

Evolution of Shape in Diatoms

Robert Bentham

Supervised by Dr Karen Page and Dr Eileen Cox

Abstract

The evolution of shape in centric diatoms is investigated by probabilistic means on phylogenetic data using the Bayes Traits software, developed by Mark Pagel. Preliminary results of this method are presented showing transitional rates between one trait group to another, the results suggest that diatom valve outlines with higher degrees of symmetry are evolutionary more unstable. Weaknesses of the method used, how it can be improved and other future possible approaches are then discussed.

§1 Introduction

Diatoms are phytoplankton that exist in most liquid environments across planet, composing a vital part of the world's ecosystem by being responsible for one-fifth of all the photosynthesis carried out[?]. Diatoms belong to the eukaryota domain and compose a major group of algae. Besides diatoms ecological importance they have a distinct and unique morphological characteristics in the form of a silica cell wall. This cell wall is developed by a process of biomineralization¹ and varies enormously across the genera of diatoms. These silica walls² have been primary researched to aid our understanding of biological morphogenesis and their possible applications in nanotechnology[?]. This work however will look at the less studied aspect of the evolution of shape of these diatoms.

Evolutionary factors of shape in diatoms

There are roughly 100,000 known extant species of diatoms, of which there is great diversity in their genetic makeup. For instance [?] describes the divergence between centric and pennate diatoms as not too dissimilar to the divergence between mammals and fish. The development of the raphe³ in some species of diatom allows a gliding locomotion enabling diatoms to exist in new environments and has been compared with the evolution of flight in birds[?].

Just what factors are important in determining the evolution of the form of the frustule of the diatom over time? Part of the great genetic variety

between the diatoms is due to bacterial horizontal gene transfer that has taken place repeatedly over time. In fact there is a great tangled nature of interactions between bacteria and diatoms, co-existing as they do in liquid environments. In [?] the authors describe how bacteria can embed themselves into a silica cell wall of the diatom species *Pseudonitzschia multiseries*. Could co-evolution of bacteria and diatom be an important factor of the evolution of diatom form?

Other factors are primary associated with advantages they entail in the lifecycle of the diatom, such as controlling buoyancy, as well as how novel new morphological forms such as the presence of the raphe create a new selective advantage. It should be noted that while genetics are the main factors in determining diatom shape, environmental factors such as the salinity density have an important effect[?].

Project aims

To study the evolution of shape in diatoms some kind of mathematical model must be formed. In Rene Thom's seminal book *Structural Stability and Morphogenesis* [?] he discusses formal models of this nature dividing them into two parts, the kinematic and the dynamic. The kinematic part aims to parameterize the forms or states of the process that is being considered, the factors involved in diatom shape evolution for our purposes. The dynamic part is concerned with describing the evolution in time of the forms or states of this process, giving the transition probabilities between two states given by two different sets of parameters.

As described the factors involved in the shape evolution are very complex. In this work, the focus will be on only the transitional probabilities between

¹The incorporation of inorganic compounds into a biological system.

²Known as frustules.

³A narrow slit in the silica cell wall of the diatom

forms, since they can be investigated using statistical means. A complete model of the evolution of shape is at present difficult to make due to lack of knowledge of the factors controlling the process, how the diatom genetic information controls shape formation is almost entirely unknown. The aim in this work is to use phylogenetic data to calculate what hopefully will be a good approximation to these transitional probabilities.

Such transitional probabilities while far from a full model, will give insight into how the evolution of shape in diatoms has occurred. Based on statistics from real data sets they will also be able to judge the veracity of any future models. However the calculations of these transitional probabilities remain unstudied and seem to be a non-trivial problem depending on the classification method for the shapes, creation of accurate phylogenies and a statistical technique to do the calculation of the probabilities. This work will make a first attempt to calculate these probabilities, and detail methods which with more time this calculation could be improved, as well as other possible approaches that can be used on this problem.

§2 Morphology of diatoms

The exact form of diatoms must be discussed, the illustration by Ernst Haeckel in Figure ?? give some indication of the variety of shape across the genera of Diatoms. The book *The Diatoms - Biology & Morphology of The genera by F.E. Round, R.M. Crawford and D.G. Mann* [?] gives a detailed account of the shape of each genera of the diatoms and was used extensively in this work.

Key components of diatoms

As the outward appearance of the diatom is the chief interest of this work, the focus will be exclusively on the components of the silica wall and ignore all the components of the interior, the protoplast. The silica wall collectively known as the frustule is made up of multiple parts, two large components called valves along with several thinner structures collectively known as the girdle.

The valves can be separated into the epivalve and the hypovalve. During cell division the valves separate, and a new valve is formed to accompany each. Thus in each diatom the older valve is known as the

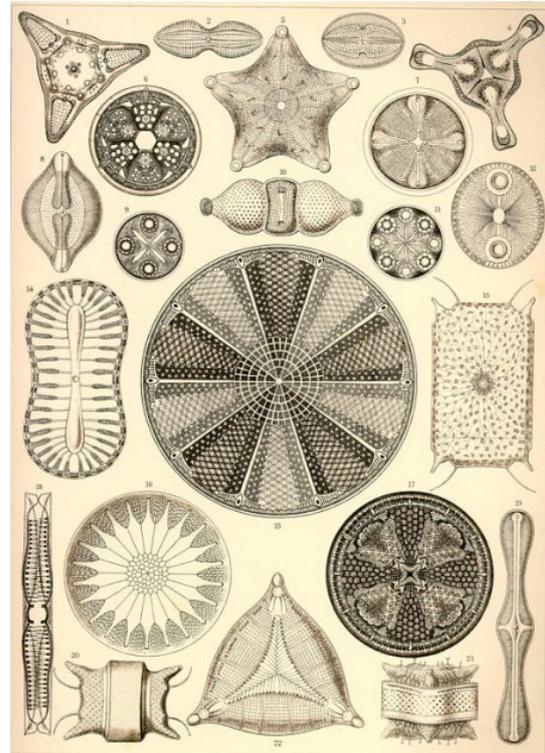


Figure 1: Ernst Haeckel, an eminent German biologist, produced these illustrations of diatoms for his 1904 book *Kunstformen der Natur* (*Art Forms of Nature*) [?]. The selection shows some of the wide variety of shapes of diatoms most likely selected by Haeckel due to their pleasing geometric properties.

epivalve while the newer valve is called the hypovalve.

There are two main ways of viewing a diatom, in valve view or in girdle view. Typically valve view is more useful in distinguishing features so the valve will be the view used to classify the diatoms in this work. The primary structure of the valve is a system of silica ribs, in fact in some genera (eg. *Papilocellulus*) this rib structure is all that makes up the silica cell wall. There are two main classes of diatoms that are distinguished by their rib structure, centric diatoms have ribs radiating from a ring while pennate diatoms have ribs extending from a longitudinal element that may also have a slit making it a raphe pennate. The evolution of shape only in centric diatoms will be investigated here for simplicity.

Even with this exclusions of all araphic and raphic pennates diatoms only studying the valves of the centrics, their exact structure can still be very complicated and difficult to formulate. Besides the

structures of the ribs, which are in no way uniform across centric diatoms, on the surface of valves there are many pores. These are often referred to as areolae and can form intricate patterns. There can be structures called occuli, areas where areolae are smaller and surrounded by a discrete rim of silica. Additionally there can also be various elevations on the valve, and even structures like horns.

The shape of a diatom forms during auxospore development, the auxospore being the form of the diatom during sexual reproduction where the cell is roughly spherical and lacks a silica cell wall. During this development a complex arrangement of silica bands and hoops form on the auxospore, this limits its expansion and causes the distinctive shape of many diatoms. The precise workings of this process are however largely unknown making it difficult to model.

Valve classification

With all these features on a centric valve, how exactly do you classify them into simple groups? The simplest method is to initially ignore all of the more complicated features of diatom shape and focus solely on valve outline. This is the method which will be used here, though other possible techniques will be discussed in Section 4.

The following classification method proposed here is based on the symmetrical properties of the diatoms:

1. Contains all diatoms whose valve outlines are circles.
2. The bipolar diatoms, more specifically diatoms with valve outline rotational symmetry of order 2, specifically these can be ellipses, sub-circular or linear.
3. Diatoms with higher degrees of symmetry of valve outline, includes triangular and square shaped diatoms.⁴

§3 Probabilistic Rules

Now a framework will be provided showing how the probabilistic rules of evolution of diatom shape can be investigated. The task requires a suitable

⁴Diatoms with higher degrees of symmetry have been grouped together due to their low occurrence in the phylogenetic data, and also to increase the simplicity of the groupings

phylogeny of the centric diatoms to be used and a statistical technique for calculating the probabilistic rules from such a phylogeny. Ideally a phylogeny would be made specifically for this task but due to an absence of time, phylogenies from the literature will be used. This has the drawback of reducing the reliability of my results but the importance of this work is to establish a framework of what can be done in the future.

The statistical analysis will be done using a piece of software called Bayes Traits [?](software can be found at www.evolution.rdg.ac.uk) developed by Mark Pagel.

Phylogenies

The first phylogeny that will be used is the section dealing with the centrics presented in [?] where full details of how the phylogeny was constructed using the PATHd8 program [?] can be found. The data for the phylogeny was kindly made available by Dr Sorhannus in chronogram form, and can be seen in Phylogeny A.

The second phylogeny used is presented in [?] where full details of how the phylogeny was constructed are given. Kindly Dr Theriot made the data available to me in phylogram form, shown in Phylogeny B.

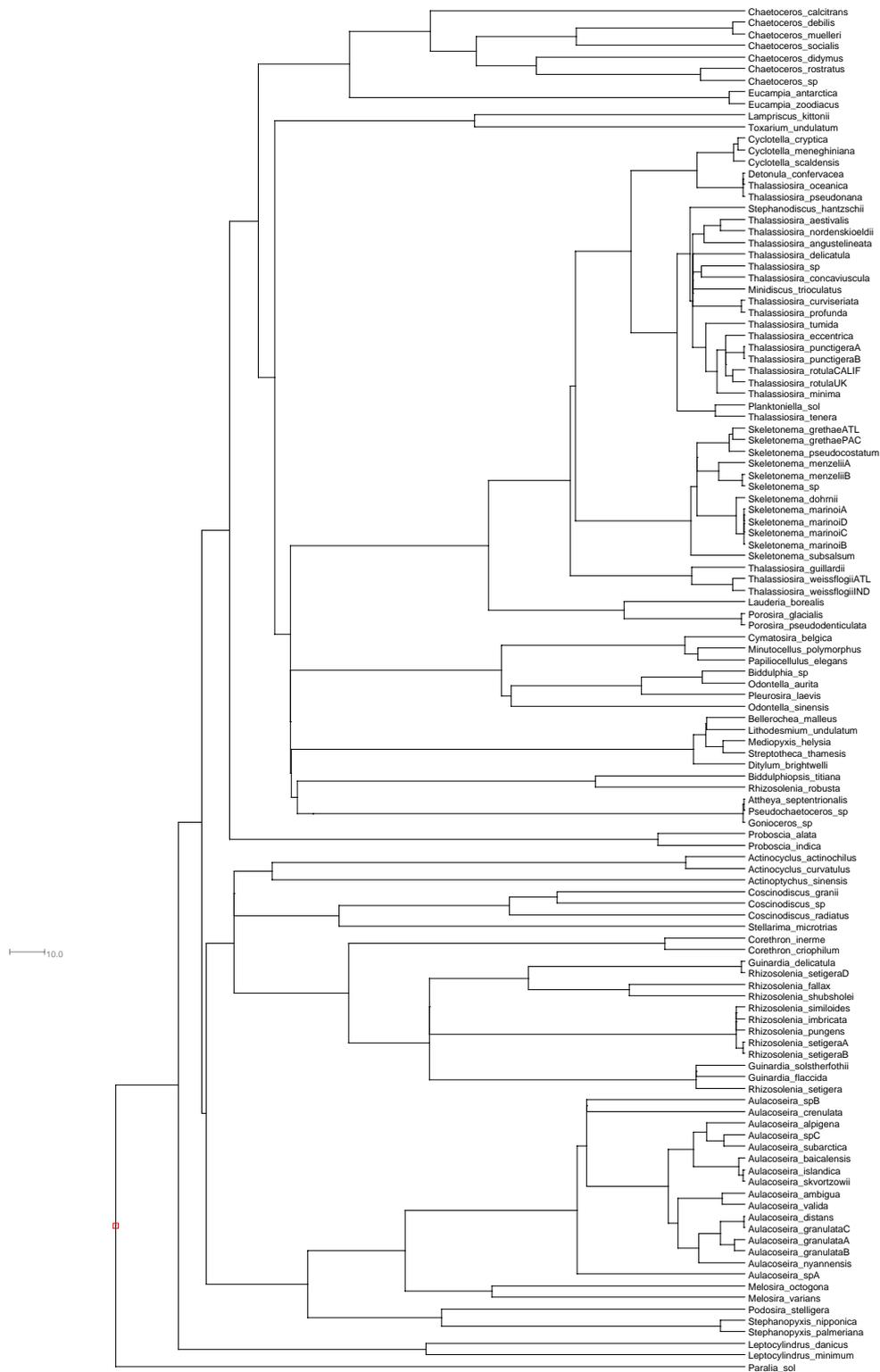
Bayes Traits

Bayes Traits [?] uses Markov Chain Monte Carlo (MCMC) methods to estimate the ancestral character states on phylogenies. There are different versions of the Bayes Traits software for continuous and discrete traits, for our current purpose only the method for discrete traits is of interest.

To model trait evolution, a continuous time Markov model is used [?], consider a binary trait that can switch between two states 1 and 0 (representing say the presence or absence of a morphological feature). The rate parameters are given by q_{01} and q_{10} measuring the rate at which trait 1 changes to trait 0 over a small time period of dt as:

$$\begin{aligned} P_{01}(dt) &= q_{01}dt \\ P_{10}(dt) &= q_{10}dt \end{aligned} \quad (1)$$

Over longer time periods these equations are no longer valid, and the model must be written in the



Phylogeny A: Drawn with Tree View software [?], of a sample of taxa from the centric diatoms from data in [?], used with the Bayes Traits software.



Phylogeny B: Drawn with Figtree software, of a sample of taxa from the centric diatoms from data in [?]. Branches are colour coded representing the trait group. Green refers to group 1, red group 2 and blue group 3.

form of a matrix Q :

$$Q = \begin{matrix} & 0 & 1 \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} -q_{01} & q_{10} \\ q_{10} & -q_{01} \end{pmatrix} \end{matrix} \quad (2)$$

Q shows the state parameters for changing states and no change of states ($-q_{01}$ and $-q_{10}$ for binary states in general given by minus the sum of the other rate coefficients in the row). Given this matrix Q the probabilities over longer intervals are found by:

$$P(t) = e^{-Qt} = \begin{bmatrix} P_{00}(t) & P_{01}(t) \\ P_{10}(t) & P_{11}(t) \end{bmatrix} \quad (3)$$

With $P_{ij}(t)$ being the probability that the state changes from state i to j over time t . Full details of this computation can be found in [?], note that while binary traits were only considered here it is a simple matter to extend these calculations to n traits. It is these rate parameters (q_{01} and q_{10}) which can be used to calculate the transitional probabilities that we seek for our model of diatom shape evolution. Pagel in [?] uses these probabilities to calculate the ancestral traits in the phylogeny, and the bulk of the software Bayes Traits provides a MCMC method to estimate these rate parameters.

Before the MCMC method used to estimate the rate parameters is explained, the assumption of this model for trait evolution should be stressed. The model is a Markov process meaning that the probability of changing from one state to the next only depends on the present state, and not the ancestor. In using this model we are also implicitly assuming that a single set of rate parameters can be accurately describe the evolution of all the centric diatoms, that is the rate parameters are universal across all branches of the phylogeny.

Using MCMC methods to estimate the rate parameters

Supposing we have a set of rate coefficients Q_i , we want to know what the probability that these parameters are correct for any given tree with given species traits. The idea being to find the rate coefficients with maximum likelihood. Using Bayes Theorem the joint posterior probability of Q_i with a dataset of traits D and tree T is:

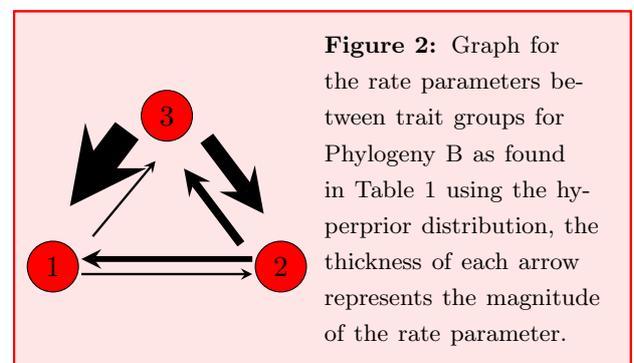
$$p(Q_i|D, T) = \frac{p(D|Q_i)p(Q_i)}{\int_Q p(D|Q)p(Q)dQ} \quad (4)$$

$p(D|Q_i)$ is the probability of the dataset of traits given a particular set of rate coefficients (which can be easily calculated) and $p(Q_i)$ is the prior distribution of Q . The part in Equation ?? which is difficult to calculate is the integral. For this the MCMC method is essentially a means of calculating this integral using Monte Carlo integration with samples selected from an appropriately chosen Markov Chain.

Notice that with this method, the entire posterior probability distribution of $p(Q_i|D, T)$ is obtained, from which the means of the rate parameters can be extracted. A little detail of how the Markov Chain is chosen should be given, making an appropriate chain turns out not to be too difficult using the Metropolis-Hastings algorithm which is fully described in [?]. Essentially the M-H algorithm produces a markov chain which stationary distribution is precisely the posterior distribution of interest ($p(Q_i|D, T)$ in the Bayes Traits software). After an initial burn time the Markov chain reaches its stationary distribution and from there can be sampled to perform the Monte Carlo integration required.

To use the Bayes traits software there are several issues to be aware of, first is the choice of the prior distribution $p(Q_i)$ which should be chosen such that it does not dominate the posterior result. As well as this the Markov chain must be run for long enough after reaching its stationary distribution for the Monte Carlo integration to produce an accurate result.

Results



The Bayes Traits software can be used to estimate the rate parameters in the Phylogeny A and B. The taxa were divided into 3 groups, each corresponding to one of the traits described in the previous

section⁵. Once every species had been classified, problems of the sampling immediately arose. The Phylogeny A is heavily biased to centric diatoms which are circular, while the Phylogeny B is described in [?] as attempting to give an even sampling across the diatoms. Despite this problem with poor sampling, Bayes Traits was run on both trees to show what effect this biased sampling has.

With three possible traits in this model there are six rate constants for Bayes Traits to find. The complete results can be seen in Table 1.

It is hard to intuitively see from Table 1 just what these provisional results mean regarding diatom shape evolution. For this reason Figure 1 is helpful to visualize how stable each trait group is. According to these results Group 3, the diatoms with high degrees of symmetry, is the most evolutionary unstable, meaning that there is a higher probability that during evolutionary time diatoms of this type will evolve to circular or bipolar shapes.

How true are these results to actual diatom evolution? Clearly there are some problems in the analysis used, with only two phylogenies being used of which there are clear sample biases. Also some large branches on Phylogeny B have no variation in trait group suggesting either a sampling bias or that within those branches the valve shape has become stable and a continuous time markov model no longer applies. Evolution can certainly lead to some traits where the selective advantage is high enough to become fixed. However the main result that suggests that diatoms with higher orders of valve outline symmetry are more evolutionary unstable makes some sense, and would account for the relatively few diatoms of this type observed in nature.

§4 Further Work

So far this had by no means been an exhaustive look at diatom shape evolution, and there have been many areas that have been untouched, in this section future areas this research could take will be discussed, as well as weaknesses in the previous analysis and how the results could be improved on.

⁵Phylogeny B has been colour coded according to these groups.

Constructing Phylogenies

One of the main issues with the previous results was acquiring a phylogeny to run the Bayes Traits software. Due to time restraints, phylogenies from the literature were used, and as constructed for a different purpose were not ideal for the trait analysis. Specifically a large unbiased sample of taxa to make up the phylogeny is required to generate accurate results.

There are numerous ways of constructing phylogenies from genomic data. Most experimental biologist would use a computational package such as PHYLIP[?] to convert the various genetic data into a phylogenies, a detailed overview of the different methods used in PHYLIP can be read in [?]. Most approaches requires all the relevant genetic data for each taxa of diatom properly aligned which may prove difficult. One way some of this work would be to construct the phylogeny from unaligned data. Such a method is presented in [?] and has a package available for use in PHYLIP.

In the paper describing the methods behind the Bayes Traits software[?], the authors propose a method of constructing phylogenies based on bayesian inference using MCMC methods, the appeal of this method is by considering the entire topological space of all possible trees the uncertainty of the trees can be taken into account. The disadvantage is that the MCMC method needs to be designed in such a way that the prior does not distort the results.

A sample with the highest genetic variability should be the most unbiased, thus given a total number of taxa with genetic data, a subsample can be found that maximises this quantity. It could be also possible that Bayesian inference techniques can also be used and a MCMC method constructed to select subsamples of similar genetic variability. This could take the uncertainty of the sample bias into account.

Alternative classification methods

The method used before to divide diatom into three different classes is quite limited. Diatoms show a rich variety of structure not only limited to valve outlines, but including pore patterning creating structures such as oculi, the division of diatoms valves into ‘sectors’ such in the genera *Aseterophalus* and *Actinoptychus*. Diatoms such as

Phy	Prior	q_{12} : Av	q_{12} :Var	q_{13} :Av	q_{13} :Var	q_{21} :Av	q_{21} :Var
A	Uniform	13.60	44.87	3.602	8.569	63.39	608.6081
B	Uniform	1.250	0.7760	1.927	8.541	2.496	3.015
B	hyperprior exp	0.8928	0.5063	0.8716	0.9081	3.216	2.247

Phy	Prior	q_{23} : Av	q_{23} : Var	q_{31} :Av	q_{31} :Var	q_{31} :Av	q_{32} :Var
A	Uniform	20.6081	310.4	67.77	597.8	65.95	642.7
B	Uniform	5.476	29.54	18.96	180.0	13.67	201.4
B	hyperprior exp	3.283	3.216	11.07	46.03	6.157	26.82

Table 1: Average and Variance of the rate parameters from Bayes Traits analysis using phylogenies A and B and prior distributions. The prior was originally chosen as a uniform distribution between 0 and 100 for both phylogenies. With this, Phylogeny A gave poor results with very high variances. With the uniform prior for phylogeny B the results are better with smaller variances, but still become quite large. Since the number of species listed in the phylogeny is small and only 75, a uniform distribution is most likely not suited as a prior and a properly chosen stronger prior should be used. For the rate parameters, we know that they must be positive, and they are not expected to be very large, thus an exponential distribution should fit. To avoid the difficulty of choosing the right parameters for this exponential function, the mean of the exponential is seeded with a uniform distribution known as a hyperprior. The ancestral state of the root node though calculated has been omitted from the table. This is due to the pennate branch of diatoms being absent from the phylogeny and hence any result not being reliable.

Attheya even have horns. Thus dividing diatoms shape into three classes does not nearly represent features that have evolved and this simplistic analysis needs to be extended to provide deeper insight. Here two possible methods will be described in detail.

Method 1.-Single genus phylogenies

To fully describe the variety of forms of diatom shape more classification groups which are more specific are needed. However there is a limit on how many groups are used, due to enough taxa being required in each group for meaningful statistical results. Instead of creating very large phylogenies over all the genera of diatoms, it should be possible to focus on only one smaller branch. This would give a detailed account of evolution and could be used to show how features like sectors have evolved in genera such as *Asterophalus*.

Method 2.- Continuous models

Instead of using discrete classification groups of traits, a continuous measure of traits could be used. One way of achieving this would be by using Fourier Descriptors. Fourier descriptors are fully described in [?] and in brief are a set of numbers obtained from a Fourier series that completely de-

fine a simple closed curve. This technique has already been applied to diatoms, in [?] the first 30 fourier descriptors were used to describe the valve shape. Due to the high number of parameters of this new data set [?] uses principle component analysis (PCA) to reduce the dimensionality and classify the diatoms into groups.

Once fourier descriptors have been obtained either PCA could be used to classify the diatoms into discrete groups then the Bayes Traits analysis could be run as previously, or the continuous model of Bayes Traits could be used. The continuous model uses a constant but unknown variance random walk (brownian motion) model and is described in detail in [?], a generalised least square (GLS) approach is used to estimate the trait value of each species $t\sigma^2$ where t is the path length to the root of the tree and σ^2 is the variance. There are other continuous quantities besides fourier descriptors which can be used, for instance [?] shows how the curvature of a diatom valve can be obtained and thus be used to distinguish diatoms.

Other possible models

The continuous markov model used previously may not be the best model to use to explain diatom shape evolution, due to the assumptions it makes. While certainly useful in many cases, other

models could be created that may also be useful in gaining insight. Features such as valve outline for the pennates seem to change little in [?] the exact nature of morphological change over the geological range of the diatom *Nitzschia jouseae* was found to be of evolutionary stasis. This suggests that high degrees of morphogenetic stability can exist in the diatom taxa, which are not represented in a continuous markov model.

Alternatively an evolutionary model of diatom shape could be constructed. The genome would be represented by a set of fourier descriptors for the outline of the diatom at various heights, in a way that would describe its entire 3D shape. Mutations in the genome will correspond to shape changes, and with a suitably chosen fitness function the evolution of the shape could be studied over generations. The choice of the fitness function here is key, it is difficult to imagine such a simple mathematical function that would reproduce all the complexities seen in diatom structure, but a basic function containing various physical constraints and possible interactions with bacteria would be interesting to examine.

§5 Final Discussion

There is much more work to do on the study of the evolution of diatom shape, the controlling feature of determining the form of the silica shell is genetic, but how this functions is still unknown. The fact that simple rules in mathematics often give rise to complex behaviour gives hope that the rules that determine diatom shape in all its wide degree of forms could one day be found and understood.

Concerning the rate parameters studied in this work it is interesting to speculate on the effect diatom evolution has had. Pennate diatoms differ from centrics in having ribs that extend from a straight sternum, while centrics radiate from a central annulus. Pennate diatoms also seem to have a high degree of stability concerning their valve shape. It seems clear that the evolution of this sternum changed the rate parameters studied, and made pennate diatoms with respect to general valve outline extremely morphological stable.

One major weakness in the approach of the type of modelling used is the lack of ability to find novel forms. An example of this would be the evolution of the raphe in pennate diatoms. This new mor-

phological feature and others like it that cause unexpected evolutionary results can not be expressed in a model that switches between a limited set of traits. Novel features in diatoms could evolve like the raphe as an advantageous new form of locomotion or as causing some other beneficial interaction in their environment such as with a certain kind of bacteria. This provides much difficulty for how exactly can you predict and model the evolution of such novel unexpected traits without knowing the full details a priori?

What this short work hopefully has shown is the potential to use various mathematical techniques to investigate this problem. In a brief period of time the result was obtained that valves with higher degrees of symmetry statistically are most unstable, while this result is not as statistically rigorous as would be liked, it shows the direction future work can and hopefully will head in.

Acknowledgements

Thanks to Dr. Ulf Sorhannus and Dr. Edward Theriot for kindly sharing their phylogenetic data with me in NEWICK format and thus saving me a substantial amount of time. Special thanks also to my supervisors Dr. Karen Page and Dr. Eileen Cox for their help throughout.

References

- [1] E. Virginia Armbrust. The life of diatoms in the world's oceans. *Nature*, 459:185–192, 2009.
- [2] Rebecca J. Bixby and Erik C. Zeek. A simple method for calculating valve curvature, journal = Proceedings of the Academy of Natural Science of Philadelphia, volume = 160, number = 1, page=73, year = 2010.
- [3] Tom Britton, Cajska Lisa Anderson, David Jacquet, Samuel Lundqvist, and Kåre Bremer. Estimating divergence times in large phylogenetic trees. *Systematic Biology*, 56:741–752, 2007.
- [4] C. Bowler et al. The phaeodactylum reveals the evolutionary history of diatom genomes. *Nature*, 456:239–244, 2008.
- [5] Joseph Felsenstein. *Inferring Phylogenies*. Sinauer Associates Inc., Massachusetts, 2004.

- [6] J. Felsenstein. Phylip – phylogeny inference package.
- [7] Eugene J. Fenster, Ulf Sorhannus, Lloyd H. Burckle, and Antoni Hoffman. Patterns of morphological change in the neogene diatom nitzschia jouseae burckle. *Historical Biology*, 2(3):197–211, 1989.
- [8] W.R. Gilks, W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. *Markov chain Monte Carlo in practice*. Chapman & Hall, 1996.
- [9] Ernst Haeckel. *Kunstformen der Natur*. 1904.
- [10] I. et al Kaczmarek. Diversity and distribution of epibiotic bacteria on pseudonitzschia multiseriata (bacillariophyceae) in culture, and comparison with those on diatoms in native seawater. *Harmful Algae*, 4:725–741, 2005.
- [11] N. Kröger and N. Poulsen. Diatoms- from cell wall biogenesis and nanotechnology. *Annual Review of Genetics*, 42:83–107, 2007.
- [12] Daiqing Mou and Eugene F. Stoermer. Separating tabellaria (bacillariophyceae) shape groups based on fourier descriptors. *Journal of Phycology*, 28(3), 1992.
- [13] Roderic D.M. Page. Tree view: An application to display phylogenetic trees on personal computers. *Computer applications in the biosciences : CABIOS*, 12(4):357–358, 1996.
- [14] M. Pagel. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proceedings of the Royal Society (B)*, 255:37–45, 1994.
- [15] M. Pagel, A. Meade, and D. Barker. Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology*, 53:673–684, 2004.
- [16] Mark Pagel. Inferring the historical patterns of biological evolution. *Nature*, 401, 1999.
- [17] Eric Persoon and King-Sun Fu. Shape discrimination using fourier descriptors. *Systems, Man and Cybernetics, IEEE Transactions on*, 7(3):170–179, 1977.
- [18] F.E. Round, R.M. Crawford, and D.G. Mann. *The Diatoms: biology & morphology of the genera*. Cambridge University Press, 1990.
- [19] Patricia A. Sims, David G. Mann, and Linda K. Medlin. Evolution of the diatoms: insights from fossil, biological and molecular data. *Phycologia*, 45(4):361–402, 2006.
- [20] U. Sorhannus. A nuclear-encoded small-subunit ribosomal rna timescale for diatom evolution. *Marine Micropaleontology*, 65:1–12, 2007.
- [21] Gary W. Stuart, Karen Moffett, and Steve Baker. Integrated gene and species phylogenies from unaligned whole genome protein sequences. *Bioinformatics*, 18(1):100–108, 2001.
- [22] Edward C. Theriot, Matt Ashworth, Elizabeth Ruck, Teofil Nakov, and Robert K. Jansen. A preliminary multigene phylogeny of the diatoms (bacillariophyta): challenges for future research. *Plant Ecology and Evolution*, 143(3):278–296, 2010.
- [23] R. Thom. *Structural stability and morphogenesis: an outline of a general theory of models*. W. A. Benjamin, 1975.
- [24] Engel G. Vrieling, Qian Yao Sun, Mingwen Tian, Patricia J. Kooyman, Winfried W. C. Gieskes, Rutger A. van Santen, and Nico A. J. M. Sommerdijk. Salinity-dependent diatom biosilicification implies an important role of external ionic strength. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25):10441–10446, 2007.