Stellarima, Stictodiscus*), Baldauf and Monjanel (1989) (*Actinoptychus, Cestodiscus, Coscinodiscus, Craspedodiscus*), Round et al. (1990) (*Actinoptychus, Arachnoidiscus, Asterolampra, Brightwellia, Coscinodiscus, Craspedodiscus, Rocella, Stellarima, Stictodiscus*), Fourtanier (1991) (*Craspedodiscus, Stellarima, Stictodiscus*), Harwood and Maruyama (1992) (*Asterolampra, Cestodiscus, Coscinodiscus, Rocella*), Barron et al. (2004) (*Cestodiscus, Coscinodiscus*), Bohaty et al. (2011) (*Actinoptychus, Stictodiscus*), Oreshkina (2012) (*Brightwellia*), Gladenkov (2012) (*Arachnoidiscus*).

2.1. Morphological complexity

Measurement of morphological complexity is based on 3D surface models (Pappas, 2005a, 2005b, 2008; Pappas and Miller, 2013) of centric diatom taxa listed above. A 3D surface is a proxy for the phenotype of an organism (Pappas and Miller, 2013). When viewing a diatom using a compound light or scanning electron microscope, focusing up and down allows one to see the 3D quality of the surface. That is, the heights of processes reflect a lighter tint or shade than the depths of pores which are darker. The shades and tints are qualities of the 3D surface texture. Changes from one morphological surface structure to another are changes in geometry of peaks, valleys and saddles over the entire surface. On the 3D surface, a single point can be selected. A line drawn through this point, touching only that point, is a slope of the surface. Each point on the surface has a slope, and the slopes are calculated from geometrically-created models of 3D diatom surfaces via the

Pythagorean Theorem. These slopes are a numerical representation of the morphological features of the 3D diatom surface.

To create the models, parametric 3D equations make use of 3D coordinates (x, y, z) via parameters (u, v) in a Euclidean space (e.g., Pappas and Miller, 2013). Such equations enable the depiction of a 3D form where any point on the surface is characterized by the parameters in the context of the coordinates. In this way, one can move from point to point on the 3D surface to see how the slopes change. The slopes on the 3D surface are calculated via first partial derivatives of the parametric 3D equations and solved on an interval defining the boundary values of those derivatives. The result is a Jacobian matrix (i.e., Jacobian) of values of all the slopes on the 3D surface. A summary value of the matrix, the Jacobian determinant, is calculated to represent morphological complexity of the 3D surface. The higher the number calculated, the higher the number of slopes on the surface, and the finer the texture of the surface morphology, resulting in a more morphologically complex 3D surface. For more details on methods of calculation, see the Appendix A.

2.2. Functional complexity

Diatoms have amorphous opaline biogenic silica frustules as a result of synthesis and polymerization of silaffins, long-chain polyamines, and other proteins (e.g., frustulins) (Kröger and Poulsen, 2008). Diatom frustules exhibit many patterns and ornamentation (Round et al., 1990) and consist of nanoscale silica spheres (Hamm et al., 2003; Hamm, 2007) arranged in closely packed layers of "honeycomb" sheets which give the diatom frustule a lightweight appearance (Hamm, 2005). Different shaped holes in each layer overlapping each other and providing a sort of scaffolding configuration strengthen the frustule where the layers act together in concert to resist breakage (Hamm, 2005). Pores on the frustule are lined with nanoscale amounts of monosaccharides such as glucuromannans (Tesson and Hildebrand, 2013).

The whole diatom is encased in a polysaccharide mucilage tube that provides a shield in terms of predation, turgor pressure, cell division and development, and photosynthesis (Round et al., 1990; Hamm et al., 2003). Collisions with sand or other sedimentary particles, ice, and debris may produce abrasions or deformation of the diatom frustule (Hamm, 2005). Diatoms are able to withstand osmolality changes as a result of changes in salinity (Round et al., 1990). Diatoms are a rich source of food for organisms of all sizes (Round et al., 1990), including microcrustaceans and those worms that use mandibles or teeth to crack planktonic diatoms for ingestion of their contents (e.g., Hamm, 2005). Gastropods scrape diatoms from substrate, using a filing motion with their radulae to induce stress to the diatom frustule (Hamm, 2007).

The relationship between form and function is important in the evolution of organisms (Wainwright, 2007). Many forms of organisms have evolved for relatively few identifiable and redundant functions (Adami et al., 2000; Adami, 2002). Measuring functional complexity involves the deformation of a 3D surface as a continuum in form and position when changing from an original to deformed state. Deformation occurs in which the principal axes of strain are assumed to remain parallel and the density and stiffness of the materials remain unchanged (Heinbockel, 2001).

Diatom frustules can undergo small deformations, exhibiting elasticity as a result of the impact of forces (Hamm et al., 2003; Hamm, 2005). From empirical tests, the more ornamented the surface, the more resistant it is to breakage (Hamm et al., 2003; Hamm, 2005, 2007). At the micrometer scale, diatom frustules undergo creep as the application of a force to the frustule travels around, not through the nanospheres of silica (Hamm et al., 2003). The behavior of layers of amorphous silica alternating and infused with glucuromannans occurs approximately as a linear elastic system where internal forces are minimal.

The degree to which stress and strain affect the diatom frustule surface depends on the degree of ornamentation present which is reflected in the amount of surface area. Size of the diatom frustule affects surface
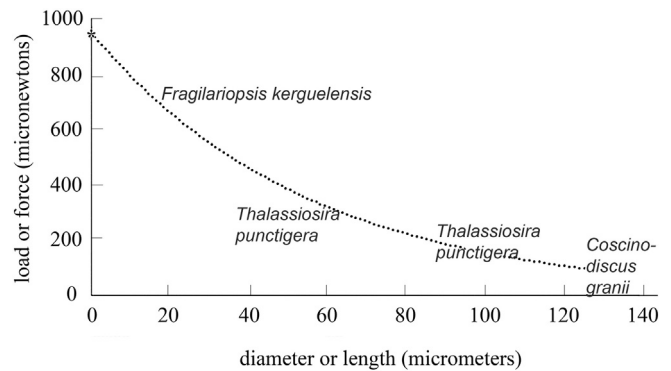

Fig. 1. Load vs. diameter after Hamm et al.'s (2003) data. *Coscinodiscus granii* is used as a reference point from with respect to the best-fit curve. Taxa plotted from left to right are: *Fragilariopsis kerguelensis*, *Thalassiosira punctigera* (50 μm), *T. punctigera* (100 μm), and *Coscinodiscus granii* where $y = 950.83e^{-0.018x}$ and $R^2 = 0.9345$.

area as well. Valve face surface area can be obtained from measurement of the radius (or length of non-circular diatom shapes), and surface area can be used in calculating stress and strain.

Published data from Hamm et al. (2003) are used for force values and their relationship to particular taxa and their maximum diameter or length. A plot of this relationship was devised, and a negative exponential function is used as a least squares best-fit curve of $y = 950.83e^{-0.018x}$ with $R^2 = 0.9345$ where *Coscinodiscus granii* Gough is used as the reference taxon at which the function must converge (Fig. 1). From Fig. 1, diatoms with heavier ornamented, rougher textured surfaces require more force to break than smoother surfaces with fine features. That is, texture is 3D surface quality in terms of roughness, graininess, smoothness, or any other descriptor about the peaks, valleys and saddles that are geometrically characteristic of the surface of the diatom. The surface area of a largest possible patch on the diatom valve face is calculated via the results obtained for morphological complexity. Stress on the surface area is calculated using Hooke's Law and is used as a measurement of functional complexity of the 3D surface. For more details on methods of calculation, see the Appendix A.

2.3. Multivariate complexity analysis

Matrices of morphological and functional complexity values, are aggregated using external unfolding analysis (Coombs, 1964; Heiser, 1987) based on multidimensional scaling (MDS) (Heiser, 1987) to create complexity gradients. The multivariate technique folds different types of data onto their ideal points so that the map that results depicts approximate distances among taxon complexities. Trade-offs among taxa and their complexity attributes are calculated as transformed proximities to produce a low dimensional complexity spatial map as complexity gradients of taxa over time. Eocene and Oligocene taxa can be combined as average proximities to assess morphological, functional and total complexity. The resultant gradients will rank the morphological, functional and total complexities as the change in complexity from the Eocene to Oligocene as a vector of taxa over time. For more details, see the Appendix A.

2.4. Magnitude and rate of complexity change

Using total complexity results as initial conditions, complexity change can be simulated to occur over a long period of time culminating in a steady-state system via a finite discrete Markov chain. At equilibrium the expectation is that probability values will attain the same value across all taxa. If differences between initial and final states are detected, then changes of increasing or decreasing probability values are an indication of increasing or decreasing complexity. The hypotheses to be

tested are: H_0: Complexity does not change over time; H_{A1}: Complexity increases over time; H_{A2}: Complexity decreases over time.

Taxon complexity shifts occur as a finite number of probable transitions. At each step, change in state (= taxon) is represented by the probability that the new state (= another taxon) occurs only with reference to the just current state. The history of previous changes are not a factor in a Markov process, and the expected outcome is a mean number of steps to achieve each state where ultimately over a long period of time, equilibrium (stationary probability distribution) is reached. In the long run, magnitude of complexity change at equilibrium is measured as a distance metric between initial and final states of the Markov chain. In the short run, mean first passage time as the magnitude of complexity change is the time it takes to get from the initial complexity state to each subsequent complexity state, culminating in arrival at the equilibrium complexity state. The length of time is a measure of the short term behavior between complexity states. The more transition states needed to get from one state to another, the higher the mean first passage time. The rate of complexity change can be determined by a plot of the taxon rank-ordered total complexity gradient versus total complexity transition states for each taxon. A least-squares best-fit curve is determined, and the slope of the curve is the rate of total complexity change. For more details, see the Appendix A.

2.5. Complexity change over time: driven or random trend?

To determine the mode of complexity change over time, structural equation modeling (SEM) (Pearl, 2000) is used. Originally used by Wright (1921), SEM involves a group of inferentially related variables in a graphical depiction of those relationships with respect to the causal assumptions used to devise the model. Causal analysis is used to model changing events and the actual and potential conditions under which such events might occur (Pearl, 2009). In causal inference, confounding bias as untested assumptions, unobserved factors, or non-experimental data can be taken into account (Pearl, 2000, 2009). SEM can include independent variables becoming dependent variables as well as the reverse (Pearl, 2000) and be both a cause and an effect in the same model (e.g., Bowen and Guo, 2012; Pearl, 2000).

Causal inference can be used to determine whether a driven or random evolutionary trend is exhibited by centric diatom genera at the EOT. Each type of trend can be characterized via probability values. For a driven trend, centric diatom taxon complexity follows a time-directed trajectory so that those taxa occurring in the Oligocene have a complexity probability of 1, and those taxa occurring in the Eocene have a complexity probability of 0. To represent a random trend, Bernoulli sampling at the 0.1, 0.25, 0.4, 0.5, 0.6, 0.75, and 0.95 levels is used to randomly assign a probability of 0 or 1 to a given taxon. The proportion of 0 s to 1 s for the entire taxon data set is dictated by the level of Bernoulli sampling.

For input data, total complexity measurements are used in SEM. The Jacobian determinant is used as morphological complexity, and the Cauchy traction data extracted from Fig. 1 are used as non-orthonormalized functional complexity measures for taxa. Each data set is square root transformed to approximate a multivariate normal distribution, and this rescaling insures that no category of data exerts undue influence in SEM. Time is represented as categorical variables so that Eocene is represented by 0 since it always precedes the Oligocene which is represented by 1. To test for a driven or random trend as potential outcomes, probabilities represent counterfactual conditionals (Pearl, 2000) of complexity change over time.

Counterfactuals enable the consideration of potential outcomes which can be characterized by "What if…" statements (Pearl, 2000) alongside actual outcomes (Neyman, 1923; Rubin, 1974). "Driven" or "Random" are counterfactual to each other in SEM. Counterfactuals are conditional statements that can be focused into hypotheses to be tested with respect to evolutionary trends: H_0: Complexity change over time is

a random evolutionary trend; H_A: Complexity change over time is a driven evolutionary trend.

In SEM, the different types of variables are combined into a single analysis and depicted as a path diagram representing the structural part of the model. Such graphs can be models of predictions and interventions (Pearl, 2000). Predictions involve cause and direct effect. Interventions produce indirect effects, including insertion or deletion of observations, action or observation exchange, insertion or deletion of actions (Pearl, 2000), and act as confounding variables or paths (e.g., Pearl, 2000, 2009). Connections representing predictions and interventions are represented by unidirectional arrows as causes, and bi-directed arrows representing associations (Pearl, 2000). Causal diagrams structured such that the effects are identifiable will have a minimal number of path connections that are stable (Pearl, 2000). An acyclic directed graph is a causal diagram that has mutually independent error variances and is a well-structured model having the Markovian property (Pearl, 2000).

Causal connections can be made between endogenous and exogenous variables which may be observed or unobserved. Endogenous variables are potential or realized effects and are influenced by other variables. Exogenous variables not influenced by other variables such as error variances that serve as background influences on the outcomes of endogenous variables (Pearl, 2000). Observed variables are directly measured and are also indicator variables. Unobserved variables and their relationship to observed variables constitutes the measurement part of SEM (e.g., Bowen and Guo, 2012).

In graphical SEM, vertices are drawn as rectangles representing endogenous observed variables and ovals representing endogenous unobserved variables. Exogenous unobserved variables represented by small circles are attached at the base of a unidirectional arrow. Means, variances, regression weights, and intercepts can be assigned values as constraints in SEM or left to vary freely, depending on the minimality and stability levels achieved in attempting to fit a structural model using maximum likelihood or a least-squares approach (Bowen and Guo, 2012). Constraints are used to produce a causal path diagram that enables the estimation of interventions and/or counterfactuals for predictions (Pearl, 2000). Interventions and counterfactuals can be represented either as quantities via variables or actions as arrows (Pearl, 2000).

The counterfactual conditional, "Random," representing random outcome as an evolutionary trend is represented by the statement, "Evolutionary trend would be random had complexity changes from the Eocene to the Oligocene occurred by chance." The counterfactual "Driven" representing driven outcome as an evolutionary trend is represented by the statement, "Evolutionary trend would be driven had complexity changes from the Eocene to the Oligocene occurred in one direction." In each model, the probability values for "Driven" indicate time-dependence where Eocene = 0 and Oligocene = 1 and for "Random" indicate Bernoulli sampling, viz. 0 or 1 at levels from 0.1 to 0.95 that are assigned to each taxon. To decide if the null hypothesis should be rejected, correlation coefficients calculated for causally inferred relationships among variables will be evaluated according to Table 1, where the largest correlation provides the decision on potential outcome. A positive correlation coefficient indicates a driven and non-random trend, a zero value indicates no relation between variables, and a negative value indicates a non-driven and random trend. Non-random and non-driven trends may be fractal or chaotic and are not evaluated at

Table 1
Correlation coefficient categories for evaluation of results from structural equation modeling (SEM).

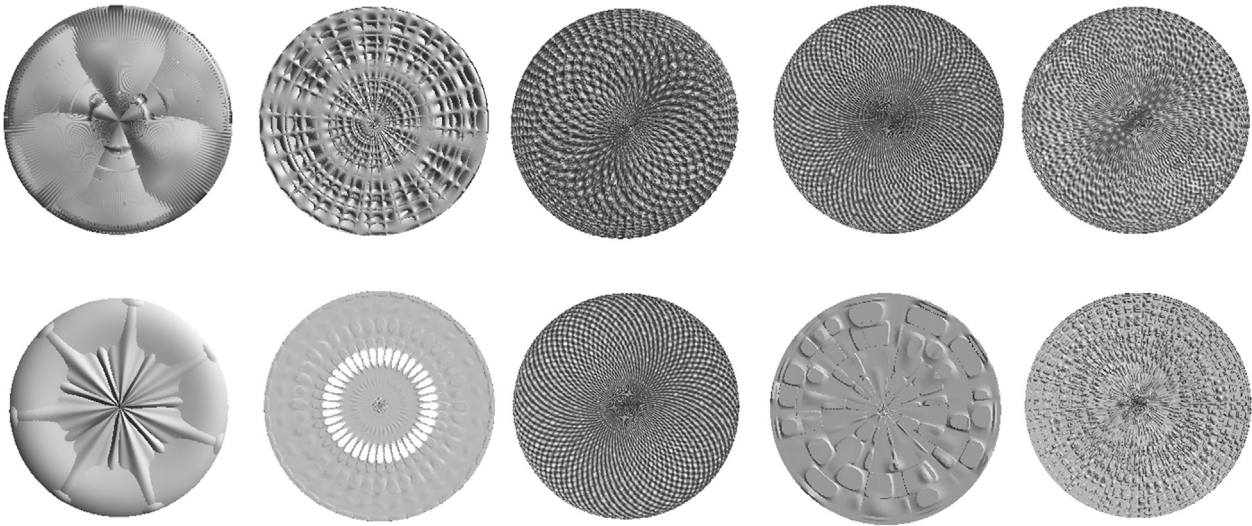| Correlation	Evolutionary trend	
Positive correlation coefficient	Driven	Not random
Zero correlation coefficient	No relation between variables	No relation between variables
Negative correlation coefficient	Not driven	Random

Fig. 2. Diatom 3D surface models of *Actinoptychus*, *Arachnoidiscus*, *Craspedodiscus*, *Cestodiscus*, *Stellarima* (top row), *Asterolampra*, *Brightwellia*, *Coscinodiscus*, *Rocella*, *Stictodiscus* (bottom row).

this time. Only a driven or random trend will be evaluated. For more details on structural equation modeling, see the Appendix A.

3. Results

A total of 313 3D surface models were created for centric diatom genera from the Eocene and Oligocene. Each genus was represented by 31–33 models; examples are depicted in Fig. 2. Jacobians were calculated for each model from which each matrix was made square asymmetric using the cross product of the columns. Jacobian determinants were calculated and averaged for each genus as measurement of morphological complexity. Rank ordering the taxa indicates that *Stictodiscus* is the most morphologically complex followed by *Actinoptychus*, *Rocella*, *Asterolampra*, *Brightwellia*, *Coscinodiscus*, *Craspedodiscus*, *Cestodiscus*, *Arachnoidiscus*, and *Stellarima*, the least morphologically complex (Fig. 3). The five most morphologically complex taxa have more irregular features on their valve faces than those with more regular valve face features. Morphological complexity change over time is mixed.

Functional complexity was calculated for each 3D diatom surface model as the Cauchy stress tensor using surface area of the valve face and interpolated scalar values of the Cauchy traction vector from Hamm et al.'s (2003) empirical load data (Fig. 4; least squares best-fit curve $y = 950.62e^{-0.018x}$ with $R^2 = 0.9982$). The results were analyzed using unidimensional external unfolding analysis (Fig. 5) where the solution was non-degenerate, having non-zero Kruskal's and Young's

STRESS values, a low De Sarbo's Index (1.2), and a Shepard's Index of 72% (Table 2 and Fig. 6) and final coordinates were averaged for each genus. Rank ordering indicated that highest functional complexity occurs with *Asterolampra*, followed by *Arachnoidiscus*, *Rocella*, and *Stictodiscus*. These taxa exhibit multilayered structures on their valve face surfaces and have an irregularity to their surface features. Next in line going from more to less functional complexity are *Cestodiscus*, *Coscinodiscus*, *Craspedodiscus* and *Stellarima* (Fig. 5). These genera have a more regular, repeating cross-hatching valve face surface, with finer cross-hatching occurring from more to less complex forms, respectively. The last two genera in line have very different valve face surfaces, with *Actinoptychus* having large smooth, flatter sections and *Brightwellia* having a ring of holes surrounded by finely textured regular features. For functional complexity, four of the Oligocene taxa occur near the head of the time arrow.

Total complexity was determined as the aggregation of morphological and functional complexity values using external unfolding analysis. The solution was non-degenerate, having non-zero Kruskal's and Young's STRESS values, a low De Sarbo's Index (2.0), and a Shepard's Index of 74% (Table 3 and Fig. 7), and unidimensional final proximities were averaged for each diatom genus. Rank ordering of total complexity values follows a gradient similar to functional complexity with two exceptions (Fig. 8). *Brightwellia* and *Actinoptychus* trade places so that total complexity is least for *Actinoptychus* (Fig. 8). On a taxon-by-taxon basis, functional complexity contributes to total complexity change over time.

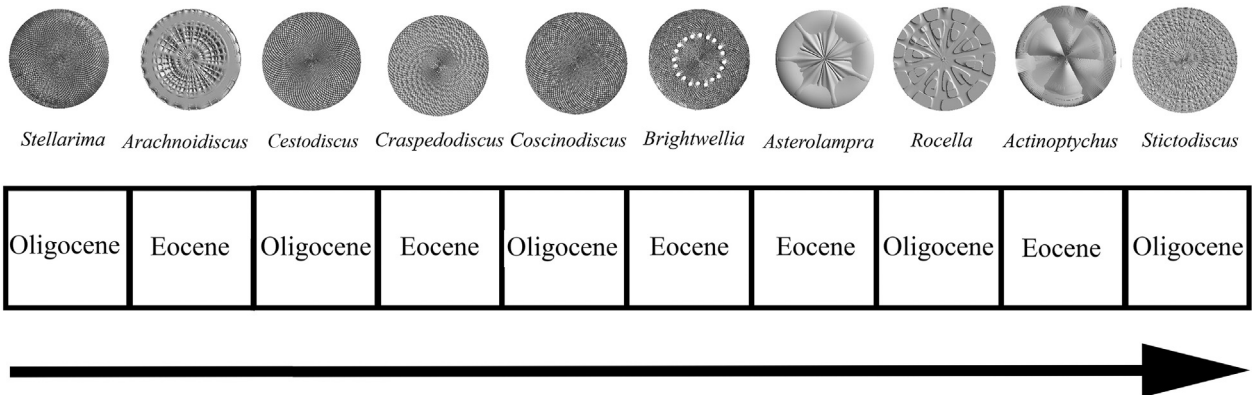|									
Stellarima	*Arachnoidiscus*	*Cestodiscus*	*Craspedodiscus*	*Coscinodiscus*	*Brightwellia*	*Asterolampra*	*Rocella*	*Actinoptychus*	*Stictodiscus*
Oligocene	Eocene	Oligocene	Eocene	Oligocene	Eocene	Eocene	Oligocene	Eocene	Oligocene

Fig. 3. Morphological complexity based on the average Jacobian determinant for each diatom 3D surface model. Arrow points in the direction of least to most morphologically complex taxa from *Stellarima* to *Stictodiscus*, respectively, representing the Eocene or Oligocene.
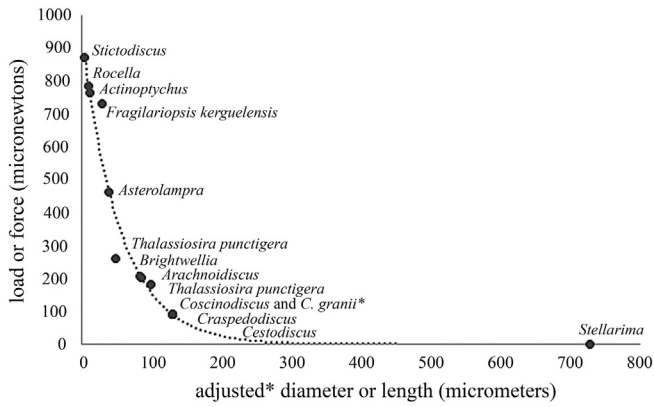
Fig. 4. Combined force data with adjusted values for diatom 3D surface models. The least squares best-fit curve $y = 950.62e^{-0.018x}$ with $R^2 = 0.9982$ is based on the original curve (Fig. 1) using *Coscinodiscus granii* as the reference point (130 μm diameter) to convert patch surface areas to diameters based on the face of a cylinder with height = 1.

Taxa were combined into an Eocene or Oligocene time bin for morphological, functional and total complexity using averaged proximities from external unfolding analysis. The percentages of the amount of contribution to each type complexity indicate that the Oligocene bin contains more complex taxa combined of which morphological complexity

contributes positively to this result (Fig. 9). Functional complexity indicates the opposite by contributing negatively to the Oligocene bin (Fig. 9).

3.1. Magnitude and rate of complexity change

Using the results from external unfolding total complexity analysis, an initial probability transition probability matrix was calculated (Table 4) as input to produce a finite discrete Markov chain of ten centric diatom genera from the Eocene or Oligocene (Fig. 10). The Markov chain is acyclic with no connection between *Actinoptychus* at the start and *Asterolampra* at the end (Fig. 10). The Markov chain is irreducible, transient, non-absorbing, and reversible, and as all states are aperiodic and positive recurrent, the Markov chain is ergodic. As a result and using the Chapman–Kolmogorov equation relating n-transition probabilities among all taxa, the stationary probability states are equivalent to the limiting probability states as the values of the Markov chain at equilibrium (Table 5).

The stationary probability distribution indicates that transition from state to state as taxon shifts is exponential (Fig. 11). Equilibrium is achieved and verified by the normalized left eigenvector elements that sums to one for the eigenvalue equal to one. From the stationary distribution, initial complexity state at *Actinoptychus* is 0.072. Subsequently, the complexity transition values range from 0.073 for *Brightwellia* to 0.08 for *Stellarima*, *Craspedodiscus*, and *Coscinodiscus* to approximately 0.1 for *Cestodiscus*, *Stictodiscus*, and *Rocella* to 0.13 for
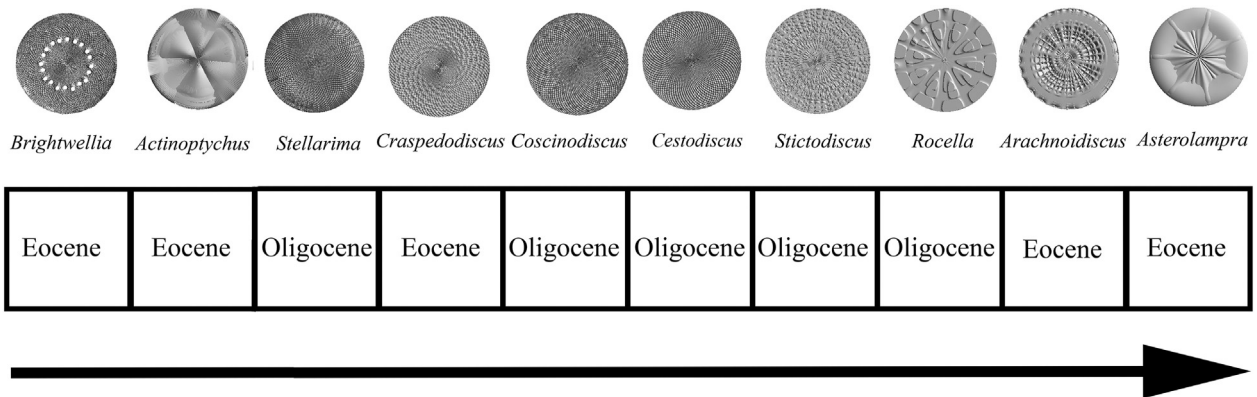

Fig. 5. Functional complexity based on Cauchy stress tensor values for each diatom genus from the average final proximities from unidimensional external unfolding. Arrow points in the direction of least functionally complex, *Brightwellia*, to most functionally complex, *Asterolampra*, representing the Eocene or Oligocene.

Table 2
Diagnostics for 1D and 2D external unfolding analysis of functional complexity values.

| Diagnostic		1D PREFSCAL unfolding	2D PREFSCAL unfolding
Iterations		5000	5000
Penalized stress		.659788856752	.522988312586
	Stress part	2.276719864819	.019482682569
	Penalty part	7.856230504374	14.038968922059
Badness of fit	Normalized stress	.417915502470	.000379565312
	Kruskal's Stress-I	.646463844673	.019482435988
	Kruskal's Stress-II	2.485296272389	.037385585377
	Young's S-Stress-I	.742137743428	.020417442640
	Young's S-Stress-II	.878613655667	.027157055296
Goodness of fit	Dispersion accounted for	.582084497530	.999620434688
	Variance accounted for	.131812933801	.998859015639
	Recovered preference orders	.678548607942	.985993204524
	Spearman's Rho	.313899873360	.973227315183
	Kendall's Tau-b	.255336290642	.948715666859
Variation coefficients	Variation proximities	2.468256953469	2.468256953469
	Variation transformed proximities	1.008911927748	.736608150986
	Variation distances	.332685604713	.720682107040
Degeneracy indices	Sum-of-squares of DeSarbo's intermixedness indices	1.523355677308	1.236062642009
	Shepard's rough non-degeneracy index	.722133475328	.727192048278
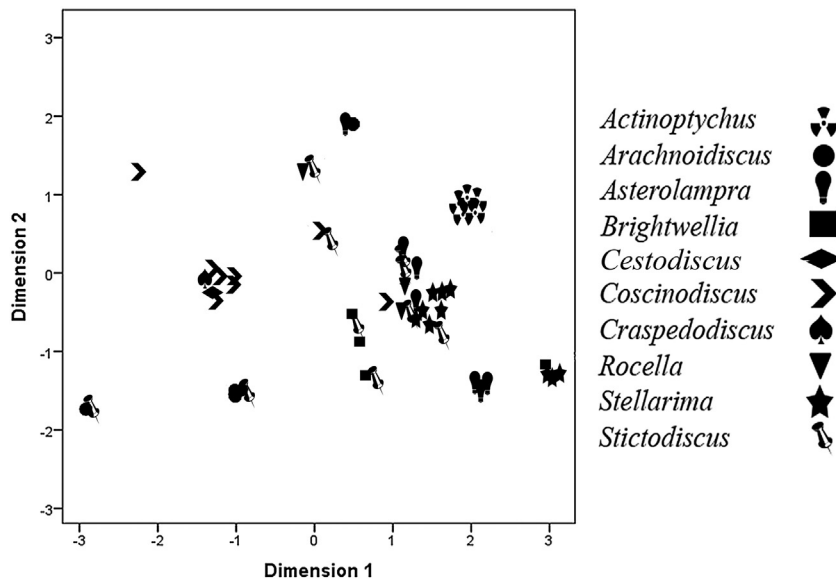
Fig. 6. 2D external unfolding non-degenerate functional complexity space. See Table 2 for diagnostic measures.

Arachnoidiscus to 0.16 for *Asterolampra*. Overall, approximately twice as much time is spent in *Arachnoidiscus* and *Asterolampra* states as initial state *Actinoptychus*. At equilibrium, the more complex taxa, *Arachnoidiscus* and *Asterolampra* persist longer than the less complex taxon *Actinoptychus*. When comparing Oligocene taxa to Eocene taxa, there is an increase in complexity of approximately 1.1 to 1.5 times the initial state of *Actinoptychus* (Fig. 11, Table 6).

The distance between initial transition and final stationary (limiting) probability matrices is calculated as a Frobenius norm of 0.019053. The

Table 3
Diagnostics for 1D and 2D external unfolding analysis of total complexity values.

Diagnostic		1D PREFSCAL unfolding	2D PREFSCAL unfolding
Iterations		5000	5000
Penalized stress		2.605252264042	.351576153530
	Stress part	.637575933383	.009468416789
	Penalty part	10.645538835326	13.054536411464
Badness of fit	Normalized stress	.388466827967	.000089642742
	Kruskal's Stress-I	.623271071017	.009467985096
	Kruskal's Stress-II	2.578092032429	.016322959925
	Young's S-Stress-I	.715364037772	.009229561479
	Young's S-Stress-II	.868877463194	.011255972417
Goodness of fit	Dispersion accounted for	.611533172033	.999910357258
	Variance accounted for	.123482129344	.999748753499
	Recovered preference orders	.693283127948	.946569298646
	Spearman's Rho	.364306944583	.943470778234
	Kendall's Tau-b	.316687454314	.906600015476
Variation coefficients	Variation proximities	2.749852632077	2.749852632077
	Variation transformed proximities	.939219618337	.846069714793
	Variation distances	.301793435185	.829860339418
Degeneracy indices	Sum-of-Squares of DeSarbo's intermixedness indices	1.943820752045	2.022062731149
	Shepard's rough non-degeneracy index	.720056798012	.738516151935

small positive change indicates an increase in total complexity that occurred in the long run for centric diatom genera across the EOT.

Mean first passage time decreases from least to most complex taxa (Fig. 12). More transitions are necessary, and therefore more time is used to get from the initial state, *Actinoptychus*, to two Eocene taxa—*Brightwellia* and *Craspedodiscus*—and three Oligocene taxa—*Stellarima*, *Coscinodiscus*, and *Cestodiscus*—in contrast to the remaining taxa. The shortest transition period occurs from initial state *Actinoptychus* to more complex states, *Stictodiscus*, *Rocella*, *Arachnoidiscus* and *Asterolampra* (Fig. 12). Transitioning from less to another less complex taxon is more time consuming than transitioning from less to more complex taxa. Taxon rank-ordered total complexity vs. total complexity transition states is depicted in Fig. 13. Least squares best-fit curve is $y_{trans} = 1.4081x + 85.087$ ($R^2 = 0.9543$), with slope = 1.4081 as the rate of complexity change over time.

From initial to final state at equilibrium, total complexity increases over time in the long run via discrete finite Markov chain analysis. More complex centric diatom genera are more likely to occur in the short run via mean first passage time, therefore complexity increases. From these analyses, H_0 and H_{A2} are rejected.

3.2. Complexity change over time: driven or random trend?

A sample size of 313 3D surface models representing taxa from 10 centric diatom genera were used in SEM to determine mode of complexity trend. The theoretical causal structure of the model is characterized by an evolutionary trend as an effect of complexity change which occurs as an effect over time. Unweighted least squares was used to calculate the discrepancy (χ^2) function. Of the 100+ different models that were created, 7 admissible models described the theoretical structure of the relationship of the variables had a stability index between 0 and 1. Minimality as minimum discrepancy (CMIN) yielded a best-fit model with lowest root mean square residual (RMR) and highest goodness-of-fit (GFI; AGFI; PGFI) statistics, reflecting a less penalized solution that is more parsimonious than others.

The best-fit structural model resulted in a path diagram of a recursive directed acyclic graph (Markovian) based on unweighted least squares representing the causal inference model (Fig. 14). Minimality and stability were achieved, and a permutation test of 500 iterations yielded $p = 0.002$. Diagnostics indicated that Bernoulli sampling = 0.5 was the best-fit model (Table 6). All models tested were
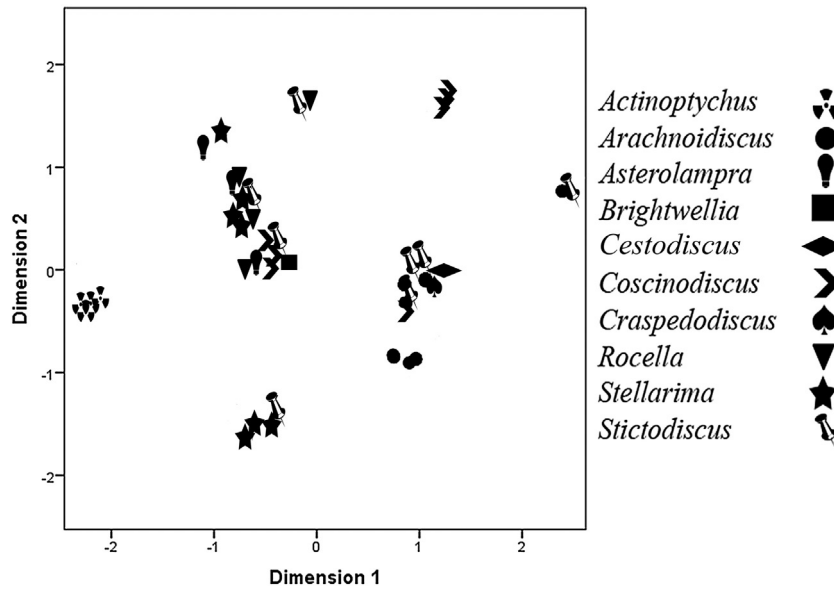
Fig. 7. 2D external unfolding non-degenerate total complexity space. See Table 3 for diagnostic measures.

unconstrained with regard to arrows between unobserved and observed endogenous variables.

The best-fit model (Table 6) resulted in 18 variables of 6 observed and 12 unobserved variables of which nine were endogenous and nine were exogenous (Fig. 14). There were 26 parameters where 17 had fixed weights and nine error variances were set equal to one as input to constrain the model to be linear among unobserved endogenous variables. Forty-two sample moments were calculated with 6 unconstrained estimated parameters and 36 degrees of freedom.

Regression weights as standardized estimates are labeled on each arrow in the path diagram (Fig. 15) and are an indicator of the stability of the model. Unobserved endogenous variables were highly correlated with direct effects from Time to Complexity (0.707) and from Complexity to Evolutionary Trend (0.816). Predictors explained the variance at 50% for Complexity and 66.7% for Evolutionary Trend (Fig. 15). Time was correlated with indirect effects on Evolutionary Trend (0.577). Variance was not predicted for Time (Fig. 15). Using Table 1, an overall evolutionary trend can be causally inferred from these results.

Regression weights between observed and unobserved endogenous variables from the path diagram indicated direct effects where Complexity and Functional complexity were highly positively correlated (0.965), while Complexity was highly negatively correlated with Morphological complexity (−0.831). Time indirectly was highly correlated with Functional complexity (0.682) and highly negatively correlated

with Morphological complexity (−0.588). Predictors explained the variance at 69% for Morphological complexity and 93.1% for Functional complexity. Using Table 1, these results indicate that morphological and functional complexity are causally opposing evolutionary trends.

Eocene had a low negative correlation (−0.198) and Oligocene had a low positive correlation (0.113) with Time (Fig. 15). Driven had a low positive correlation with a small direct effect on Evolutionary Trend (0.099), while Random had a correlation near zero (0.012) and a negligible direct effect. Predictors for Eocene, Oligocene, Driven, and Random were in the range of 0 to 4%. Indirect effects for Driven on Time and Complexity were negligible (0.057 and 0.081, respectively) and for Random were near zero (0.007 and 0.010, respectively). Using Table 1, an overall evolutionary trend as directly driven cannot be causally inferred.

Residual covariances among observed endogenous variables were [−1, 1] and are sufficiently small so that implied correlations of all variables are approximately equal to actual correlations expected of the input variables indicating the minimality of the model. Implied correlations showed that Evolutionary Trend was highly positively correlated with Functional complexity (0.788) and highly negatively correlated with Morphological complexity (−0.679). Morphological complexity was highly negatively correlated to Functional complexity (−0.802) reinforcing the earlier result regarding regression weights. Observed endogenous variables implied correlations with each other had negligible values ranging from −0.14 to 0.12. Using Table 1, implied
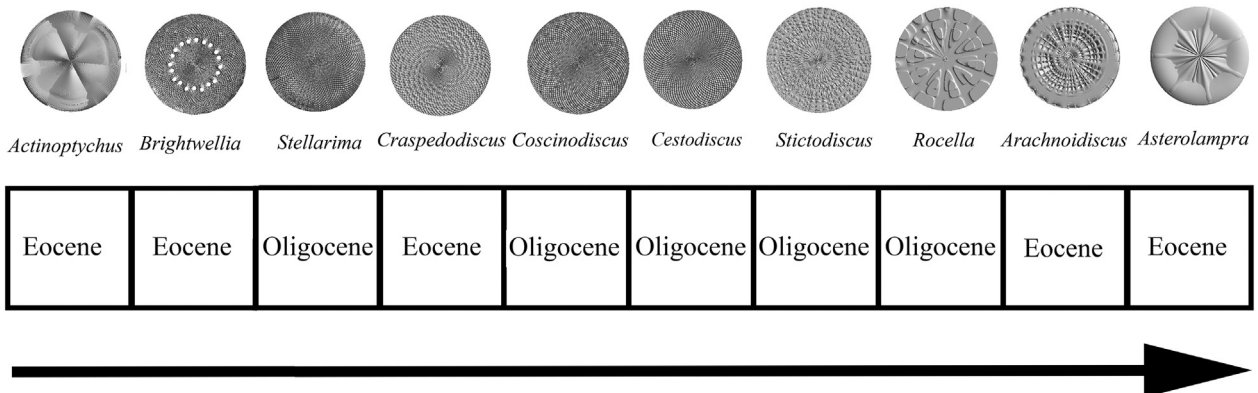

Fig. 8. Total complexity for each diatom genus from the average final proximities from unidimensional external unfolding. Arrow points in the direction of least complex taxon, *Actinoptychus*, to most complex taxon, *Asterolampra*, representing the Eocene or Oligocene.
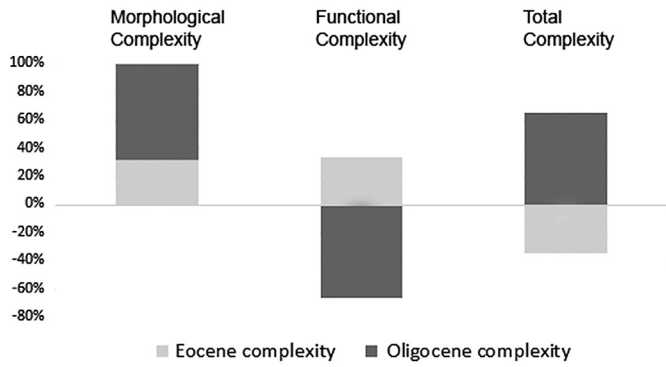
Fig. 9. Comparison of combined taxa Eocene and Oligocene bins for morphological, functional and total complexity.

correlations indicate morphological and functional complexity as causally inferred opposing evolutionary trends, reinforcing earlier results.

Subtracting the residual covariance matrix from the implied covariance matrix results in the sample matrix. From this, the sample correlation matrix indicated that Morphological and Functional complexity were highly negatively correlated (-0.838), Oligocene and Driven were highly positively correlated (1.000), and Eocene was highly negatively correlated (-0.818) with Oligocene and Driven. Functional complexity was positively correlated (0.205) while Morphological complexity had a smaller positive correlation with Driven (0.085). Functional complexity was negatively correlated with Eocene (-0.356) and positively correlated with Oligocene (0.202). Morphological complexity had a correlation near zero with Eocene (0.001) and small positive correlation with Oligocene (0.085). Correlations of Morphological complexity, Functional complexity, Eocene, Oligocene, and Driven with Random were very small values (-0.068, 0.011, 0.073, -0.022, -0.022, respectively). Using Table 1, sample correlations indicate morphological and functional complexity to be inferred as weakly driven evolutionary trends, and the null hypothesis can be rejected.

4. Discussion

Initially, 3D surfaces of centric diatom genera were measured deterministically to assess morphological, functional, and total complexity. As such, parts did not have to be defined, and there was no need to refer to particular descriptors, negation terms or qualifiers. Rather than relying on quantity via counting, quality of 3D surface morphological features was used to measure complexity and provide a natural link between form and function. It was shown that complexity is measurable as a multivariate quantity. Each type of complexity is multivariable, enabling multiple measurements to be consolidated into specific complexity values in determining n-dimensional complexity.

Measurement of the magnitude and mode of complexity as an evolutionary trend was accomplished within probability (Markovian) frameworks. Finding complexity to be a dynamical, ergodic system could be used as predictive values to enable projection of trends beyond the results given. Given the results of this study, for example, increasing total complexity and functional complexity (unlike morphological complexity) as a driven trend may be projected to be true throughout the Cenozoic to the recent. Probabilistic frameworks presented here could be used with additional taxa to test whether the trends are maintained, and if not, what hypotheses may emerge as relevant for testing.

Additional diatom taxa could be used in an extension of the current study. From diatom biostratigraphy of the Late Oligocene to the Quaternary, *Cestodiscus* and *Coscinodiscus* are prevalent along with morphologically similar *Actinocyclus* Ehrenberg (Barron, 1985), and *Asteromphalus* Ehrenberg is present (Barron, 1985) which is morphologically similar to *Asterolampra*. The very large genus of *Thalassiosira* P. T. Cleve is well represented, especially in the Pliocene onward

Table 4
Transition probability matrix of total complexity for construction of finite discrete Markov chain of 10 centric diatom taxa.

|	Actinoptychus	Brightwellia	Stellarima	Craspedodiscus	Coscinodiscus	Cestodiscus	Stictodiscus	Rocella	Arachnoidiscus	Asterolampra
Actinoptychus	0.166667	0.166897	0.173131	0.17617	0.177159	0.178564	0	0	0	0
Brightwellia	0.165975	0.166205	0.172412	0.175439	0.176424	0.177823	0.000809	0.000826	0.000881	0.001037
Stellarima	0.147999	0.148204	0.153739	0.156438	0.157316	0.158564	0.021852	0.022291	0.023782	0.028002
Craspedodiscus	0.139694	0.139887	0.145112	0.14766	0.148488	0.149666	0.031574	0.032209	0.034362	0.040459
Coscinodiscus	0.137055	0.137244	0.14237	0.14487	0.145682	0.146838	0.034664	0.03536	0.037724	0.044418
Cestodiscus	0.133353	0.133538	0.138525	0.140957	0.141748	0.142872	0.038996	0.03978	0.04244	0.04997
Stictodiscus	0.0938	0.09393	0.097438	0.099149	0.099705	0.100496	0.085297	0.087011	0.092828	0.1093
Rocella	0.085386	0.085504	0.088697	0.090254	0.090761	0.091481	0.095147	0.097059	0.103548	0.121922
Arachnoidiscus	0.059148	0.05923	0.061442	0.06252	0.062871	0.06337	0.125861	0.12839	0.136974	0.161278
Asterolampra	0	0	0	0	0	0	0.195099	0.19902	0.212325	0.25
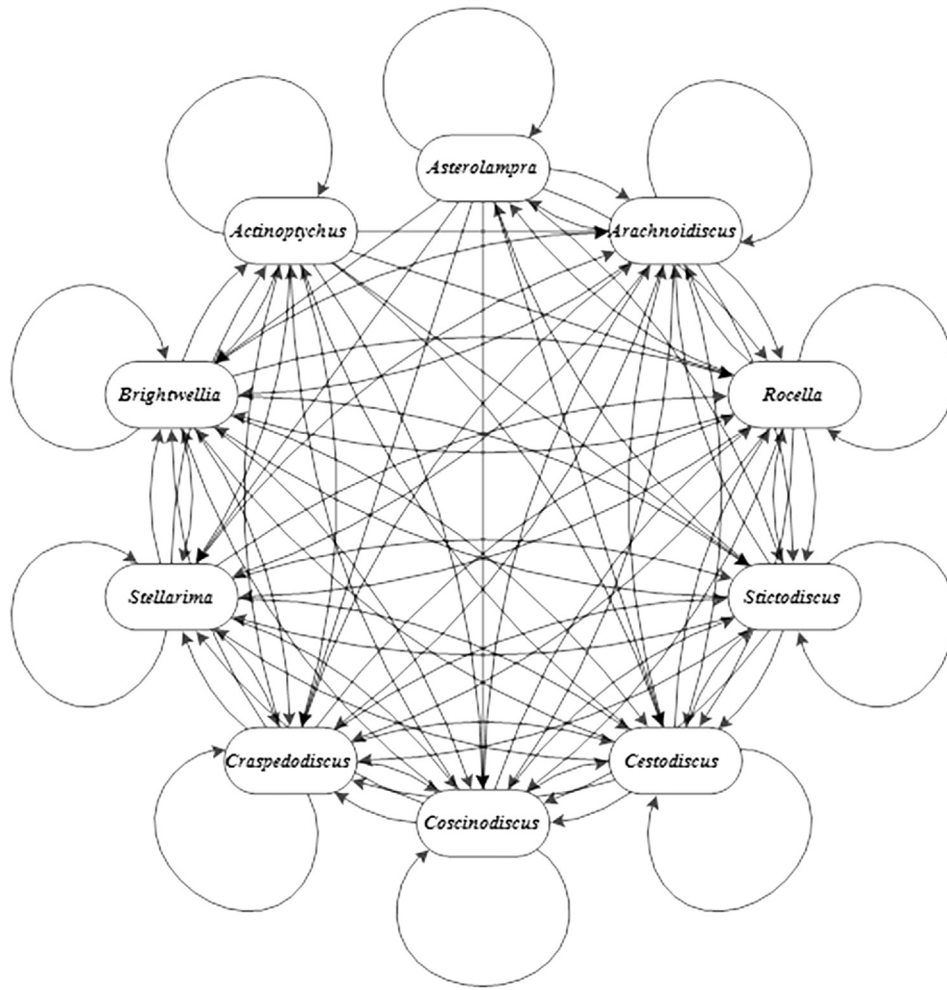
Fig. 10. Finite discrete Markov chain of centric diatom genera with arrows representing conditional probabilities from the limiting probability matrix. Each probability represents the chance that a taxon will become more complex than the preceding taxon in the chain. All taxa in the Markov chain are self-looping, indicating ergodicity.

(Barron, 1985), and somewhat similar in valve face geometric texture to *Stellarima*.

Throughout the Cenozoic into the Quaternary, other stratigraphically important taxa (e.g., Barron et al., 2015), both centric and pennate, include *Hemiaulus* Ehrenberg, *Clavularia* Greville, *Cavitatus jouseana* (Sheshukova-Poretskaya) Williams, *Triceratium* Ehrenberg, *Rhizosolenia* T. Brightwell, *Nitzschia* Hassall, *Rouxia* J. Brun and Héribaud-Joseph in Héribaud (Fenner, 1985; Barron, 1982/1983; Barron, 1985). The expectation is that functional complexity of the pennate taxa would be similar to *Fragilariopsis kerguelensis* with their position nearby on the load curve (Fig. 1). For *Triceratium*, the expectation of functional complexity is a position above *Thalassiosira* and below the pennates on the load curve (Fig. 1). From the result of this study, inferences about any diatom taxon position could be made, and whether any of these taxa would bolster or detract from the results of the current study would need to be tested.

Taxa such as *Hemiaulus* (Eocene) and *Rhizosolenia* (Miocene) have protrusions which may be construed to be morphologically complex attributes of frustule structure. As single cells, *Rhizosolenia* is lightly silicified compared to the more robustly silicified *Hemiaulus* (Round et al., 1990) which may be viewed as more functionally complex. Protrusions are primarily used to form chains, and with both taxa, these protrusions are easily broken (Round et al., 1990). Predation resistance is unlikely to be functionally significant for these taxa, and therefore functional complexity would require a different measure with respect to ability to form colonies.

Complexity change was determined by looking at gradients of taxa per time as well as time bins of combined taxa. On a taxon-by-taxon basis, functional complexity was more important in total complexity (Figs. 5 and 8). By contrast, when looking at combined taxa in time bins, morphological complexity contributed more to total complexity (Fig. 9). Functional and morphological complexity occur in opposing ways, and these results were supported by causal inference analysis (Fig. 15). That morphological and functional complexity may exhibit opposing trends has been determined in previous studies (e.g., McShea and Hordijk, 2013).

Functional complexity was found to be a driver of increase in complexity change as an evolutionary trend. An increase in functional complexity may induce a decrease in environments selecting for increasing morphological complexity (Auerbach and Bongard, 2014). Morphological complexity may not increase in complex environments if the cost is too high (Auerbach and Bongard, 2014), and higher morphological complexity is not necessarily equivalent to evolutionary complexity and success (McNamara, 2006).

Many surface morphologies can have the same functionality (Wainwright, 2007). Fine details and attributes of phenotype in combination determine functional properties which are related to the underlying design or body plan of an organism. An increase in structural duplication and the greater number of elements in a given form, the more morphological dimensions are required to account for that form (Wainwright, 2007). Redundancy in design elements enables the expression of different functional capabilities for each element, resulting

Table 5
Limiting probability matrix of total complexity finite discrete Markov chain of 10 centric diatom taxa.

|	Actinoptychus	Brightwellia	Stellarima	Craspedodiscus	Coscinodiscus	Cestodiscus	Stictodiscus	Rocella	Arachnoidiscus	Asterolampra
Actinoptychus	0.147614	0.147613	0.147614	0.147613	0.147614	0.147613	0	0	0	0
Brightwellia	0.147001	0.147001	0.147001	0.147001	0.147001	0.147001	0.001286	0.001286	0.001286	0.001286
Stellarima	0.13108	0.13108	0.13108	0.13108	0.13108	0.13108	0.034725	0.034725	0.034725	0.034725
Craspedodiscus	0.123724	0.123724	0.123724	0.123725	0.123724	0.123724	0.050173	0.050173	0.050173	0.050173
Coscinodiscus	0.121387	0.121386	0.121386	0.121387	0.121386	0.121386	0.055083	0.055083	0.055083	0.055083
Cestodiscus	0.118108	0.118108	0.118108	0.118108	0.118108	0.118108	0.061968	0.061968	0.061968	0.061968
Stictodiscus	0.083077	0.083077	0.083077	0.083077	0.083077	0.083077	0.135543	0.135543	0.135542	0.135543
Rocella	0.075624	0.075624	0.075624	0.075624	0.075624	0.075624	0.151195	0.151195	0.151195	0.151196
Arachnoidiscus	0.052386	0.052386	0.052386	0.052386	0.052386	0.052386	0.200002	0.200001	0.200002	0.200001
Asterolampra	0	0	0	0	0	0	0.310025	0.310026	0.310026	0.310025
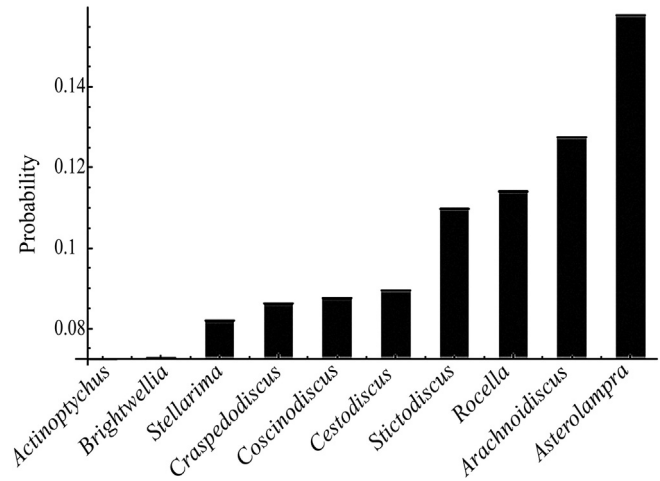
Fig. 11. Stationary probability distribution for the total complexity finite discrete Markov chain.

in functional separation and the potential for novelties in organismal form (Alfaro et al., 2004; Wainwright, 2007). Highly different morphologies can produce the same functionality (Alfaro et al., 2004; Wainwright, 2007) which facilitates phenotypic diversification (Wainwright, 2007). Centric diatoms from the Eocene and Oligocene may be exhibiting this pattern.

Diatom morphological changes have been closely linked to environmental changes (Round et al., 1990), and distributional patterns were affected by environmental changes during the EOT (Jordan and Stickley, 2010; Gladenkov, 2014). In the Early Cretaceous, diatoms were more environmentally tolerant, had more complicated and variable life cycles (Harwood et al., 2007), and by the Early Eocene were cosmopolitanly distributed (Fenner, 1985; Harwood et al., 2007). Oceanic circulation was influenced by equatorial patterns with relatively warm bottom waters and broad upwelling events with locally high biogenic silica deposition in upwelling regions (Barron and Baldauf, 1989; Baldauf and Barron, 1990). In the Middle Eocene, species with robust, heavily silicified frustules were present, and widespread biogenic silica deposition occurred because of an intensive increase in oceanic circulation via increased tectonic activity (Baldauf and Barron, 1990). As a result, surface water productivity increased with increasing coastal and equatorial upwelling activity (Barron and Baldauf, 1989; Baldauf and Barron, 1990), and diatom assemblages generally became more provincial (Fenner, 1985). By the late Middle Eocene, polar cooling accelerated, latitudinal thermal gradients became more pronounced, and increased stratification occurred (Barron and Baldauf, 1989; Baldauf and Barron, 1990; Zachos et al., 2008).

Continued cooling occurred in the earliest Oligocene, and biogenic silica deposits increased in the Southern Oceans and decreased in the middle latitudes (Barron and Baldauf, 1989; Baldauf and Barron, 1990). In the low latitudes, upwelling events became well established, oceanic circulation continued to intensify, and diatoms continued to be provincial (Fenner, 1985). In the Late Oligocene, biogenic silica deposition increased in some areas, but sediments in the equatorial Pacific were devoid of diatom frustules (Barron and Baldauf, 1989). An increase in taxa with finely silicified frustules emerged as heavily silicified forms decreased (Fenner, 1985; Barron and Baldauf, 1989) as biogenic silica deposition changed in magnitude from the Atlantic to Pacific and Southern Oceans (Barron and Baldauf, 1989).

Turbulence in the oceanic water column for brief periods spurred diatom productivity to ensure that diatoms dominated the Cenozoic (Katz et al., 2004). Rapid evolutionary response (Jordan and Stickley, 2010) via selectivity led to the rise and success of diatoms (Katz et al., 2004). Episodic nutrient availability was capitalized on by diatoms because of their ability to acquire and store resources (Katz et al., 2004). During

Table 6
Best-fit diagnostics for SEM models. All models are unconstrained with Bernoulli sampling used to assign probability 0 or 1 to each taxon. Best-fit parameters measured are as follows: NPAR is the number of estimated parameters, CMIN is the minimum discrepancy, RMR is root mean residual, GFI is goodness-of-fit, AGFI is adjusted goodness-of-fit, and PGFI is parsimonious goodness-of-fit. Best-fit model used for analysis is Bernoulli sampling = 0.5.

Model	NPAR	CMIN	RMR	GFI	AGFI	PGFI
Bernoulli sampling = 0.1	6	508.457	.375	.988	.987	.847
Bernoulli sampling = 0.25	6	489.740	.367	.989	.987	.848
Bernoulli sampling = 0.4	6	471.892	.359	.989	.988	.848
Bernoulli sampling = 0.5	6	468.681	.358	.989	.988	.848
Bernoulli sampling = 0.6	6	470.861	.359	.989	.988	.848
Bernoulli sampling = 0.75	6	483.943	.364	.989	.987	.848
Bernoulli sampling = 0.95	6	522.560	.380	.988	.986	.847

the Eocene and Oligocene, major diatom turnover events occurred to change the composition of assemblages in response to changing climates (Barron and Baldauf, 1989; Baldauf and Barron, 1990), sea levels, ice advancement, and atmospheric CO_2 levels (Jordan and Stickley, 2010). Diatoms prospered in shallow waters on continental margins and have changed the structure of marine food webs (Katz et al., 2004).

From the current study, the composition of diatom taxa that occurred throughout the Eocene and Oligocene changed in complexity. *Actinoptychus*, a morphologically complex genus (Fig. 3), occurred in the Late Cretaceous into the Eocene (Baldauf and Barron, 1987; Harwood and Nikolaev, 1995). Along with the less morphologically complex *Craspedodiscus* (Fig. 3) in the Early to Middle Eocene, morphologically complex *Asterolampra* and *Brightwellia* (Fig. 3) were prevalent (Fenner, 1985). *Arachnoidiscus*, a morphologically less complex genus having a regular surface pattern (Fig. 3), was abundant as well during this time (Fenner and Mikkelsen, 1990). By the end of the Middle Eocene, the less common species of *Craspedodiscus* disappeared, and by the Middle to Late Eocene, although *Asterolampra*, *Brightwellia*, and *Craspedodiscus* still dominated, many common and typical Eocene taxa disappeared in the Middle Oligocene (Fenner, 1985). In the Early Oligocene, cosmopolitan species were still present, and at the Eocene–Oligocene boundary, the less complex genus, *Cestodiscus* (Fig. 3) had species that became endemics. Cosmopolitan species of the morphologically less complex genus, *Coscinodiscus*, and the more complex genus *Rocella* (Fig. 3) first appear and became common during the Middle Oligocene (Fenner, 1985).

With increasing cooling and oceanic turbulence along with increased tectonic activity producing increases in productivity and diversification from the Eocene to the Oligocene, environmental and climate change resulted in more complex ecological habitats for diatoms (Katz et al., 2004; Jordan and Stickley, 2010). Environmental specificity and adaptability of diatoms (Harwood et al., 2007) may indicate ecospace differentiation and expansion to accommodate the change

from cosmopolitan to more provincial and endemic taxa. Taking advantage of the expansion of ecospace induces an increase in ecological complexity (Knoll and Bambach, 2000) which may be a response to changing latitudinal, bathymetric, hydrodynamic, and temperature gradients. The cost of even slightly increasing complexity may be acceptably borne in terms of ability to acquire nutrients and resources in changing environmental and climate conditions. Such costs may induce the adaptation of specific diatom taxa to specialized environments.

Selectivity of predation resistance is evident in centric diatom morphologies (Smetacek, 2001). Selecting for predation resistance should cause more complex morphologies to evolve than what might happen by chance (Auerbach and Bongard, 2014). However, mechanical complexity decreases in environments that select for increasing morphological complexity (Auerbach and Bongard, 2014) so that morphological complexity may decrease as a response to increasing functional complexity (McShea and Hordijk, 2013). This may be inferred with centric diatom genera from the Eocene and Oligocene by the current study. By the time of the turnover event at the EOT, *Brightwellia* became extinct as did many species of *Actinoptychus*, *Arachnoidiscus*, *Asterolampra*, and *Craspedodiscus*, so that these Eocene taxa were unavailable as food for predators. Oligocene taxa *Rocella*, *Stictodiscus* and many species of *Cestodiscus*, and *Stellarima* became extinct as another turnover event from the Middle to Late Miocene occurred with the proliferation of pennate diatom taxa (Strelnikova, 1990). The net result is a loss of morphologically complex centric taxa over time, yet more predation resistance is necessary because increasing ecological complexity provided an
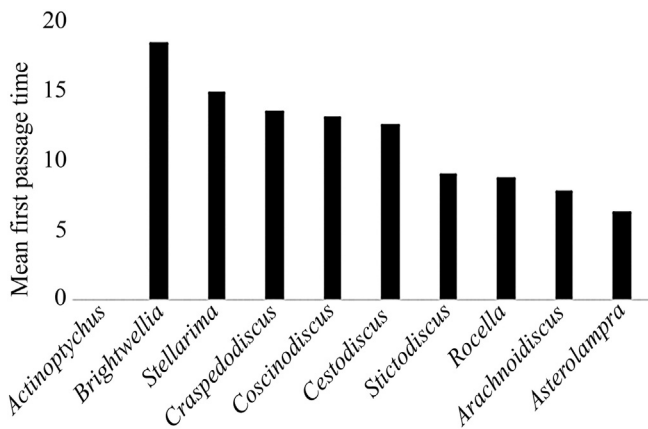

Fig. 12. Mean first passage time from the initial state, *Actinoptychus*, to each succeeding state (= taxon) in the process of changing complexity. Each successive state has fewer transitions than the preceding state.
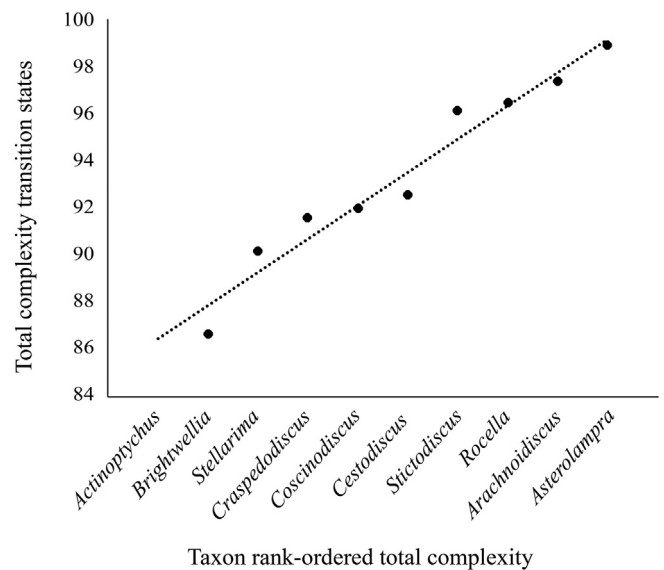

Fig. 13. Plot of taxon rank-ordered total complexity vs. total complexity transition states from mean first passage times from discrete finite Markov chain. Least squares best-fit curve is $y_{trans} = 1.4081x + 85.087$ with $R^2 = 0.9543$; the slope = 1.4081 and is the rate of complexity change over time.
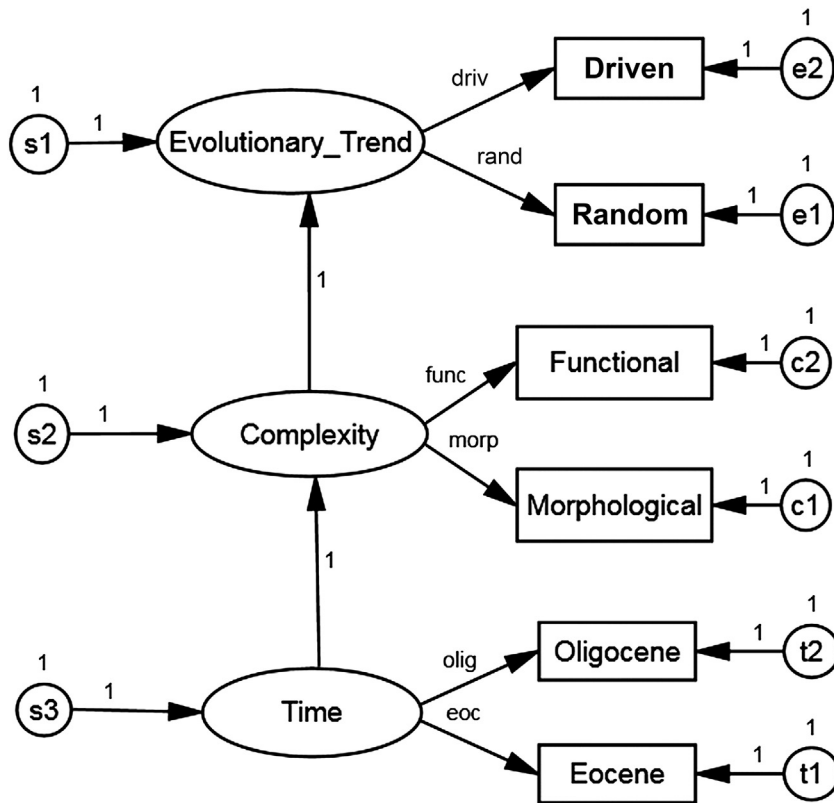
Fig. 14. SEM of labeled causal inference diagram for centric diatom genera and evolutionary trend of complexity. Path from unobserved endogenous variables Time to Complexity to Evolutionary Trend is constrained to be linear (regression weights = 1). All exogenous variables are set to error variances = 1 and constrained to be linearly related to unobserved endogenous variables. Observed endogenous variables are unconstrained. Driven and Random are probability vectors that are counterfactual to each other.
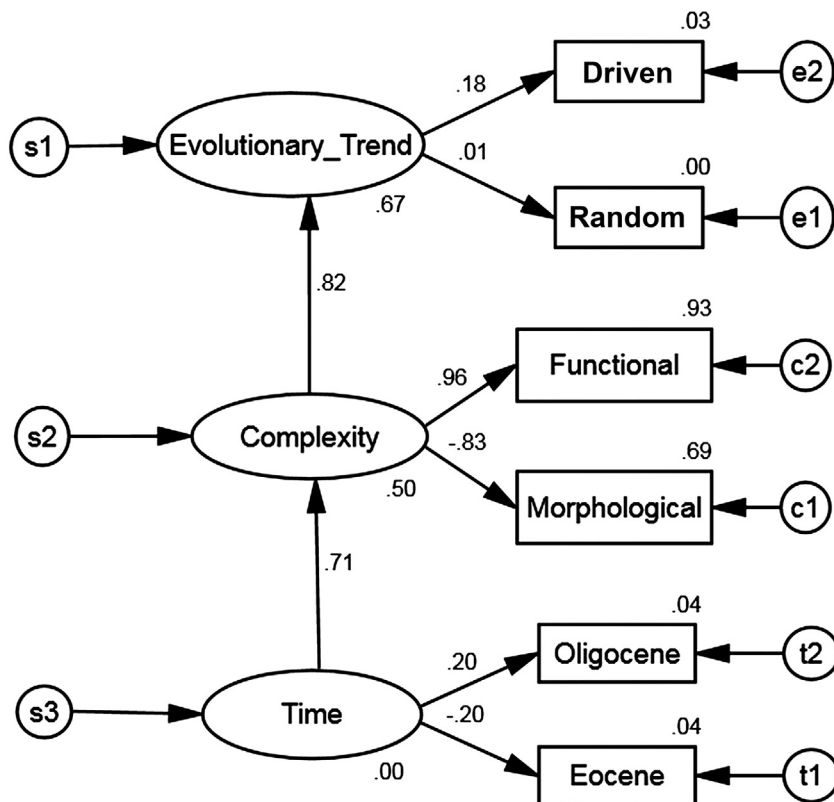

Fig. 15. SEM of resultant causal inference diagram for centric diatom genera and evolutionary trend of complexity. Standardized implied and sample correlation coefficients for unconstrained observed and standardized implied correlation coefficients for unobserved endogenous variables.

Table 7

‡Terms with parameters and coefficients of the z-equation for each centric diatom genus. For all parameters, $u, v \in [0, 2\pi]$.

Taxon	z-Equation[†]
Actinoptychus	$\gamma \sin 3u - m \cos tv \sin \varsigma u \cos \omega u \sin \Xi v \pm \cos\ 3v + \cos(10u - 3) - k \cos \Psi u \sin \rho u$
Arachnoidiscus	$\psi\ , \cos \beta v^4 B \left\{ {, \sin \atop , \cos} \right\} t u^4$
Asterolampra	$\varepsilon \sin 7v^3\ s \sin 0.37u^3\ m \cos 1.5u + \mu \cos 2.7u^2\ \kappa \sin 4.2u^2 + \Gamma \sin 1.9u^2\ N \cos \Lambda u^2 + J \sin lu^3\ Q \sin \beta u^3\ 1.7 \cos 1.5u$
Brightwellia	$\sin \Lambda u^\Xi 2 \sin K u^2 2 \sin 14 u^l\ \cos tu\ \ 3 \sin \xi v^\psi - \cos \varsigma u\ \ \cos \omega v$
Cestodiscus	$3 \sin\ \phi u^5 - \cos \varsigma u\ \ \cos \omega v$
Coscinodiscus	$\varsigma \sin\ \phi u^\Omega - \cos \varsigma u\ \ \cos \omega v$
Craspedodiscus	$\psi \sin\ \phi u^\Omega - \Lambda \cos \varsigma u\ \ \cos \omega v$
Rocella	$2 \cos \Xi v^3 N \cos q u^k W \sin N v^3 3 \cos \beta v 3 \sin 0.1 u^{12} 3 \sin M u 3 \cos C v$
Stellarima	$\Omega\ , \cos t u^k -\ , \cos \varsigma u\ \ , \cos \omega v +\ , \cos \varepsilon u^\tau\ \ , \sin \psi \left\{ {v \atop u} \right\}$
Stictodiscus	$\sin \Lambda u^\circledast \cos \varsigma u^\tau\ \ \sin \xi v\ \ \cos \omega v^W\ \ \sin m u^\Xi\ \ \cos \rho u$

‡ See text for x- and y-equations.

† Coefficients: $0 \le \varepsilon \le 2, 0 \le \beta \le 4.3, 0.1 \le \mu \le 0.4, 0.1 \le \kappa \le 1, 0.1 \le q \le 2, 0.2 \le m \le 1.6, 0.2 \le \Lambda \le 4, 0.4 \le \Gamma \le 0.7, 0.5 \le \gamma \le 2, 0.6 \le \phi \le 1, 0.7 \le J \le 2.5, 0.7 \le L \le 8, 1 \le \tau \le 2, 1 \le W \le 3, 1 \le \psi \le 4.1, 1 \le t, \Omega \le 5, 1 \le \Xi \le 6, 1 \le k \le 8, 1 \le M \le 10.3, 2 \le l \le 3, 2 \le s \le 4, 2 \le C \le 10, 2.7 \le N \le 10, 8 \le K \le 10, 8 \le \xi \le 31, 20 \le \Psi \le 60, 30 \le \varsigma \le 590, 43 \le \rho \le 85, 50 \le \omega \le 1050.$

avenue for increasing abundances of Cenozoic marine invertebrates, which in turn induced large changes in marine food webs (Katz et al., 2004).

Limiting nutrient conditions in the oceans have changed the degree of silicification of diatom frustules but has not prevented diatom diversification (Harwood et al., 2007; Jordan and Stickley, 2010). In general, diatom genera richness is highly correlated with Cenozoic invertebrate richness (Katz et al., 2004). Increased quality and quantity of primary producers during the Cenozoic has influenced marine invertebrate evolution (Katz et al., 2004). Coevolution of diatom taxa and some predators may be evidence of a driven response (Smetacek, 2001; Michels et al., 2012). For some copepods, silica coating on their chitinous teeth is a coevolved trait with diatoms (Katz et al., 2004; Michels et al., 2012). Few studies exist on predation resistance at the scale of diatoms, and invertebrate grazers are inferred to be species specific in terms of head size relative to diatom size (Tall et al., 2006a, 2006b).

For centric diatoms across the EOT, all have differing morphologies yet all have some degree of predation resistance (Hamm, 2005). Plankton evolution is generally driven by predator–prey interactions rather than competition (Smetacek, 2001). The variety of morphologies of the diatom frustule is a reflection of predation resistance as a unifying functionality rather than competitive advantage among diatom taxa (Smetacek, 2001). However, as diatoms inhabit more specific habitats (e.g., Round et al., 1990; Lowe, 2011) over the long term, this situation may place limitations on the kinds of diatoms that are available to predators in a given area. Nutritional value of protein, lipids and carbohydrates vary in different diatom taxa (Lowe, 2011). Predators may occupy specific habitats which may restrict their access to highly nutritious diatoms (Lowe, 2011). Some diatoms may avoid predators by vertically descending in the water column at night to acquire nutrients, then vertically ascending during daylight to photosynthesize (Katz et al., 2004). All of these factors could put selective pressure on predators.

Rapid evolution of diatoms during the Eocene and Oligocene occurred in response to environmental and climate conditions, increasing diatom environmental specificity and the proliferation of endemic and provincial taxa. Selecting for predation resistance might occur as a constraint on the kinds of predators adapted to the specific habitats now occupied by diatoms during this time. Functional efficiency (McShea and Hordijk, 2013) may be more a result of changing ecological complexity rather than morphological complexity. Ecological complexity over the long term increases as a directional pattern, and functional complexity should do so as well

(Knoll and Bambach, 2000). Morphological complexity changes in response to ecological complexity, yet need not be increasing or driven. Morphological and functional complexity change may respond in opposing, unconnected ways to ecological complexity.

5. Conclusions

Using 3D surface models of centric diatom genera, morphological and functional complexity was analyzed as multivariate quantities. Markov chain analysis enabled a probabilistic framework in which to measure multivariate total complexity as increasing across the EOT. Functional complexity on a taxon-by-taxon basis contributed to increasing total complexity. Combined taxa in Eocene or Oligocene bins indicated that morphological complexity contributed to increasing total complexity. Functional complexity does not need to increase with increasingly complex morphologies or vice versa. From causal inference analysis, morphological and functional complexity exhibited weakly driven evolutionary trends. Changes in environmental and climate conditions across the EOT may have induced predation resistance as an ecological response differently when looking at morphological in contrast to functional complexity as macroevolutionary patterns over time.

Appendix A

A.1. Methods

A.1.1. Selection procedure of taxa for inclusion in complexity analysis

I define all possible surface morphological patterns to be the pool from which to select samples, s. Prior knowledge, **Z**, consists of types of n-surface morphological patterns for Coscinodiscales, Stictodiscales, Arachnoidiscales, and Asterolamprales given as Z_1, Z_2, Z_3, Z_4, respectively. The values to be determined are the genera, **Y**, that have the i, 1, 2, …, n-surface morphological patterns.

The selection indicator variable for samples s_1, s_2, s_3, s_4 is $\mathbf{A}_s = (A_1, A_2, …, A_n)^T$. The sampling scheme to select genera from $\mathbf{A}_s$ is $\left\{ {A_i = 1, i \in s \atop \text{otherwise}, A_i = 0, i \in \bar{s}} \right.$. Sampling is accomplished according to the rule used to evaluate A and is given as $f(\mathbf{A}_S |\ \mathbf{Z}, \mathbf{Y}; \Phi)$ where Φ is an unknown parameter (Smith, 1983). From **Z**, the rule for genus selection is: the valve face shape approximates circularity and has radiating areolae reflecting radial symmetry. The sampling scheme is based only on **Z** so that any genus, Y_i, can be selected via $\mathbf{A}_s$. That is in probability terms, $P(Y = y, A = a\ |\ Z = z) = P(Y = y\ |\ Z = z)\ P(A = a\ |\ Z = z)$. In effect, the rule becomes $f(\mathbf{A}_s\ |\ \mathbf{Z}; \Phi)$, and the sampling selection process satisfies conditional independence and the ignorability criterion (Dawid, 1979; Smith, 1983; Pearl, 2000, 2009).

A.2. Morphological complexity

Partial derivatives numerically solved from parametric 3D equations with boundary conditions $[0, 2\pi]$ become elements of the Jacobian matrix (Jacobian) (Pappas, 2005a, 2005b, 2008; Pappas and Miller, 2013). Parametric 3D equations are a vector mapping $F(u, v) = f[x(u, v), y(u, v), z(u, v)]$ with $F(x, y, z) = F[f(u, v), g(u, v), h(u, v)]$ so that tangent planes $F(u, v) = F(x, y, z)$ for one dimension. Similarly, tangent planes for the other two dimensions as $G(u, v) = G(x, y, z)$ and

$H(u,v) = H(x,y,z)$ can be defined. Since 3D surface models are size invariant, and rotation and reflection involve only a switch in parameters and a change in sign, respectively (Pappas and Miller, 2013), numerical solutions of this affine mapping for all forms can be compared with respect to morphological complexity.

Implicit equations $x = f(u,v)$, $y = g(u,v)$, and $z = h(u,v)$ are the basis for a diatom 3D surface model where x, y, z are independent. These equations represent a modified capped cylinder where the cap is the diatom valve face surface with a circular shaped boundary where x and y are identical equations except for a cosine function in the x-direction and a sine function in the y-direction. Changes in the 3D valve face occur in the z-direction so that z is a function of parameterized x and y and is given as $x = g(u,v)$, $y = h(u,v)$, and $z = f(x,y)$.

The generalized system of parametric 3D equations for cylindrical centric diatoms is: $a\cos v\left(1 + \left\{\begin{smallmatrix}\sin bu^2\\\cos bu^2\end{smallmatrix}\right\}\right)c\left\{\begin{smallmatrix}\cos ju^2\\\sin ju^2\end{smallmatrix}\right\}$, $a\sin v\left(1 + \left\{\begin{smallmatrix}\sin bu^2\\\cos bu^2\end{smallmatrix}\right\}\right)c\left\{\begin{smallmatrix}\cos ju^2\\\sin ju^2\end{smallmatrix}\right\}$, and the z-equation varies for each centric diatom genus as given in Table 7‡. Coefficients for the x- and y-equations are $6 \leq a \leq 40$, $0.5 \leq b \leq 2$, $1 \leq p \leq 2$, $0.2 \leq c \leq 1$, $0.5 \leq j \leq 1$, $0 \leq s \leq 30$, $0 \leq t \leq 90$, and $2 \leq k \leq 8$; coefficients for the z-equation are given in Table 7‡.

Each column of the Jacobian represents each parameter, u, v, and each parameter is a vector and is given as $J = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \\ u_3 & v_3 \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{bmatrix}$.

The 3×2 asymmetric matrix is made square (3×3 matrix) by calculating the magnitude of the cross product of row elements for each taxon which becomes the third column in the matrix. The cross product of the two columns is $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\|\,\|\mathbf{v}\|\sin\theta$, where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$, and $\|\mathbf{u} \times \mathbf{v}\|$ is the vector of magnitude. The Jacobian becomes

$\mathbf{M}_{u,v} = \begin{bmatrix} u_1v_1(u_2v_3 - u_3v_2)\mathbf{i} \\ u_2v_2(u_3v_1 - u_1v_3)\mathbf{j} \\ u_3v_3(u_1v_2 - u_2v_1)\mathbf{k} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \left(\frac{\partial y}{\partial u}\times\frac{\partial z}{\partial v} - \frac{\partial z}{\partial u}\times\frac{\partial y}{\partial v}\right) \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \left(\frac{\partial z}{\partial u}\times\frac{\partial x}{\partial v} - \frac{\partial x}{\partial u}\times\frac{\partial z}{\partial v}\right) \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \left(\frac{\partial x}{\partial u}\times\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\times\frac{\partial x}{\partial v}\right) \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$.

The Jacobian determinant is $\det(\mathbf{M}_{u,v}) = m_{11}[(m_{22}m_{33}) - (m_{23}m_{32})] - m_{12}[(m_{21}m_{33}) - (m_{23}m_{31})] + m_{13}[(m_{21}m_{32}) - (m_{22}m_{31})]$, and is a summary value of 3D surface morphological complexity. Values for the Jacobian determinants for each diatom genus are averaged and rank-ordered to produce a morphological complexity gradient. The higher the Jacobian determinant, the higher the number of tangent lines and planes on the valve face that are more closely spaced together representing a finer textured surface.

A.3. Functional complexity

As with morphological complexity measurement, surface geometry is measured using parametric 3D equations that are solved numerically as Jacobians (Pappas and Miller, 2013). For the diatom as a capped cylinder, the sides of the cylinder are fixed so that variation of the valve face—a capped end of the cylinder—is represented as a patch by keeping x- and y-directions constant while varying the z-direction in the parametric 3D equations. In this way, z becomes a function of x and y so that z is now dependent on x and y (Taylor and Mann, 1983). From

Eq. (3), $\begin{bmatrix} j_1 \\ j_2 \\ j_3 \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial y}{\partial u}\times\frac{\partial z}{\partial v} - \frac{\partial z}{\partial u}\times\frac{\partial y}{\partial v}\right) \\ \left(\frac{\partial z}{\partial u}\times\frac{\partial x}{\partial v} - \frac{\partial x}{\partial u}\times\frac{\partial z}{\partial v}\right) \\ \left(\frac{\partial x}{\partial u}\times\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\times\frac{\partial x}{\partial v}\right) \end{bmatrix} = \begin{bmatrix} m_{13} \\ m_{23} \\ m_{33} \end{bmatrix}$. From the cross products and the Pythagorean Theorem, the patch surface area is calculated as $S_{patch} = \sqrt{j_1^2 + j_2^2 + j_3^2}\,\Delta u\,\Delta v$ with boundary conditions $u, v \in [0, 2\pi]$; total whole 3D surface area is $S_{total} = \int_0^v \int_0^u \sqrt{j_1^2 + j_2^2 + j_3^2}\,du\,dv$. The relation between diameter and surface area of a cylinder is $cylinder_{surface\;area} = 2\pi r + 2\pi rh$.

Stress and strain on the diatom frustule surface can be calculated on a patch that is representative of that surface. Newton's second law describes the change from a diatom valve face at rest to its deformed configuration as $\nabla \cdot \sigma + \mathbf{F}$ as a governing equation of motion in terms of stress σ. Displacement as deformation is given as the governing equation $\varepsilon = \frac{1}{2}\left[\nabla\mathbf{u} + (\nabla\mathbf{u})^T\right]$ in terms of strain ϵ. A force acting on the valve face surface will cause changes in multiple directions. A continuous linear elastic material such as the amorphous silica-polysaccharide matrix exhibits stress and strain related to each other as the constitutive equation given by Hooke's Law of $\sigma = -c\varepsilon$, where σ is the stress tensor, ε is the strain tensor, and c is the elasticity (or stiffness) tensor, and all tensors are contravariant. The stress tensor is expressed as $\sigma = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix}$, and for $x = 1$, $y = 2$, $z = 3$, then $\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$. For the strain tensor, $\varepsilon = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix}$. Hooke's Law becomes $\sigma_{ij} = -\sum_{k=1}^{3}\sum_{l=1}^{3} c_{ijkl}\varepsilon_{kl}$ where i and $j = 1, 2, 3$, and in particular for continuous materials, σ is the Cauchy stress tensor and ε is the infinitesimal strain tensor. The Cauchy stress tensor is a measure of true stress relative to the actual deformation that occurs at a given time when stress is small, elasticity is linear, and it completely defines stress at each point internally. While the stress and strain tensors are second order, the elasticity (or stiffness) tensor is fourth order and is given as $c_{ijkl} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{41} & c_{42} & c_{43} & c_{44} & c_{45} & c_{46} \\ c_{51} & c_{52} & c_{53} & c_{54} & c_{55} & c_{56} \\ c_{61} & c_{62} & c_{63} & c_{64} & c_{65} & c_{66} \end{bmatrix}$.

For small deformations with linear elasticity, the Cauchy stress tensor is related to surface traction as the Cauchy traction vector expressed as $T_{ij}^{(n)} = \sigma_{ij}n_i$ or $\mathbf{T}^{(n)} = \boldsymbol{\sigma} \cdot \mathbf{n} = \frac{dF_i}{dS}$ where F_i is force and S is surface area. Surface traction is everywhere on the surface boundary so that strains and displacements are minimal. The Cauchy stress tensor becomes $\sigma = \begin{bmatrix} \sigma_{xx} & \tau_{xy} & \tau_{xz} \\ \tau_{yx} & \sigma_{yy} & \tau_{yz} \\ \tau_{zx} & \tau_{zy} & \sigma_{zz} \end{bmatrix}$. According to the Euler–Cauchy stress principle, action of one body on another body is the system of distributed forces internally as a vector field and is the surface traction, $\mathbf{T}^{(n)}$, that depends only on $\mathbf{n}$. Externally, forces on the surface are governed by Euler's equations of motion so that the laws of conservation of mass, energy and momentum hold; contact forces that act on the surface are balanced by internal forces in the material that is being affected (Heinbockel, 2001).

From a given point, all the stress vectors are acting on all planes of a continuous amorphous silica-polysaccharide diatom surface. Endpoints are defined on an ellipsoid surface oriented to have coordinate axes in the same direction as principal axes that are the principal stress vectors, $\sigma_1, \sigma_2, \sigma_3$. Normal stresses are $\sigma_{xx}, \sigma_{yy}, \sigma_{zz}$ and calculated for each diatom model via patch surface area adjusted to diameter value for x-values and finding Cauchy traction vector ($\mathbf{T}^{(n)}$) scalar force values using the y-intercept and slope from the curve fit of Hamm et al.'s (2003) data.

Each 3D diatom valve surface will dictate the direction of forces on that surface; stresses normal to the surface are orthogonal. To calculate the tensor, numerical representation of the 3D diatom valve face surface is necessarily diagonalized and orthonormalized. The 3×3 asymmetric Jacobian matrix is diagonalized via a Gram matrix on an inner product space as $G(X) = \begin{bmatrix} X_1 \cdot X_1 & X_1 \cdot X_2 & X_1 \cdot X_3 \\ X_2 \cdot X_1 & X_2 \cdot X_2 & X_2 \cdot X_3 \\ X_3 \cdot X_1 & X_3 \cdot X_2 & X_3 \cdot X_3 \end{bmatrix}$ or $G(x_1, \ldots, x_3) = \begin{bmatrix} \langle x_1, x_1 \rangle & \langle x_1, x_2 \rangle & \langle x_1, x_3 \rangle \\ \langle x_2, x_1 \rangle & \langle x_2, x_2 \rangle & \langle x_2, x_3 \rangle \\ \langle x_3, x_1 \rangle & \langle x_3, x_2 \rangle & \langle x_3, x_3 \rangle \end{bmatrix}$. That is, the inner matrix

product of the Jacobian matrix is the Gram matrix. The Gramian determinant will have a non-singular value if linear independence holds true; that is, $|G| > 0$. Next, the Gram matrix is orthonormalized as $G(\frac{X}{\|X\|}) =$

$$
\begin{bmatrix}
\frac{X_1 \cdot X_1}{\|X_1 \cdot X_1\|} & \frac{X_1 \cdot X_2}{\|X_1 \cdot X_2\|} & \frac{X_1 \cdot X_3}{\|X_1 \cdot X_3\|} \\
\frac{X_2 \cdot X_1}{\|X_2 \cdot X_1\|} & \frac{X_2 \cdot X_2}{\|X_2 \cdot X_2\|} & \frac{X_2 \cdot X_3}{\|X_2 \cdot X_3\|} \\
\frac{X_3 \cdot X_1}{\|X_3 \cdot X_1\|} & \frac{X_3 \cdot X_2}{\|X_3 \cdot X_2\|} & \frac{X_3 \cdot X_3}{\|X_3 \cdot X_3\|}
\end{bmatrix}
$$
. The product of the scalar value of the Cauchy traction vector and the orthonormalized Gram matrix for each 3D diatom surface model will yield the Cauchy stress tensor matrix entries. The Cauchy stress tensor is calculated as $\boldsymbol{\sigma} = \mathbf{T}^{(n)} \cdot \mathbf{n}$ and acts as representation of diatom 3D surface functional complexity.

A.4. Multivariate complexity analysis

External unfolding is instituted using iterative or stress majorization (De Leeuw, 1988; Heiser, 1995; Bussing et al., 2005; Martins et al., 2009) in which stepwise minimization occurs between a majorizing function and original loss function with respect to transformed proximities of the original data points (De Leeuw, 1988; Heiser, 1995). This is termed the sandwich inequality (De Leeuw, 1977) for original loss function, $f(x)$, and majorizing function, $g(x)$, for k steps and is $f(x^{k+1}) \leq g(x^{k+1}) < g(x^k) = f(x^k)$. The majorizing function coincides with the original function at minimally one supporting point (Heiser, 1995; De Leeuw and Michailidis, 2000). At each step in the minimization process, the differences between the original and majorizing functions are updated and assessed until a local convergence is achieved usually at a linear or sometimes a quadratic rate (De Leeuw and Michailidis, 2000).

Degeneracy is a problem in distance-based multivariate techniques in which optimal distances are found that relate to trivial spatial solutions as results are independent of the original data (Kruskal and Carroll, 1969; Bussing, 2010). To avoid degeneracy in external unfolding analysis, particular loss functions based on STRESS (standardized residual sum-of-squares) (Kruskal, 1964), are instituted to assess badness-of-fit and goodness-of-fit. Penalties are necessarily imposed (Bussing, 2010), and a large penalty means that arbitrary or trivial STRESS values occur as almost equal distances between original points. STRESS functions used to deal with degeneracy assessment include Kruskal's STRESS-1 (raw STRESS—see Kruskal, 1964) and STRESS-2 (normalized STRESS—see Kruskal and Carroll, 1969) and Young's STRESS-1 and STRESS-2 (Young and Lewyckyj, 1979) based on squared distances. A penalty based on the coefficient of variation will occur when the transformations of the proximities have small variances in distances and result in an arbitrary spatial map (Bussing et al., 2005). A penalty on the intercept or slope is included in the transformation process when the transformations allow approximately equal distances or have a zero slope (Bussing, 2006).

The PREFSCAL algorithm (Bussing et al., 2005; Borg and Groenen, 2005; Bussing, 2006; Martins et al., 2009) uses a penalized STRESS loss function with implicit normalization (Bussing, 2010), p-STRESS, and the squared penalized STRESS (Bussing, 2010), expressed as $\sigma_p^2(\Gamma, \mathbf{X}, \mathbf{Y}) = \sigma_r^{2\lambda}(\Gamma, \mathbf{X}, \mathbf{Y})\mu(\Gamma)$ where $\sigma_r^2(\Gamma, \mathbf{X}, \mathbf{Y})$, is the raw STRESS, $\mu(\Gamma)$ is the penalty function, and λ is a penalty parameter in $0 < \lambda \leq 1$. As λ increases, $\sigma_r^2(\Gamma, \mathbf{X}, \mathbf{Y})$ influences the outcome less; i.e., as λ approaches zero, a stronger penalty is instituted. The result is a minimization of the penalized STRESS function (Borg and Groenen, 2005). The penalty function is $\mu(\Gamma) = 1 + \frac{\omega}{\nu^2(\Gamma)}$ where $\nu^2(\Gamma)$ is the coefficient of variation, and ω is a penalty parameter with values of $\omega \geq 0$. As ω increases, a concomitant increase in the coefficient of variation results in a decreasing penalty function and an increase in non-degeneracy (Borg and Groenen, 2005; Bussing, 2010).

A.5. Magnitude and rate of complexity change

To calculate the transitions between all taxa, the total complexity vector is used to calculate a tensor product. From averaged proximities of total complexity, $\mathbf{A}$, an outer product matrix is calculated as $\mathbf{W} = \mathbf{A} \otimes \mathbf{A}^T$ that represents the square paired product matrix of diatom taxa. Each row of $\mathbf{W}$ is then linearly transformed on the interval $[0, 1]$ and the matrix becomes $\mathbf{T}(a^>) = \begin{bmatrix} a_{1i} & \cdots & a_{1n} \\ \vdots & & \\ a_{mi} & \cdots & a_{mn} \end{bmatrix}$. Next, $\mathbf{T}(a^>)$ is normalized using the basis vector $\begin{bmatrix} e_i \\ \vdots \\ e_n \end{bmatrix}$ so that the resultant row elements in columns i to n sum to one for rows 1 to m. That is, the elements of each row of the matrix are divided by the sum of each transposed row's elements so that each row now represents a probability distribution of each taxon. This right stochastic matrix is the transition probability matrix, P, where each row is the taxon's conditional probability states with respect to every other taxon's state.

To satisfy the Markov property, probability P of a sequence of random variables (taxa), X, is of the form $P\{X_{n+1} = x_{n+1} | X_n = x_n\}$ and are conditional probabilities of stepwise transitions from one state to another in state space, S, where only the previous state affects the next succeeding state; the transitions in time occurring from n to $n + 1$ are $p_{ij}^{(n)} = P\{X_{n+1} = j | X_n = i\}$, and the probabilities are independent (Stewart, 1994). For a homogenous system, n-transition states are calculated as a transition probability matrix according to the recursive formula from Chapman–Kolmogorov given as $p_{ij}^{(n)} = \sum_{all\ k} p_{ik}^{(l)} p_{kj}^{(n-l)}$ where $0 < l < n$ for conditional probabilities $p_{ij}^{(n)}$ between genera, with $0 \leq p_{ij} \leq 1$ for row elements, i, and for all i, $\sum_j p_{ij} = 1$. The matrix entries are conditional probabilities of complexity.

Let x_n be the random variables of a homogeneous discrete time Markov chain in state space, S. We want a probability distribution of complexity states from the transition probability matrix, P, to represent equilibrium or steady-state, resulting in a stationary distribution as probability vector, π. This vector is obtained via convergence of a system of homogeneous difference equations calculated for each step via the iterative power method, with initial state distribution, $\pi(0)$, and given as $\pi_{ij}^{(n)} = \sum_{all\ i,j} P_{ij}^{(n)}(\{X_{n+1} = j | X_n = i\}) \pi_{ij}(0) = \sum_{all\ i} p_{ij}\pi_i$. For each state in the transition probability matrix, each π_j is calculated as the dot product of π with the jth column of P. The initial state distribution is independent of the limiting distribution, and the stationary distribution will converge to a limiting distribution if $\lim_{n \to \infty} \pi^n = \pi$ exists.

For an ergodic Markov chain, since $P^{(n)}$ and $\pi^{(n)}$ converge to the same distribution, $\lim_{n \to \infty} p_{ij}^{(n)} = \pi_j$ for all i. P is a positive square stochastic matrix; therefore, there is a unique largest positive eigenvalue associated with a positive eigenvector viz. π (Stewart, 1994). The stationary probability vector is obtained iteratively by a linear combination of the eigenvectors, e, with conditional probabilities from P. Because P is stochastic, $Pe = e$ and $e = (1, 1, \ldots, 1)^T$ with $\pi = \pi P = \lambda \pi$. For $(P' - I)\pi = 0$, the e of $\mathbf{e}$ that corresponds to the eigenvalue $\lambda = 1$ for the transpose of P is a left eigenvector, $\mathbf{e}_L$, that is π of the Markov chain (Stewart, 1994). For $\mathbf{e}_L P = \lambda_L \mathbf{e}_L$ and $P\mathbf{e}_R = \lambda_R \mathbf{e}_R$, $\lambda_L = \lambda_R$ but $\mathbf{e}_L \neq \mathbf{e}_R$. The left eigenvector associated with $\lambda = 1$ is the equilibrium of the system that represents steady-state complexity. The subdominant right eigenvectors associated with $|\lambda|\ \text{mod} < 1$ are interpretable as the number of transitions needed for a given level of complexity to reach equilibrium. For an aperiodic, irreducible Markov chain, if there is a value of $n \geq 1$ such that P^n has all positive entries, then the Perron–Frobenius theorem holds true, and π is unique (Stewart, 1994).

Magnitude of complexity change is measured as a distance metric between initial and final states of the Markov chain.

This metric is the Frobenius norm and is calculated as $\|Fr_{\mathbf{T}(a^>)_e - \mathbf{T}(b^>)_e}\| = \left[tr\left((\mathbf{T}(a^>)_e - \mathbf{T}(b^>)_e) * (\mathbf{T}(a^>)_e - \mathbf{T}(b^>)_e)'\right)\right]^{1/2}$, where $\mathbf{T}(a^>)_e$ and $\mathbf{T}(b^>)_e$ are the initial transition and stationary (limiting) probability matrices, respectively, and tr is the trace. In the short run, complexity change is measured using mean first passage time (M) as the mean number of transitions from one complexity state to another for states $j \neq k$ as $M_{jk} = 1 + \sum_{i \neq k} p_{ji} M_{ik}$; mean recurrence times are $M_{kk} = 1 + \sum_{i \neq k} p_{ki} M_{ik}$ (Stewart, 1994). Total complexity transition states are calculated as $TC_{trans} = \sum_{i \neq k}^{taxon_{final}} M_{jk} - M_{ij}$, where $j \neq k$ and $i = 1, ..., taxon_{final}$.

A.6. Complexity change over time: driven or random trend? Counterfactuals

The counterfactual conditional for each model, M, is given as $Y_x(t) \underline{\Delta} Y_{M_x}(t)$, where Y represents effects as evolutionary trends (ET) with "Driven" (Dr) and "Random" (Ran) as particular effects, x represents causes as particular kinds of complexity, t represents the particular time from the Eocene to the Oligocene. The probability vectors Dr and Ran for particular probabilities dr and ran, respectively as counterfactual conditionals $Prob(Dr_{ET} = dr)$ and $Prob(Ran_{ET} = ran)$ are $Prob[(ET = et) \square \rightarrow (Dr = dr)]$ and $Prob[(ET = et) \square \rightarrow (Ran = ran)]$ where $Dr_{ET}(t) \underline{\Delta} Dr_{Model_{ET}}(t)$ and $Ran_{ET}(t) \underline{\Delta} Ran_{Model_{ET}}(t)$.

A.6.1. Structural equations

A generalized functional set of equations representing the structural model are $x_i = f_i(pa_i, \varepsilon_i)$ where pa_i is the set of variables as causes of X_i, E_i are background exogenous variables (error variances) that are mutually independent, and $i = 1, ..., n$ (Pearl, 2000). For nine variables, functional measurement equations are $F_i = f_F(C, \xi_F) = \alpha_F + \Lambda_F \eta + \varphi$, $M_i = f_M(C, \xi_M) = \alpha_M + \Lambda_M \eta + \gamma$, $E_i = f_E(T, \xi_E) = \alpha_E + \Lambda_E \eta + \delta$, $O_i = f_O(T, \xi_O) = \alpha_O + \Lambda_O \eta + \zeta$, $Dr_i = f_{Dr}(ET, \xi_{Dr}) = \alpha_{Dr} + \Lambda_{Dr} \eta + \varsigma$, $Ran_i = f_{Ran}(ET, \xi_{Ran}) = \alpha_{Ran} + \Lambda_{Ran} \eta + \psi$, $C_i = f_C(\xi_C)$, $T_i = f_T(\xi_T)$, and $ET_i = f_{ET}(C_i, T_i, \xi_{ET})$, where F and M are vectors of observed endogenous complexity variables functional and morphological, respectively, E and O are vectors of categorical time variables Eocene and Oligocene, respectively, Dr and Ran are vectors of probability evolutionary trend values Driven and Random, respectively, α is the intercept, Λ is the matrix of the square roots of predictor variables given as squared multiple correlations, η is the vector of unobserved endogenous (indicator) variables, ξ are the error variances for C, T, and ET, and given as $\varphi, \gamma, \delta, \zeta, \varsigma, \psi$ for F, M, E, O, Dr, Ran, respectively, and $i = 1, ..., n$. Indicators of the vector of unobserved endogenous variables are expressed as the structural equation $\eta = \alpha_\eta + B_\eta + \Gamma_\xi + \vartheta$ where B is the matrix of interactions among the unobserved endogenous variables on each other, Γ is the matrix of coefficients of error variances ξ with respect to observed endogenous variables impact on η, and ϑ are the error variances with respect to unobserved endogenous variables.

References

Adami, C., 2002. What is complexity? BioEssays 24, 1085–1094.
Adami, C., Ofria, C., Collier, T., 2000. Evolution of biological complexity. Proc. Natl. Acad. Sci. 97, 4463–4468.
Alfaro, M.E., Bolnick, D.I., Wainwright, P.C., 2004. Evolutionary dynamics of complex biomechanical systems: an example using the four-bar mechanism. Evolution 58, 495–503.
Armbrust, E.V., 2009. The life of diatoms in the world's oceans. Nature 459, 185–192.
Arney, J.E., McGonigal, K.L., Ladner, B.C., Wise Jr., S.W., 2003. Lower Oligocene to middle Miocene diatom biostratigraphy of ODP site 1140, Kerguelen plateau. In: Frey, F.A., Coffin, M.F., Wallace, P.J., Quilty, P.G. (Eds.), Proceedings of the Ocean Drilling Program, Scientific Results 183. Ocean Drilling Program, College Station, Texas, pp. 1–21.

Auerbach, J.F., Bongard, J.C., 2014. Environmental influence on the evolution of morphological complexity in machines. PLoS Comput. Biol. 10 (1), e1003399. http://dx.doi.org/10.1371/journal.pcbi.1003399.
Baldauf, J.G., Barron, J.A., 1987. Oligocene marine diatoms recovered in dredge samples from the navarin basin province, Bering sea. U.S. Geological Survey Bulletin 1765. U.S. Geological Survey, Department of the Interior, United States Government Printing Office, Washington, D.C.
Baldauf, J.G., Barron, J.A., 1990. Evolution of biosiliceous sedimentation patterns — Eocene through quaternary: paleoceanographic response to polar cooling. In: Bleil, U., Thiede, J. (Eds.), Geological History of the Polar Oceans: Arctic Versus Antarctic. Kluwer Academic Publishers, The Netherlands, pp. 575–607.
Baldauf, J.G., Monjanel, A.-L., 1989. 20. An Oligocene diatom biostratigraphy for the Labrador sea: DSDP site 112 and ODP hole 647A. In: Srivastava, S.P., Arthur, M.A., Clement, B., Aksu, A., Baldauf, J., Bohrmann, G., Busch, W., Cederberg, T., Cremer, M., Dadey, K., De Vernal, A., Firth, J., Hall, F., Head, M., Hiscott, R., Jarrard, R., Kaminski, M., Lazarus, D., Monjanel, A.-L., Nielsen, O.B., Stein, R., Thiebault, F., Zachos, J., Zimmerman, H. (Eds.), Proceedings of the Ocean Drilling Program, Scientific Results 105. Ocean Drilling Program, College Station, Texas, pp. 155–170.
Barber, H.G., Haworth, E.Y., 1981. A guide to the morphology of the diatom frustule. Freshwater Biology Association Science Publication 44. The Ferry House, Ambleside, Cumbria, UK (112 pp.).
Barron, J.A., 1982/83. Latest Oligocene through early middle Miocene diatom biostratigraphy of the eastern tropical Pacific. Mar. Micropaleontol. 7, 487–515.
Barron, J.A., 1985. 16. Miocene to Holocene planktic diatoms. In: Bolli, H.M., Saunders, J.B., Perch-Nielsen, K. (Eds.), Plankton Stratigraphy. Cambridge University Press, Cambridge, UK, pp. 763–809.
Barron, J.A., Baldauf, J.G., 1989. Tertiary cooling steps and paleoproductivity as reflected by diatoms and biosiliceous sediments. In: Berger, W.H., Smetacek, V.S., Wefer, G. (Eds.), Productivity of the Ocean: Present and Past, Dahlem Workshop Reports. John Wiley & Sons Ltd., New York, pp. 341–354.
Barron, J.A., Baldauf, J.G., 1995. Cenozoic marine diatom biostratigraphy and applications to paleoclimatology and paleoceanography. In: Blome, C.D., Whalen, P.M., Reed, K.M. (Convenors), Siliceous Microfossils, Paleontological Society Short Courses in Paleontology, The Paleontological Society Papers, No. 8. University of Tennessee, Knoxville, TN, pp. 107–118.
Barron, J.A., Mahood, A.D., 1993. Exceptionally Well-Preserved Early Oligocene Diatoms from Glacial Sediments of Prydz Bay, East Antarctica, USGS Staff — Published Research, Paper 266. U.S. Geological Survey DigitalCommons@University of Nebraska — Lincoln. (Available at: http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1266&context=usgsstaffpub)
Barron, J.A., Fourtanier, E., Bohaty, S.M., 2004. Oligocene and earliest Miocene diatom biostratigraphy of ODP leg 199 site 1220, equatorial Pacific. In: Wilson, P.A., Lyle, M., Janecek, T.R., Firth, J.V. (Eds.), Proceedings of the Ocean Drilling Program, Scientific Results 199. Ocean Drilling Program, College Station, Texas, pp. 1–25.
Barron, J.A., Stickley, C.E., Bukry, D., 2015. Paleoceanographic, and paleoclimatic constraints on the global Eocene diatom and silicoflagellate record. Palaeogeogr. Palaeoclimatol. Palaeoecol. 422, 85–100.
Bohaty, S.M., Kulhanek, D.K., Wise Jr., S.W., Jemison, K., Warny, S., Sjunneskog, C., 2011. Age assessment of Eocene–Pliocene drill cores recovered during the SHALDRIL II expedition, Antarctic peninsula. In: Anderson, J.B., Wellner, J.S. (Eds.), Tectonic, Climatic, and Cryospheric Evolution of the Antarctic Peninsula, Special Publication 063. American Geophysical Union, Washington, D.C., pp. 63–113.
Borg, I., Groenen, P.J.F., 2005. Modern Multidimensional Scaling: Theory and Applications. second ed. Springer, New York.
Bowen, N.K., Guo, S., 2012. Structural Equation Modeling. Oxford University Press, Oxford, UK (217 pp.).
Bussing, F.M.T.A., 2006. Avoiding degeneracy in metric unfolding by penalizing the intercept. Br. J. Math. Stat. Psychol. 59, 419–427.
Bussing, F.M.T.A., 2010. Advances in Multidimensional Unfolding. Gildeprint Drukkerijen, Enschede, The Netherlands (276 pp.).
Bussing, F.M.T.A., Groenen, P.J.K., Heiser, W.J., 2005. Avoiding degeneracy in multidimensional unfolding by penalizing on the coefficient of variation. Psychometrika 70, 71–98.
Conley, D.J., 2002. Terrestrial ecosystems and the global biogeochemical silica cycle. Glob. Biogeochem. Cycles 16, 1121, doi:http://dx.doi.org/10.1029/2002GB001894.
Coombs, C.H., 1964. A Theory of Data. Wiley, New York.
Darwin, C., 1859. On the Origin of Species by Means of Natural Selection. Murray, London.
Dawid, A., 1979. Conditional independence in statistical theory. J. R. Stat. Soc. Ser. B 41, 1–31.
De Leeuw, J., 1977. Applications of convex analysis to multidimensional scaling. In: Barra, J.R., Brodeau, F., Romier, G., Van Cutsem, B. (Eds.), Recent Developments in Statistics. North Holland Publishing Company, Amsterdam, The Netherlands, pp. 133–146.
De Leeuw, J., 1988. Convergence of the majorization method for multidimensional scaling. J. Classif. 5, 163–180.
De Leeuw, J., Michailidis, G., 2000. Majorization methods in statistics. J. Comput. Graph. Stat. 9, 26–31.
Falkowski, P.G., Katz, M.E., Knoll, A.H., Quigg, A., Raven, J.A., Schofield, O., Taylor, F.J.R., 2004. The evolution of modern eukaryotic phytoplankton. Science 305, 354–360.
Fenner, J., 1984a. 49. Middle Eocene to Oligocene planktonic diatom stratigraphy from the deep sea drilling sites in the south Atlantic, equatorial Pacific, and Indian oceans. In: Hay, W.W., Sibuet, J.-C., Barron, E.J., Brassel, S.C., Dean, W.E., Huc, A.Y., Keating, B.H., McNulty, C.L., Meyers, P.A., Nohara, M., Schallreuter, R.E.L., Steinmetz, J.C., Stow, D.A.V., Stradner, H., Boyce, R.E., Amidei, R. (Eds.), Initial Reports of the Deep Sea Drilling Project 75. United States Government Printing Office, Washington, D.C., pp. 1245–1271.

Fenner, J., 1984b. Eocene–Oligocene planktic diatom stratigraphy in the low latitudes and the high southern latitudes. Micropaleontology 30, 319–342.
Fenner, J., 1985. Late Cretaceous to Oligocene planktic diatoms. In: Bolli, H.M., Saunders, J., Perch-Nielsen, K. (Eds.), Plankton Stratigraphy. Cambridge University Press, Cambridge, UK, pp. 713–762.
Fenner, J., Mikkelsen, N., 1990. Eocene–Oligocene diatoms in the western Indian Ocean: taxonomy, stratigraphy, and paleoecology. In: Duncan, R.A., Backman, J., Peterson, L.C., Baker, P.A., Baxter, A.N., Boersma, A., Cullen, J.L., Droxler, A.W., Fisk, M.R., Greenough, J.D., Hargraves, R.B., Hempel, P., Hobart, M.A., Hurley, M.T., Johnson, D.A., Macdonald, A.H., Mikkelsen, N., Okada, H., Rio, D., Robinson, S.G., Schneider, D.A., Swart, P.K., Tatsumi, Y., Vandamme, D., Vilks, G., Vincent, E., Barbu, E.M. (Eds.), Proceedings of the Ocean Drilling Program, Scientific Results 115. Ocean Drilling Program, College Station, Texas, pp. 433–463.
Fourtanier, E., 1991. 7. Diatom biostratigraphy of equatorial Indian Ocean site 758. In: Peirce, J.W., Weissel, J.K., Taylor, E., Dehn, J., Driscoll, M.W., Farrell, J.W., Fourtanier, E., Frey, F.A., Gamson, P.D., Gee, J.S., Gibson, I.L., Janecek, T.R., Klootwijk, C.T., Lawrence, J.R., Littke, R., Newman, J.S., Nomura, R., Oven, R.M., Pospichal, J.J., Rea, D.K., Resiwati, P., Saunders, A.D., Smit, J., Smith, G.M., Tamaki, D., Weis, D., Wilkinson, C., Dearmont, L.H., Mazzullo, E.K., Stewart, N.J. (Eds.), Proceedings of the Ocean Drilling Program, Scientific Results 121. Ocean Drilling Program, College Station, Texas, pp. 189–208.
Gladenkov, A.Y., 2012. Middle Eocene diatoms from the marine Paleogene stratigraphic key section of northeast Kamchatka. Aust. J. Earth Sci. 105 (1), 72–76.
Gladenkov, A.Y., 2014. Bipolar distribution of some earliest Oligocene marine diatoms. In: Kociolek, J.P., Kulikovskiy, M.S., Witkowski, J., Harwood, D.M. (Eds.), Diatom Research over Time and Space: Morphology, Taxonomy, Ecology and Distribution of Diatoms —from Fossil to Recent, Marine to Freshwater, Established Species and Genera to new Ones; Celebrating the Work and Impact of Nina Strelnikova on the Occasion of her 80th Birthday. Nova Hedwigia Beiheft 143, pp. 337–368.
Gombos, A.M., 1983a. Middle Eocene diatoms from the south Atlantic. In: Ludwig, W.J., Krasheninnikov, V.A., Basov, I.A., Bayer, U., Bloemendal, J., Bornhold, B., Ciesielski, P.F., Goldstein, E.H., Rovert, C., Salloway, J.C., Usher, J.L., vonder Dick, H., Weaver, F.M., Wise Jr., S.W., Usher, J.L., Blakeslee, J.H., Lee, M. (Eds.), Initial Reports of the Deep Sea Drilling Project 71. United States Government Printing Office, Washington, D.C., pp. 565–581.
Gombos, A.M., 1983b. Survey of diatoms in the upper Oligocene and lower Miocene in holes 515B and 516F. In: Barker, P.F., Johnson, D.A., Carlson, R.L., Cepek, P., Coulbourn, W.T., Gamboa, L.A., Hamilton, N., de Melo, U., Pujol, C., Shor, A.N., Suzyumov, A.E., Tjalsma, L.R.C., Walton, W.H., Whalen, E. (Eds.), Initial Reports of the Deep Sea Drilling Project 72. United States Government Printing Office, Washington, D.C., pp. 793–804.
Gombos, A.M., 1987. Middle Eocene diatoms from the north Atlantic, deep sea drilling project site 605. In: Van Hinte, J.E., Wise Jr., S.W., Biart, B.N.M., Covington, J.M., Dunn, D.A., Haggerty, J.A., Johns, M.W., Meyers, P.A., Moullade, M.R., Muza, J.P., Ogg, J.G., Okamura, M., Sarti, M., von Rad, U., Blakeslee, J.H., Whalen, E. (Eds.), Initial Reports of the Deep Sea Drilling Project 93. United States Government Printing Office, Washington, D.C., pp. 793–799.
Gombos, A.M., Ciesielski, P.F., 1983. Late Eocene to early Miocene diatoms from the southwest Atlantic. In: Ludwig, W.J., Krasheninnikov, V.A., Basov, I.A., Bayer, U., Bloemendal, J., Bornhold, B., Ciesielski, P.F., Goldstein, E.H., Rovert, C., Salloway, J.C., Usher, J.L., von der Dick, H., Weaver, F.M., Wise Jr., S.W., Usher, J.L., Blakeslee, J.H., Lee, M. (Eds.), Initial Reports of the Deep Sea Drilling Project 71. United States Government Printing Office, Washington, D.C., pp. 583–634.
Hajós, M., 1976. IV. Upper Eocene and lower Oligocene Diatomaceae, Archaeomonadaceae, and Silicoflagellatae in southwestern Pacific sediments, DSDP leg 29. In: Hollister, C.D., Craddock, C. (Eds.), Initial Reports of the Deep Sea Drilling Project 35. United States Government Printing Office, Washington, D.C., pp. 817–883.
Hamm, C.E., 2005. The evolution of advanced mechanical defenses and potential technological applications of diatom shells. J. Nanosci. Nanotechnol. 5, 108–119.
Hamm, C.E., 2007. Interactions between biomineralization and function of diatom frustules. In: Bäuerlein, E. (Ed.), Handbook of Biomineralization: Biological Aspects and Structure Formation. Wiley-VCH Verlag GmbH, Weinheim, Germany, pp. 83–94.
Hamm, C.E., Merkel, R., Springer, O., Jurkojc, P., Maier, C., Prechtel, K., Smetacek, V., 2003. Architecture and material properties of diatom shells provide effective mechanical protection. Nature 421, 841–843.
Harwood, D.M., Maruyama, T., 1992. 38Middle Eocene to Pleistocene diatom biostratigraphy of southern oceans sediments from the Kerguelen plateau, leg 120. In: Wise Jr., S.W., Schlich, R., Palmer Julson, A.A., Aubry, M.-P., Bergren, W.A., Bitschene, P.R., Blackburn, N.A., Breza, J.R., Coffin, M.F., Harwood, D.M., Heider, F., Holmes, M.A., Howard, W.R., Inokuchi, H., Kelts, K., Lazarus, D.B., Mackensen, A., Maruyama, T., Munschy, M., Pratson, E.L., Quilty, P.G., Rack, F.R., Salters, V.J.M., Sevigny, J.H., Storey, M., Takemura, A., Watkins, D.K., Whitechurch, H., Zachos, J., Barbu, E.M. (Eds.), Proceedings of the Ocean Drilling Program, Scientific Results 120. Ocean Drilling Program, College Station, Texas, pp. 683–733.
Harwood, D.M., Nikolaev, V.A., 1995. Cretaceous diatoms: morphology, taxonomy, biostratigraphy. In: Blome, C.D., Whalen, P.M., Reed, K.M., (Convenors), Siliceous Microfossils, Paleontological Society Short Courses in Paleontology, The Paleontological Society Papers, No. 8. University of Tennessee, Knoxville, Tennessee, pp. 81–106.
Harwood, D.M., Nikolaev, V.A., Winter, D.M., 2007. Cretaceous records of diatom evolution, radiation, and expansion. In: Starratt, S.W. (Ed.), From Pond Scum to Carbon Sink: Geological and Environmental Applications of the Diatoms, Paleontological Society Short Course, The Paleontological Society Papers, No. 13. Paleontological Society, Knoxville, Tennessee, pp. 33–59.
Heinbockel, J.H., 2001. Introduction to Tensor Calculus and Continuum Mechanics. Trafford, Victoria, B.C., Canada.

Heiser, W., 1987. Joint ordination of species and sites: the unfolding technique. In: Legendre, P., Legendre, L. (Eds.), Developments in Numerical Ecology. Springer-Verlag, Berlin, Heidelberg, Germany, pp. 189–221.
Heiser, W., 1995. Convergent computation by iterative majorization: theory and applications in multidimensional data analysis. In: Krzanowski, W.J. (Ed.), Recent Advances in Descriptive Multivariate Analysis. Oxford University Press, Oxford, UK, pp. 157–189.
Hopkinson, B.M., Dupont, C.L., Allen, A.E., Morel, F.M.M., 2011. Efficiency of the CO2-concentrating mechanism of diatoms. Proc. Natl. Acad. Sci. U. S. A. 108, 3830–3837.
Jordan, R.W., Stickley, C.E., 2010. Diatoms as indicators of paleoceanographic events. In: Smol, J.P., Stoermer, E.F. (Eds.), The Diatoms: Applications for the Environmental and Earth Sciences, second ed. Cambridge University Press, Cambridge, UK, pp. 424–453.
Katz, M.E., Finkel, Z.V., Grzebyk, D., Knoll, A.H., Falkowski, P.G., 2004. Evolutionary trajectories and biogeochemical impacts of marine eukaryotic phytoplankton. Annu. Rev. Ecol. Evol. Syst. 35, 523–556.
Knoll, A.H., 2003. Biomineralization and evolutionary history. Rev. Mineral. Geochem. 54, 329–356.
Knoll, A.H., Bambach, R.K., 2000. Directionality in the history of life: diffusion from the left wall or repeated scaling of the right? Paleobiology (Supplement) 26, 1–14.
Kröger, N., Poulsen, N., 2008. Diatoms—from cell wall biogenesis to nanotechnology. Annu. Rev. Genet. 42, 83–107.
Kruskal, J.B., 1964. Nonmetric multidimensional scaling: a numerical method. Psychometrika 29, 115–129.
Kruskal, J.B., Carroll, J.D., 1969. Geometrical models and badness-of-fit functions. In: Krishnaiah, P.R. (Ed.)Multivariate Analysis vol. 2. Academic Press, New York, pp. 639–671.
Lamarck, J.B.P.A.M., 1809. Zoological Philosophy. University of Chicago Press, Chicago, Illinois.
Lowe, R.L., 2011. The importance of scale in understanding the natural history of diatom communities. In: Seckbach, J., Kociolek, J.P. (Eds.), The Diatom World, Cellular Origin, Life in Extreme Habitats and Astrobiology 19. Springer Science + Business Media B.V, New York, pp. 293–311.
Martins, A.A.A.F., Cardoso, M.G.M.S., Pinto, I.M.S., 2009. Multidimensional unfolding and the use of alternative input scales. J. Target. Meas. Anal. Mark. 17, 195–204.
McNamara, K.J., 2006. Evolutionary trends. Encyclopedia of Life Sciences (eLS). John Wiley & Sons, Ltd., pp. 1–7 http://dx.doi.org/10.1038/npg.els.0004136.
McShea, D.W., 1991. Complexity and evolution: what everybody knows. Biol. Philos. 6, 303–324.
McShea, D.W., 1992. A metric for the study of evolutionary trends in the complexity of serial structures. Biol. J. Linn. Soc. 45, 39–55.
McShea, D.W., Brandon, R.M., 2010. Biology's First Law: The Tendency for Diversity and Complexity to Increase in Evolutionary Systems. The University of Chicago Press, Chicago and London (170 pp.).
McShea, D.W., Hordijk, W., 2013. Complexity by subtraction. Evol. Biol. 40, 504–520.
McShea, D.W., Venit, E.P., 2001. What is a part? In: Wagner, G.P. (Ed.), The Character Concept in Evolutionary Biology. Academic Press, Elsevier, Waltham, Massachusetts, pp. 259–283
Michels, J., Vogt, J., Gorb, S.N., 2012. Tools for crushing diatoms—opal teeth in copepods feature a rubber-like bearing composed of resilin. Sci. Rep. 2, 465–470.
Miconi, T., 2008. Evolution and complexity: the double-edged sword. Artif. Life 14, 325–344.
Neyman, J., 1923. On the application of probability theory to agricultural experiments. Essay on Principles. Section 9. Roczniki Nauk Rolniczych Tom X, pp. 465–480 ([in Polish]; translated in Statistical Science 5).
Oreshkina, T.V., 2012. Evidence of late Paleocene — early Eocene hyperthermal events in biosiliceous sediments of western Siberia and adjacent areas. Aust. J. Earth Sci. 105, 145–153.
Pappas, J.L., 2005a. Geometry and topology of diatom shape and surface morphogenesis for use in applications of nanotechnology. J. Nanosci. Nanotechnol. 5, 120–130.
Pappas, J.L., 2005b. Theoretical morphospace and its relation to gomphonemoid-cymbelloid diatom (Bacillariophyta) lineages. J. Biol. Syst. 13, 385–398.
Pappas, J.L., 2008. More on theoretical morphospace and its relation to freshwater gomphonemoid-cymbelloid diatom (Bacillariophyta) lineages. J. Biol. Syst. 16, 119–137.
Pappas, J.L., Miller, D.J., 2013. A generalized approach to the modeling and analysis of 3D surface morphology in organisms. PLoS One 8 (10), e77551.
Pearl, J., 2000. Causality: Models, Reasoning, and Inference. Cambridge University Press, Cambridge, UK (384 pp.).
Pearl, J., 2009. Causal inference in statistics: an overview. Stat. Surv. 3, 96–146.
Round, F.E., Crawford, R.M., Mann, D.G., 1990. The Diatoms, Biology & Morphology of the Genera. Cambridge University Press, Cambridge, UK (747 pp.).
Rubin, D.B., 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. J. Educ. Psychol. 66, 688–701.
Saito, T., Barron, J.A., Sakamoto, M., 1988. 68. An early late Oligocene age indicated by diatoms for a primitive desmostylian mammal Behemotops from eastern Hokkaido, Japan. Proc. Jpn. Acad. Ser. B 64, 269–273.
Scherer, R., Bohaty, S.M., Harwood, D.M., 2000. Oligocene and lower Miocene siliceous microfossil biostratigraphy of Cape Roberts project core CRP-2/2A, Victoria land basin, Antarctica. Papers in the Earth and Atmospheric Sciences, Paper 282 (Available at: http://digitalcommons.unl.edu/geosciencefacpub/282).
Scherer, R., Gladenkov, A., Barron, J., 2007. Methods and applications of Cenozoic marine diatom biostratigraphy. In: Starratt, S.W. (Ed.), From Pond Scum to Carbon Sink: Geological and Environmental Applications of the Diatoms, Paleontological Society Short Course, The Paleontological Society Papers, No. 13. Paleontological Society, Knoxville, Tennessee, pp. 61–83.
Smetacek, V., 2001. A watery arms race. Nature 411, 745.
Smith, T.M.F., 1983. On the validity of inferences from non-random samples. J. R. Soc. Ser. A (Gen.) 146, 394–403.

Stewart, W.J., 1994. Introduction to the Numerical Solution of Markov Chains. Princeton University Press, Princeton, New Jersey (539 pp.).

Strelnikova, N.I., 1990. Evolution of diatoms during the Cretaceous and Paleogene periods. In: Simola, H. (Ed.), Proceedings of the Tenth International Diatom Symposium. Koeltz Scientific Books, Koenigstein, Germany, pp. 195–204.

Tall, L., Cloutier, L., Cattaneo, A., 2006a. Grazer-diatom size relationships in an epiphytic community. Limnol. Oceanogr. 51, 1211–1216.

Tall, L., Cattaneo, A., Cloutier, L., Dray, S., Legendre, P., 2006b. Resource partitioning in a grazer guild feeding on a multilayer diatom mat. J. N. Am. Benthol. Soc. 25, 800–810.

Taylor, A.E., Mann, W.R., 1983. Advanced Calculus. third ed. John Wiley & Sons Inc., New York (732 pp.).

Tesson, B., Hildebrand, M., 2013. Characterization and localization of insoluble organic matrices associated with diatom cell walls: insight into their roles during cell wall formation. PLoS One 8 (4), e61675.

Theriot, E.C., Ashworth, M., Ruck, E., Nakov, T., Jansen, R.K., 2010. A preliminary multigene phylogeny of the diatoms (Bacillariophyta): challenges for future research. Plant Ecol. Evol. 143, 278–296.

Theriot, E.C., Ruck, E., Ashworth, M., Nakov, T., Jansen, R.K., 2011. Status of the pursuit of the diatom phylogeny: are traditional views and new molecular paradigms really that different? In: Seckbach, J., Kociolek, J.P. (Eds.), The Diatom World, Cellular Origin, Life in Extreme Habitats and Astrobiology 19. Springer Science + Business Media B.V, New York, pp. 119–142

Valentine, J.W., 2000. Two genomic paths to the evolution of complexity in bodyplans. Paleobiology 26, 513–519.

Wainwright, P.C., 2007. Morphological diversity in macroevolution. Annu. Rev. Ecol. Evol. Syst. 38, 381–401.

Wicken, J.S., 1979. The generation of complexity in evolution: a thermodynamic and information-theoretical discussion. J. Theor. Biol. 77, 349–365.

Williams, D.M., Kociolek, J.P., 2007. Pursuit of a natural classification of diatoms: history, monophyly and the rejection of paraphyletic taxa. Eur. J. Phycol. 42, 313–319.

Witkowski, J., Harwood, D.M., 2011. *Blochia* and *Nikolaevia*: new Cretaceous diatom genera related to *Stellarima* Hasle & Sims. Diatom Res. 25, 445–458.

Witkowski, J., Bohaty, S.M., Edgar, K.M., Harwood, D.M., 2014. Rapid fluctuations in mid-latitude siliceous plankton production during the middle Eocene climatic optimum (ODP site 1051, western north Atlantic). Mar. Micropaleontol. 106, 110–129.

Wright, S., 1921. Correlation and causation. J. Agric. Res. 20, 557–585.

Yaeger, L., Griffith, V., Sporns, O., 2008. Passive and Driven Trends in the Evolution of Complexity (Available at: arXiv:1112.4906 [cs.NE]).

Young, F.W., Lewyckyj, R., 1979. ALSCAL-4 Users Guide. second ed. Data Analysis and Theory Associates, Carrboro, North Carolina.

Zachos, J.C., Dickens, G.R., Zeebe, R.E., 2008. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. Nature 451, 279–283.