UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

School of Ocean and Earth Sciences

High-resolution evolution: calibration and application of fossil foraminifera in evolutionary time series

by

Jenneke F.A. Brombacher

Thesis for the degree of Doctor of Philosophy

August 2017

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Doctor of Philosophy

Thesis for the degree of Doctor of Philosophy

HIGH-RESOLUTION EVOLUTION: CALIBRATION AND APPLICATION OF FOSSIL FORAMINIFERA IN EVOLUTIONARY TIME SERIES

By Jenneke Fopke Antonia Brombacher

Many evolutionary processes are well-studied on generational and macroevolutionary time scales, but much less is known about the processes bridging short- and long-term biotic change. This is commonly due to a lack of sufficiently-high-resolution fossil records over long microevolutionary time scales. The work presented in this thesis describes new calibrations and applications of planktonic foraminifera in evolutionary biology. The high-resolution fossil archives of planktonic foraminifera enable construction of continuous, long-term microevolutionary time series of large numbers (>10,000) of individuals.

In **Chapter 2** I study the repeatability of traits commonly used in studies describing foraminifera evolution. The results show that some traits are reliable, whereas others are very susceptible to small mounting-induced errors and should be used with caution.

Chapter 3 deals with various representations of foraminifera body size, and whether these proxies remain accurate in a lineage undergoing morphological change. This was shown to be the case for foraminifera shell area as measured from a two-dimensional image, but less so for shell diameter.

In **Chapter 4** I study within- and among-population allometries during an interval of global climatic upheaval. When climate remains constant the within-population allometries predict evolutionary change from one time-step to the next. However, the evolutionary allometry measured across step-wise environmental change deviates significantly from the static evolutionary allometries.

Changes in biodiversity are often linked to climate change, usually represented by global temperature. However, climate consists of many interacting variables, and species likely respond to the entire climate system as opposed to individual variables. In **Chapter 5** I show that evolutionary response in two species of planktonic foraminifera is indeed best explained by combinations of environmental parameters.

Chapter 6 presents evolutionary time series of two species of planktonic foraminifera from six sites along an Atlantic transect over 600,000 years. The results show that temporal dynamics do not match spatial variation, implying that care should be taken when extrapolating one population's predicted response to another location.

Table of Contents

List of Tables

List of Figures

Supporting material

Due to the size of the dataset generated in this thesis, all data is submitted in the form of a CD with the submission of the revised thesis. Additionally, parts of the dataset have been deposited online as supplementary material accompanying published/submitted manuscripts of Chapters 3 and 4:

Chapter 3: http://dx.doi.org/10.5061/dryad.8jf2k

This dataset will be made publicly available after final publication of the paper in *The American Naturalist*

Chapter 4: https://figshare.com/s/**9db6657150242fb8a593**

This dataset will be made publicly available after acceptance of the manuscript currently submitted to *Proceedings of the Royal Society B*

Academic Thesis: Declaration of Authorship

I, **Jenneke Fopke Antonia Brombacher** declare that this thesis and the work presented in it are my own and have been generated by me as the result of my own original research.

High-resolution evolution: calibration and application of fossil foraminifera in evolutionary time series

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

   Brombacher, A., Wilson, P.A., Bailey, I., Ezard, T.H.G. (2017), The breakdown of allometry during climatic upheaval. The American Naturalist, 190(3).

   Brombacher, A., Wilson, P.A., Ezard, T.H.G. (2017), Recommendations of foraminiferal test size and shape measures based on repeatability. Marine Micropaleontology, 133, 21-27.

Signed: .....

Date: .....

Acknowledgements

First and foremost, I would like to thank my supervisors Tom Ezard, Paul Wilson and Ian Bailey, without whom this project would not have been possible. They have been fantastic mentors and it has been (and will continue to be) a joy to work with them. Tom's limitless enthusiasm about evolution and statistics has made me even more passionate about the former, and much less nervous about the latter. In our discussions I always felt treated as a peer rather than a student, and his occasional gentle nudges away from my comfort zone have helped me gain confidence as a scientist. Paul's guidance has been essential in keeping my interdisciplinary project together. We also enjoyed an on-going exchange of cultural values, where inside knowledge of British subtleties was traded for Dutch chocolate sprinkles. Ian's vast knowledge of everything Pliocene helped me put evolutionary change into a climatic context and he was the first to suggest I look at *Globoconella puncticulata*, which became one of the two key species in this thesis.

It has been a privilege and pleasure to collaborate with leading foraminifera experts both in the UK and abroad. I am very grateful to Pincelli Hull and Leanne Elder for hosting me at Yale and letting me work with their awesome 3-dimensional foram analysis system, and to Kirsty Edgar for her extensive help with foram taxonomy and in-depth discussions about my work. I would also like to thank Paul Pearson, Bridget Wade, Andy Purvis, Tracy Aze and Gene Hunt for many inspiring discussions about my data, and foraminifera evolution in general. The work presented in this thesis has also greatly benefited from discussions with my panel chair, Sam Gibbs.

I gratefully acknowledge Megan, Yuxi and their team of undergraduate student helpers for all round lab assistance, and washing a mountain of samples for me. I would also like to thank Walter Hale and Alex Wuebers for their assistance during sampling visits to the Bremen Core Repository.

The past four years would not have been the same without the NOCS Palaeo team. Richard, James S and Wendy: thanks for the tea, coffee, chats/rants, and all other forms of procrastination in office 784/15. I greatly enjoyed the Palaeo Paper Discussion sessions with the aforementioned and other Palaeo students and postdocs: Marina, James F, Rachael, Harriet, Gabriella, Rosie, Max, Gwen, Vicky, Matt, Anya, Claire, Diederik and Tom.

Attending conferences and workshops all over the world has been made even more fun with my fellow (inter)national foram friends Isabel, Rehemat, Lyndsey, Jana and Manuel. Here's to many more AirBnB shares, Brazilian cocktails, midnight foram chats and jetlagged sunrise walks.

Another big thank you goes to everyone not directly involved in my PhD, but equally important. My housemates Marina, Helen, Freya, Tianya, Matt (and Amy), Stuart, Jesse, Jon, my Utrecht undergrad friends Annemarie, Auke, Bjorn, Hannelore, Hilco, Max, Olaf, Sanne, Thijs, everyone at FolkSoc, and Susanne: you were crucial in maintaining a healthy work-life balance.

I’d like to thank Mum, Dad, Lilian and Joost for their support throughout my PhD. Thank you for always being there for me, supporting me in my decision to move abroad, the pep talks through the ups and downs, the weekend visits, the airport pick-ups, and providing me with a steady supply of decent cheese.

And finally, thank you Darren, for sharing it all with me.

Definitions and Abbreviations

Yr	Years
Kyr	Thousands of years
Myr	Millions of years
Ma	Millions of years ago
NHG	Northern Hemisphere Glaciation
iNHG	Intensification of Northern Hemisphere Glaciation
SEM	Scanning Electron Microscopy
CAS	Central American Seaway
NAC	North Atlantic Current
MOC	Meridional Overturning Circulation
AMOC	Atlantic Meridional Overturning Circulation
MIS	Marine Isotope Stage
IRD	Ice rafted debris
SST	Sea surface temperature
$\delta^{13}C$	Ratio of ^{12}C and ^{13}C isotopes
$\delta^{18}O$	Ratio of ^{16}O and ^{18}O isotopes
DSDP	Deep Sea Drilling Program
ODP	Ocean Drilling Program
IODP	Integrated Ocean Drilling Program
GLOW	GLObal Warming cruise
GAM	Generalised Additive Model
GLM	Generalised Linear Model
PCA	Principle Component Analysis
LLR	Lines of Least Resistance
MSE	Mean Squared Error analysis
ANOVA	Analysis of Variance
ANCOVA	Analysis of Covariance
CT	Computer Tomography

Chapter 1: Introduction

Global diversity faces an enormous challenge to deal with anthropogenic climate change. The recent rate of species extinctions far exceeds neutral extinction rates as documented from the fossil record (Barnosky et al., 2011; Wake and Vredenburg, 2008), even when compared to the most conservative estimates of past species extinction rates (Ceballos et al., 2015). This trend is expected to continue, with the extent of diversity decline depending on the magnitude of anthropogenic pressures (Maclean and Wilson, 2011; Thomas et al., 2004). However, large uncertainties exist within and among models trying to predict the extent of biodiversity loss (Araújo and New, 2007; Pereira et al., 2010; Thuiller, 2004). The model results and uncertainties differ among species (Thuiller et al., 2009) with varying predicted range shifts and resulting extinctions under future climate scenarios (Bellard et al., 2012; He and Hubbell, 2011). Combinations of biodiversity range and niche shift models have been reported to provide the most consistent results, although large error margins remain (Fukami and Wardle, 2005; Thuiller, 2003).

The models rely on modern species distribution and abundance data, using present-day species occurrence and abundance data to model species' response to past (Kitchener and Dugmore, 2000; Ruegg et al., 2006) and future (Huntley et al., 1995; Thomas et al., 2004; Thuiller et al., 2005) climatic conditions. These models assume that species response to environmental change is constant across space and through time. However, populations are likely to encounter varying environmental pressures at different locations throughout the species' range (Hargreaves et al., 2014), and climate is not constant over time, resulting in constantly changing abiotic forcing. In particular, when projected climate change exceeds the range of environmental variability observed in the dataset it is impossible to predict any non-linear species responses. Therefore, studies of evolutionary response to environmental change require long time series.

Phanerozoic (0 - 451 Myr) species richness has been shown to covary with global temperature (Erwin, 2009; Mayhew et al., 2008). Cenozoic diversity patterns of mammals (Blois and Hadly, 2009; Figueirido et al., 2012), plants (Jaramillo et al., 2006; Nyman et al., 2012), insects (Nyman et al., 2012) and plankton (Ezard et al., 2011; Lazarus et al., 2014) correlate with the high latitude climate signal recorded in the $\delta^{18}O$ composition of benthic foraminifera (Zachos et

al., 2001), which is widely used to represent global Cenozoic temperatures. Harnik et al. (2012) argued that simultaneous changes in multiple environmental parameters drove most Phanerozoic extinction events. These studies imply that global biodiversity is driven by global climate change over geological time scales.

To understand the driving mechanisms of biodiversity over long time scales we need to study the factors responsible for evolution of individual species (Arnold et al., 2001; Erwin, 2000). Do all species respond to multivariate climate change in a similar way, or do individual species respond to a specific single parameter? To answer this question, evolutionary trajectories of individual species should be studied as a response to multivariate climate change.

1.0.1 Internal trait constraints

Empirical and model evidence has shown widely varying responses among species to environmental change, due to varying degrees of adaptation and dispersal potential (Holt, 1990; Thuiller, 2003). Genetic and phenotypic variance and covariance patterns in particular can greatly affect the direction of evolutionary change. Many traits do not evolve independently: selection on one trait can influence the response to selection in others (Lande, 1979; Lande and Arnold, 1983). These constraints among traits can have large impacts on the direction of evolution, either facilitating evolution in the case of positive covariances (Gavrilets and Losos, 2009), or constraining adaptation when covariances are negative (Agrawal and Stinchcombe, 2009).

From one generation to the next, the mean change in the n-dimensional genotype or phenotype can be predicted by the multivariate Breeder's equation (Lande, 1979):

$$\boldsymbol{z} = \boldsymbol{G}\beta \tag{1}$$

Here $\boldsymbol{z}$ is an n-dimensional vector representing the change in n trait means, $\mathbf{G}$ is the additive genetic variance-covariance matrix with genetic variances on the diagonal elements and covariances as non-diagonal elements, and β is an n-dimensional vector consisting of directional selection gradients (Lande, 1979; Lande and Arnold, 1983). Repeated over many generations, the phenotype is expected to evolve in the direction of the dominant eigenvector of $\mathbf{G}$, which has been defined as the line of least genetic resistance $\mathbf{g}_{max}$ (Schluter, 1996). On generational time scales species have indeed been shown to follow the direction of trait covariances (Bégin et al., 2003; Lande, 1979; Lande and Arnold, 1983).

However, given sufficient time, genes can combine in any way (Futuyma, 2010), potentially eroding the lines of least resistance. Hunt (2007) and Renaud et al. (2006) applied the concept of Schluter's (1996) lines of least resistance to the fossil record, and found that speciation over macroevolutionary time scales (>1 million years) also tended to occur in the direction of maximum phenotypic variation of the ancestor population.

Unfortunately we lack robust data on how trait covariances within species evolve in deep time. Firmat et al. (2014) showed that within-species evolution followed the lines of least resistance over 600,000 years, but it is unclear how a changing environment would distort long-term microevolutionary patterns. This is best visualised by the concept of the adaptive landscape (Simpson, 1944; Wright, 1931, 1932) which uses a two-dimensional space to represent all possible combinations of any two traits. The fitness of trait combinations is indicated by elevation contours. External perturbations acting on the population can cause the fitness optimum to shift, and under such a displaced optimum the population might be expected to move away from the initial direction of maximum genotypic or phenotypic variation (Estes and Arnold, 2007). Hunt (2007) showed that the lines of least resistance in an ostracode clade eroded after a few million years, and Renaud et al. (2006) found that among-species evolution departed from the lines of least resistance following pronounced environmental change. However, the sampling resolution of these studies did not allow for the analysis of individual species trajectories over time as a response to global environmental change.

1.0.2 Response to climate change

When quantifying the effects of environmental factors on evolutionary processes, the species response is generally modelled as a response to a single, supposedly dominant, environmental parameter representing the global climate. However, climate consists of many interacting variables, and species likely respond to the entire climate system as opposed to separate variables. Harnik et al. (2012) showed that biodiversity decline during most Phanerozoic extinction events was indeed driven by combinations of environmental factors, but it is unclear if this pattern also holds at the species level. Individual species could respond to all environmental parameters and their interactions in roughly similar ways, or specific species could respond to specific drivers. In either case, studying

evolutionary response to a single environmental driver could severely underestimate biotic response to global climate change.

Additionally, different traits could respond differently to environmental change. Ecological traits like species distribution and abundance might be able to respond faster to environmental change than morphological traits, which could take thousands of years to adapt to a shift in the peak of the adaptive landscape (Bell et al., 2006). Therefore, to accurately quantify the effects of climate change on individual species' evolution, multivariate morphological change should be studied as a response to multivariate environmental change. This approach requires high-resolution records of multivariate species' evolution, as well as reconstructions of multiple local environmental parameters. However, due to the resolution of the fossil record this is rarely feasible, and no such studies exist to date.

1.1 Planktonic foraminifera

In this thesis, I aim to study evolutionary processes within individual species over microevolutionary time scales (>100,000 years), using the fossil record of planktonic foraminifera. This group of calcifying single-celled zooplankton build calcium carbonate ($CaCO_3$) tests that after death rain down on the sea floor and accumulate in the sediment to form sequences of relatively undisturbed sediments that go back millions of years. The high abundance and global distribution of planktonic foraminifera in the world's oceans produce a high-resolution fossil record that is at least as complete as the best-preserved genus-level records of macro-invertebrates (Ezard et al., 2011). As an additional advantage, environmental records such as temperature, ocean circulation and nutrient availability can be generated from the fossil medium. The preservation of full-body specimens, combined with the possibility to generate multiple local records of environmental change allows for high-resolution comparisons between multivariate evolution and multivariate climate change.

1.1.1 Ecology

Planktonic foraminifera grow to 100-1000 μm in diameter and live in the mixed layer and upper thermocline throughout the world's oceans. Currently 45 morphologically defined species exist, which are characterised by widely varying

morphologies with different test ornamentations and chamber configurations (Hemleben et al., 1989). Additionally, many morphospecies have been reported to consist of a number of cryptic species with distinct genotypes (Darling et al., 1999; de Vargas et al., 1999) which are suggested to have adapted to specific biogeographic zones (Darling and Wade, 2008).

Most planktonic foraminifera species are omnivorous and feed on a combination of other zooplankton, phytoplankton, algae and sinking phytodetritus (Anderson et al., 1979; Hemleben et al., 1989). Several species host dinoflagellate or chrysophyte photosymbionts as an additional source of carbon. These species tend to be found in the upper mixed layer of the ocean to maximise the photosynthetic potential of their symbionts. Not much is known about species preying on planktonic foraminifera: test remains have been found in the gut contents of larval fishes, copepods and pteropods (Conover, 1982), but the frequency of predation by these species remains largely unknown.

Planktonic foraminifera reproduce sexually, so for gametes to fuse other individuals of the same species need to be close by. Species density in the upper water column has been estimated to 0.1-10 individuals per cubic meter of sea water (Hemleben et al., 1989) so synchronised development would optimise the chance of reproduction. For several species the number of juvenile individuals has indeed been shown to peak shortly after the full moon, suggesting that reproduction occurs around the full moon, although the mechanisms that trigger individuals to release their gametes remain unknown (Bijma et al., 1990; Erez et al., 1991; Hemleben et al., 1989). However, a study from the South Atlantic showed that for most species, lunar periodicity in abundance could not be proven (Lončarić et al., 2005), and a more recent study showed that although lunar cycles were present in some species, a significant part of the variance in abundance was not explained by cyclical reproduction (Jonkers et al., 2015).

1.1.2 Evolution

After death, the empty foraminifera tests sink to the bottom of the ocean, where they are preserved in the seafloor sediment. Although planktonic foraminifera make up only a minor portion of the overall zooplankton biomass, they contribute to 30-80% of the total open ocean sedimentary calcium carbonate budget, depending on local dissolution and sedimentation rates (Schiebel, 2002). Typically, a cubic centimetre of sea floor sediment contains >1000 foraminifera tests. With open ocean sedimentation rates of 3-5 cm/kyr, this produces

continuous, high-resolution archives of foraminiferal tests over millions of years, enabling biostratigraphic age controls by well-defined first and last species occurrences, as well as detailed reconstructions of their evolutionary history.

The phylogeny of planktonic foraminifera has recently been completely revised (Aze et al., 2011) and its fossil record is so complete that the chance of species detection of fossil planktonic foraminifera is at least as high as the chance of genus detection in the macro-invertebrate fossil record (Ezard et al., 2011; Foote and Sepkoski, 1999). Planktonic foraminifera evolved from benthic foraminifera during the early Jurassic (Darling et al., 1997; Hart et al., 2003), but only two macroperforate species survived the end-Cretaceous extinction event. Rapid subsequent diversification resulted in a peak standing diversity of 45 species during the middle Eocene. Another decrease in diversity occurred across the Eocene-Oligocene boundary, although this reduction was not as severe as the end-Cretaceous extinction. This time, diversity gradually increased to reach another peak during the early Pliocene, after which diversity declined again. Typically, the species surviving large extinction events or biodiversity decreases tend to be small generalist species with little shell ornaments, which subsequently give rise to more specialist species. Keeled species died out at every extinction event, but evolved again afterwards from unkeeled species (Norris, 1991b), as did other morphotypes such as spherical and stellate forms. Similarly, species hosting photosymbionts radiated during the Paleocene and Eocene, declined during the Eocene-Oligocene, and re-established during the Miocene (Norris, 1996). These studies suggest that foraminifera diversity is indeed linked to environmental change. However, to understand the processes responsible for the observed patterns individual species' responses to environmental change need to be investigated.

Due to its high resolution the fossil record of planktonic foraminifera has previously been used to study the general speed of evolutionary processes in individual species (Arnold, 1983; Malmgren et al., 1983; Malmgren and Kennett, 1981; Wei and Kennett, 1988), and the mechanisms behind Cope's Law (increasing species' body increases relative to their ancestors (Cope, 1887)). Ezard et al. (2011) used the fossil record of all Cenozoic macroperforate planktonic foraminifera to disentangle the relative importance of biotic and abiotic drivers of evolutionary change (Barnosky, 2001; Van Valen, 1973a). Biotic factors include interactions among species, such as resource competition and predation, and variation within species, whereas abiotic factors include extrinsic,

environmental processes. Both have been shown to be important drivers of evolutionary processes, although the relative importance of each was unknown due to the low resolution of most fossil records (Stenseth and Maynard Smith, 1984). Benton (2009) proposed a verbal framework with different processes acting on different time scales: biotic interactions typically shape evolutionary processes on shorter (<10,000) time scales, whereas environmental factors are more important on longer, macroevolutionary time scales. This framework was tested by Ezard et al. (2011) using the fossil record of all Cenozoic planktonic foraminifera. They found that the interplay between biotic and abiotic drivers is indeed crucial to explain evolutionary change. In particular, they showed that speciation was mainly driven by biotic processes such as diversity dependence, whereas extinction events were more strongly driven by environmental change.

Using the global benthic $\delta^{18}O$ stack (Zachos et al., 2001) to represent global climate, (Ezard and Purvis, 2016; Ezard et al., 2011) showed that environmental change is an important driver of evolution. The benthic $\delta^{18}O$ values mainly represent average global temperature and ice volume. Climate is however a complex system with a high degree of local variability, and many additional environmental factors have been proposed to influence species' diversity (Garcia et al., 2014; Stigall and Saupe, 2013). Changes in nutrient availability have been argued to affect diversity throughout the food web (Allmon and Martin, 2014) causing both speciation (Marx and Uhen, 2010) and extinction events (Edgar et al., 2013; Wade, 2004; Wade and Pearson, 2008). Landmass configuration can cause isolation of terrestrial populations and alter global ocean circulation, influencing diversity by changes in local temperature, nutrient availability and water column stratification (Pearson et al., 2008; Schmidt et al., 2004b; Steeman et al., 2009; Wade and Pearson, 2008; Wei and Kennett, 1986). Furthermore both temperature and nutrient availability separately have been proposed to drive Cetacean and diatom diversity in the Cenozoic (Lazarus et al., 2014; Steeman et al., 2009). However, none of these studies have quantified the combined effects of multiple environmental parameters on diversity, whereas evidence from mass extinctions implies that large-scale evolutionary processes might be better explained when the multi-faceted nature of global climate change is taken into account (Harnik et al., 2012). Considering multiple climate drivers and their interactions could therefore mean a significant addition to studies of evolution. As the deep-sea sediments containing fossil foraminifera also yield information on local climate (Brassell et al., 1986; Emiliani, 1955; Nürnberg et al., 1996), the

fossil record of planktonic foraminifera has the potential to provide direct comparisons between morphological and environmental change.

1.1.3 Climate reconstructions

The chemical composition of fossil foraminifera tests has been extensively used to create high-resolution, continuous records of past climate. Emiliani (1954) first employed the ratio of the stable oxygen isotopes ^{16}O and ^{18}O ($\delta^{18}O$) in the test calcite as a proxy for environmental change: isotopic fractionation in the foraminiferal calcite decreases with increasing temperature, resulting in lower $\delta^{18}O$ values of calcite precipitated at higher temperatures. Additionally, carbon isotopes in the foraminiferal calcite have been used to reflect productivity regimes. At the surface, lighter ^{12}C isotopes are preferentially taken up during photosynthesis, leaving the surface waters enriched in heavier ^{13}C isotopes. At depth, the sinking organic matter remineralises and releases light ^{12}C isotopes to the deep water, resulting in lower $\delta^{13}C$ values. Therefore, the difference in isotopic composition of calcite formed at surface and deep water can be used to reconstruct the strength of the biological pump. More recently, the use of trace metals incorporated in the foraminiferal calcite has been used for environmental reconstructions. Mg/Ca ratios in the foraminiferal tests reflect the sea surface temperature at the time the test was formed (Nürnberg et al., 1996), Cd/Ca ratios reflect the distribution of nutrients in the water column (Rosenthal et al., 1997) and Neodymium isotopes in the Fe-Mn oxyhydoride coatings on foraminifera tests formed during burial are used to reconstruct past ocean circulation by assessing the origin of water masses bathing a site at the time of their deposition (Vance and Burton, 1999).

The records generated from foraminiferal calcite have been used to reconstruct long-term climate records of important periods in the development of the Earth's climate system. For instance, benthic foraminiferal $\delta^{18}O$ stacks spanning the Cenozoic (also called the 'Zachos-curve', (Zachos et al., 2001; Zachos et al., 2008)) and late Neogene (Lisiecki and Raymo, 2005) have revealed climate variability operating on multiple time scales, including gradual, long-term (10^5 – 10^7 years) warming and cooling trends, cyclic patterns driven by orbital processes lasting 10^4 – 10^6 years such as the onset of Northern Hemisphere glaciation (Mudelsee and Raymo, 2005) and rapid transient events lasting 10^3 - 10^4 years such as the Paleocene-Eocene Thermal Maximum (Zachos et al., 2008).

1.2 Environmental settings

Global temperatures have been gradually declining since the Middle Eocene Climatic Optimum (Zachos et al., 2001), likely due to increased silicate weathering which reduced atmospheric CO_2 (Pagani et al., 2011; Ruddiman and Kutzbach, 1989). Antarctic ice caps first started to develop during the Late Eocene (Ehrmann, 1998) but East Antarctica only became fully glaciated during the Eocene-Oligocene boundary at ~33.7 Ma (Coxall et al., 2005). The presence of sea-ice diatoms indicate the periodic sea ice from ~47 Ma onwards (Stickley et al., 2009), and records of ice rafted debris (IRD) deposition proximal to Greenland indicate that (small) ice caps existed on this continent between ~40-33 Myr ago. Atmospheric CO_2 concentrations were, however, still too high (>> 280 μatm) during this time for permanent, large-scale ice sheets to develop in the Northern Hemisphere (DeConto et al., 2008). The Antarctic ice sheet expanded periodically during the Oligocene and Miocene, and particularly so around 23.1 Ma in an event known as Mi-1 (Liebrand et al., 2011). This interval was characterised by a 200-kyr interval of low amplitude variability in obliquity combined with minimum precessional forcing, that may have also favoured ice sheet growth in the Northern Hemisphere. The onset of major Northern Hemisphere glaciation however did not occur until the Late Pliocene (Mudelsee and Raymo, 2005). The first IRD pulses from the Greenland ice sheet were observed around 3.3 Ma (Jansen et al., 2000), followed by major IRD fluxes starting at Marine Isotope Stage (MIS) G6 (2.72 Ma) from marine-based Greenland and European ice sheets, and MIS G2 (2.64 Ma) from the North American ice sheet (Bailey et al., 2013). Atmospheric CO_2 concentrations first fell below the 275 μatm threshold proposed by DeConto et al. (2008) for full scale Northern Hemisphere glaciation at ~2.8 Ma (Martínez-Botí et al., 2015). From this time Northern Hemisphere ice sheets grew periodically and linearly in accordance to 41-kyr variations in Earth's orbital obliquity (Ruggieri et al., 2009). These cycles continued to dominate the pacing of glacial-interglacial cycles until ~1 Ma ago, when during the Mid Pleistocene Transition, Earth's climate system transitioned to a state that was dominated by quasi-100-kyr glacial-interglacial cycles that persist until today (McClymont et al., 2013).

Global ocean circulation influences meridional heat and moisture transport. It has therefore always been tempting to implicate changes in ocean circulation that increase moisture flux to the high latitude a mechanism for Cenozoic ice-sheet expansion. Indeed, the closure of the Central American

Seaway (CAS) during the late Neogene has been proposed to have had a major influence on Meridional Overturning Circulation (MOC) (Keigwin, 1982) and on Northern Hemisphere glaciation (Bartoli et al., 2005; Haug and Tiedemann, 1998). The timing of the CAS closure (Molnar, 2008, 2017; Montes et al., 2015; O'Dea et al., 2016), its impact on MOC (Bell et al., 2015; Karas et al., 2017a; Karas et al., 2017b) and whether any increase in moisture flux to the high northern latitudes would have driven glaciation in any case (Lunt et al., 2008) are controversial. Nevertheless, its proposed closure from ~4.6 Ma, has been suggested to have preconditioned the Northern Hemisphere ice sheet growth (Haug and Tiedemann, 1998). Alternatively, hypothesised 'final stages' of CAS closure between ~3-2.5 Ma ago, and a subsequent strengthening of MOC, are also thought to have been important in driving ice-sheet growth in the Northern Hemisphere during the late Pliocene (Bartoli et al., 2005; Haug et al., 2001). State-of-the-art coupled climate-ice-sheet models indicate, however, that it is ice-sheet summer melt and not ice-accumulation that drives positive ice-sheet mass balance (DeConto et al., 2008) and evidence for any spin-up in MOC during the intensification of Northern Hemisphere glaciation (iNHG) has yet to be found (e.g. Lang et al., 2016).

Instead, a lowering of atmospheric CO_2 is the most likely candidate to have driven orbitally-paced glacial expansions during the late Neogene through several feedback mechanisms. Since MIS G6, increased ocean stratification has been observed at both Northern and Southern high latitudes (Haug et al., 1999; Sigman et al., 2004; Woodard et al., 2014). Water column stratification reduces exchange between deep water masses and the atmosphere, and prevents CO_2 from remineralised organic matter to reach the atmosphere. Additionally, increased ice cover in the Southern Ocean has been suggested to further reduce ocean-atmosphere gas exchange, increase CO_2 storage in deep water masses and lower atmospheric CO_2 (Hillenbrand and Cortese, 2006). A recent study by Lang et al. (2016) showed that southern sourced water extended into the North Atlantic during prominent glacials from MIS G6 onwards, and may have helped to amplify glacial cycles on orbital timescales through increased deep water CO_2 storage. However, no secular change in North Atlantic water mass structure was observed during 3.3-2.4 Ma, implying that changes in Atlantic Ocean circulation likely acted as a positive feedback on glaciation, rather than a driving mechanism at this time.

The reorganisation of global climate during iNHG provides an opportunity to study biotic response to multiple changing environmental factors simultaneously. Planktonic foraminifera prefer specific temperature ranges

(Lombard et al., 2011; Lombard et al., 2009), productivity regimes (Hemleben et al., 1989), and ocean pH ranges (Hemleben et al., 1989) among others. Therefore, environmental change related to iNHG enables comparisons of single species' evolutionary trajectories to multivariate environmental change.

1.2.1 Biotic response

The intensification of Northern Hemisphere glaciation provides an interesting setting to study biotic response to environmental change, as it marks both a secular change to lower temperatures and atmospheric CO_2 levels, as well as the onset of cyclical glacial-interglacial environmental change. Biotic responses to these changing conditions have been observed across all trophic levels. Increased turnover rates have been reported for South American (Tonni et al., 1992) and African mammals (Vrba, 1995), although neither were as severe as late Pleistocene mammal extinctions. Temperate tree diversity declined steeply in Europe as a result of lower temperatures, but not in North America and eastern Asia (Svenning, 2003). Bivalve populations were decimated in the Mediterranean, west Atlantic and North Sea due to decreasing temperatures (Raffi et al., 1985; Stanley, 1986) and Caribbean reef corals declined in abundance (Budd et al., 1996). In the Indian Ocean, benthic foraminifera diversity decreased, likely due to changes in deep water circulation as result of increased glaciation (Kawagata et al., 2006). Most of these extinction events only occurred locally, implying that environmental change associated with the intensification of Northern Hemisphere glaciation did not result in a major global extinction event.

Similarly, planktonic foraminifera diversity decreased during the Pliocene and Pleistocene (Wei and Kennett, 1986). No sudden extinction events have been reported, but the subsequent extinctions of 14 species gradually lowered foraminiferal diversity (Aze et al., 2011). All but one of the species in the genus *Menardella* survived into the Pleistocene (Chaisson, 2003). Preceding their extinction the abundance of several *Menardella* species decreased, and others migrated to shallower depths during the expansion of Northern Hemisphere ice sheets (Pfuhl and Shackleton, 2004). Similarly, *Globoturborotalia nepenthes, Globoturborotalia bolli* and *Dentoglobigerina altispira* have been reported to retreat to lower latitudes as a result of lowering sea surface temperatures, followed by diachronous population extinctions with latitude (Jenkins, 1992). Wei (1994a) proposed that increased water column stratification due to changes in global ocean circulation restricted the habitat of *Globoconella puncticulata*

directly preceding its extinction. Interestingly, he also noted that the trait combinations necessary for the species to adapt to the environmental changes were unlikely to happen due to internal trait constraints: decreasing test size and roundness (i.e. more rectangular tests) would have been beneficial for the species to counteract changes in water column stratification by decreasing its settling velocity, but these two traits covaried inversely, implying that it would have been difficult for both to decrease at the same time (Wei, 1994a). These results imply that comparing trait variance and covariance directly to regional environmental change can shed new light on the driving mechanisms behind extinction.

1.3 Research objectives

In this thesis, I will study the evolutionary responses of *G. puncticulata* and the surviving, ecologically similar species *Truncorotalia crassaformis* to the intensification of Northern Hemisphere glaciation. I use populations from the North Atlantic Ocean: the iNHG is best recorded at the Northern Hemisphere, and generally foraminifera are better preserved in the Atlantic than in the Pacific due to the Pacific's deeper depth and dissolution of $CaCO_3^{2}$ material at the sea floor. The approaches presented here to study high-resolution evolutionary change over long time scales require a large (>10,000 individuals) dataset. To enable this, I first developed a high-throughput method to automatically analyse multiple foraminifera shells (see section 1.4), and tested it for repeatability (Chapter 2) and accuracy (Chapter 3). When the reliability of the results was established, trait variance, covariance and response to environmental change could be quantified. This thesis comprises three main research objectives:

1. **How do trait constraints shape within-species evolution during intervals of pronounced environmental change?**
   Trait constraints can have profound effects on a species' potential to adapt to a changing environment. Here, using 12,633 individuals of two species of planktonic foraminifera, I study the consistency of lines of least resistance (Schluter, 1996) during the most recent great climate transitions on Earth (Mudelsee and Raymo, 2005): the intensification of Northern Hemisphere glaciation. I focus on a special case of trait covariance: the allometry between shell size and shape. Allometries describe covariations

of given traits with body size according to a power relationship (Huxley, 1932). They are considered exemplary lines of least evolutionary resistance because allometric constraints imply that internal growth regulators restrict trait evolution (Pélabon et al., 2014).

2. **How does multivariate climate change affect species evolution?**
   Changes in biodiversity are often linked to climate change, usually represented by global temperature. A global environmental driving mechanism of biodiversity is implied by the strong correlation between the Cenozoic diversity patterns of a wide variety of fauna and flora and changes in global climate as recorded by the oxygen isotope composition of benthic foraminifera. Yet climate consists of many interacting variables, and species likely respond to the entire climate system as opposed to individual variables. Here, both ecological and morphological traits of 12,629 individuals of two species of planktonic foraminifera will be used with similar ecologies but contrasting evolutionary outcomes. Many different environmental factors have been proposed to influence the evolution of foraminifera, such as sea surface temperature (Edgar et al., 2013; Wade and Pearson, 2008), productivity (Wade, 2004) and ocean stratification (Wade and Pearson, 2008; Wei and Kennett, 1986). However, none of these studies analysed the evolutionary impacts of the interplay of multiple climate drivers, as implicated in mass extinction events (Harnik et al., 2012).

3. **How do species responses vary across space and through time?**
   Spatial species variation and distribution patterns are often used to predict temporal change. Methods such as bioclimatic and species diversity models (reviewed in (Guisan and Zimmermann, 2000)) use present-day species occurrence and abundance data to model species' response to past (Kitchener and Dugmore, 2000; Ruegg et al., 2006) and future (Huntley et al., 1995; Thomas et al., 2004; Thuiller et al., 2005) climatic conditions. These models assume that species response to environmental change is consistent across space and through time. However, populations are likely to encounter varying biotic and abiotic pressures at different locations throughout the species' range. Here, I have constructed a dataset consisting of 29,438 individuals (the largest single-study microfossil dataset to date) to study populations' responses across latitudes and

compare within- and among population trait covariation patterns as well as response to local environmental change, to investigate whether spatial species variation predicts temporal change.

1.4 Methods

1.4.1 Study species

To study evolutionary responses to climate change through time and across space, study species should be highly abundant in the fossil record to allow for high-resolution morphological reconstructions, and be widely distributed to enable comparisons of biotic response across a wide variety of environmental settings. Ideally, the species' stratigraphical range should include time intervals characterised by profound environmental change to capture the full range of (non-linear) biotic responses to abiotic drivers. *Globoconella puncticulata* (Figure 1.1a) is a globally distributed thermocline-dwelling species that originated around 4.7 Ma and became extinct at 2.4 Ma (Wei, 1994b), shortly after the intensification of Northern Hemisphere glaciation (iNHG). Its extinction has been linked to iNHG (Chapman et al., 1998; Scott et al., 2007; Wei, 1994a), implying a high sensitivity to environmental change related to increased glacial intensity. Therefore, this species provides an excellent opportunity to study morphological response to global climate upheaval preceding extinction. *Truncorotalia crassaformis* (Figure 1.1b), another thermocline-dwelling species originated around 5.7 Ma but survived iNHG and is still alive today. Its morphological similarity and comparable distribution and depth habitat in thermocline waters to *G. puncticulata* (Aze et al., 2011; Kennett and Srinivasan, 1983) enable comparisons of biotic response to abiotic change under contrasting evolutionary scenarios.

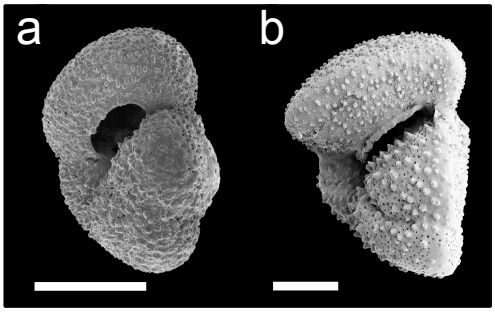

Figure 1.1 Study species *Globoconella puncticulata* (a) and *Truncorotalia crassaformis* (b). The white bars represent 100 μm.

1.4.2 Study sites

To enable high-resolution reconstructions of morphological change over time, well-preserved fossil samples are required to allow for analysis of complete foraminifera tests. Additionally, study sites should be characterised by high and consistent sedimentation rates to allow for sufficient spacing between samples to avoid time averaging by bioturbation between consecutive samples, and existing orbitally resolved age models are required for each site to enable comparisons of biotic change among populations. IODP Site U1313 (Figure 1.2) meets all these criteria. Its high sedimentation rates and excellent preservation of fossil foraminifera tests allows for high-resolution morphological reconstructions and is here used to study the evolution of trait covariations during the last great climate transition on Earth: the intensification of Northern Hemisphere glaciation (iNHG) (Chapter 4). Furthermore, existing site-specific reconstructions of sea surface temperature, productivity and nutrient sources allow for direct comparisons between morphology and multivariate environmental change (Chapter 5). Finally, five additional sites over a latitudinal Atlantic transect (ODP Sites 606, 659, 925 and 981 and IODP Site 1264, see (Table 1.1) were analysed to study the consistency of biotic response among populations under different environmental settings through time and across space (Chapter 6).

Site	Latitude	Longitude	Average sed. rate (cm/kyr)	Sample resolution (cm)	Average time resolution (kyr)	Sample width (cm)	Average time in sample (kyr)	Sample volume (cc)
981	55.48	-14.65	9.29	25	12.6	1.5	0.16	20
U1313	41.00	-32.96	4.69	30	6.4	2	0.43	20
606	37.34	-35.50	3.05	15	5.0	1.5	0.49	20
659	18.08	-21.03	2.90	15	4.9	2	0.69	10
925	4.20	-43.49	2.85	15	5.2	2	0.70	10
1264	-28.53	2.85	0.90	15	18.3	1	1.10	10

Table 1.1 Sampling specifics for all studied sites

All sites were sampled from cores stored at the Bremen Core Repository (BCR), except for Sites 1264 and U1313, for which I used a subset of the samples used by Bell et al. (2014) and Bolton et al. (2010) respectively. Depending on site-specific sedimentation rates samples were taken at 15-30 cm intervals, resulting in time resolutions of 5-10 kyr for Site U1313, and 10-20 kyr for the other sites (Table1.1). Where possible cores were sampled using 2 cm-wide plastic scoops. In several cases however the sediment was too dry for the scoops to penetrate, and 1.5-cm slices were taken instead. Samples were washed using deionised water and sieved over a 63-μm mesh sieve. The residues were dried in an oven set to 50 °C and stored in glass jars.

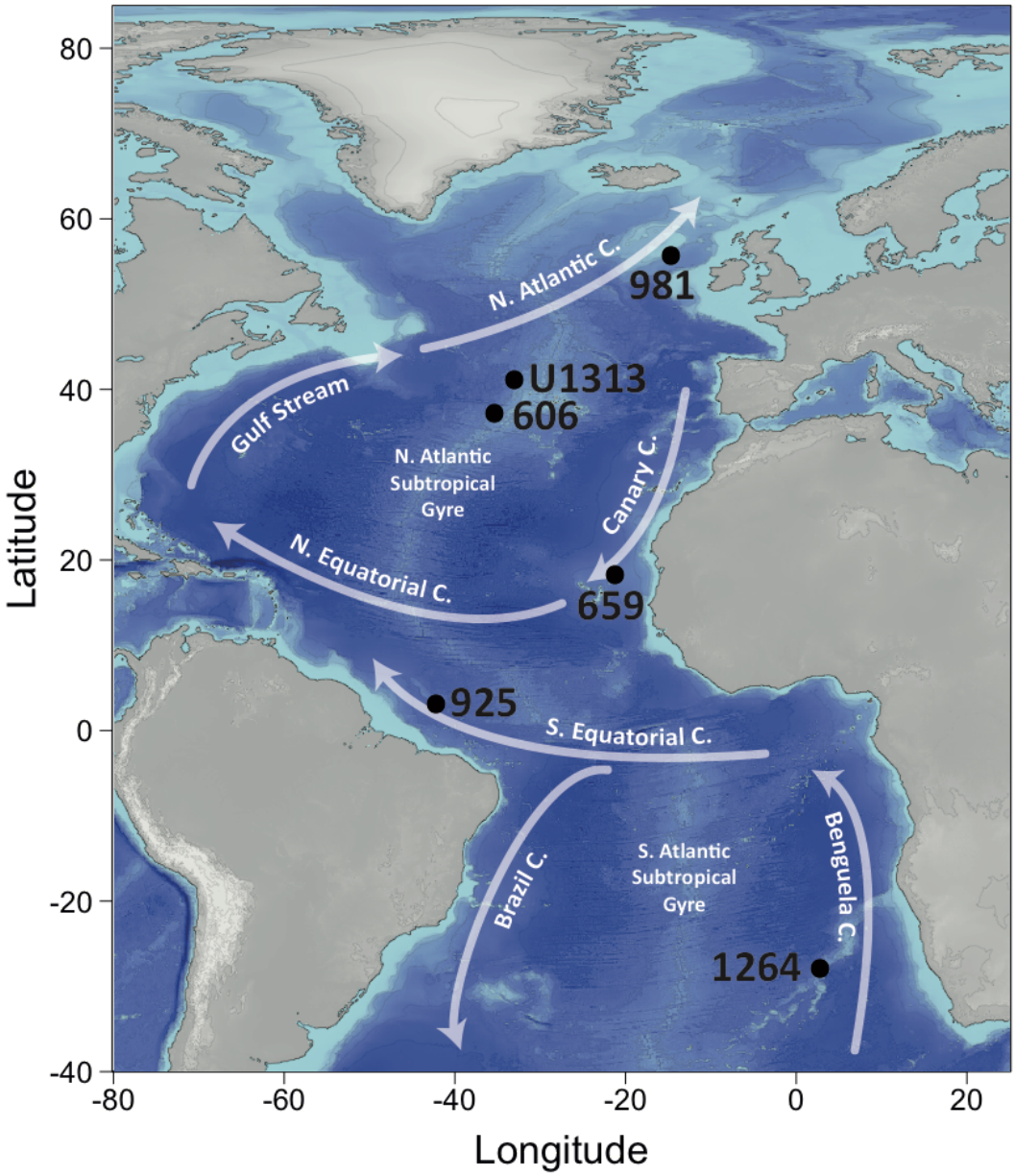

Figure 1.2 Position of all six study sites and major ocean currents in the Atlantic Ocean

1.4.3 Analysis

Foraminifera specimens were picked from the samples' >150 μm size fraction to exclude juvenile individuals from analysis. To capture the full extent of morphological variability within each species over time, 50-150 individuals were picked per species per sample. Samples were split into smaller fractions using a microsplitter, which prevents size bias within single fractions, until a single split contained the target number of specimens of either species. All individuals present were picked from the split to avoid size bias in the picked specimens. To calculate the total number of individuals present in the sample, the number of individuals in the split was multiplied by the split fraction, and total abundance was calculated as number of individuals per gram course fraction of the sediment per 1000 years.

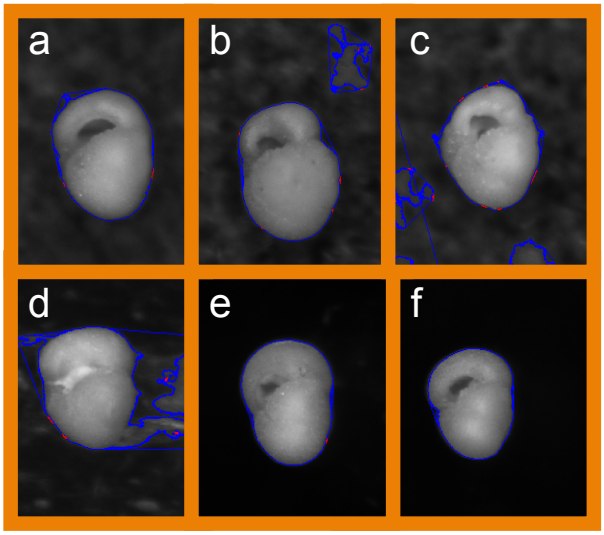

Figure 1.3 Foraminifera tests mounted on cardboard (a-c) and glass (d-f) slides using Pritt stick (a,d), transparent particle glue (b,e) and transparent double-sided tape (c,f). Blue outlines indicate the tests as recognised by the Image Pro Premier software.

Foraminifera were imaged using a Lumenera Infinity 3 camera attached to an Olympus SZX10 light microscope. Specimens were imaged in side view to capture the aperture and degree of chamber inflation. Images were analysed using Image Pro Premier 9.2, which separates objects from their background using a threshold brightness level. Therefore, to accurately analyse foraminifera tests they need to be clearly identifiable from the media they are mounted on. To ensure optimum specimen detection two different types of slides and three different adhesives were tested (see also Chapter 2). Gridded cardboard slides allow for easy

specimen identification with one individual per numbered square, however small white background imperfections could result in parts of the cardboard being analysed as belonging to the foraminifera test. Transparent glass slides provide a homogenous dark background when illuminated from above, but specimens are harder to attach to the smooth surface. Pritt stick is easy to apply on both types of slides but leaves opaque strands of glue that could be identified as part of the foraminifera test; transparent particle glue which leaves less traces but dries out quickly, giving little time to mount tests; and transparent double-sided sticky tape which is easy to apply and does not dry out quickly, but small imperfections in the tape could compromise measurement accuracy similar to the cardboard slides. Both types of slides were tested using all three adhesives to find the best slide-adhesive combination (Figure 1.3).

As expected, the glass slides provide a darker background in the images than the cardboard slides, which appear dark grey when exposed to the microscope's high light intensity (Figure 1.3a-c). On the cardboard slides there were several cases where the image analysis software could not distinguish between foraminifera tests and white strands in the cardboard (Figure 1.3b,c). Of the three tested adhesives Pritt stick left most traces, with white glue strands occasionally being analysed as part of the foraminifera tests (Figure 1.3d). Both the particle glue and the double-sided tape left little to no traces (Figure 1.3b,c,e,f), however the glue did not stick well to the glass slides whereas the tape was easy to apply and did not dry out, providing an easy medium to orient tests in the desired position. Therefore, the combination of glass slides and double-sided tape was chosen to mount and analyse foraminifera tests on.

Traits on individual foraminifera tests were automatically measured from the outline generated by the Image Pro Premier software, which provides a wide range of measures to describe object size and shape. As the tests were manually mounted in side view, small orientation errors could influence the measurements. Prior to analyses, I tested the effects of manually induced orientation errors on four widely used measures of test size and shape: area, aspect ratio (the height: width ratio), perimeter and roundness (π*perimeter^2/area). I found that both test area and aspect ratio are repeatable measures of test size and shape, whereas roundness and perimeter were not repeatable for the species studied here (Chapter 2). Additionally, I checked whether test area as measured from a two-dimensional representation is a reliable proxy for test volume for different test shapes and sizes, and found this to be the case (Chapter 3).

Chapter 2: Calibration of the repeatability of foraminiferal test size and shape measures with recommendations for future use

This chapter is a reproduction of an article published in *Marine Micropaleontology*: Brombacher, A., Wilson, P.A., Ezard, T.H.G. (2017), Recommendations of foraminiferal test size and shape measures based on repeatability. Marine Micropaleontology, 133, 21-27.

2.1 Abstract

The fossil record of planktonic foraminifera is ideally suited to defining stratigraphic age controls and exploring fundamental questions in evolutionary biology due to its excellent preservation potential that yields continuous, high-resolution fossil archives of large numbers of individuals. For full morphometric analyses foraminifera tests are generally mounted, oriented and imaged manually, while data are processed using standard software such as ImageJ or Image Pro. However, manually induced orientation errors are a source of potential bias in trait measurements even when quantified using the same computational subroutine. Here I test the repeatability of four measures of foraminiferal test shape on six morphologically distinct species and present a calibration (power analysis) of the number of individuals needed to determine a given percentage change in these traits. I mounted and measured every individual twice and analysed the difference between the two measurements to determine the effects of small orientation changes on the studied traits. I show that measurements of test area and aspect ratio are statistically indistinguishable between runs for all species studied, and a power law calibration suggests that between 25 and 50 individuals are needed to detect at least a 10% in- or decrease in either trait. However, despite mounting tests on glass slides to clarify perimeter outlines, test perimeter was only repeatable in the spherical species *Orbulina universa*, and test roundness was not repeatable for three out of six studied species. These

results recommend the use of lengths and avoidance of perimeters and their dependent metrics to reduce orientation induced bias.

2.2 Introduction

The planktonic foraminifera bequeath one of the most complete fossil records known to science. The accumulation in deep sea sediments of well-preserved shells of vast numbers of individuals make the planktonic foraminiferal fossil record uniquely suited for both biostratigraphic age controls (Berggren et al., 1995b; Blow, 1969; Bolli et al., 1989; Wade et al., 2011), and for answering fundamental questions in evolutionary biology (e.g. Alizon et al., 2008; Ezard et al., 2011; Hull and Norris, 2009; Norris, 1991a; Wei and Kennett, 1988). The preservation of complete specimens allows for the construction of multivariate trait datasets, which can be used to distinguish between species in a quantitative manner and pinpoint the exact timing of speciation and extinction (Aze et al., 2011; Pearson and Ezard, 2014; Wei, 1994b), and allow for high-resolution reconstructions of species' evolutionary trajectories over millions of years (Kucera and Malmgren, 1998; Pearson et al., 2014; Wade and Olsson, 2009).

The reliability of morphometric records depends on the precision with which individual traits can be measured. A good measurement system allows for rapid processing, is repeatable between runs and produces reliable results. Planktonic foraminifera are most often measured from two-dimensional images taken by a camera attached to a microscope with individual tests mounted in a given orientation, and trait measurements are extracted from the image's 2D test representation. This set up has the potential to introduce bias in two main ways: manually measuring traits on the imaged specimens involves subjectivity, and hand mounting of individual specimens introduces error in the orientation of the tests. The first issue can be addressed using automated specimen detection and trait measurements with fixed magnification and light intensity, which reduces subjective human input. Measurement biases induced by mounting errors can be reduced by mounting tests on a rotatable hemispherical stage which is manually adjusted to fine-tune specimen orientation prior to imaging (Knappertsbusch, 2007; Knappertsbusch et al., 2009; MacLeod and Carter, 1984; Pearson and Ezard, 2014). However, specimen adjustment in this setup still relies on subjective human input, and the time-consuming nature of adjusting, imaging

and analysing each individual separately makes this approach suboptimal for large (>10,000 specimens) datasets that are increasingly produced (Brombacher et al.; Hsiang et al., 2016; Knappertsbusch, 2000; Malmgren and Kennett, 1981; Pearson and Ezard, 2014). I focus, therefore, on hand-mounted individuals using a fixed stage, but the effects of mounting errors need to be estimated by repeatedly mounting, measuring, remounting and re-measuring individual tests.

Here I test the repeatability of four measures of foraminifera test size and shape: test area, perimeter, aspect ratio and roundness. Test area represents the individual's body size, an ecologically important trait (Hecht, 1976b; Schmidt et al., 2004a) that is easy to measure and a useful first estimate to distinguish between species. Test perimeter is often used in multivariate morphometric studies to assign the position of landmarks (Biolzi, 1991; Wei, 1994b; Wei and Kennett, 1988). Aspect ratio and roundness are two measures of test shape, further enabling species identification as well as quantifying the test area-to-volume ratio, an important measure in terms of metabolic processes. Both metrics are routinely generated by popular software such as ImageJ or Image Pro. Together, these four traits form an important part of describing planktonic foraminifera morphologies. Therefore, quantifying their precision is crucial to the task of interpreting species morphometric records for both biostratigraphic and evolutionary purposes.

2.3 Material and Methods

2.3.1 Study species

Here I present repeated measurements on the tests of six species of planktonic foraminifera with distinct shell morphologies. All taxonomic descriptions given here are from Kennett and Srinivasan (1983) and references therein.

Orbulina universa (Figure 2.1a). The adult stage consists of a single spherical final chamber enveloping the earlier part of the test. In this study only adult tests are used.

Globoconella inflata (Figure 2.1b). Low trochospiral tests with a broadly rounded axial periphery and an extraumbilical-umbilical aperture. Chambers more inflated on the umbilical side than the spiral side and increase slowly in size as added.

Globoconella puncticulata (Figure 2.1c). Low trochospiral tests with a flattened spiral side, highly vaulted umbilical side and bluntly rounded axial periphery. Chambers are angular and increase slowly in size as added. The extraumbilical-umbilical aperture is a high interiomarginal arch.

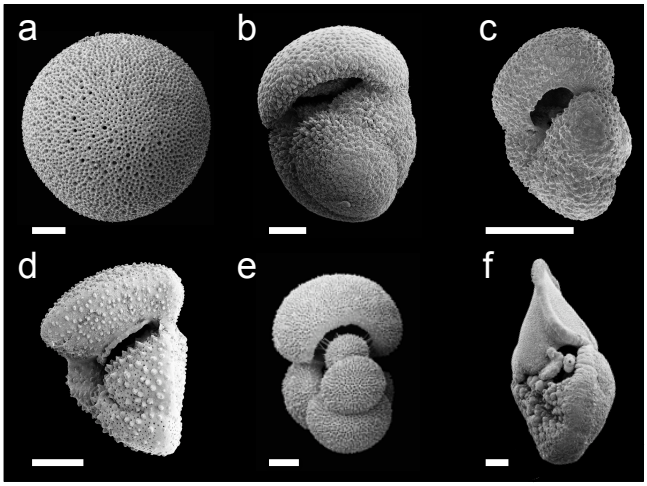

Figure 2.1 SEM images of a) Orbulina universa, b) Globoconella inflata, c) Globoconella puncticulata, d) Truncorotalia crassaformis, e) Globigerinella siphonifera and f) Globorotalia tumida. Scale bars represent 100 µm.

Truncorotalia crassaformis (Figure 2.1d). Low trochospiral tests with a flat spiral side, strongly convex umbilical side and planoconvex, subacute axial periphery. Chambers are compressed and increase rapidly in size as added. The extraumbilical-umbilical aperture is a low-arched slit bordered by a lip.

Globigerinella siphonifera (Figure 2.1e). Adult tests are evolute and planispiral, with a rounded axial periphery and a wide arched interiomarginal aperture. Chambers are globular and increase rapidly in size as added.

Globorotalia tumida (Figure 2.1f). Tests are trochospiral and convex, with the spiral side more convex than the umbilical side and a narrow extraumbilical-umbilical aperture. The axial periphery is acute with a heavy keel. The chambers are wedge-shaped and increase rapidly in size as added. The extraumbilical-umbilical aperture is covered by a plate-like lip.

2.3.2 Analysis

Specimens of *O. universa, G. siphonifera* and *G. tumida* were picked from a box core sample collected by the GLObal Warming (GLOW) cruise at station GLOW 5 (-8.9 °N, 41.5 ‘W). Individuals of *G. inflata, G. puncticulata* and *T. crassaformis* were collected from IODP Site U1313 (Leg 306, 41 °N, 32.5 ‘W, see Figure 1.2). *G. puncticulata* and *T. crassaformis* were picked from sample 306-U1313C-12H-4, 22-24 cm and specimens of *G. inflata* were picked from sample 306-U1313B-10H-4, 45-47 cm.

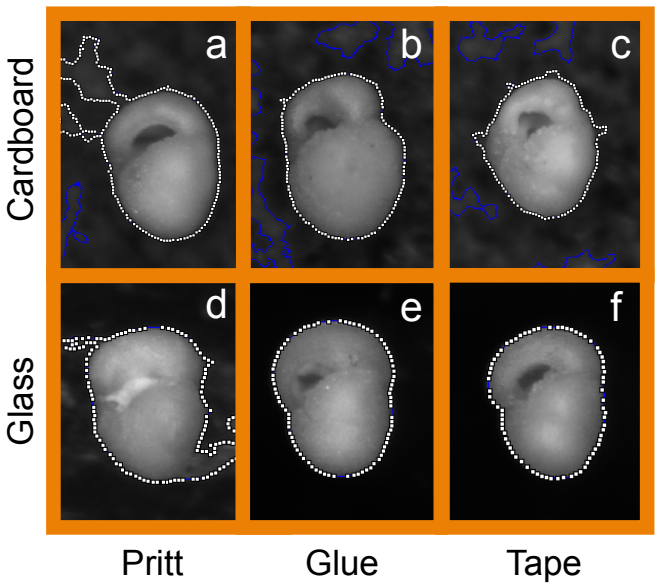

Figure 2.2 Foraminifera tests mounted on cardboard (a-c) and glass (d-f) slides using Pritt stick (a,d), transparent particle glue (b,f) and transparent double-sided tape. White dotted outlines indicate the tests as recognised by the Image Pro Premier software. Blue outlines indicate background imperfections also picked up by the software. Object recognition was best using foraminifera mounted on glass slides using double sided tape (f) and therefore this setup was used here to test trait repeatability.

I picked and mounted the first 100 specimens encountered of each species. To minimise measurement errors from background imperfections two types of slides were tested using three different adhesives to find the most homogenous background. Gridded cardboard slides allow for easy specimen identification with one individual per numbered square, however small white background imperfections in the cardboard result in parts of the cardboard slide being mistakenly identified as belonging to the foraminifera test (Figure 2.2a-c). Transparent glass slides provide a homogenous dark background when illuminated from above (Figure 2.2d-f). Pritt stick is easy to apply on both types of slides but leaves opaque strands of glue that are identified as part of the foraminifera test (Figure 2.2a,d); transparent particle glue leaves less traces but dries out quickly, allowing too little time to mount tests carefully (Figure 2.2b,e). Transparent double-sided sticky tape is easy to apply and does not dry out so quickly (Figure 2.2f). Therefore, tests were mounted on glass slides using transparent double-sided tape (Figure 2.2f).

Individuals were oriented in side view on a fixed stage. This orientation generally provides the best view of the test aperture and related, ecologically relevant landmarks and is often used in morphometric analyses of planktonic foraminifera (Knappertsbusch, 2007; Kucera and Malmgren, 1998; Lohmann and Malmgren, 1983; Pearson and Ezard, 2014; Wei, 1994b). However, depending on the acuteness of the axial periphery this is also the least stable position, which potentially increases measurement errors because the test sits obliquely on the slide and more or less of the perimeter is visible. Tests were imaged with fixed light intensity and camera exposure time using an Infinity 3 Lumenera camera mounted on an Olympus SZX10 light microscope, with illumination from above. Test area, aspect ratio (ratio between maximum test height and width), perimeter length and roundness (π*Perimeter^2/Area) were extracted from the images using an automated image analysis macro in the Image Pro Premier software. Individuals were mounted, measured, carefully removed from the slides to avoid damaging the test, and remounted and re-measured once to provide an upper bound on trait repeatability.

Here I study the differences between the first and second measurements of a given trait on the same individuals. Measurements of the first and second run are plotted in Figure 2.3 using continuous frequency distributions analogous to histograms (kernel density estimates with a guassian kernel and bandwidth $h = 1.06*sn^{-1/5}$ following Silverman (1986), with s the standard deviation of trait

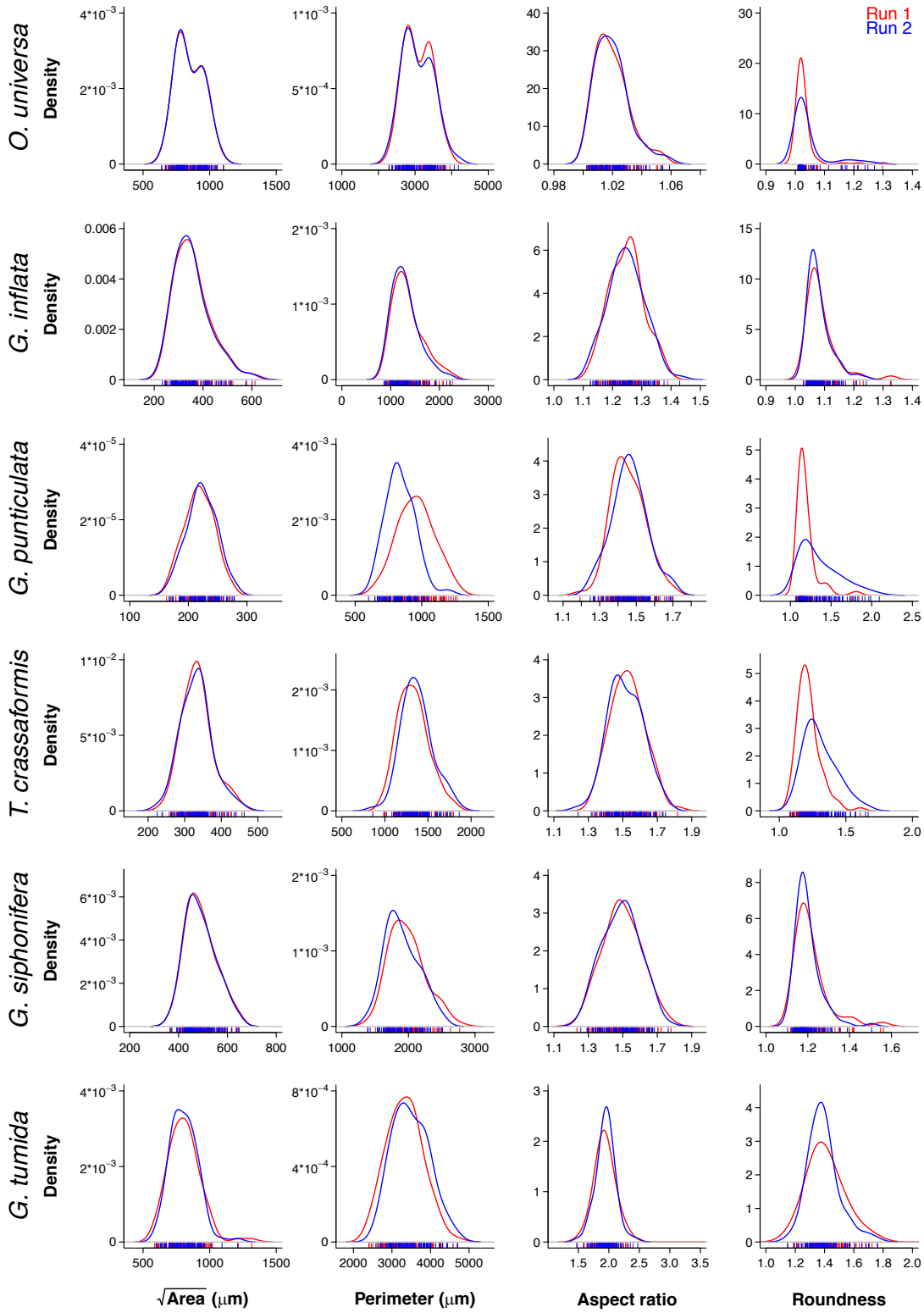

Figure 2.3 Kernel density plots of first (red) and second (blue) set of morphometric measurements on *Orbulina universa*, *Globoconella inflata*, *Globoconella puncticulata*, *Truncorotalia crassaformis*, *Globigerinella siphonifera* and *Globorotalia tumida*. Rug plots on the horizontal axes indicate individual data points.

measurements per species and *n* the number of analysed individuals). To determine which traits are repeatable in which species a Wilcoxon signed-rank test was performed in R (R Core Team, 2013), which is a non-parametric test that compares repeated measurements on a single sample to assess whether their mean ranks differ. If the first and second runs are statistically similar the measurement differences within individuals are expected to be centred around 0 and show little error. Differences significantly deviating from 0 indicate a systematic offset between measurements in the first and second run, implying high sensitivity to small differences in test orientation and low trait repeatability.

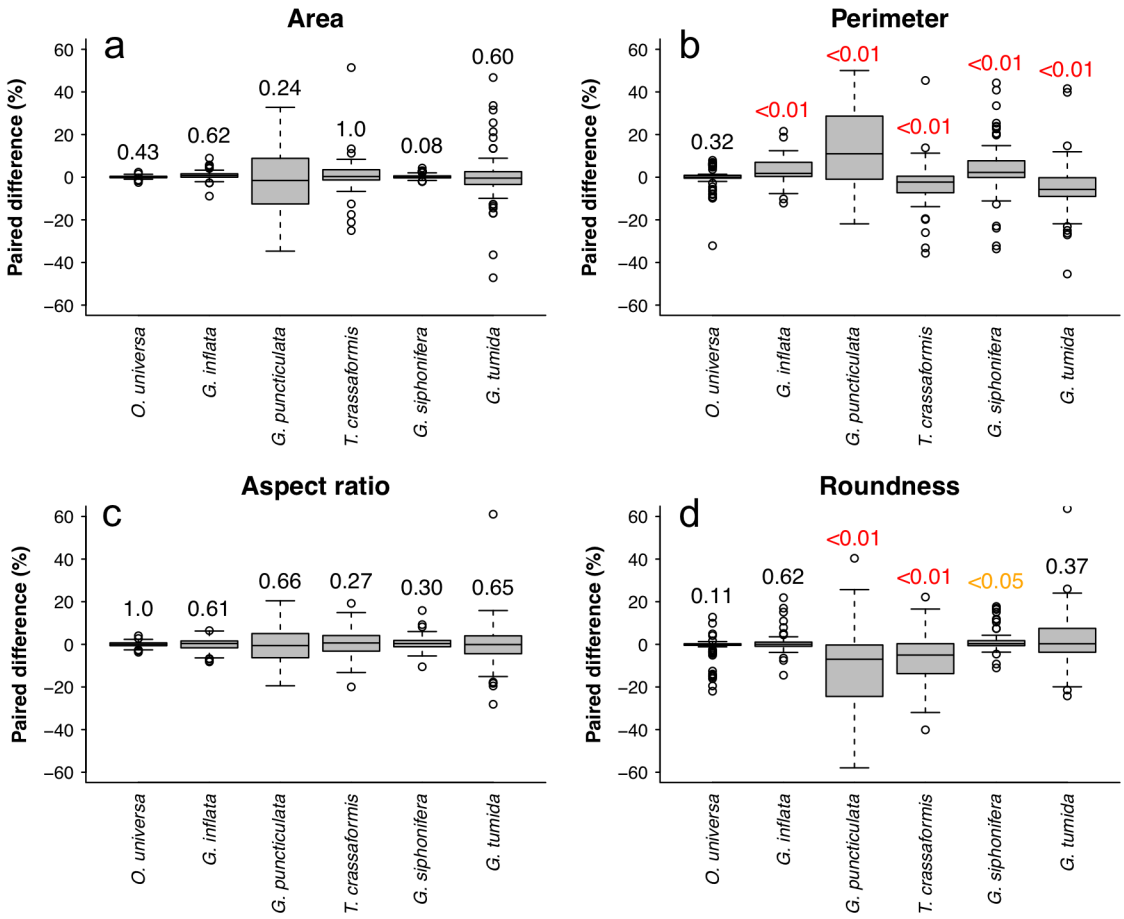

Figure 2.4 Barplots of the difference between repeated measurements on a) area, b) perimeter, c) aspect ratio and d) roundness on the same individual (paired difference) expressed as percentage of the individual's trait mean. p-values of the Wilcoxon signed-rank test performed on subsequent measurements on the same individuals are given for every species, with p-values smaller than 0.01 indicating significant differences between runs shown in red.

If a trait is shown to be repeatable, I determine the number of individuals required to reliably detect a change in a given trait using power analysis (Cohen, 1988). This number is influenced by both the natural variability within a species, with higher variability requiring a larger number of individuals to detect a given trend, as well as variation induced by small mounting errors. Both kinds of variability are present in our dataset, but because it is impossible to perfectly mount specimens it is not possible to separate these effects in our dataset. However, high repeatability suggests that most of the observed variation is due to natural trait variability as opposed to procedural errors, whereas it delimiting the two becomes more troubling when repeatability is low. I apply power analysis with the 'pwr' package in R (R Core Team, 2013). When variation in a population is known, power analysis calculates the sample size required to detect a specified trend (effect size) for a given power (probability of finding a true effect) and significance level (probability of finding that an effect that is not there). I use power analysis to determine the sample size required to detect a trait change of 5, 10, 15, 20, 25 and 30% for varying power values with a significance level set to 0.01.

2.4 Results & Discussion

2.4.1 Area

Measurements of test area vary little between runs (Figure 2.3). Differences between first and second measurements per individual were very small with the Interquartile Range of the differences (the distance between the 25th and 75th percentiles) reaching less than 5% away from the species' mean for all studied species except *G. puncticulata* (Figure 2.4a) and no significant differences were detected between runs for all studied species (Wilcoxon signed-rank tests, see Figure 2.4a for species-specific p-values). These results imply that foraminifera body size is a repeatable measure not dependent on small mounting errors. Because of this trait's high repeatability, relatively small size changes can be detected reliably (Figure 2.5): only 50 individuals are needed to detect an increase in test area of 10% or larger for power, i.e., the probability of detecting an effect that is present in the data, >0.9 and a confidence level, i.e. probability of a false

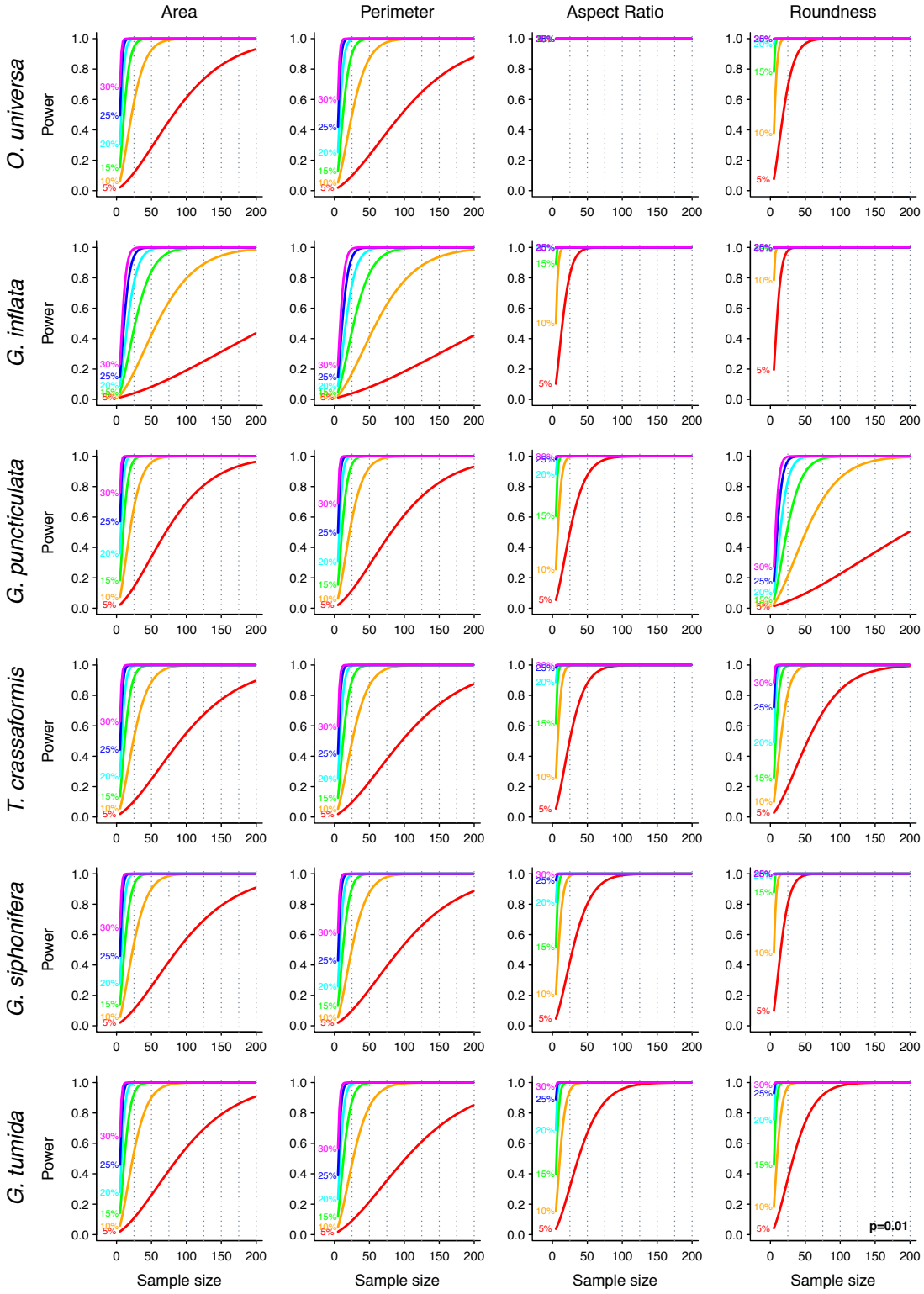

Figure 2.5 Calibration of the number of individuals required to detect a given trait change in individual species. Here power is plotted against the number of individuals needed to detect changes in trait values by 5% (red), 10% (orange), 15% (green), 20% (cyan), 25% (blue) and 30% (magenta). The significance level is set to p=0.01.

positive, of p=0.01 (Figure 2.5). The only exception is *G. inflata*, for which the same number of individuals would only enable a detection of a >15% change in size. In principle, this lower sensitivity in *G. inflata* can be explained by either higher natural and/or higher mounting-induced variability in this species, but given that this species is relatively easy to mount because of its rounded periphery, and mounting induced errors are very low in *G. inflata* size (Figure 2.4a), I conclude that a high natural size variability due to species-specific morphological variability such as the existence of different morphotypes (Kennett and Srinivasan, 1983) is the most likely explanation for the observed differences.

2.4.2 Aspect ratio

Measurements of aspect ratio are similar between runs for all species (Figure 2.3, third column) and repeated measurements on the same individuals are statistically indistinguishable (Figure 2.4c), Wilcoxon signed-rank tests), implying that aspect ratio is a repeatable measure of foraminifera test shape. Results from power analysis show that only 25 individuals are needed to detect a 10% increase in aspect ratio for power>0.9 and a confidence level of 0.01 for all species (Figure 2.5).

2.4.3 Perimeter

Measurement distributions of test perimeter vary between runs for all species (Figure 2.3). The differences between first and second measurements on individuals deviate significantly from 0 in all species except the spherical *O. universa* (Wilcoxon singed-rank tests, see Figure 2.4a for species-specific p-values), implying that test perimeter is not a repeatable measure in the other five species. These results underline the need for species-specific error quantification when using this measure, especially when used in a full morphometric approach where landmarks for other traits are assigned at specific points on the test outline. When this approach is used the repeatability of each landmark should be quantified separately, because uncertainty in the perimeter can also influence the repeatability of the individual landmarks and their associated traits.

2.4.4 Roundness

The similarity of the measurement distribution for roundness varies among the studied species, with most repeatable measurements for *O. universa*, *G. inflata*

and *G. siphonifera* (Figure 2.3). The individual differences between first and second measurements deviate significantly from 0 in *G. puncticulata, T. crassaformis* and *G. siphonifera*, implying that roundness is not a repeatable trait in these species Figure 2.4d). In *O. universa*, *G. inflata* and *G. tumida*, however, I found no significant mounting-induced errors. For each of these three species, power analysis shows that fewer than 25 individuals are required to detect a 10% change in roundness with power>0.9 and a significance level of p=0.01 (Figure 2.5).

The varying degrees of repeatability in roundness could reflect the composite nature of the trait: roundness is defined as $\pi * Perimeter^2/Area$. Because test perimeter is greatly influenced by orientation errors, its biases are also expected to influence test roundness. These results imply that extra care should be taken when analysing more complex composite traits, and that the reliability of all separate traits should be quantified prior to interpreting any changes in roundness in the fossil record.

2.5 Conclusion and Recommendations

In this chapter I report a test of the repeatability of four measures of planktonic foraminifera: size, shape and outline and the sample size required to pick up trends of a given magnitude. This chapter present a novel mounting technique using a glass slides that reduces background imperfections and increases accuracy of trait capture (Figure 2.2). Both test area and aspect ratio are repeatable measures of test size and shape, whereas roundness is a repeatable measure for *O. universa*, *G. inflata* and *G. tumida* but not *G. puncticulata*, *T. crassaformis* and *G. siphonifera*, while perimeter is not repeatable for any of our non-spherical species (Figure 2.4). Our results underline the need for measurement error quantification in individual species' traits prior to interpreting their morphological records. In particular, test perimeter and the other composite traits it influences should be applied with extreme caution. Work is needed to investigate the repeatability of individual landmarks on test outline before they are applied for evolutionary or biostratigraphical purposes.

Results from the power analyses show that between 25 and 50 individuals are needed to detect a 10-15% change in the repeatable traits, which is well within the scope of most species of planktonic foraminifera. A significance threshold of

p=0.01 was used because of the abundance of the microfossil record: it is recommended that micropalaeontologists target lower significance levels (e.g., $p<0.01$ rather than $p<0.05$), particularly in common species, to reduce the probability of reporting false positive results. The sample size required to detect statistically significant trait changes depends on the magnitude of change, and should therefore be determined at the start of each experiment separately to ensure efficient data collection protocols. Focussing on repeatable traits will also ensure that statistical outputs like effect size, which are arguably more informative than the level of statistical significance in inferring the ecological role of trait changes, can be estimated more accurately.

Chapter 3: Calibrations of test diameter and area measured from two-dimensional images as proxies for body size in the planktonic foraminifer *Globoconella puncticulata*

This chapter is in preparation for submission to the *Journal of Foraminiferal Research* as: Brombacher, A., Elder, L., Hull, P.M., Wilson, P.A., Ezard, T.H.G., Calibrations of test diameter and area measured from two-dimensional images as proxies for body size in the planktonic foraminifer *Globoconella puncticulata*

3.1 Abstract

Body size is one of the most commonly measured traits in ecology and evolution because it covaries with environmental (e.g., temperature, latitude, degree of population isolation) and life-history (metabolic rate, generation time) traits. However, the driving mechanisms of body size variation in deep time are less well known and complicated by partial specimen recovery, limited population-level sampling, and the use of linear measurements as proxies for three-dimensional volumetric size data. How much information are we missing by focussing on sub-optimal size metrics? Here I examine this question in an evolving lineage of planktonic foraminifera. I measure test diameter and area of over 500 individuals of the species *Globoconella puncticulata* using two-dimensional images, and compare the results to measurements of test volume of the same individuals as measured by a recently developed high-throughput method for analysing three-dimensional morphometrics, as well as high-resolution three-dimensional computed tomography scanning. Our results show that even in a lineage showing substantial morphological change, test area can provide consistent proxies for body size measured volumetrically. Approximating volumes with linear size measurements is more problematic, systematically over- and underestimating for the smallest and largest tests, respectively. Shape only explained marginally more variation when included in the regressions. The use of light microscopy introduces a small degree of scatter in the data, but the number of individuals

necessary to detect trends in body size with sufficient statistical power is comparable to the sample size required for other traits. These results imply that even in an evolving lineage undergoing substantial morphological change, cross-sectional area can provide consistent proxies for body size.

3.2 Introduction

Body size is an ecologically important trait that scales with many morphological, functional and life history traits (Huxley, 1932; Peters, 1983). It has been studied extensively in modern populations, showing correlations with ecological parameters such as latitude (Bergmann's Rule, (Bergmann, 1848; James, 1970)) and energetic limitations (Clavel and Morlon, 2017; Foster, 1964; Van Valen, 1973b). Additionally, body size has been argued to increase along species branches in deep time (Cope's Rule (Cope, 1887)). However, due to the low resolution of most fossil records and poor preservation of complete specimens, factors influencing body size in deep time are less well known.

Planktonic foraminifera are ideally suited to study evolutionary processes due to the availability of temporally resolved deposits of complete specimens from sites throughout the global ocean. Their abundance, distribution, and relatively continuous deposition in sea floor sediments enable studies of both micro- and macroevolutionary processes in deep time. Body size can be accurately measured in three dimensions with methods like micro-computed tomography (micro-CT) scanning or Synchrotron radiation X-ray tomographic microscopy (SRXTM) (Caromel et al., 2016; Schmidt et al., 2013), but the expensive and time-consuming nature of these methods constrain their utility for generating large datasets to study of evolutionary change. As a result, specimen body size is generally estimated using test diameter or area as measured from two-dimensional images. However, condensing complex shell morphology to one- or two-dimensional parameters results inevitably in a loss of information. Morphological changes over a species' existence could distort the diameter or area to body size ratio, for example by changing chamber inflation (Coxall et al., 2000; Norris and Hull, 2012) and this covariation is not constant through time (Brombacher et al.). Therefore, to reliably use two-dimensional representations of shell morphology the loss of information needs to be quantified.

Here I measure test diameter, two-dimensional area, and three-dimensional volume of the same 539 individuals of the species *Globoconella puncticulata*. This

species has been reported to undergo drastic size and shape change during the 500,000 years preceding its extinction (Brombacher et al. (, see also Figure 3.1), providing an excellent opportunity to study the robustness of one-, two-, and three-dimensional representations of body size with changing morphology. Test diameter and two-dimensional area were analysed from two-dimensional images. Those results were then compared to measurements of test volume of the same individuals using a recently developed high-throughput method for analysing three-dimensional morphometrics from vertically stacked images (Hsiang et al., 2016), as well as CT-scans (Schmidt et al. 2013).

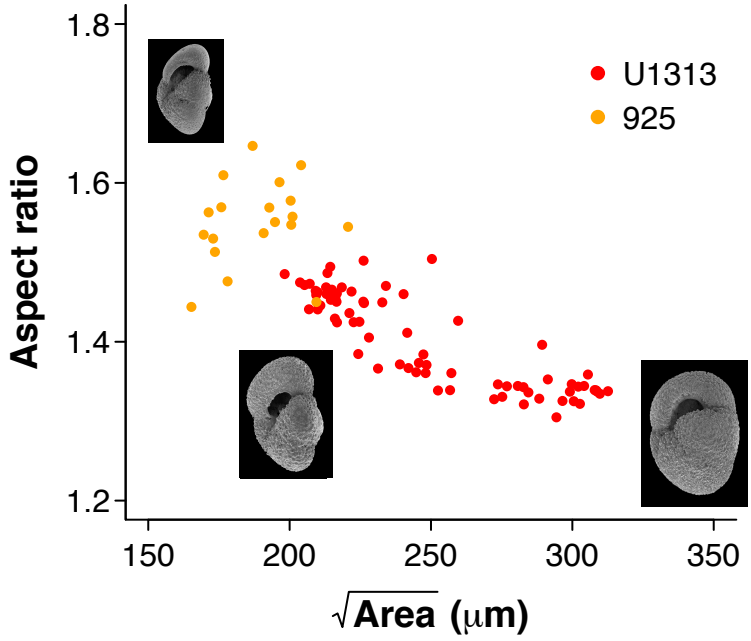

Figure 3.1 Mean size and shape per time slice of *Globoconella puncticulata* at Sites U1313 and 925, as measured from two-dimensional test representations. U1313 data from Brombacher et al. (.

3.3 Methods

3.3.1 Material

To test the relationship of one-, two-, and three-dimensional representations of body size I measured test diameter, area and volume of the planktonic foraminifer *Globoconella puncticulata*. This thermocline-dwelling species is characterised by a low trochospiral test with a flattened spiral side and highly vaulted umbilical side (Kennett and Srinivasan, 1983).

Samples were obtained from two sites in the Atlantic Ocean: equatorial Ocean Drilling Program (ODP) Site 925 and mid-latitude Integrated Ocean Drilling Program (IODP) Site U1313 (Figure 3.2). The specimens at Site U1313 become smaller and more rectangular-shaped in the interval leading up to their extinction. Samples from IODP Site U1313 were selected from three different time slices at 2.5, 2.7 and 2.8 Ma to include the full range of morphological variability. Individuals at Site 925 were smaller and less abundant than at Site U1313. No changes in morphology were observed at Site 925 (unpublished data), so 10 equally spaced samples were analysed over the interval 2.4 – 2.6 Ma. Combining data from both sites provides a wider size spectrum of individuals with which to compare consistency among one-, two- and three-dimensional proxies.

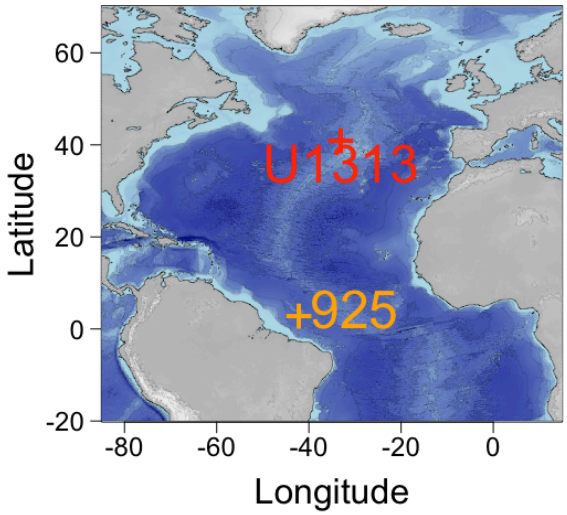

Figure 3.2 Location of ODP Site 925 and IODP Site U1313 in the Atlantic Ocean

All samples were dry-sieved over a >150 μm² mesh sieve. Due to the high abundance of *G. puncticulata* at Site U1313 these samples were split using a microsplitter until a single split contained 70-150 specimens. Full splits were picked to avoid manual selection bias, resulting in 471 individuals picked from this site. All individuals of *G. puncticulata* present in the samples of Site 925 were picked, yielding another 68 specimens.

3.3.2 Measurements

Individuals were mounted on glass slides using transparent double-sided tape (Brombacher et al., 2017). For the extraction of test diameter and area from two-dimensional images, specimens were mounted in side view to capture chamber

inflation, imaged using a Lumenera Infinity 3 camera attached to an Olympus SZX10 light microscope and analysed using the Image Pro Premier software. Specimens were automatically identified by setting a brightness threshold that detects the white foraminifera tests from the dark background, and diameter (1D) and area (2D) were measured from the objects. Additionally, aspect ratio (length/width) was measured as a repeatable proxy for test shape (Brombacher et al., 2017).

Test volume was measured from the same individuals using the AutoMorph software package developed by Hsiang et al. (2016). Slides were imaged with transmitted light using a Leica DFC450 camera attached to a Leica Microsystems DM6000M compound microscope with a motorized *x*-*y* scanning stage. Each slide scan generates a stack of raw slide images (called planes) at different *z*-axis heights. All slides in this study were imaged with a *z*-step size (distance between planes) of 31.1µm and an imaging range of 800µm resulting in 25 focused planes. AutoMorph then isolates the focused planes for each individual foraminifera using the module `segment`. AutoMorph's `run3dmorph` module creates a three-dimensional representation of the specimen's visible upper half, while the invisible lower half is represented by a dome, cylinder or cone shape (see Hsiang et al. (2016) for further details). To compare volume estimates of foraminifera tests mounted in different positions (umbilical, spiral and edge view) and with different lower half representations, size measurements of six specimens acquired with this technique were compared to volume estimates of the same individuals acquired through CT-scanning.

3.3.3 Analysis

To determine the degree of scatter between test diameter and area as measured from two-dimensional images, and to test the consistency of test diameter and area as proxies for foraminifera body size I regressed all measures using Generalised Linear Models (GLMs) in R (R Core Team, 2013). Comparing one-dimensional to two-dimensional objects a polynomial relationship of the form $y = ax^2$ is expected, with increasing variation with the mean, which was incorporated in the GLM via a power link function (Figure 3.3a). To test the consistency of test diameter and area as proxies for foraminifera body size I compared both to the volume measurements performed on the same tests using a GLM with the expected power functions ($y = bx^3$ for one-dimensional diameter to three-dimensional volume, and $y = cx^{3/2}$ for two-dimensional area to three-dimensional

volume respectively) as link fucntions, and the variance set to increase linearly with the mean. To study the influence of varying test shapes on the regression, both GLMs were repeated with shape (as represented by aspect ratio) and the interaction between size and shape included in the models. R^2 values were then compared between models with and without aspect ratio to calculate the added variance explained by test shape. Power analysis (Cohen, 1988) was employed to calculate the number of individuals required to detect changes in body size given a certain power (probability of finding a true effect) and significance level (probability of a false positive).

Table 3.1 Mean Squared Error and mean deviation from the CT-scan estimates of body size for tests mounted in edge, spiral and umbilical view with dome, cylinder and cone underside representations.

Orientation	Underside	Mean Squared Error ($*10^{12}$)
Edge	Dome	7.5
Edge	Cylinder	14.5
Edge	Cone	6.2
Spiral	Dome	7.4
Spiral	Cylinder	13.0
Spiral	Cone	4.7
Umbilical	Dome	6.8
Umbilical	Cylinder	7.1
Umbilical	Cone	7.7

3.4 Results & Discussion

Mean Squared Error analysis showed that individuals oriented in spiral view, with a cone shape replacing the invisible underside came closest to the CT-scan values (Table 3.1). Therefore, this orientation will be used to represent test volume in this study. Same-shell measurements of diameter and area from two-dimensional images relate to their corresponding volume measurements as expected. Plotted against shell volume, diameter follows a power function of the form $y = bx^3$ (GLM, power(1/3) link function and a μ error distribution that allows the variance to increase with the mean, $p<0.001$, $R^2 = 0.71$, Figure 3.3b), although low diameter values are overestimated by the power law regression and high diameter values underestimated. Area is described by the function $y = cx^{3/2}$ (GLM, power(2/3) link function and a μ error distribution, $p<0.001$, $R^2 = 0.73$, Figure 3.3c) and provides a consistent estimate for volumetric body size across the data set (Figure 3.3c).

These results suggest that, as a two-dimensional representation, test area provides the most consistent representation of foraminiferal size.

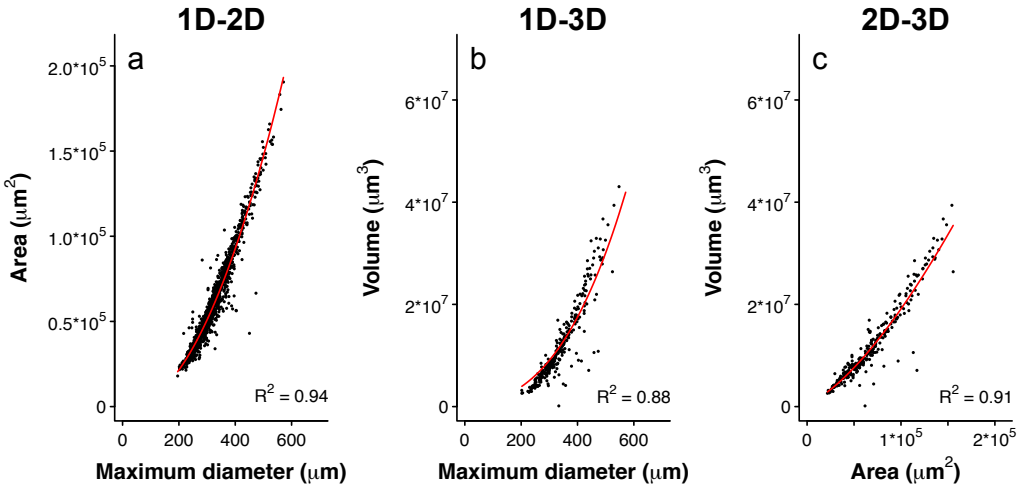

Figure 3.3 Scatter between measurements of diameter and area (a), diameter and volume (b) and area and volume (c) of *Globoconella puncticulata*. Red lines indicate the Generalised Linear Model regressions with specific power link functions of the form $y = ax^2$ (1D-2D), $y = bx^3$ (1D-3D) and $y = cx^{2/3}$ (2D-3D).

Test shape only explains up to 2% extra variance in both datasets (GLM, $R^2 = 0.89$ and $R^2 = 0.91$ for test diameter and area respectively, Table 3.2). These results imply that test diameter and area as measured from two-dimensional images are consistent proxies for body size for the varying test sizes and shapes observed along the lineage. The remaining ~10% of the variance could be explained by scatter inherent to both methods. Comparisons between diameter and area measured from two-dimensional images using the same software show 6% of unexplained variance (Figure 3.3a).

Table 3.2 Results of Generalised Linear Models for 1D to 2D, 1D to 3D and 2D to 3D comparisons, with and without test shape. Note equivalent statistical significance ($p < 0.001$) but quite different extents of variance explained.

Model	*p*	R^2
Area~Diameter	<0.001	0.94
Volume~Diameter	<0.001	0.87
Volume~Diameter*Aspect Ratio	<0.001	0.89
Volume~Area	<0.001	0.91
Volume~Area*Aspect Ratio	<0.001	0.91

To quantify the influence of the observed scatter on evolutionary time series power analysis (Cohen, 1988) was employed to calculate the number of individuals required to detect trends in body size, using the individuals mounted in side view with a cone-shaped lower half. The results show that 15 individuals are needed to detect a 5% change in size for power >0.95 at a confidence level of 0.01. This sample size is well within the feasible range for microfossils and is comparable to the number of individuals needed to detect changes in various other traits (Brombacher et al., 2017).

The regressions presented here are specific to *G. puncticulata*, and most readily extrapolatable to similar morphologies. The orientation and lower half representation that best describes an object in 3dmorph depends on the test shape and complexity of individual species and should therefore be determined prior to comparisons between one-, two- and three-dimensional test measurements. However, establishing the best orientation and shape for any given morphology requires only a handful of specimens for CT-scanning, and the ~500 individuals to quantify the regression between diameter or size and volume measurements can be measured by Automorph within a few days. Therefore, quantifying the accuracy of body size as represented by varying measures in any given species can be performed relatively quickly using the setup presented here.

3.5 Conclusion

Our results show that even in a lineage showing substantial morphological change (Brombacher et al.), test area can provide consistent proxies for body size measured volumetrically (Figure 3.3). Approximating volumes with linear size measurements is more problematic, as even though the R^2 value remains high, the smallest and largest tests are systematically over- and underestimated respectively (Figure 3.3b). The use of light microscopy introduces a small degree of scatter in the data, however the sample size required to detect trends in body size is comparable to the number of individuals needed to detect changes in other traits, implying that no higher-resolution methods are needed to accurately quantify body size. The method presented here can be readily repeated on other species of planktonic foraminifera, validating their potential to study the driving mechanisms of changing body size in deep time.

Chapter 4: The breakdown of static and evolutionary allometries during climatic upheaval

This chapter is a reproduction of an online early article in *The American Naturalist*: Brombacher, A., Wilson, P.A., Bailey, I., Ezard, T.H.G. (2017), The breakdown of static and evolutionary allometries during climatic upheaval. The American Naturalist, 190, 3. *Online early.*

4.1 Abstract

The influence of within-species variation and covariation on evolutionary patterns is well established for generational and macroevolutionary processes, most prominently through genetic lines of least resistance. However, it is not known whether intraspecific phenotypic variation also directs microevolutionary trajectories into the long term when a species is subject to varying environmental conditions. Here I present a continuous, high-resolution bivariate record of size and shape changes among 12,633 individual planktonic foraminifera of a surviving and an extinct-going species, over 500 thousand year time interval. This time interval spans the late Pliocene to earliest Pleistocene intensification of Northern Hemisphere glaciation, an interval of profound climate upheaval that can be divided into three phases of increasing glacial intensity. I found that within each of these three Plio-Pleistocene climate phases the within-population allometries predict evolutionary change from one time-step to the next, and that the among-population (i.e. evolutionary) allometries match their corresponding static (within-population) allometries within specific climate phases. However, the evolutionary allometry across the three climate phases deviates significantly from the static and phase-specific evolutionary allometries in the extinct-going species. Although intraspecific variation leaves a clear signature on mean evolutionary change from one time-step to the next, our study suggests that the link between intraspecific variation and longer-term micro- and macroevolutionary phenomena is prone to environmental perturbation that can force evolutionary change away from the direction of trait constraints induced by within-species trait covariation.

4.2 Introduction

Intraspecific variation is essential for evolution. Phenotypic variation is the target for natural selection, with the possibilities for phenotypic change determined by the extent of genetic variation. Many traits do not evolve independently: selection on one trait can influence the response to selection in others (Lande, 1979; Lande and Arnold, 1983). These constraints among traits can have large impacts on the direction of evolution, either facilitating evolution in the case of positive covariances (Gavrilets and Losos, 2009), or constraining adaptation when covariances are negative (Agrawal and Stinchcombe, 2009). Lande (1979) used a multivariate quantitative genetics framework to describe brain: body size allometry in mammals. From one generation to the next, the multivariate Breeder's equation (Lande, 1979) predicts the change in mean n-dimensional phenotypes as:

$$\boldsymbol{z} = \boldsymbol{G}\beta \tag{1}$$

Here $\boldsymbol{z}$ is an n-dimensional vector representing the change in n trait means, **G** is the additive genetic variance-covariance matrix with genetic variances on the diagonal elements and covariances as non-diagonal elements, and β is an n-dimensional vector consisting of directional selection gradients (Lande, 1979; Lande and Arnold, 1983). Repeated over many generations, the phenotype is expected to evolve in the direction of the dominant eigenvector of **G**, which has been defined as the line of least genetic resistance $\mathbf{g}_{max}$ (Schluter, 1996). Although populations have been shown to evolve along lines of least resistance on generational time scales (Bégin et al., 2003; Lande, 1979; Lande and Arnold, 1983), as well as among modern species and genera (Lande, 1979; Schluter, 1996), we lack robust data on how trait covariance evolves along species branches in deep time.

Allometries present specific examples of trait covariations, with a given trait covarying with body size according to a power relationship (Huxley, 1932). They are considered exemplary lines of least evolutionary resistance (Pélabon et al., 2014) because allometric constraints imply that internal growth regulators restrict trait evolution. Marroig and Cheverud (2005) showed that the direction of evolutionary change in New World Monkeys mainly occurred along within-population allometric slopes of body size and cranial features, using morphological data of extant species supplemented by modelled reconstructions of ancestral morphology. One of the challenges to empirical studies of how trait

covariance evolves in free-living populations, and an explanation for the "dismally few" empirical tests (Estes and Arnold, 2007), is that studies using recent species are restricted to a comparison of relatively few matrices on contemporaneous populations (Arnold et al., 2008). Empirical studies of phenotypic change over long time scales require the fossil record, which typically lacks genetic information. Quantitative morphological traits represent a multi-locus representation of intraspecific variation, and have been shown to influence adaptive evolution (Simpson, 1953). Over long timescales, assuming a constant selection regime (Arnold et al., 2008) (and mutational pleiotropy notwithstanding (Cheverud, 1996)), evolutionary constraints due to phenotypic trait covariation provide a good approximation of the adaptive landscape (Arnold et al., 2001). While pleiotropy is often present in mutation input, this observation implies that it is instructive to study changes in the relationships among traits.

Hunt (2007) and Renaud et al. (2006) applied the concept of Schluter's (1996) lines of least resistance to the fossil record. Hunt (2007) found that speciation in an ostracode clade spanning ~37 Myrs tended to occur in the direction of maximum phenotypic variation of the ancestor population. Renaud et al. (2006) showed that dental patterns in rodents were channelled along the direction of greatest intraspecific variation over an 11 Myr-interval. Additionally, Haber (2016) showed that traits evolve fastest if divergence is aligned with the phenotypic covariance matrix in ruminant skulls. Phylogenetic comparative methods have identified changing trait covariance as a key diagnostic among distinct subclades during the origin of birds (Puttick et al., 2014). These four studies focus on changing trait covariances on macroevolutionary time scales, but no observations have been reported of how trait covariance evolves during a species' existence.

The allometric constraint hypothesis states that within-population allometries typically remain constant, shaping evolutionary constraints over longer time scales (Pélabon et al., 2014). Firmat et al. (2014) showed that static (within-population) allometries predict evolutionary (i.e. among-species) allometries on <1 million-year time scales, implying consistent alignment between **G** and the adaptive landscape (Arnold, 1992; Cheverud, 1984; Lande, 1980). However, a shift in evolutionary optima would upset this alignment, forcing the individuals and population into an alternative selective regime. Under such a displaced optimum (Estes and Arnold, 2007), the consistency between within-population and evolutionary allometries might be expected to break down.

Renaud et al. (2006) reported that among-species morphological variation in one of their two studied lineages departed from the lines of least resistance following pronounced environmental change. Their sampling resolution, however, did not allow for the analysis of the evolution of static allometries within species over time. Additionally, Hunt (2007) showed that evolution in an ostracode clade did follow the lines of least resistance initially, but the effect eroded after a few million years. To study changes in static allometries as a response to changing selective gradients and their effect on evolutionary allometries, high-temporal resolution records of individual species are needed.

Here, I analyse size and shape allometries from 12,633 individuals in two ecologically and environmentally similar planktonic foraminifera species found at Integrated Ocean Drilling Program (IODP) Site U1313 (~41 ° N) situated in the mid-latitude North Atlantic Ocean. Planktonic foraminifera are sexually reproducing protists distributed in high abundance throughout the world's oceans. The large population size, global distribution and excellent preservation potential of their calcite shells make them uniquely suited for continuous, high-resolution morphological reconstructions over millions of years. I explore the temporal consistency of allometries within and among populations during the most recent great climate transition in Earth's history: the late Pliocene to earliest Pleistocene intensification of Northern Hemisphere glaciation (iNHG) (Mudelsee and Raymo, 2005). I investigate two main questions. First, do allometric lines of least phenotypic resistance predict evolutionary change within species over thousands of generations? Second, do within-sample (static) and among-species (evolutionary) allometries become decoupled during climate upheaval?

4.3 Material & Methods

4.3.1 Study species

This chapter focusses on two ecologically similar species, *Truncorotalia crassaformis* and *Globoconella puncticulata* (Figure 4.1, Figure 4.2 (inserts)), characterised by low trochospiral tests with flattened spiral sides, inflated umbilical sides and umbilical-extraumbilical apertures (Kennett and Srinivasan, 1983). Both inhabit similar habitats with highest abundances in thermocline waters at middle and low latitudes (Aze et al., 2011; Kennett and Srinivasan,

1983). *T. crassaformis* originated around 5.7 Ma and survives to the present day. *G. puncticulata* first appeared around 4.6 Ma and became extinct at 2.41 Ma (Wei, 1994b), shortly after the onset of significant Northern Hemisphere glaciation at 2.72 Ma (Bailey et al., 2013). Wei (1994a) has reported allometric changes in the *G. puncticulata* lineage on macroevolutionary time scales. In particular, he noted that *G. puncticulata* gave rise to its only descendant species, *G. inflata,* through shifts in ontogenetic allometric relationships starting around 3.5 Ma in the southwest Pacific (Wei, 1994a, b). *G. inflata* did not occur in the North Atlantic until ~2.09 Ma (Berggren et al., 1995a; Chapman et al., 1998), ~30,000 years after the extinction of *G. puncticulata*. The planktonic foraminifera biostratigraphy of Site U1313 reports no overlap between the local stratigraphic ranges of *G. puncticulata* and *G. inflata* (Channell et al., 2006b) and no individuals of *G. inflata* were found in our higher resolution sample set of the study's target time interval. The extirpation of *G. puncticulata* from the North Atlantic is therefore not attributable to replacement by a descendant species.

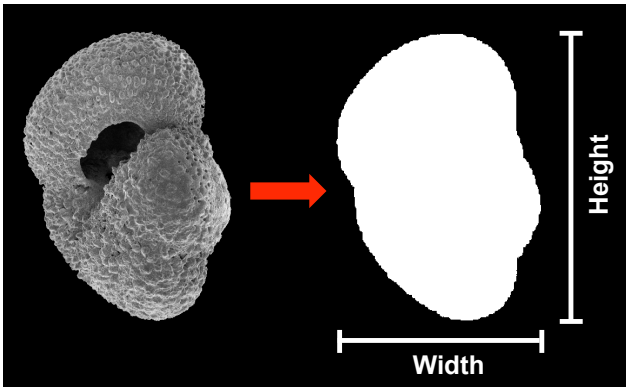

Figure 4.1 Analysed traits on foraminifera shells. Specimens were imaged in edge view (left), and area, height and width were measured on the 2-dimensional shell representation (right).

4.3.2 Material

IODP Site U1313 is located in the mid North Atlantic Ocean at the base of the upper western flank of the Mid-Atlantic Ridge at a water depth of 3426 m (41 °N, 32.5 'W), at the northern edge of the North Atlantic Gyre. In this oceanographic environment, surface currents are driven by clockwise wind circulation in the high-pressure system overlying the subtropical North Atlantic. Deflected by the Coriolis force, these currents form a basin-wide rotating circulation system with little movement of surface water currents into or out of the gyre, preventing large-scale migration of plankton and generating a well-mixed eco-evolutionary

system. Site U1313 was drilled during IODP Expedition 306 in 2005 and constitutes a reoccupation of Deep Sea Drilling Project (DSDP) Site 607 (Channell et al., 2006b; Raymo et al., 1990) using advanced piston coring techniques to provide continuous records of evolutionary change (Channell et al., 2006b). The Site U1313 record is characterised by consistently high sedimentation rates (~5 cm/kyr) for the past 5 Myr (Channell et al., 2006b; Lisiecki and Raymo, 2005), a demonstrably continuous record of sedimentation for iNHG (Bolton et al., 2010) and exceptionally well-preserved microfossil carbonate (Lang et al., 2014).

The time interval studied here spans 2.4 to 2.9 Ma and captures the intensification of Northern Hemisphere Glaciation (iNHG). We identify three distinct climate states in the studied interval characterised by stepwise increases in glacial state in the face of constant interglacial (background) conditions (Figure 4.2e,f). Before Marine Isotope Stage (MIS) G6 at 2.72 Ma, here called the Initial Phase, Northern Hemisphere ice sheets, as tracked by benthic foraminiferal oxygen isotopes ($\delta^{18}O$), were small (Lisiecki and Raymo, 2005; Mudelsee and Raymo, 2005). From MIS G6 onwards Northern Hemisphere climate became dominated by 41-kyr glacial-interglacial cycles (here called the Transition Phase, 2.54-2.72 Ma), but glacial ice sheets did not reach their full iNHG size until 2.54 Ma with Marine Isotope Stages 96, 98 and 100 (Bailey et al., 2013) (here called the Glacial Phase, <2.54 Ma). Comparing within-population allometries to evolutionary allometries within these three climate phases, as well as over the entire study interval, provide new insights into the evolution of allometric relationships during times of global climatic upheaval.

I used 75 of the samples at every ~30 cm (~5-kyr-resolution) from the shipboard primary splice studied originally by Bolton et al. (2010): 20 samples in the Initial Phase, 30 in the Transition Phase and 25 in the Glacial Phase. The site's average bioturbation depth is estimated to be 2-3 cm (Channell et al., 2006b) implying no time averaging of foraminifera populations between consecutive samples. For age control, I used the orbital-resolution benthic foraminiferal $\delta^{18}O$ stratigraphy for Site U1313 generated by Bolton et al. (2010). The samples were dry-sieved over a >150 μm^2 mesh sieve and split using a microsplitter until a single split contained 70-150 specimens of *T. crassaformis* or *G. puncticulata*. The splits were picked for all specimens of both of these species, resulting in a total of 12,629 specimens (6058 specimens of *T. crassaformis* and 6575 of *G. puncticulata*) over the studied interval. While picking for *T. crassaformis* extra care was taken to exclude specimens from the closely related and

morphologically similar species *Truncorotalia oceanica*, *Truncorotalia ronda*, *Truncorotalia viola* and *Truncorotalia hessi* (Kennett and Srinivasan, 1983; Stewart, 2003). Foraminifera shells were mounted on glass slides in groups of 20 individuals using double-sided adhesive tape with the apertures facing upwards. Groups were imaged using an Infinity 3 Lumenera camera mounted on an Olympus SZX10 light microscope. Because I aim to evaluate changes in traits as a consequence of climate change, only traits with known functional morphologies were analysed. Shell size is known to reflect a species' ecological optimum (Hecht, 1976a; Schmidt et al., 2004a) and shell shape impacts the test volume to surface area ratio, influencing respiration and metabolic processes (Caromel et al., 2014). Shell area and aspect ratio (the ratio between test height and width) were extracted from the images using an automated image analysis macro in the Image Pro Premier software (Figure 4.1). To assess measurement consistency a subset of 100 individuals per species was remounted and reanalysed. Measurement errors were defined as the difference between two repeated measurements on the same individual, divided by the mean of those two measurements. Results from both sets of measurements were highly consistent with average trait errors less than 5% (Table 4.1).

Table 4.1 Measurement errors for size and shape in *Globoconella puncticulata* and *Truncorotalia crassaformis*. Measurement errors were defined as the difference between two repeated measurements on the same individual, divided by the mean of the two measurements.

Species	Measure	Error
G. puncticulata	Area	5.0%
G. puncticulata	Aspect ratio	3.3%
T. crassaformis	Area	2.0%
T. crassaformis	Aspect ratio	2.4%

4.3.3 Analysis

The R package 'paleoTS' (Hunt, 2006) allows analysis of paleontological time series using maximum likelihood models, which was used to test whether univariate trait evolution in the different climatic states is best described by stasis, directional evolution or a random walk. To test whether trait evolution

occurred more slowly than would be consistent with genetic drift, I calculated Lynch's delta metric (Lynch, 1990). Given the observed within-population phenotypic variance, this measure decribes the range of evolutionary rates that would be consistent with neutral evolution (10^{-3} - 10^{-5}). Higher values imply directional selection, whereas lower values support hypotheses of stabilising selection. To quantify trait responses to changes in Northern Hemisphere climate linear models were used to compare the morphology records to the LR04 benthic $\delta^{18}O$ stack representing Northern Hemisphere glacial-interglacial climate cycles (Lisiecki and Raymo, 2005) after taking first differences to account for temporal autocorrelation and for consistency with the random walk approach of the paleoTS analyses.

To study trait covariation over time, phenotypic variance-covariance matrices $\mathbf{P}_t$ were constructed using individual measurements of shell size and shape at time *t* scaled to unit variance to produce measurements in comparable units. The dominant eigenvector of $\mathbf{P}_t$, here called $\mathbf{p}_{max}$, describes maximum phenotypic variation (Hunt, 2007; Schluter, 1996) and, in our case, represents the static allometric size-shape relationship at time *t*. Evolutionary allometries were calculated in the same way using the sample means of size and shape to reconstruct the variance-covariance matrix over the entire study interval (total evolutionary allometries) as well as from separate climate phases (phase-specific evolutionary allometries). To test whether populations preferentially evolve along static allometries from one time step to the next, I measured the angle θ between $\mathbf{p}_{max}$ at time *t* and the direction of evolutionary divergence $\boldsymbol{z}$ from the sample mean at time *t* to the sample mean at time *t+1* (Figure 4.3a), and compared θ to a simulated distribution of angles between two randomly chosen vectors generated using Knuth's Algorithm (Knuth, 1969). A Wilcoxon rank-sum test was performed to check if the distribution of all ϑ differed from the randomly generated angles. A set of angles significantly smaller than the randomly generated angles implies that population evolution from one time step to the next is constrained by its within-population allometries. A paired Wilcoxon test was applied to test if both species were similarly constrained by their allometric lines of least resistance.

Over longer time scales and under a constant selection regime, species are expected to preferentially evolve along $\mathbf{p}_{max}$ (Firmat et al., 2014; Haber, 2016; Hunt, 2007; Schluter, 1996), resulting in evolutionary allometries similar to the within-population allometries (Firmat et al., 2014; Voje et al., 2014). To determine whether within-population allometries also predict evolutionary allometries during

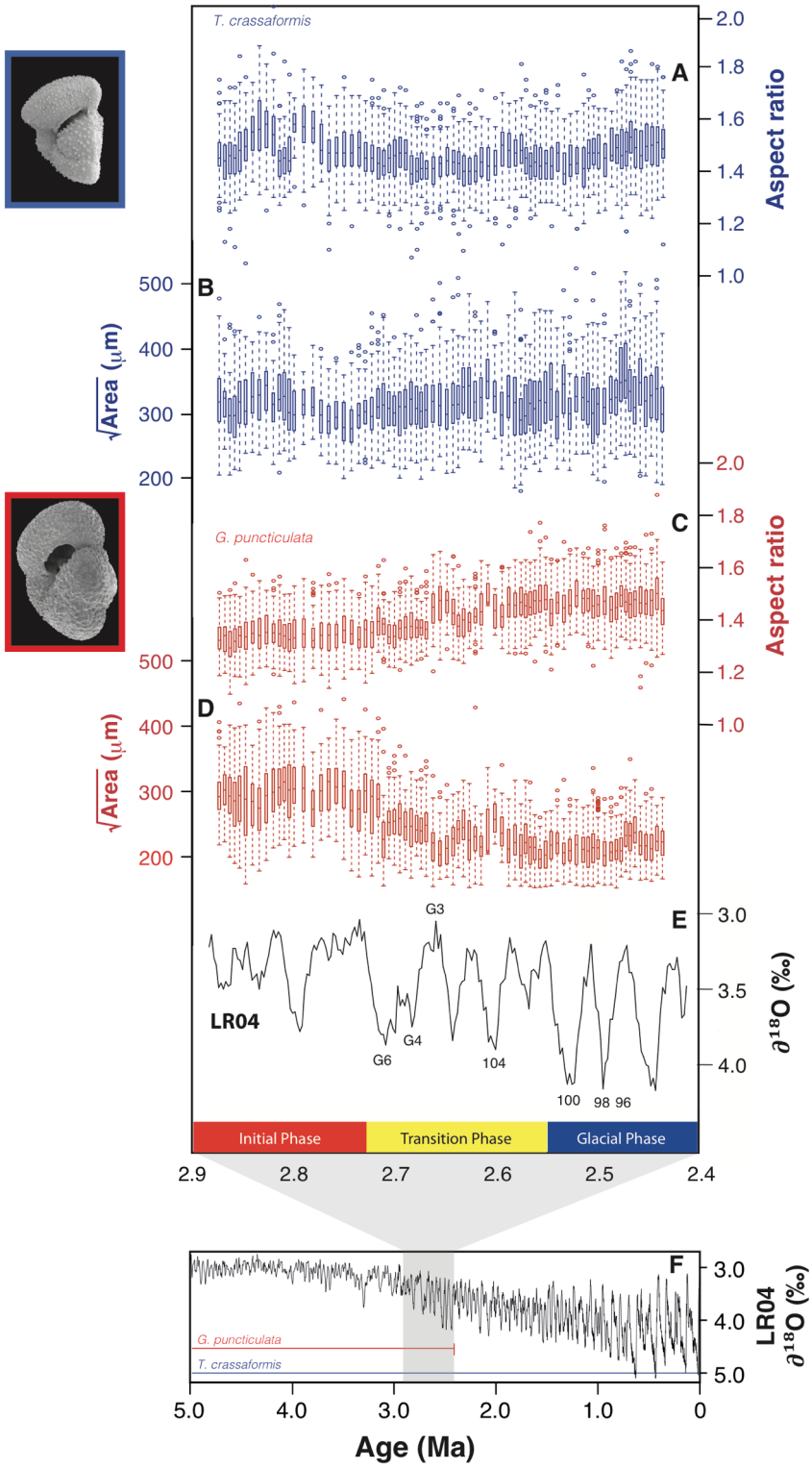

Figure 4.2 Box and whisker plots of a) shape and b) size of *Truncorotalia crassaformis* and c) shape and d) size of *Globoconella puncticulata* at Site U1313 over time. e) and f) represent the global oxygen isotope benthic stack (Lisiecki & Raymo, 2004) for the study interval and the past 5 million years respectively, with key Marine Isotope Stages in e).

intervals of global environmental change, the within-population allometries within climate phases were compared to both the phase-specific evolutionary allometries as well as the evolutionary allometry across the entire studied interval (see Figure 4.3b for an example).

4.4 Results

Time series of size and shape dynamics in both species are presented in Figure 2a-d. We particularly note an abrupt decrease in size by 30±9% in *G. puncticulata* at the start of the Transition Phase during MIS G6. No significant relationships were found between any of the traits and $\delta^{18}O$ (Linear Model, $p = 0.61$, $R^2 = 0.0037$ and $p = 0.99$, $R^2 = 2.2*10^{-4}$ for size and shape of *G. puncticulata* and $p = 0.64$, $R^2 = 0.0031$ and $p = 0.77$, $R^2 = 0.0012$ for size and shape of for size and shape of *T. crassaformis*, see also Figure 4.4), arguing against genetic and plastic species response to climate change. Maximum likelihood models implemented in the paleoTS package provide approximately 60-70% support for a directionless random walk as compared to directional evolution (20-30%) or stasis (0-20%) for all time series over the entire studied interval (Table 1). Analysed within separate climate states, size of *G. puncticulata* is best described by stasis in the Initial

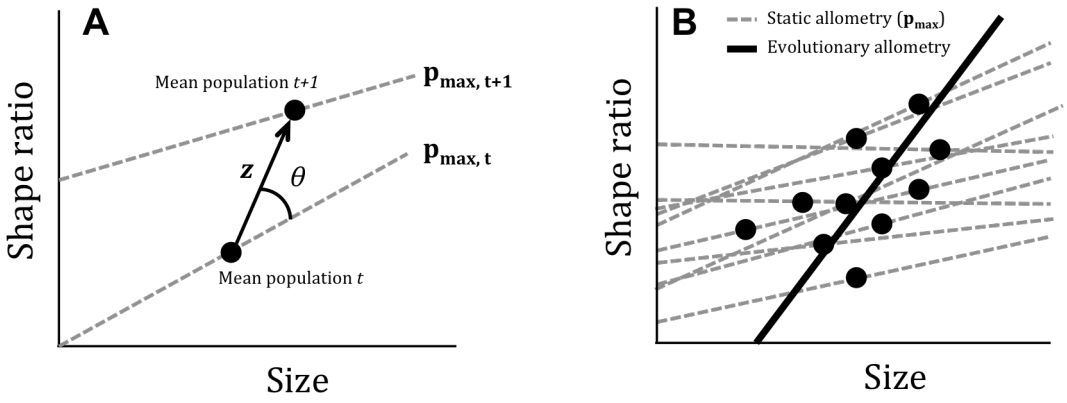

Figure 4.3 Schematic of (a) the angle θ between p_{max} at time *t* and the direction of evolutionary divergence (z) from the sample mean at time *t* to the sample mean at *t+1*. Here p_{max} represents the within-population allometry. (b) Evolutionary allometries were defined as p_{max} of the phenotypic variance-covariance matrix of the sample means, and do not necessarily line up with the within-population allometries of individual samples.

Phase and by a random walk in the Transition and Glacial phases, whereas shape is best described by stasis in the Initial and Glacial Phases, and by a random walk in the Transition Phase. Size of *T. crassaformis* most resembles a random walk in the Initial and Transition Phases and stasis in the Glacial Phase, while shape is best described by a random walk throughout (Table 4.2). Additionally, for all studied traits Lynch's delta values fall well outside the 10^{-3} – 10^{-5} range proposed by Lynch (1990) to represent neutral evolution (see Table 2), implying a form of stabilising selection acting on all traits.

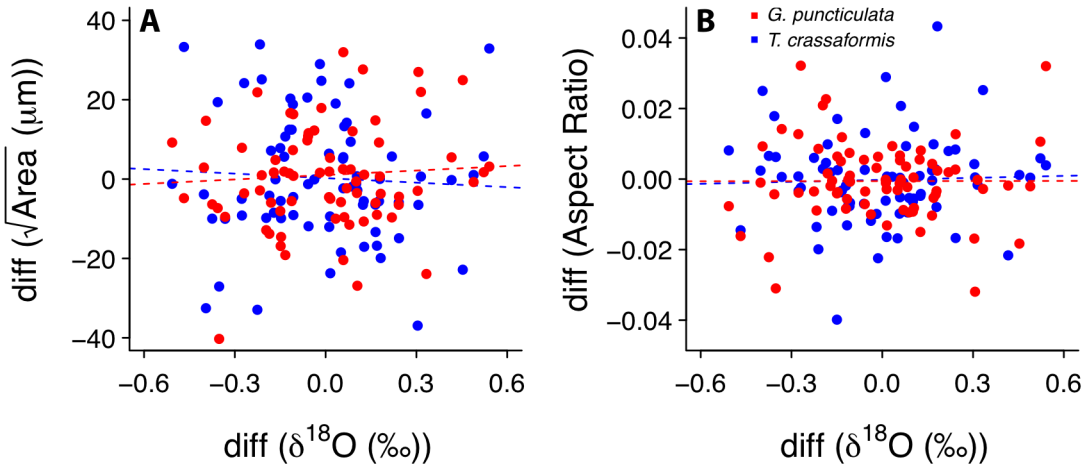

Figure 4.4 Linear regressions of first differences of a) mean size and b) mean shape of *Globoconella puncticulata* (red) and *Truncorotalia crassaformis* (blue) against first differences of the global benthic $\delta^{18}O$ stack (Lisiecki and Raymo, 2005). Correlations between traits and $\delta^{18}O$ were not significant for either species (Linear Model, p = 0.61, R^2 = 0.0037 and p = 0.99, $R^2 = 2.2*10^{-4}$ for size and shape of *G. puncticulata* and p = 0.64, R^2 = 0.0031 and p = 0.77, R^2 = 0.0012 for size and shape of for size and shape of *T. crassaformis*).

T. crassaformis occupies a similar position in size-shape space throughout the studied interval with a strong overlap among the separate climate phases (Figure 4.5a). Sample means of *G. puncticulata* go from large and square-shaped in the Initial Phase, to decreasing shell area and increasing aspect ratio in the

Table 4.2 AICc values and Akaike weights for stasis, unbiased random walks and directional evolution in size and shape of *Globoconella puncticulata* and *Truncorotalia crassaformis* over both the separate climate phases and the entire studied interval, analysed using the PaleoTS package in R.

Species	Trait	Phase	Stasis		Unbiased random walk		Directional evolution	
			AICc	Akaike weight	AICc	Akaike weight	AICc	Akaike weight
G. puncticulata	Size	Initial	194.0	0.532	194.8	0.359	197.2	0.109
		Transition	252.1	0.000	235.9	0.740	238.0	0.260
		Glacial	144.3	0.122	140.8	0.683	143.3	0.194
		All	**743.7**	**0.000**	**602.9**	**0.685**	**604.5**	**0.315**
	Shape	Initial	-141.3	0.963	-134.0	0.026	-132.3	0.011
		Transition	-91.2	0.000	-121.7	0.735	-119.7	0.265
		Glacial	-98.4	1.000	-80.8	0.000	-78.3	0.000
		All	**-205.5**	**0.000**	**-334.4**	**0.712**	**-332.6**	**0.288**
T. crassaformis	Size	Initial	202.2	0.006	192.5	0.763	194.9	0.231
		Transition	229.7	0.399	229.5	0.454	231.7	0.147
		Glacial	166.9	0.945	173.1	0.043	175.6	0.012
		All	**623.3**	**0.245**	**621.7**	**0.559**	**623.8**	**0.195**
	Shape	Initial	-74.0	0.002	-86.1	0.756	-83.8	0.242
		Transition	-126.9	0.046	-132.4	0.726	-130.1	0.228
		Glacial	-72.5	0.030	-78.9	0.753	-76.4	0.217
		All	**-251.5**	**0.000**	**-307.3**	**0.742**	**-305.2**	**0.258**

Table 4.3 Values for Lynch's delta metric for *Globoconella puncticulata* and *Truncorotalia crassaformis* size and shape. All values are less than 5% of the threshold proposed by Lynch (1990) to represent neutral evolution.

	Neutral evolution	*G. puncticulata*		*T. crassaformis*	
		Size	**Shape**	**Size**	**Shape**
Lynch's delta	$10^{-3} – 10^{-5}$	$1.28*10^{-7}$	$1.98*10^{-7}$	$1.21*10^{-6}$	$8.18*10^{-8}$

Transition and Glacial Phases, hinting at a strong size: shape allometry in this species (Figure 4.5b). However, within-population allometries are weak in both species (Figure 4.5c,d). Visual inspection of the data shows that the within-population allometries of *G. puncticulata* are grouped in their respective climate phases. In *T. crassaformis* the slopes change from weakly positive (Initial Phase) to weakly negative (Transition Phase) to neutral (Glacial Phase). In *G. puncticulata* within-population allometries show comparable slopes but different intercepts in the Initial and Glacial phases, with the Transition Phase acting as a bridge between the other two phases with more variable slopes and intercepts.

The evolutionary allometries within phases follow each phase's set of within-population allometries in both species (Figure 4.5c,d). For both *T. crassaformis* and *G. puncticulata* the phase-specific evolutionary allometries are significantly different among phases (ANOVA performed on phase-specific evolutionary slopes, $p<0.001$ for both species). The *T. crassaformis* evolutionary allometries shift from weakly positive in the Initial Phase, to weakly negative in the Transition Phase to almost neutral in the Glacial Phase, whereas those in *G. puncticulata* are weakly negative in the Initial and Glacial Phases, and strongly negative during the Transition Phase. Together, the phase-specific allometries explain 5.5% of all variance in *T. crassaformis*, as opposed to 33.6% in *G. puncticulata* (ANCOVA performed on phase-specific evolutionary slopes, $p<0.01$ for both species with

Table 4.4 Linear Model results on evolutionary allometries over the entire interval with ("Total interval with phases") and without phase-specific allometries ("Total interval"). R^2 values represent the variance explained in the total dataset by the model and *p*-values represent the significance of the slopes.

Model	Species	Phase	R^2 model	*p* slopes
Total interval	*T. crassaformis*		0.0015	<0.01
	G. puncticulata		0.198	<0.001
Total interval with phases	*T. crassaformis*	Initial		<0.001
		Transition	0.067	<0.001
		Glacial		<0.001
	G. puncticulata	Initial		0.245
		Transition	0.336	<0.001
		Glacial		<0.01

Table 4.5 p-values of the Wilcoxon rank-sum test performed on the angles ϑ of both *Globoconella* puncticulata and *T. crassaformis* and a set of randomly generated vectors.

Species	Phase	p-value (Wilcoxon rank sum test)
G. puncticulata	Initial	0.037
	Transition	0.283
	Glacial	0.020
	All	**0.007**
T. crassaformis	Initial	0.582
	Transition	0.003
	Glacial	0.020
	All	**0.003**

phase as a categorical explanatory variable, see Table 4.4. In *T. crassaformis* the evolutionary allometry over the entire studied interval (Figure 4.5c, black line) is comparable to the species' phase-specific evolutionary allometries (p<0.01, see Table 4.4. *G. puncticulata* on the other hand shows a strong negative slope (p<0.001) comparable to the Transition Phase allometry, but in contrast to the evolutionary allometries in the Initial and Glacial Phase and most of the within-population allometries. This allometric breakdown implies that, for this species on the verge of extinction, within-population and short-term evolutionary allometries do not predict long-term allometric trends during times of climatic upheaval.

The distribution of the angles θ between $\mathbf{p}_{max}$ and the direction of evolutionary divergence **z** are significantly smaller than angles drawn from a random distribution for both species (Wilcoxon rank sum test, p<0.01 and p<0.005 for *G. puncticulata* and *T. crassaformis*, respectively; Figure 4.5, Table 4.5). Despite their qualitatively different dynamics (Figure 5), the predictability of phenotypic change based on $\mathbf{p}_{max}$ does not differ between the two species (paired Wilcoxon rank sum test, p=0.78). When analysed separately the angles within separate climate phases occasionally deviate from this pattern. In *T. crassaformis* the angles in the Initial Phase are not significantly different from the randomly generated angles (Wilcoxon rank sum test, p=0.58, Table 4.5), whereas the angles in both the Transition and Glacial Phase are significantly smaller than expected by chance alone (Wilcoxon rank sum test, p<0.05 and p<0.001 respectively, Table 4.5) implying that the predictability of phenotypic change increases after the intensification of Northern Hemisphere glaciation. In *G. puncticulata* the angles θ are significantly smaller than the randomly generated angles in both the Initial and Glacial Phases (Wilcoxon rank sum test, p<0.05, Table 4.5), whereas no significant difference was detected in the Transition Phase (Wilcoxon rank sum test, p=0.283, Table 4.5). ϑ does not vary systematically with the within-sample size and shape covariance (linear regressions, p>0.05), implying that large phenotypic correlations do not increase the predictability of the lines of least resistance.

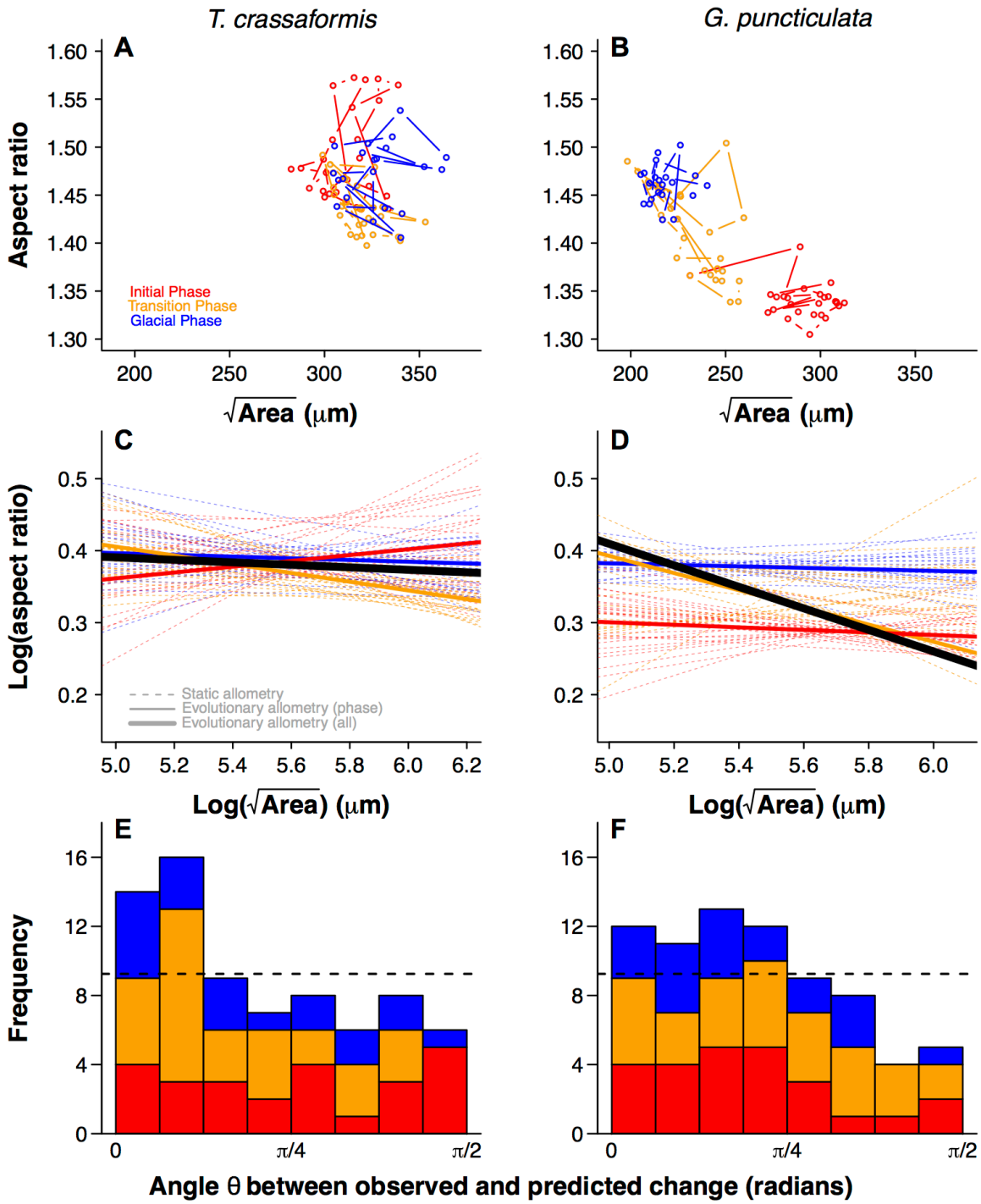

Figure 4.5 Sample means of size and shape for a) *Truncorotalia crassaformis* and b) *Globoconella puncticulata* with colours indicating the climate phases (Initial Phase: red, Transition Phase: yellow, Glacial Phase: blue). Within-population (dashed lines) and evolutionary allometries within climate phases (coloured solid lines) and over the entire study interval (solid black line) for c) *T. crassaformis* and d) *G. puncticulata*. In both species short-term evolutionary allometries are significantly different from the long-term evolutionary allometries (ANOVA,

p<0.001 for both species). Angles between predicted (p_{max}) and observed (z) one-step evolutionary change for e) *T. crassaformis* and f) *G. puncticulata*. Dashed lines indicate the expected frequency of angles when drawn from a random distribution. Angles are significantly smaller than expected from a random distribution for both species (Wilcoxon rank sum test, p<0.01 for both species, see Table 4.5 for phase-specific p-values).

4.5 Discussion

4.5.1 Allometries within climate phases

Despite low within-population (static) allometric slopes, long-term evolutionary change occurs preferentially along the within-population allometric lines of least resistance, a finding which is consistent with results of studies on generational (Schluter, 1996) and macroevolutionary time scales (Hunt, 2007; Renaud et al., 2006). In *T. crassaformis* the angles ϑ between the directions of predicted ($\mathbf{p}_{max}$) and observed ($\mathbf{z}$) evolutionary change are significantly smaller than expected by chance, implying that the lines of least resistance, to some extent, predict phenotypic change of this species through time. Traits are best described by a Random Walk (Table 1), and in all cases Lynch's Delta values indicate stabilising selection. It may seem counterintuitive that trait evolution is consistent with both a Random Walk and stabilising selection, but stabilising selection could maintain a population near a non-stationary adaptive peak. Using the divergence data compilation by Gingerich (2001), Estes and Arnold (2007) showed that such a displaced optimum had the highest explanatory power among multiple competing models of phenotypic evolution. If the optimum shifts, stabilising selection will attempt to drag the population towards the new peak (Estes and Arnold, 2007). Assuming the optimum shifts according to a Random Walk, we would expect the sample means to oscillate on or close to $\mathbf{p}_{max}$ but with zero net overall change rather than a directional pattern. Note that the random walk models (Table 1) are univariate, whereas here each foraminifera is conceptualized as bivariate when testing allometric constraints on the predictability of $\mathbf{p}_{max}$. Within each climate phase the phase-specific evolutionary allometries match the within-population allometries, implying that within-population allometries can predict evolutionary allometries over time-steps spanning ~100,000 years. In *G. puncticulata* evolution occurs significantly close to $\mathbf{p}_{max}$ in the Initial and Glacial

Phases. However, this pattern breaks down during the Transition Phase when the species' angles between **z** and $\mathbf{p}_{max}$ are randomly distributed (Table 4.5) and the within-population allometric slopes are highly variable (Figure 4.5d, yellow lines; 4F, yellow bars). In this interval the species shifts from large and square-shaped shells to smaller and more rectangular shells (Figure 4.5b), resulting in a strongly negative phase-specific evolutionary allometry (Figure 4.5d, dashed and solid yellow lines).

The primary reduction in size of *G. puncticulata* occurred during MIS G6, the first major glacial in the Transition Phase. MIS G6 marks the onset of significant widespread glaciation in the Northern Hemisphere (NHG) (Bailey et al., 2013) and a time of profound oceanographic change in the North Atlantic (Lang et al., 2016; Naafs et al., 2012). At Site U1313, sea surface temperature fell by ~6 °C as compared to the preceding interglacial stage (Friedrich et al., 2013; Naafs et al., 2010) and major increases occurred in cold stage (glacial) surface ocean primary productivity and North American-sourced aeolian dust deposition, delivering nutrients to the oligotrophic surface waters of the North Atlantic Subtropical Gyre (Lang et al., 2014; Naafs et al., 2012). It is possible, therefore, that trait response to new environmental pressures over-rode internal evolutionary constraints (Beldade et al., 2002), forcing the fitness peak of the adaptive landscape in morphospace (Arnold et al., 2001) to shift to trait combinations better adapted to glacial environments. The best supported mode of size evolution shifts from Stasis in the Initial Phase to a Random Walk in the Transition Phase, implying that this phenotypic change is driven by a new selection regime in the Transition Phase. Although there may be a role for directional evolution during parts of the Transition Phase, no strong evidence was provided for Directional Selection over the whole interval or within each phase. No significant correlation was detected between *G. puncticulata* size or shape and on-going high-latitude climate approximated by the global benthic $\delta^{18}O$ stack (Lisiecki and Raymo, 2005) (Figure 4). However, local environmental change associated with iNHG in the surface ocean might have crossed a critical threshold for the species. Planktonic foraminifera species have very specific temperature ranges for optimum growth (Lombard et al., 2011; Lombard et al., 2009) and the decrease in temperature associated with MIS G6 could have reduced the growth potential for *G. puncticulata*. Although climate shifted back to an interglacial state following MIS G6, the decrease in size also meant a loss of maximum size variance because the two covary, and the interglacial environmental conditions did not last long enough for the species to regain their initial maximum size. The species'

extinction at 2.41 Ma has been linked to iNHG (Chapman et al., 1998; Scott et al., 2007; Wei, 1994a), further pointing to a higher species-specific sensitivity to glacial conditions. However, the impact of global environmental change on the species' extinction remains to be tested among multiple populations living under different environmental conditions.

The increase in shell aspect ratio and its covariance with size is more challenging to explain. Smaller shell size reduces drag and increases settling velocity through the water column, making it more difficult for individuals to retain their preferred depth habitat. The changes in shell shape necessary to influence settling velocity are nevertheless smaller (5-6% (Caromel et al., 2014)) than the typical intraspecific variation observed (25-50% (Schmidt et al., 2004a)), while the settling velocities reported by Caromel et al. (2014) imply that shells would sink below the thermocline in several hours regardless of shape, implying that active buoyancy regulation by the individual, for example through the extensive pseudopodial network determines its position in the water column. Although it is possible that a decrease in mean shell size requires an attendant flattening of shell shape to facilitate maintenance of position in the water column, the covariation between size and shape within samples is often weak, implying limited effects of trait covariation on buoyancy.

Alternatively, we might hypothesize that the change in population trait means during the Transition Phase reflects widespread migration of morphologically distinct populations. Surface ocean currents in the North Atlantic Ocean are proposed to have undergone major reorganisations from ~2.7 Ma (Naafs et al., 2010). A southward shift of the North Atlantic Current (NAC) could have transported populations of *G. puncticulata* previously restricted to sub-polar regions into the mid latitudes. It has been suggested, based on dinoflagellate assemblage work, that the NAC was deflected south of Site U1313 only from MIS 104 onwards (Hennissen et al., 2014). However, a study of foraminiferal calcite geochemistry proposes that the position of the NAC probably lay well north of Site U1313 during the Late Pliocene to Earliest Pleistocene (Friedrich et al., 2013), implying a maintained position of Site U1313 in the North Atlantic Gyre with little changes in surface water currents. Regardless, changes in surface water currents are suggested to have occurred mainly during glacials, with currents returning to previous conditions during interglacials (Naafs et al., 2010). Thus, if migration of morphologically distinct populations of *G. puncticulata* was responsible for the changes in trait means reported here, we would expect to observe alternating

changes in morphology tracking the more intense (inter)glacial cycles during the Transition and Glacial Phases and a strong overall correlation between climate and traits, which is not the case (Figure 4.2c,d, Figure 4.4).

4.5.2 Allometries among climate phases

In *T. crassaformis* the phase-specific evolutionary allometries change from positive in the Initial Phase, to negative in the Transition Phase to almost neutral in the Glacial Phase (Figure 4.5c). Although all phase groups of evolutionary allometric slopes are significantly different, they are comparable in magnitude. The evolutionary allometry over the entire studied interval corresponds well with the phase-specific evolutionary and within-population allometries, implying that in *T. crassaformis*, within-population allometries predict evolutionary allometries on various time scales, despite pronounced global climate upheaval. In *G. puncticulata* the evolutionary allometries in the Initial and Glacial Phases have similar slopes but different intercepts, with both phases bridged by the stronger negative allometry in the Transition Phase (Figure 4.5d). However, sample means are characterised by decreasing size and increasing aspect ratios, with largest and most square-shaped shells in the Initial Phase and smallest, most rectangular shells in the Glacial Phase, which is reflected by a strong negative evolutionary allometry over the entire interval. Because the long-term evolutionary allometry deviates from the within-population and phase-specific evolutionary allometries, these results imply that environmental change can override allometric relationships over time.

4.5.3 General evolutionary implications

Projecting our bivariate size-shape matrices evokes Raup's seminal ammonoid morphospace (1966), which uses trait frequencies as fitness surrogates to provide a good approximation of the adaptive landscape under a consistent selective regime (Arnold et al., 2001). Our time-averaged samples obfuscate generation-by-generation selection estimates by amalgamation, yet do estimate the fitness optimum when the conditions are stable. The predictive ability of the lines of least resistance (Figure 4.5e,f) infers a non-symmetric ellipsoidal fitness peak, upon which the species oscillates. Our calculations of Lynch’s Delta reject drift as the dominant evolutionary mode for both traits in both species. Despite the co-occurrence of the shift in *G. puncticulata* size with the intensification of Northern hemisphere glaciation (Bailey et al., 2013), there is no evidence to

suggest that either of our two species evolves on a fine scale in response to Northern hemisphere climate cycles inferred from benthic $\delta^{18}O$. Taken together, these two results imply a role for stabilizing selection either maintaining species location around a fitness peak in the Initial and Glacial phases, or pulling the species towards a "displaced optimum" (Estes and Arnold, 2007) during the Transition Phase. The two optima for *G. puncticulata* form an apparent ridge in the evolutionary allometry when viewed across the whole interval (Figure 4.5d), but this feature is attributable to sampling through heterogeneous environmental conditions and the associated breakdown of within-population and evolutionary allometries among phases. This discrepancy between evolutionary and within-population allometry contrasts with recent results on over 300 animal taxa (Voje et al., 2014) and fossil rodents (Firmat et al., 2014), but neither of these studies sampled through a major climate transition. The phenotypic covariance might be maintained by the evolutionarily constraining mechanisms of allometries (Pélabon et al., 2014), by a particular environmental context, or a combination of both. The generality of when a changing environmental context prompts a reconfiguration of the relationship between within-population and evolutionary allometries remains to be tested across species with different macroevolutionary fates, spatial locations and allometric strengths.

4.6 Conclusion

The results presented in this chapter show that under constant environmental variability the within-population allometric lines of least resistance predict evolutionary allometries on microevolutionary time scales. However, in one of the studied species the evolutionary allometry shifted away from the within-population allometries across intervals marked by pronounced climate upheaval. We suggest that changed environmental conditions related to the intensification of Northern Hemisphere Glaciation pushed this species off its peak on the adaptive landscape, away from the direction of intraspecific phenotypic variation. Intraspecific variation left a clear signature on evolutionary change from one time-step to the next in both species studied (Figure 5), but the generality of when phenotypic covariance interacts with environmental perturbation, and how this interaction links to longer-term microevolutionary trends and macroevolutionary phenomena, remains to be tested.

Chapter 5: Temperature is a poor proxy for synergistic climate forcing of plankton evolution

This chapter is a reproduction of a manuscript that is currently in review at *Proceedings of the Royal Society B* as: Brombacher, A., Wilson, P.A., Bailey, I., Ezard, T.H.G. Temperature is a poor proxy for synergistic climate forcing of plankton evolution.

5.1 Abstract

Changes in biodiversity are often linked to climate change, usually represented by global temperature. A global environmental driving mechanism of biodiversity is implied by the strong correlation between the Cenozoic diversity patterns of a wide variety of fauna and flora and changes in global climate as recorded by the oxygen isotope composition of benthic foraminifera. Yet climate consists of many interacting variables, and species likely respond to the entire climate system as opposed to individual variables. Here, both ecological and morphological traits of 12,629 individuals of two species of planktonic foraminifera are used with similar ecologies but contrasting evolutionary outcomes to show that phenotypic and ecological changes are indeed driven by the interactions of multiple environmental factors. Models including interactions among climate variables explain at least twice as much variation in changes in size, shape and ecological abundance than models assuming that climate parameters operate independently of each other. No dominant climatic driver can be identified and temperature alone explains remarkably little variance, implying that a new approach is required to understand evolutionary response to abiotic forcing. Our results show that successful reconstruction of paleobiological dynamics will require multiple ecological and evolutionary metrics.

5.2 Introduction

Changes in biodiversity are often linked to climate change, usually temperature. Phanerozoic species richness covaries with global temperature (Erwin, 2009; Mayhew et al., 2008), and Cenozoic diversity patterns of mammals (Blois and Hadly, 2009; Figueirido et al., 2012), plants (Jaramillo et al., 2006; Nyman et al., 2012), insects (Nyman et al., 2012) and plankton (Ezard et al., 2011; Lazarus et al., 2014) correlate with the high latitude climate signal recorded in the $\delta^{18}O$ composition of benthic foraminifera (Zachos et al., 2001), implying a global environmental driving mechanism of biodiversity, or a common driver of temperature and biodiversity. Yet climate consists of many interacting variables, and species likely respond to the entire climate system as opposed to separate variables: Harnik et al. (2012) argued that simultaneous changes in multiple environmental parameters drove most Phanerozoic extinction events, while Garcia et al. (2014) show increased threats on modern biodiversity when incorporating multiple dimensions of climate change. However, the impacts of synergistic climate effects – the extent to which the impact of abiotic forcing is underestimated when only single environmental factors are assessed – on within-species evolutionary change remain largely unknown.

To understand the driving mechanisms of biodiversity over long time scales we need to study evolutionary change at the species level (Arnold et al., 2001; Erwin, 2000). To accurately quantify the link between long-term (>10,000 years) species evolution and climate change, both high-resolution fossil records of multivariate species evolution and reconstructions of multiple local environmental parameters are needed. However, such data sets are rarely available. One of the few mediums on which multivariate evolutionary and environmental change can be determined at high temporal resolution is the marine fossil record of planktonic foraminifera. The excellent preservation of this group in deep-sea sediments permits direct comparison of biotic change to high-resolution records of climate and evolution reconstructed from the same marine sediment cores. Several studies have shown individual species' responses to sea surface temperature (Edgar et al., 2013; Wade and Pearson, 2008), but also to productivity (Wade, 2004) and ocean stratification (Wade and Pearson, 2008; Wei and Kennett, 1986). However, none of these studies analysed the ecological and evolutionary impacts due to the interplay of multiple climate drivers.

Here I study species' response to multivariate climate change during the last great climate transition in Earth's history: the late Pliocene to earliest Pleistocene intensification of Northern Hemisphere glaciation (iNHG, 3.6-2.4 Ma) (Mudelsee and Raymo, 2005). This interval was characterized by major reorganizations of the global climate system. Global atmospheric CO_2 concentrations (Martínez-Botí et al., 2015) dropped below the ~280 μatm threshold for extensive NHG (DeConto et al., 2008) between 2.9-2.7 Ma (Figure 5.1c). By 2.7 Ma continental ice-sheets had expanded significantly on Greenland, Scandinavia and North America as evidenced by the onset of widespread ice-rafted debris deposition in high northern latitude oceans (Bailey et al., 2013; Shackleton et al., 1984) and an increase in the amplitude of glacial-interglacial cycles as recorded in the oxygen isotope record ($\delta^{18}O$) from benthic foraminifera (to >0.5‰) from Marine Isotope Stage (MIS) G6 (2.7 Ma) onwards (Figure 5.1a,b). In the North Atlantic Ocean this transition to deeper glacials was associated with (i) early incursions of southern-sourced deep waters (Lang et al., 2016), (ii) a major intensification of dust flux from North America carried on the westerly winds (Lang et al., 2014; Naafs et al., 2012), and (iii) increases in glacial primary productivity (Lawrence et al., 2013; Naafs et al., 2012) (Figure 5.1d,e). Modern studies of productivity in the Subtropical Gyre of the North Atlantic (Moore et al., 2006) show that iron supplied by eolian dust is an important factor controlling the dynamics of the spring bloom in this region today, although the possibility that water mass reconfigurations helped to drive these changes in productivity in the Late Pliocene North Atlantic Ocean cannot be ruled out (Naafs et al., 2012). Together, these environmental changes likely had a major impact on life in the marine realm (Jackson and Johnson, 2009). To quantify the combined effects of changes in temperature, primary productivity, dust input and atmospheric CO_2 on evolution during iNHG, I employ multivariate statistical techniques to compare ecology (abundance Figure 5.1g) and morphology (size and shape, Figure 5.1h,i) of 12,629 specimens of the ecologically similar planktonic foraminifera species *Globoconella puncticulata* and *Truncorotalia crassaformis* (Figure 1.1) to records of North Atlantic climate all generated on samples taken from Integrated Ocean Drilling Program (IODP) Site U1313 (~41˚ N) (Figure 1.2). *T. crassaformis* survived the onset of iNHG and is still alive today, whereas *G. puncticulata* became extinct shortly after 2.41 million years ago (Ma) (during MIS 96 (Wei, 1994b)). These foraminifera provide an opportunity to study species' responses to multivariate climate change under contrasting evolutionary outcomes.

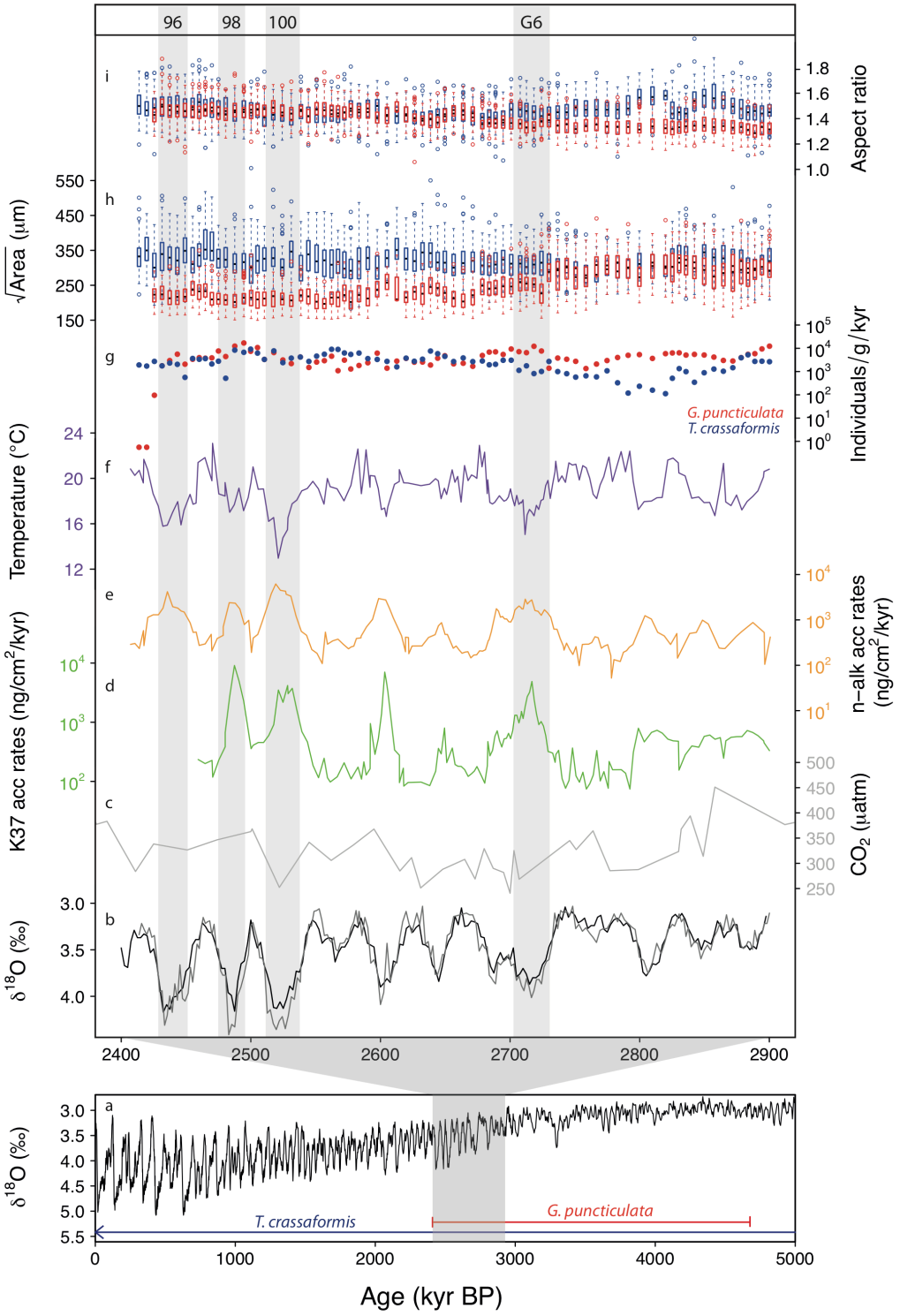

Figure 5.1 Environmental reconstructions and morphology of two planktonic foraminifera species at IODP Site U1313: Oxygen isotopes from the Lisiecki and Raymo (2005) benthic stack (a,b, black lines) and Site U1313 (Bolton et al., 2010) (b, grey line), atmospheric CO_2 reconstructed at ODP Site 999 by Martínez-Botí et al. (2015) (c), productivity (d), eolian input (e) and sea surface temperature (f) by Naafs et al. (2012), abundance (g) of *Globoconella puncticulata* (red)

and *Truncorotalia crassaformis* (blue) (this study), and size (d) and shape (e) of *G. puncticulata* and *T. crassaformis* (Brombacher et al.). Key glacial stages are indicated by grey bars.

5.3 Methods

5.3.1 Study species

Truncorotalia crassaformis and *Globoconella puncticulata* (Figure 1.1) are two ecologically similar species characterised by low trochospiral shells with flattened spiral sides, inflated umbilical sides and umbilical-extraumbilical apertures (Kennett and Srinivasan, 1983) and inhabit thermocline waters at middle and low latitudes (Aze et al., 2011; Kennett and Srinivasan, 1983). *T. crassaformis* originated around 5.7 Ma and survives to the present day. *G. puncticulata* first appeared around 4.6 Ma and became extinct at 2.41 Ma (Wei, 1994b), shortly after the onset of significant Northern Hemisphere glaciation at 2.72 Ma (Bailey et al., 2013). Our 500,000-year study interval includes the intensification of Northern Hemisphere glaciation (2.72 Ma, (Bailey et al., 2013)) and the first three major glacial stages MIS 96, 98 and 100 (Lisiecki and Raymo, 2005), and ends with the extinction of *G. puncticulata* (Wei, 1994b). Three traits are studied: mean shell size (area) and shape (aspect ratio) per time slice (data from (Brombacher et al.)) and abundance (this study) (Figure 5.1g-i). Abundance and size are indicators of ecological success (Hecht, 1976b; Schmidt et al., 2004a): abundance indicates reproductive success, and size indicates the (absence of) limiting factors on growth. Shell shape controls the area: volume ratio which influences respiratory processes.

5.3.2 Study site

IODP Site U1313 is located in the mid latitude North Atlantic Ocean at the base of the upper western flank of the Mid-Atlantic Ridge at a water depth of 3426 m (41 °N, 32.5 ‘W), on the northern edge of the North Atlantic Subtropical Gyre (Figure S2). The Site U1313 record is characterised by consistently high sedimentation rates (~5 cm/kyr) for the past 5 Myr (Channell et al., 2006a; Lisiecki and Raymo, 2005), a demonstrably continuous record of sedimentation

for iNHG (Bolton et al., 2010) and exceptionally well-preserved microfossil carbonate (Lang et al., 2014).

I used 75 samples (every 30 cm, i.e. ~5-kyr-resolution) that follow the shipboard primary splice studied originally by Bolton et al. (2010). For age control, I used the orbital-resolution benthic foraminiferal oxygen isotope ($\delta^{18}O$) stratigraphy for Site U1313 generated by Bolton et al. (2010). The samples were dry-sieved over a >150 μm^2 mesh sieve and split using a microsplitter until a single split contained 70-150 specimens of *T. crassaformis* or *G. puncticulata*. To avoid size bias all individuals from a single split were analysed, resulting in a total of 12,629 individuals (6058 specimens of *T. crassaformis* and 6575 of *G. puncticulata*) over the studied interval. The total number of specimens in the sample was estimated by multiplying the number of individuals found by the split fraction, and abundance (represented as accumulation rates) was calculated as the number of individuals divided by the coarse fraction weight, divided by the total time in the sample. Trait data are deposited in the Figshare repository at https://figshare.com/s/**9db6657150242fb8a593**.

Several published orbitally resolved environmental reconstructions are available for Site U1313, including a mean annual sea surface temperature record based on the saturation index of C_{37} alkenones, $U^{k'}_{37}$ (Naafs et al., 2012), n-alkane accumulation rates representing productivity (Naafs et al., 2010) and terrestrial plant leaf wax fluxes linked to eolian input of North American dust (Naafs et al., 2012). Naafs et al. (2012) generated two plant wax records, one using n-alkanes and the other C26-alkan-1-ol chains. Both records are highly correlated (Linear Model, R^2 = 0.91 and p < 0.001) and argued to be from a common North American origin (Naafs et al., 2012). Here I use the n-alkanes record as its values are higher by a factor ~1.5 as compared to the C26-alkan-1-ol-based record, providing the highest signal: noise ratio. Biotic responses were compared to the site-specific reconstructions of sea surface temperature, productivity and dust input (Naafs et al., 2012; Naafs et al., 2010), as well as a reconstruction of global atmospheric CO_2 concentration (Martínez-Botí et al., 2015) to represent environmental conditions (Figure 5.1c-f). All parameters would have directly influenced individual foraminifera during their lifetime: species prefer specific temperature ranges (Lombard et al., 2011; Lombard et al., 2009) and productivity regimes (Hemleben et al., 1989), and ocean pH influences calcification potential (Hemleben et al., 1989).

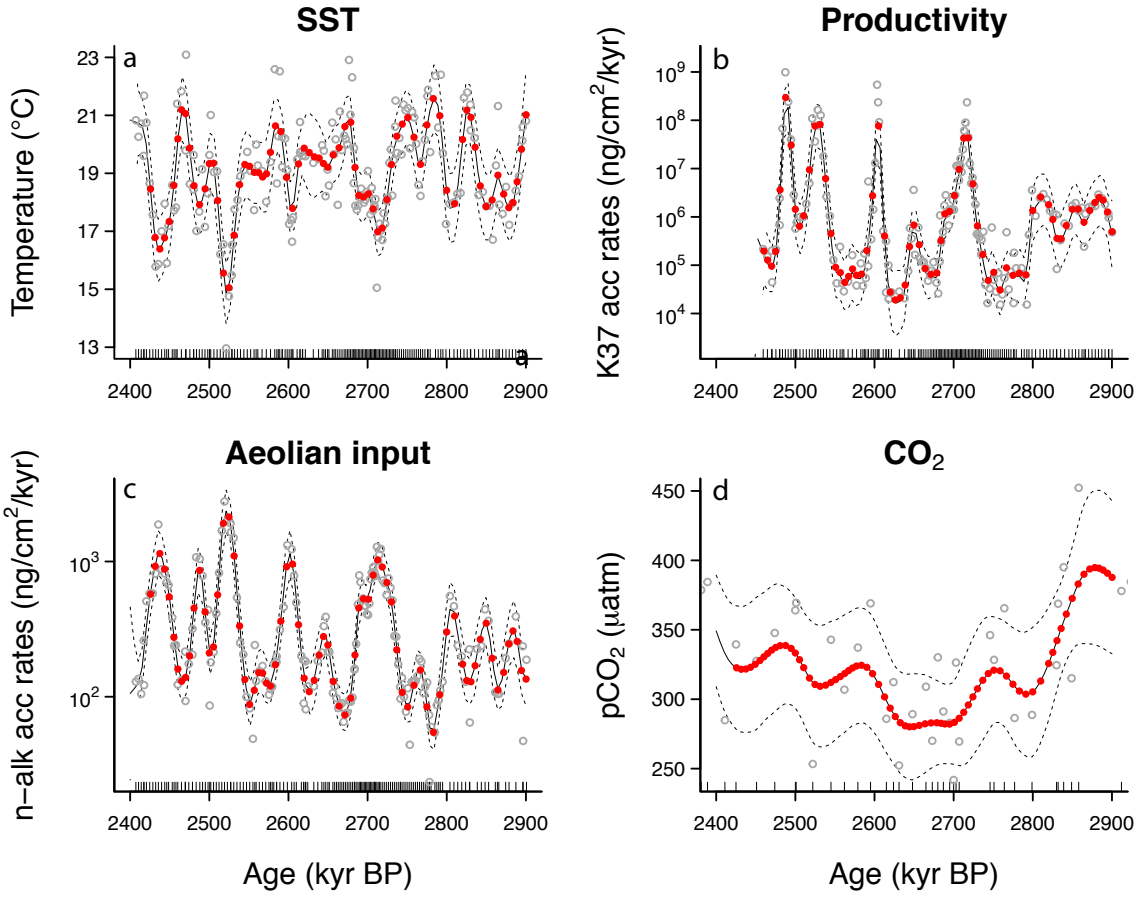

Figure 5.2 Generalised Additive Models (GAM) used to interpolate values of sea surface temperature (a), productivity (b), eolian dust input (c) and atmospheric CO_2 concentration (d) at the ages of the foraminifera samples from Site U1313 (internal tick marks on x-axis). Original data points are denoted by open circles, with solid and dashed lines representing the GAM and 95% confidence interval respectively. Estimated values are indicated by red circles.

5.3.3 Analysis

Since the environmental reconstructions of Site U1313 and the foraminifera trait data were generated using different sample sets, the climate data point ages are offset relative to the foraminifera samples. Generalised Additive Models (GAMs) were employed to interpolate the climate parameters to the ages of the foraminiferal samples. First the individual climate records were smoothed using a GAM, and the value at the age of the foraminifera samples was estimated using the smoothed curve (Figure 5.2). The biotic data were then compared to the interpolated climate records using multivariate statistical techniques. Principle

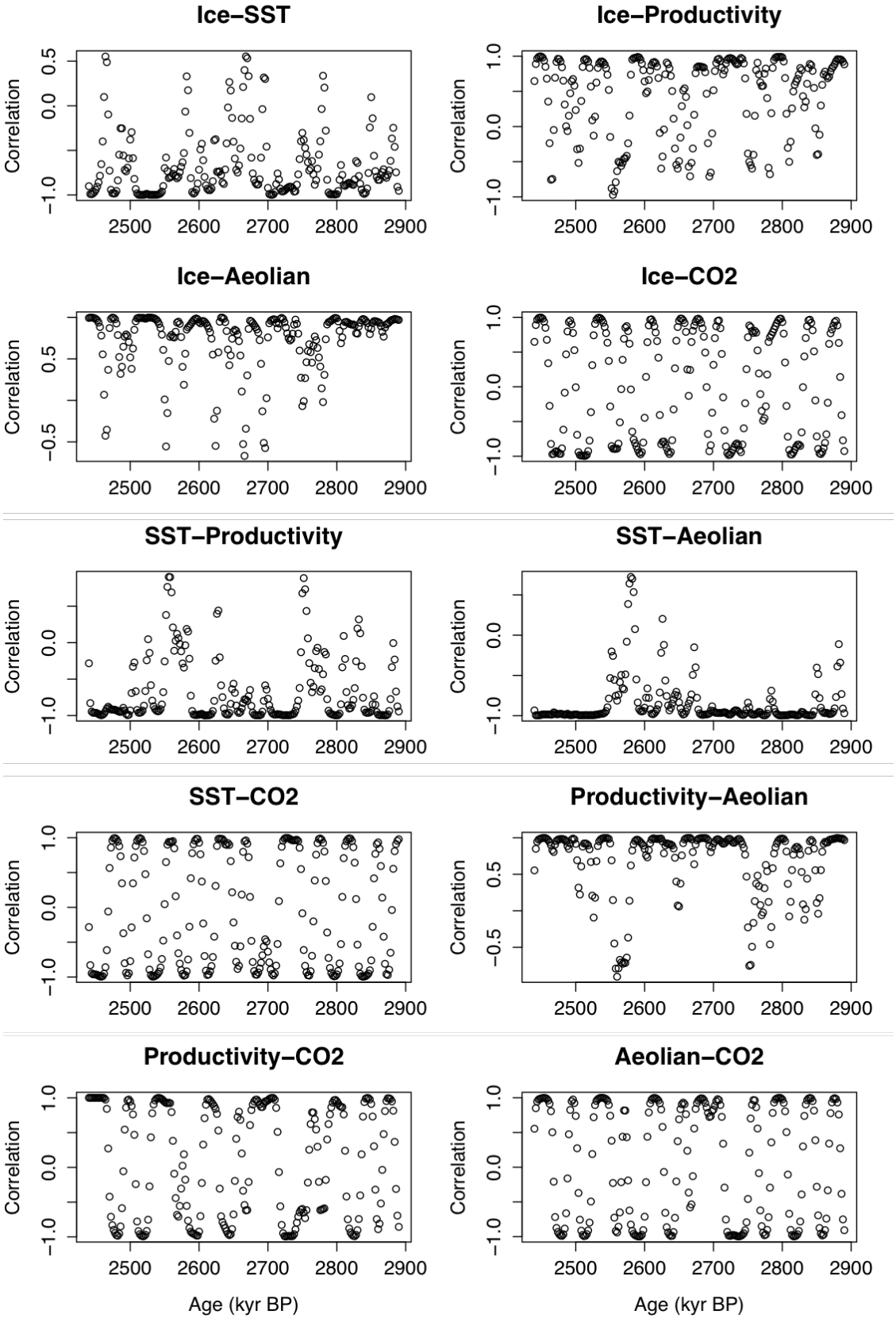

Figure 5.3 Covariances between pairs of environmental variables. For most parameter combinations the correlation coefficients are not constant through time.

Component Analysis is most often used to describe the combined effects of multiple parameters on biotic change, however this method requires constant covariance between the abiotic parameters over time, which was not the case in our study interval (Figure 5.3). Therefore, Generalised Linear Models (GLM, with a Poisson error function for the abundance data and square root link function and μ error function [variance increasing linearly with the mean] for the morphological traits)) were used instead to compare trait variance to all environmental variables and their two-way interactions (Figure 5.4). Trait variance explained by individual parameters was calculated as the variance explained by the full model (up to an including all two-way interactions), minus the variance explained by the model with each parameter removed. Another GLM without the two-way interactions between abiotic variables was compared to the full model to quantify the synergistic effects of interactions among climate variables on evolutionary change.

To assess the effects of using mean data as opposed to individual measurements per time slice the GLMs were repeated using the individual data of size and shape. The extra variance introduced by analysing individual data as opposed to means reduces the amount of variance explained by the environmental parameters to ~20%, but the overall patterns remain the same (Figure 5.5). As the response in abundance (one value per time slice) can only be compared to trait means, the GLM results based on mean values per time slice were used throughout this study.

5.4 Results

After using non-parametric Generalised Additive Models to harmonise time samples among all climate variables (Figure 5.2), mean trait records were compared to records of sea surface temperature, productivity and dust input from the same site (U1313), as well as to a global reconstruction of atmospheric pCO_2 concentration (Figure 5.4) using Generalised Linear Models (GLMs). All traits are best described by the combination of all studied environmental parameters and their interactions (Figure 5.4). The models including all variables and their two-way interactions explain 30-50% of the variance in all traits (Figure 5.4a, c-f), except for size in *T. crassaformis* (10%) (Figure 5.4b). In contrast, individual climate proxies only explain 0-30% of the variance in traits. The total model

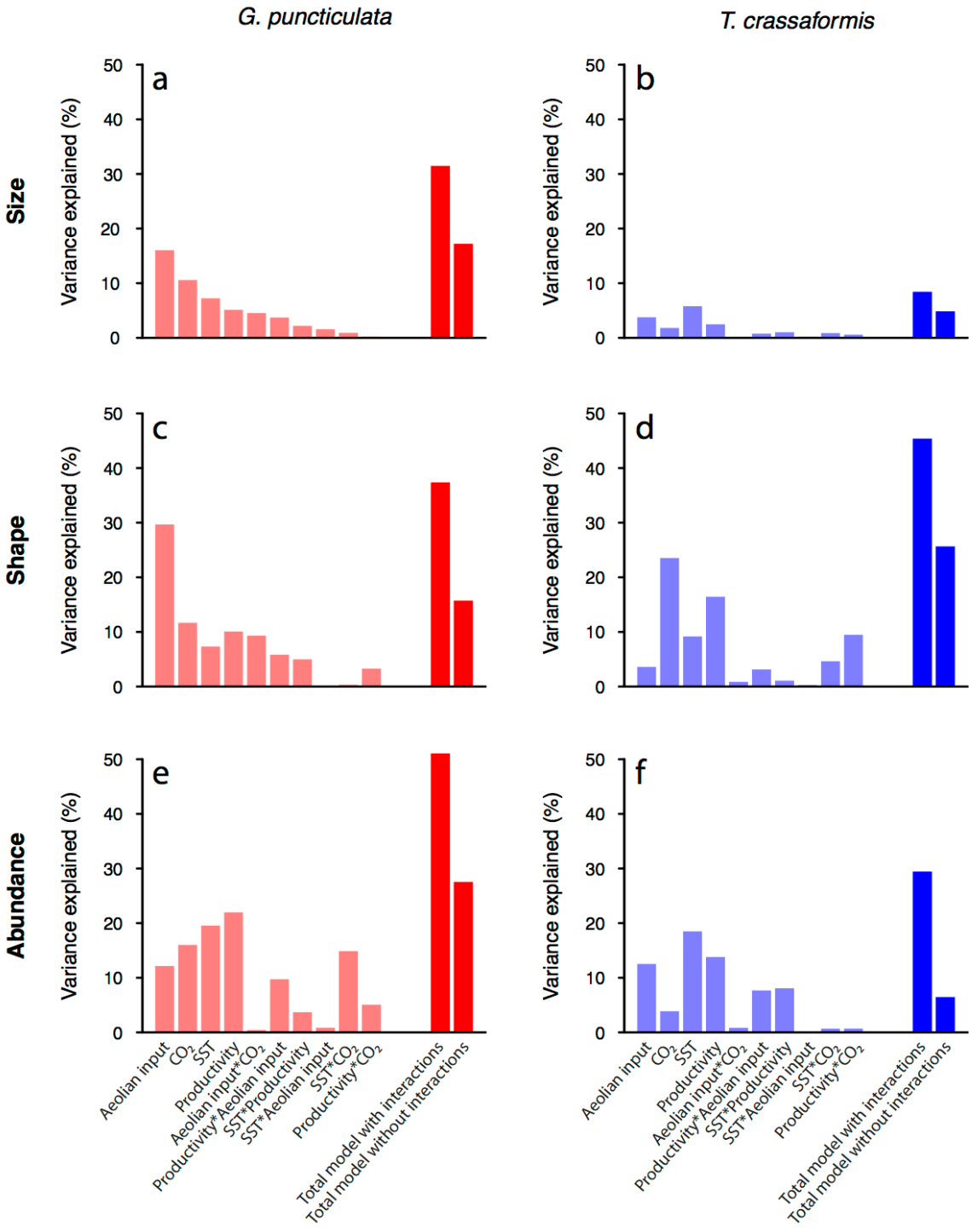

Figure 5.4 Variance explained in size (a,b), shape (c,d) and abundance (e,f) of *Globoconella puncticulata* (red) and *Truncorotalia crassaformis* (blue) from North Atlantic Site U1313 (41˚N) by the environmental parameters and their interactions using sample size and shape means.

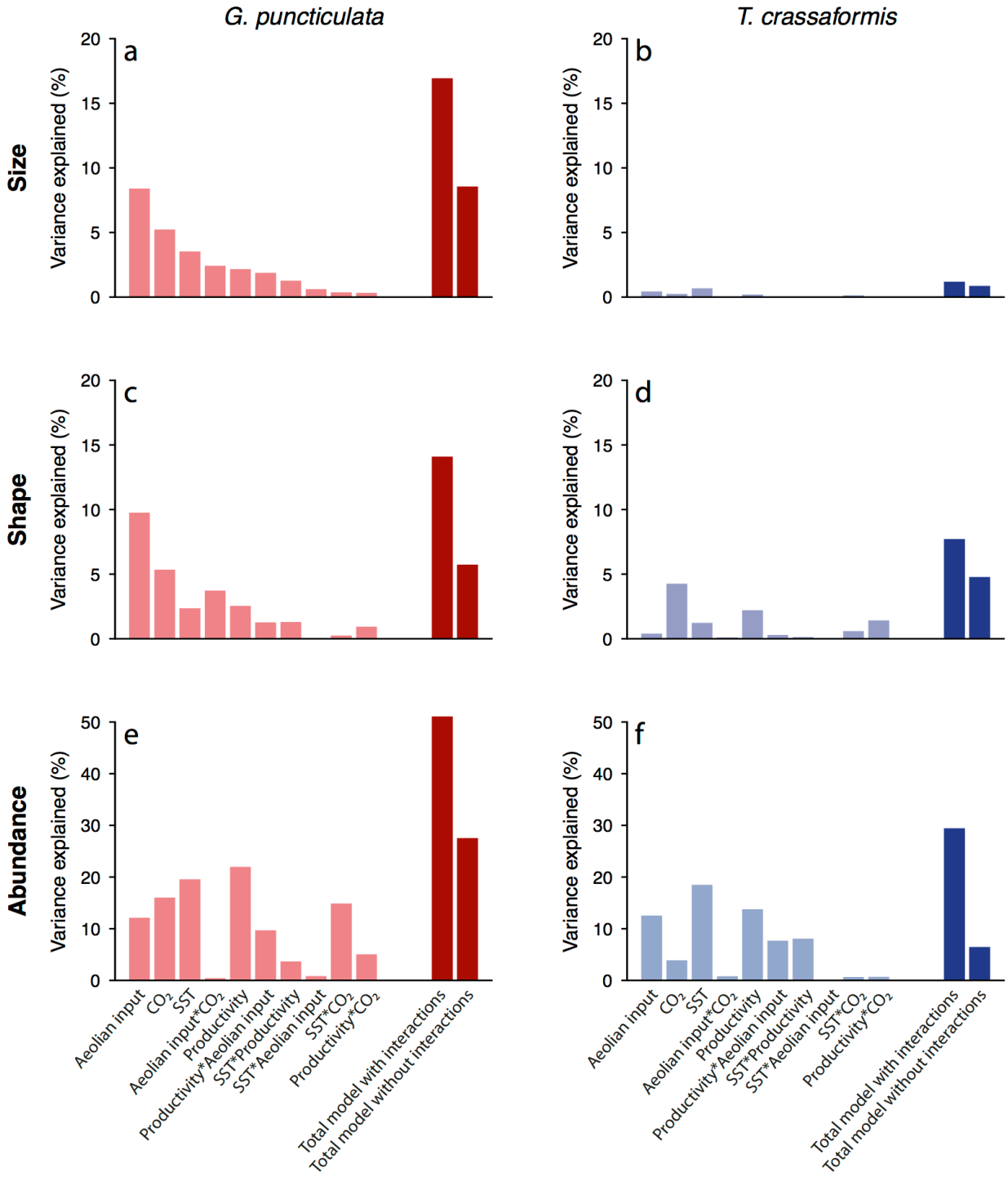

Figure 5.5 Variance explained in size (a,b), shape (c,d) and abundance (e,f) of *Globoconella puncticulata* (red) and *Truncorotalia crassaformis* (blue) of Site U1313 by the environmental parameters and their interactions, using individual measurements of shell size and shape.

including all two-way interactions provides a significantly better fit to the data than the additive model without the interactions for all traits in *G. puncticulata*, and shape and abundance in *T. crassaformis* (ANOVA, $p < 0.001$ for size and shape in both species, and $p < 0.05$ for *G. puncticulata* size). Responses also vary among traits. In both species, abundance and shape respond more strongly to

environmental forcing than size (Wilcoxon signed rank test, $p < 0.01$ for *G. puncticulata* abundance and *T. crassaformis* shape and abundance, and $p < 0.05$ for *G. puncticulata* shape). No difference is detected between the responses of shape and abundance in either species (Wilcoxon signed rank test, $p = 0.08$ and $p = 0.81$ for *G. puncticulata* and *T. crassaformis* respectively). No single parameter is found to dominate evolutionary response in all studied traits: size and shape of *G. puncticulata* are best explained by eolian dust input, whereas this species' abundance responds most strongly to changes in primary productivity. Sea surface temperature best explains variance in the size and abundance in *T. crassaformis*, and its shape is most strongly correlated to atmospheric CO_2 concentration. Overall, sea surface temperature is only the third ranked parameter in its contribution to explaining evolutionary change, behind eolian input and productivity.

5.5 Discussion

Our results show that the amount of morphological and ecological variation explained is highest when studied including interactions among multiple environmental parameters. These results imply that species' response to climate change can be underestimated when only single variables are used: here the apparent species' responses to their environment decreases by a factor ~2 if only single variables are considered (Figure 5.4). Our findings are consistent with short-term studies of modern populations that show increased mortality as a response to multiple environmental stressors (Brook et al., 2008; Godbold and Solan, 2013; Sodhi et al., 2008), as well as macroevolutionary research into the abiotic drivers of mass extinctions (Harnik et al., 2012). The response to individual parameters varies among traits. No single parameter best explains the variance in all records. Our results caution against the use of one environmental parameter to represent global climate while studying biodiversity response to

abiotic change, because a single parameter likely only explains a minor part of a species' overall response (Figure 5.4).

Our results generate an appropriately multi-faceted picture of abiotic forcing, and suggest strongly that (sea surface) temperature alone is a poor proxy for environmental change that impacts temporal ecological and evolutionary changes. These results contrast with a recent study by Fenton et al. (2016), who used the same system to show the dominance of temperature in shaping ecological processes across space. Our results imply that spatial abiotic drivers (Fenton et al., 2016) do not directly translate to those operating through time, supporting hypotheses that spatial variation is not a suitable substitute for temporal change and that fossil data are required to accurately reconstruct biodiversity dynamics over long time scales (Fritz et al., 2013; Quental and Marshall, 2010).

Neither species' responses are truly synergistic, because response to the total model describes less trait variance than the sum of the responses to single climate variables. These results are consistent with the findings of Darling and Côté (2008), who reviewed 112 published mortality experiments and found only a third showed synergistic responses to external drivers. In our case, the species' antagonistic responses (total response < sum of response to individual parameters) to abiotic change could be explained by a common driving mechanism underpinning the studied environmental variables. Late Pliocene North Atlantic sea surface temperature, productivity, aeolian dust input and CO_2 are strongly linked to iNHG (Martínez-Botí et al., 2015; Naafs et al., 2012; Naafs et al., 2010), resulting in similar trends in each record (Figure 5.1c-f) that add little extra variance explained in the biotic records. Depending on its ecological preferences, a species could respond to parameters in opposite ways: a positive response to an increase in one variable and a negative response to increase in another could lead to little net effect when both variables increase, decreasing the variance explained by the total model. However, despite this correlation the explanatory power of the total model increases when multiple variables are incorporated.

The unexplained variance could be attributable to several factors. Firstly, planktonic foraminifera have a life span of a few weeks (Hemleben et al., 1989). Individuals living in different seasons in the mid-latitude Atlantic Ocean experience temperature differences of up to 6-7 degrees Celsius (NASA, 2016). Such variability is comparable to mean annual Late Pliocene – Early Pleistocene

glacial-interglacial SST changes at our study site (Friedrich et al., 2013; Naafs et al., 2010) (Figure 5.1f) and plastic responses to these seasonal differences could increase trait variance. Secondly, some of the observed trait variance could be caused by migration of morphologically different populations from other locations. Changes in surface water currents in the North Atlantic are driven by the position of the polar front during glacials, but those changes are not maintained across successive glacial-interglacial cycles (Naafs et al., 2010). Thus, if the changes reported here were caused by migration of morphologically distinct populations, we would expect the observed changes in morphology to track most strongly the prominent (inter)glacial cycles that characterise our study interval, which is not the case (Brombacher et al.). Finally, abundance and shell shape responded more strongly to the studied environmental variables than shell size, but in reality traits are often not independent (Lande, 1979; Lande and Arnold, 1983). Such covariation can constrain evolutionary responses to environmental drivers (Agrawal and Stinchcombe, 2009). Climatic upheaval can disrupt the covariation between traits (Brombacher et al.), emphasising the need for comprehensive understanding of abiotic catalysts for biotic change.

5.6 Conclusion

The results presented in this chapter show that evolution responds to multivariate environmental change, particularly the interactions among distinct parts of global climate. No single variable was identified that best explained evolutionary change in all studied traits of both foraminifera species, implying that evolutionary response to environmental change is likely to be underestimated when only single variables such as temperature are used to represent global climate. In our study temperature is not even the most important single climate variable accounting for evolutionary change. Responses also varied among morphological and ecological traits, suggesting trait-specific sensitivities to environmental change. Our results imply that use of local temperature as a single variable to test for biotic response to climate change severely underestimates biotic response and that successful reconstruction of eco-evolutionary dynamics in deep time requires multivariate explanatory and response variables.

Chapter 6: Spatial variation does not predict temporal dynamics among 29,438 planktonic foraminifera individuals

This chapter is currently in preparation for submission to *Nature Ecology & Evolution* as: Brombacher, A., Wilson, P.A., Bailey, I., Ezard, T.H.G. Spatial variation does not predict temporal dynamics among 29,438 planktonic foraminifera individuals.

6.1 Abstract

To predict biotic responses to future climate change, a common approach is to extrapolate species' responses to current environmental conditions into the future. These projections have traditionally been justified by assuming that the species' adaptive potential is similar across its geographical range and that environmental variability among contemporary populations is a relevant predictive tool. However, no empirical data yet exists to test these assumptions. Here, I use the fossil record of planktonic foraminifera to study covariations between ecological and morphological traits in 29,438 individuals of the ecologically similar species *Globoconella puncticulata* and *Truncorotalia crassaformis* over 600,000 years spanning the intensification of Northern Hemisphere Glaciation. Size and abundance covary tightly at the species level in planktonic foraminifera, but within-population temporal dynamics had negligible explanatory power for the among-population changes observed. These results imply that morphological diversity should be tracked through both space and time since one can be a poor substitute for the other.

6.2 Introduction

Spatial species variation patterns are often used to predict temporal change. Methods such as bioclimatic and species diversity models (reviewed in (Guisan and Zimmermann, 2000)) integrate present-day species occurrence and abundance data, and model species' response to past (Kitchener and Dugmore,

2000; Ruegg et al., 2006) and future (Huntley et al., 1995; Thomas et al., 2004; Thuiller et al., 2005) climatic conditions. These models assume strongly that species' responses to environmental change are comparable across space and through time. However, populations are likely to encounter biotic and abiotic pressures that vary as a function of location throughout the species' range, causing differences in adaptation potential across populations. Different adaptive pressures have been shown to act on populations living at the edge of a species' range relative to those situated well within the range limits (Hargreaves et al., 2014), implying that these populations likely respond differently when faced with the same environmental conditions. Additionally, biotic interactions of competition and predation strongly influence spatial model results (Davis et al., 1998; Van der Putten et al., 2010), and differential evolutionary responses could emerge from the same large-scale abiotic driver such as global temperature (Ezard et al., 2011). Finally, if future climate projections exceed the range of observed environmental conditions it becomes impossible to predict biotic response. Therefore, misleading conclusions can be drawn from analyses that use spatial species variation as a proxy for expected species response to future environmental change.

Shifts in ecological optima are of particular interest when predicting species' response to future climate change scenarios, since the optimum forms the core of a species' range. Ecological optima are often defined using combinations of traits, such as mean population abundance and physiological responses (Fritz et al., 2013; Phillimore et al., 2010), or average body size (Schmidt et al., 2004a), which indicates the presence or absence of limiting factors on growth. The peak in abundance of reproductive adults implies that the species optimum also indicates the fitness optimum, as high abundance also points to maximum population growth rates. However, no empirical data exists on whether the processes that determine among-population differences in fitness also drive within-population adaptation through time following environmental change. Therefore, the extent to which individual populations follow among-population gradients remains unknown (Schmidt et al., 2003).

Covariations between morphological traits provide another useful tool for predicting species variation. Over time selection in one trait can generate a response in a covarying trait (Lande, 1979; Lande and Arnold, 1983), which in turn impacts the direction of expected evolutionary change (Schluter, 1996). At the species level, evolution has been shown to occur along the direction of

maximum variation as determined by the covarying traits (Hunt, 2007; Renaud et al., 2006; Schluter, 1996). The principles of lines of least evolutionary resistance imply that when patterns in population-specific trait covariance align across space, all populations should respond similarly when faced with the same environmental conditions. In this case spatial trait (co)variation can indeed be used to model temporal change. However, different adaptational pressures across space could lead to selection of different trait combinations and in turn produce varying trait covariation patterns among populations. In this case, spatial variation would be a poor predictor for the most likely direction of change among populations through time.

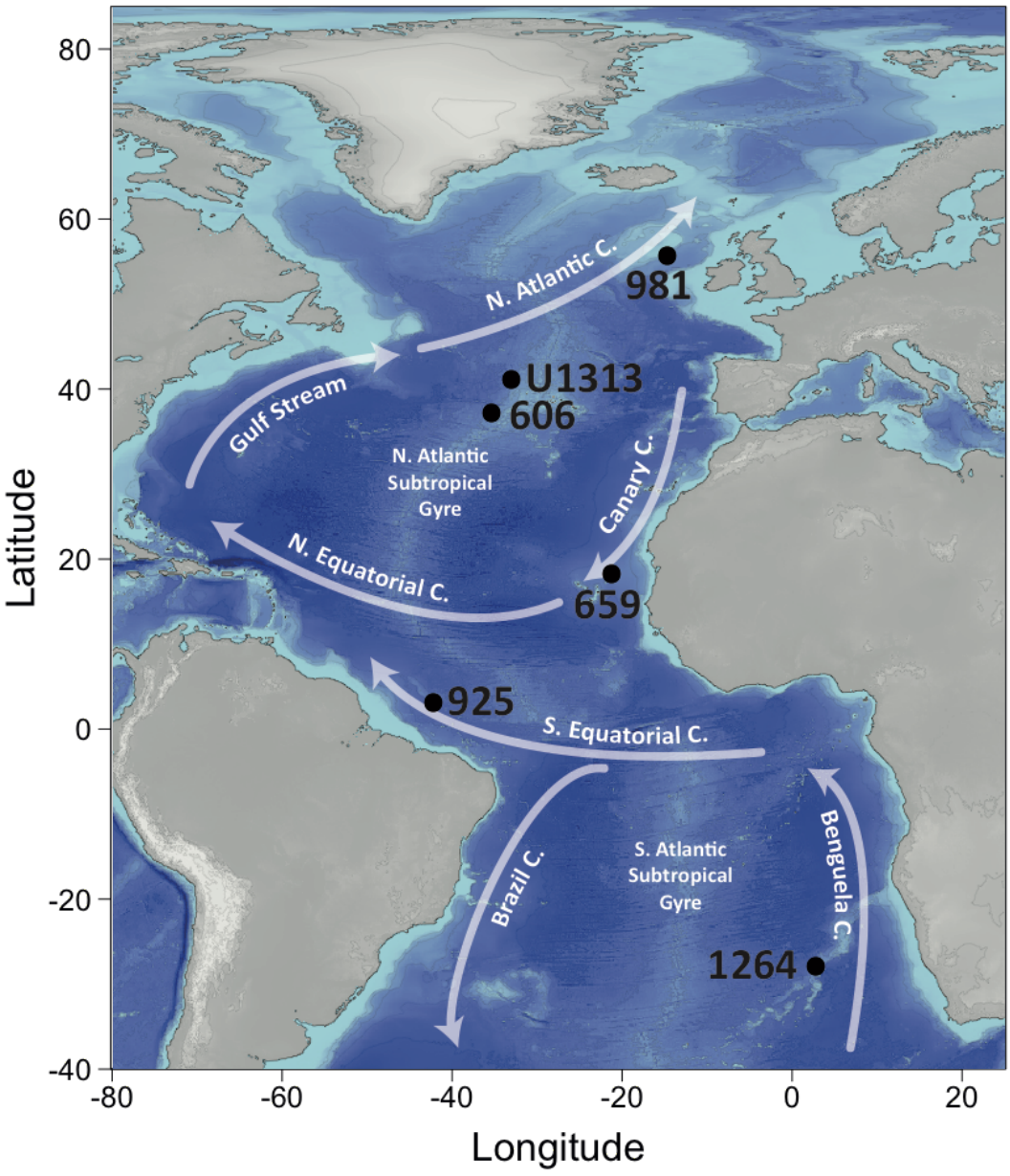

Figure 6.1 Location of study sites and main surface water currents

Here, I assemble a dataset of 29,438 individuals of planktonic foraminifera from six Atlantic sites spanning ~83˚ of latitude from the Northern Hemisphere sub-polar region to the Southern Hemisphere subtropics (Figure 6.1) that were deposited from 2.9 to 2.3 million years ago (Ma). This interval featured the last great climate transition in Earth history: the late Pliocene intensification of Northern Hemisphere glaciation (iNHG) (Mudelsee and Raymo, 2005). Planktonic foraminifera are sexually reproducing protists distributed in high abundance throughout the world's oceans. Their large population size, global distribution and the excellent preservation potential of their calcite shells allow for continuous, high-resolution morphological reconstructions at the population level over millions of years. To investigate whether variation within populations predicts total species variation through time, I study covariances between population abundance, test size and test shape in the ecologically similar species *Globoconella puncticulata* and *Truncorotalia crassaformis.* If the trait covariations among populations line up, all populations are expected to follow the same trajectories when faced with similar external drivers. However, if populations differ, their predicted responses to external drivers will vary independently. This dataset allows us to study whether within- or among population variation is more predictive of future trait dynamics during an interval of global climatic upheaval.

6.3 Methods

6.3.1 Study species

I study the morphologically and ecologically similar planktonic foraminifera species *G. puncticulata* and *T. crassaformis* (Figure 1.1). Both were globally distributed during the Pliocene, and were most abundant in thermocline waters at middle and low latitudes (Aze et al., 2011; Kennett and Srinivasan, 1983). *T. crassaformis* originated around 5.7 Ma and survives to the present day, whereas *G. puncticulata* first appeared around 4.6 Ma and became extinct at 2.41 Ma (Wei, 1994b), shortly after the onset of significant Northern Hemisphere glaciation, 2.72 Ma (Bailey et al., 2013). *G. puncticulata* gave rise to one descendant, *Globoconella inflata* around 3.5 Ma in the southwest Pacific (Wei, 1994a, b). This species was initially restricted to the Southern Hemisphere and did not migrate north until ~2.09 Ma (Berggren et al., 1995a; Chapman et al., 1998), ~300,000 years after the extinction of *G. puncticulata*.

6.3.2 Material

To study within- and among population dynamics six populations of *T. crassaformis* and *G. puncticulata* are compared over a 600,000-year interval that ends with the last local extinction of *G. puncticulata* at ~2.3 Ma (Wei, 1994b). Specimens were picked from sediments recovered from six Atlantic sites spanning ~83˚(Figure 6.1):

- Ocean Drilling Program (ODP) Site 981 (Jansen et al., 1996): a sub-polar Northern Hemisphere site (55˚ N) currently situated in the North Atlantic Current.
- Integrated Ocean Drilling Program (IODP) Site U1313 (Channell et al., 2006b): a mid-latitude Northern Hemisphere site (41˚ N) situated at the northern edge of the North Atlantic Subtropical Gyre during the Late Pliocene.
- Deep Sea Drilling Program (DSDP) Site 606 (Ruddiman et al., 1987): a mid-latitude Northern Hemisphere site located several degrees south of Site U1313 at 37˚ N, within the North Atlantic Subtropical Gyre.
- ODP Site 659 (Ruddiman et al., 1988): a subtropical Northern Hemisphere site (18˚ N) situated in the Canary Current near the West African margin.
- ODP Site 925 (Curry et al., 1995): an equatorial site (3˚ N) located in the Southern Equatorial Current transporting South Atlantic surface waters fed from the Benguela Current into the North Atlantic.
- ODP Site 1264 (Zachos et al., 2004): a subtropical Southern Hemisphere site (28˚ S) located at the eastern edge of the South Atlantic Subtropical Gyre in the Benguela Current. Of all studied sites, this is the only one with regular occurrences of *G. puncticulata*'s descendant species *G. inflata* (Zachos et al., 2004).

Samples from ODP Sites 606, 659, 981 and 925 were acquired from the IODP Bremen Core Repository in Germany. The samples from IODP Sites 1264 and U1313 were originally used in (Bell et al., 2014) and (Bolton et al., 2010) respectively. For age control I used previously published age models (Bell et al., 2014; Bickert et al., 1997; Bolton et al., 2010; Franz and Tiedemann, 2002; Mc Intyre et al., 1999; Sikes et al., 1991; Tiedemann et al., 1994) based on tuning of site-specific orbital-resolution benthic foraminiferal $\delta^{18}O$ stratigraphies to the LR04 $\delta^{18}O$ stack (Lisiecki and Raymo, 2005). To avoid time averaging of foraminifera due to bioturbation, I sampled each site at a minimum depth-resolution of 15 cm, resulting in 37 samples for Site 981 (every 45 cm, corresponding to a ~10 kyr resolution), 75 samples for Site U1313 (every 30 cm, corresponding to a ~5 kyr resolution), 36 samples at Site 606 (every 60 cm, corresponding to a ~20 kyr resolution), 57 samples at Site 659 (every 60 cm, corresponding to a ~10 kyr resolution), 19 samples at Site 925 (every 60 cm, corresponding to a ~20 kyr resolution) and 29 samples at Site 1264 (every 15 cm, corresponding to a ~20 kyr resolution). Each sample was dry-sieved over a >150 μm^2 mesh sieve and split using a microsplitter until a single split contained 50-100 individuals of *G. puncticulata* or *T. crassaformis*, all of which were picked to avoid size biases. For Site 925, occasionally, entire samples had to be picked for *G. puncticulata* due to the species' low abundance at this site. In total 29,438 individuals were analysed (16,091 specimens of *G. puncticulata* and 13,347 of *T. crassaformis*).

Three traits were studied: abundance, shell size (area) and shell shape (aspect ratio, the ratio between shell height and width). Abundance reflects a population's success, shell size has been shown to reflect a species' ecological optimum (Hecht, 1976a; Schmidt et al., 2004a) and shell shape impacts the test volume to surface area ratio, influencing respiration and metabolic processes (Caromel et al., 2014). Maximum abundance and size indicate a combination of peak fitness and absence of limiting factors on growth, which can be used to locate a species' ecological optimum (Renaud and Schmidt, 2003; Schmidt et al., 2004a), whereas the covariation between size and shape reflects the direction of expected evolutionary change in these two traits (Schluter, 1996). Species abundance per time slice was calculated by dividing the estimated total number of specimens in a sample (the number of individuals found in an analysed split multiplied by the split fraction) by the sample's coarse fraction weight, divided by the total time contained in the sample. To measure morphometrics, foraminifera shells were mounted on glass slides in groups of 20 individuals using transparent

double-sided adhesive tape with the apertures facing upwards. Groups were imaged using an Infinity 3 Lumenera camera mounted on an Olympus SZX10 light microscope, and shell area and aspect ratio (the ratio between test height and width) were extracted from the images using an automated image analysis macro in the Image Pro Premier software. Populations' positions in the species ranges (e.g. core, transition or edge) were determined by the size and abundance of

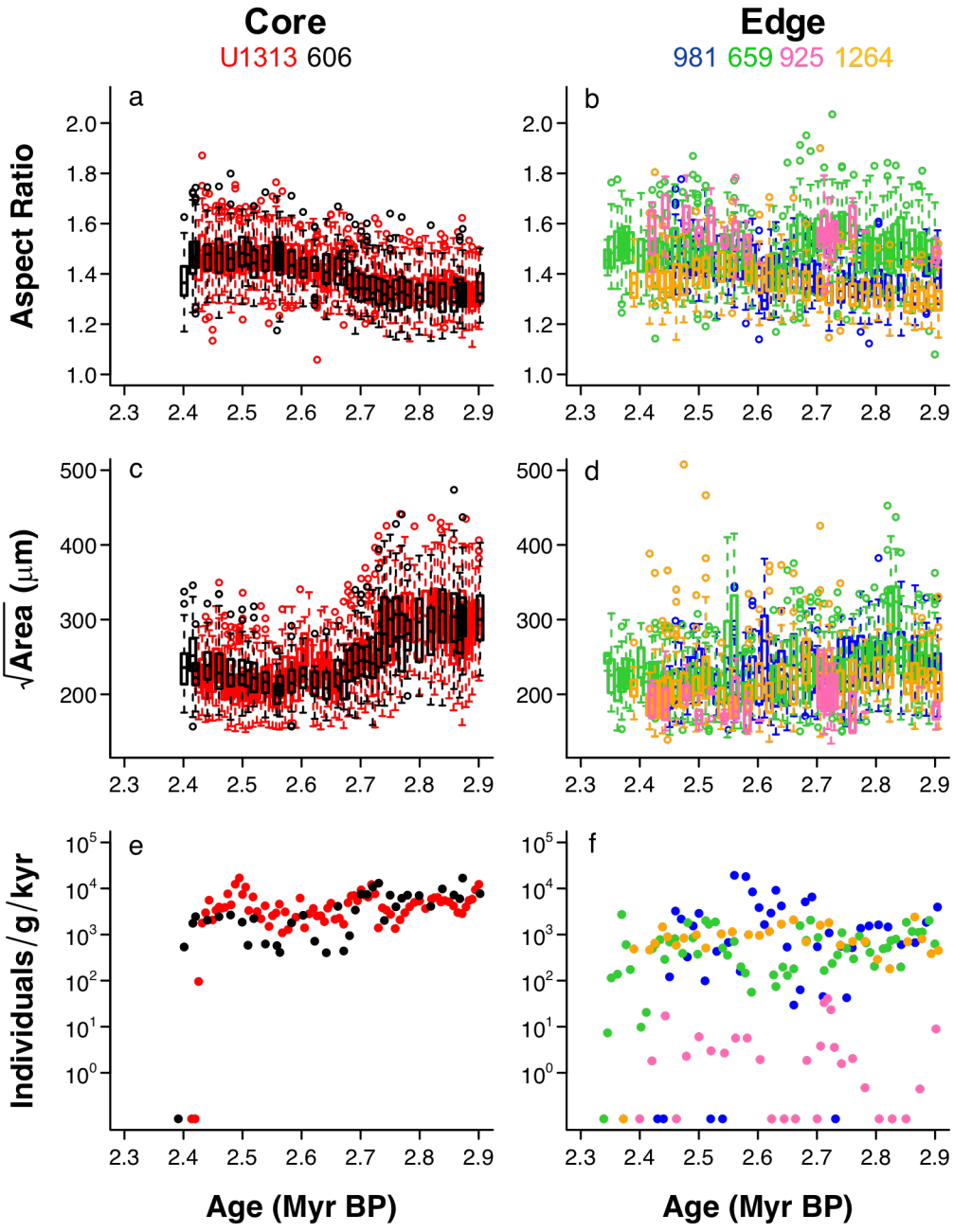

Figure 6.2 Shape (a-b), size (c,d) and abundance (e,f) through time (Myr BP) in six Atlantic populations of *Globoconella puncticulata*. Core populations indicates peak abundance and size, whereas Edge populations are characterized by lower numbers and smaller sized individuals.

individuals through time. Populations with largest individuals and highest abundance were classified as being in the core of the species' range, whereas smaller populations with reduced shell size are positioned towards the range edge.

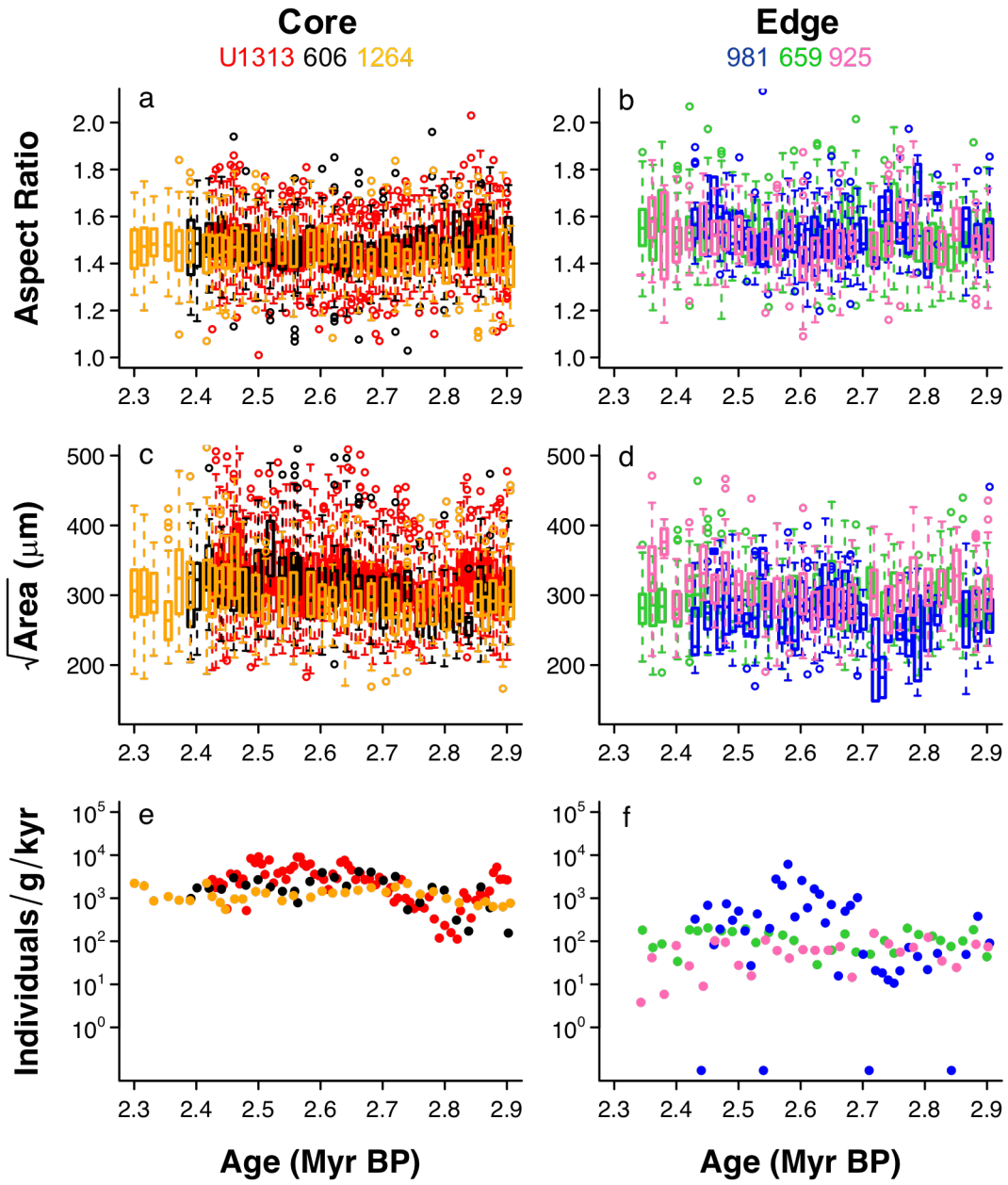

Figure 6.3 Shape (a,b), size (c,d) and abundance (e,f) through time (Myr BP) in six Atlantic populations of *Truncorotalia crassaformis*. Core populations indicates peak abundance and size, whereas Edge populations are characterized by lower numbers and smaller sized individuals.

6.3.3 Analysis

Linear Models were employed to quantify trait covariances between abundance and size, and size and shape. To study the strength of trait covariations within- and among populations, variance-covariance matrices were constructed for each population using sample abundance and mean sample shell size and shape data throughout the study interval. Measurements were scaled to unit variance to produce values in comparable units. The dominant eigenvectors of these matrices ($\mathbf{p}_{max}$) describes the direction of maximum variation within a population through time, and represents the expected direction of evolutionary change (Hunt, 2007; Schluter, 1996). The subdominant eigenvector describes the amount of variation in the direction orthogonal to $\mathbf{p}_{max}$. Together, the ellipse determined by both eigenvectors describes the variation within the population. To test whether the direction of maximum variation $\mathbf{p}_{max}$ aligns among populations I employed linear mixed effect models on the total of all sites' sample means using the 'lme4' package (Bates et al., 2014) in R (R Core Team, 2013), with 'site' as a random effect. Four models were compared: random intercepts and a zero fixed slope (m1); random intercepts and a non-zero fixed slope (m2); random slopes and intercept with a zero fixed slope (m3); random slopes and intercept with a non-zero fixed slope (m4). Models with parallel slopes (m1 and m2) support the null hypothesis of fixed trait covariation patterns among populations, whereas models supporting random slopes (m3 and m4) imply significant differences in trait covariation slopes among populations. I tested among models using the AIC, which describes the variance explained by the model minus a penalty for additional parameters to obtain a parsimonious optimal model.

6.4 Results

Abundance, shell size and shape vary widely across populations of both species (Figure 6.2, Figure 6.3). The Northern Hemisphere mid-latitude core populations of *G. puncticulata* (at ODP Site 606 and IODP Site U1313, see Figure 6.1) show highest abundance with largest shell size, although a shift in size distribution to smaller shells occurs around 2.7 Ma, shortly after the onset of significant Northern Hemisphere glaciation (Figure 6.2c,e). Lower abundance and smaller shell sizes in both sub-polar (ODP Site 981) and (sub)tropical populations (ODP Sites 659 and 925, and IODP Site 1264, Figure 6.2d,f) indicate populations in sub-optimal conditions. Frequent species disappearances in the tropical population at

ODP Site 925 (Figure 6.2f) suggest this population lies closest to the species' range edge. Differences between populations were less pronounced in *T. crassaformis*. Highest abundance and largest shells were found at the Northern Hemisphere mid-latitude and southern hemisphere subtropics populations (Sites 606, 1264 and U1313, Figure 6.3c,e), with smaller and more variable populations at the Northern Hemisphere sub-polar, subtropical and tropical sites (Figure 6.3d,f).

The Linear Models performed across all sites show significant covariations between abundance and size, and size and shape for both species (Table 6.1). Results from the Linear Mixed Effect Regressions show that models with site-specific non-parallel (random) slopes provide a significantly better fit to the data

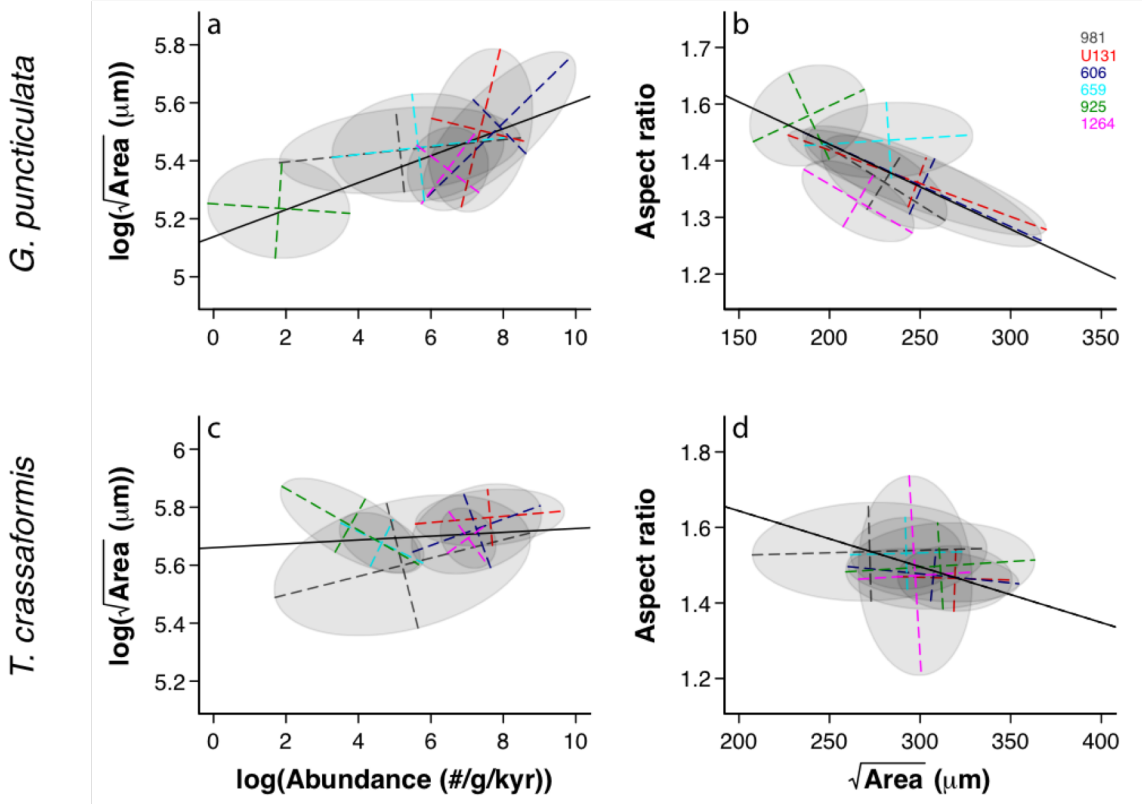

Figure 6.4 Ellipses determined by the first and second Principle Component Axes of trait combinations in all six Atlantic populations of *Globoconella puncticulata* (a,b) and *Truncorotalia crassaformis* (c,d). The black lines represent the linear regression through the all sites' time slice means (see Table 6.1 for regression coefficients). Parameters of the Linear Mixed Effect Models describing model fit with and without random slopes included are given in Table 6.2.

than models with parallel slopes for both trait covariations in *G. puncticulata*, and the combination of size and abundance in *T. crassaformis* (Table 6.2, Figure 6.4). In both trait combinations studied in *G. puncticulata*, the models with parallel site-specific slopes and a fixed among-population slope (m4) fits the data better than the models without among-population slopes (m3) (Table 6.2). Although a trend is present between size and shape in *T. crassaformis*, Mixed Effects models, including the site-specific random slopes, did not significantly improve the models' fit to the data.

Table 6.1 Parameters of Linear Models performed on all sites' means per time slice combined.

	G. puncticulata		*T. crassaformis*	
	R^2	*p*	R^2	*p*
Abundance-size	0.28	<0.001	0.10	<0.001
Size-shape	0.30	<0.001	0.05	<0.001

6.5 Discussion

Trait response varies widely across populations of both species, likely as a result of site-specific environmental and biotic pressures. Climate is distinctly variable by region, so variations in trait combinations across populations could be the result of responses, including adaptation, to local environmental conditions. Given the number of generations separating each sampling point, the observed differences in the patterns of among-population trait covariation likely represent the optimal trait expression – a peak on the adaptive landscape (Simpson, 1944) – under the prevailing environmental conditions. If all populations' covariances aligned, a shift in the peak of the adaptive landscape (Arnold et al., 2001; Simpson, 1944) towards new trait combinations would then leave the entire species in suboptimal conditions to which it would take a long time to adapt. In contrast, a species with higher among-population variability in trait covariation slopes is more likely to contain the necessary trait combinations in at least one population to move easily towards the new fitness peak. The remaining populations might then be able to catch up either through gene flow or adaption.

This way, higher within-species variability in trait covariation patterns increases morphological variability, thereby enhancing a species' evolvability (Hansen and Houle, 2008; Hansen et al., 2003).

Table 6.2 AIC scores and p-values of the linear mixed effect models. Model 1 (m1) represents random intercepts and a zero fixed slope; model 2 (m2) represents random intercepts and a non-zero fixed slope; model 3 (m3) represents random slopes and intercepts, with a zero fixed slope; model 4 (m4) represents random slopes and intercepts, with a non-zero fixed slope.

	Model					*G. puncticulata*	*T. crassaformis*
Traits		Site slopes	Total slope	AIC	*p*	AIC	*p*
Abundance -Size	m1	-	zero fixed	-372.77		-530.53	
	m2	-	Non-zero fixed	-394.58	<0.001	-532.32	0.0514
	m3	random	Zero fixed	-396.50	<0.05	-547.48	<0.001
	m4	random	Non-zero fixed	-400.20	<0.05	-545.72	0.827
Size-Shape	m1	-	zero fixed	-910.35		-782.87	
	m2	-	Non-zero fixed	-1026.64	<0.001	-791.41	0.462
	m3	random	Zero fixed	-1070.23	<0.001	-789.02	1.000
	m4	random	Non-zero fixed	-1072.79	<0.05	-787.83	0.367

In three out of four trait combinations studied here, within-population variation does not predict among-population patterns on microevolutionary time scales (>100,000 years) (Figure 6.4). Earlier results on the covariations between size and abundance among foraminifera species show that covariations among traits are present across species (Schmidt et al., 2004a), a pattern that is also observed here in the covariation of site-specific mean size and abundance of *G. puncticulata*. However, the within-population trait covariations deviate from the among-population trend as observable because the population ellipses are not parallel. These results suggest that site-specific selection drives alternative combinations of traits within populations, and thus alternative orientations of fitness peaks on the adaptive landscape. If populations are adapted to site-specific combinations of environmental parameters, their within-population trait

covariances do not therefore necessarily translate directly to populations living at other latitudes. These results corroborate earlier work on planktonic foraminifera, which showed that different combinations of environmental drivers are responsible for ecological and morphological patterns across space (Fenton et al., 2016) and through time (Brombacher et al., 2017a).

Through time, local adaptation pressures to biotic and abiotic factors likely cause differential population responses to environmental change. During my study interval, there was a time transgressive equatorward expansion of the subpolar gyres and contraction of the subtropical gyres in response to equatorward movement of westerlies (Bolton et al., 2011; Friedrich et al., 2013; Naafs et al., 2012; Naafs et al., 2010). This resulted in regional changes in sea surface temperature and nutrient concentrations throughout the Atlantic Ocean (Herbert et al., 2015; Herbert et al., 2010; Lawrence et al., 2013; Naafs et al., 2012). If populations are adapted to site-specific combinations of environmental parameters, local factors could result in preferential selection of new trait combinations, forcing a population away from the larger latitudinal trend evidenced through all population means.

Traits are hypothesised to respond to multiple environmental drivers and their interactions (Brombacher et al.), which do not always covary: sea surface temperature, for example, scales largely with latitude, whereas nutrient availability depends on the presence or absence of upwelling or downwelling, river runoff, surface water nutrient usage and aeolian dust input. Additionally, populations do not only need to adapt to site-specific mean environmental conditions, but also to latitudinally dependent variations in environmental variability on orbital (>10,000 years) timescales. The amplitude of local environmental variability also increases with latitude: Late Pliocene glacial-interglacial temperatures in the tropics varied by only 2-3 °C (Lawrence et al., 2006), as compared to 5 °C in the mid-latitudes [18] and up to 7 °C in the sub-polar regions [19] of the North Atlantic. Model simulations show that in constant environments, and given sufficient among-generation variation, specialist strategies are favoured over generalists (Gilchrist, 1995). As the core of the studied species' ranges (i.e. mid-latitudes) is more environmentally stable than the margins (sub-polar regions), different ratios of generalists to specialists will be selected for.

Biotic interactions also vary with location, so site-specific competition and predation pressures are expected to affect populations differently. The degree of

competition directly influences species abundance, and might, by reducing nutrient availability for a focal species, also affect body size. Because planktonic foraminiferal diversity is lowest in high latitude oceans and highest in the nutrient-poor subtropical gyres (Rutherford et al., 1999), varying degrees of competition pressure among planktonic foraminifera species exist along the latitudinal transect studied here. Competition with other groups of zooplankton for food sources such as phytoplankton and algae (Anderson et al., 1979) also likely varies spatially as zooplankton diversity varies globally (Irigoien et al., 2004). Site-specific predation pressures directly affect a population's abundance and size structure. Predation pressure selects for certain ranges of body size: the ratio of specialist to generalist predator species has been shown to affect plankton community structure through the degree of size-selective feeding (Fuchs and Franks, 2010). Therefore, local adaptation likely increases among-population trait variation.

Although the results presented here are based on the study of planktonic foraminifera, they are also likely to remain true for other species: individuals within populations contain varying combinations of genetic information and plastic species' responses, and need to adapt to local environmental settings, including predation and competition. These results contradict the assumption of constant trait covariation among populations, instead implying that intraspecific variation across space is a poor predictor for specific populations' predicted response to environmental change through time (Figure 6.3). Different covariation patterns among populations likely cause different responses to environmental change: selection in one trait will generate a different response in the covarying trait among populations. Given that trait covariation is predictive of morphological changes through time (Brombacher et al.) and across ancestor-descendent species pairs (Hunt, 2007), the explanatory power of spatially explicit models could be improved by including the within-population covariance structure.

6.6 Conclusion

The results presented in this chapter show that population variation across space is a poor predictor for variation through time. The within-population trait covariance patterns do not match with the variation across populations, which is

likely explained by site-specific selection pressures due to climatic or biotic processes, or genetic variation. As many other species experience different environments and biotic interactions among populations, it is likely that the mismatch between within- and among population variation holds up for other species as well. Particularly, the higher variance introduced by population-specific variance might contribute to the species' adaptation potential and therefore its chances of survival. However, our results caution against the unfettered substitution of spatial for temporal processes in model input.

Chapter 7: Conclusion

The work presented in this thesis describes new calibrations and applications of planktonic foraminifera in evolutionary biology. First, the repeatability of often-used morphological traits, the number of individuals required to reliably detect trends of varying magnitude, and the reliability of one- and two-dimensional representations of foraminiferal shells as a proxy for body size were assessed. Following that, repeatable traits were measured over a 600,000-year interval across six Atlantic populations to study trait variance and covariance through time and across space, during an interval of pronounced global climate upheaval. The results are summarized below.

Chapter 2: trait repeatability

The high-resolution fossil archives of planktonic foraminifera enable construction of evolutionary time series of large numbers (>10,000) of individuals. Foraminifera tests are generally mounted, oriented and imaged manually, while data can be processed semi-automatically using standard software such as ImageJ or Image Pro. However, manually induced orientation errors are a source of potential bias in trait measurements even when quantified using the same computational subroutine. I tested the repeatability of four measures of foraminiferal test shape on six morphologically distinct species using a novel mounting technique in which shells are imaged from glass slides to reduce background imperfections and increase the accuracy of trait capture. I also presented a calibration (power analysis) of the number of individuals needed to determine a given percentage change in these traits. Both test area and aspect ratio were found to be repeatable measures of test size and shape, whereas roundness was a repeatable measure for only half of the species assessed while perimeter was not repeatable for any of our non-spherical species. Results from the power analyses show that between 25 and 50 individuals are needed to detect a 10-15% change in the repeatable traits, which is well within the scope of most species of planktonic foraminifera. Similarly, a change in trait measurements of 10-15% is at the lower range of the changes generally reported (Brombacher et al.; Wade and Olsson, 2009; Wade et al., 2016). However, these numbers strongly depend on the morphological variation present in the species.

My results underline the need for measurement error quantification in individual species' traits prior to interpreting their morphological records. In

particular, test perimeter, and the other composite traits it influences, should be used with extreme caution in applications of geometric morphometrics. More work is needed to investigate the repeatability of individual landmarks on the test outline before they are applied for evolutionary or biostratigraphical purposes.

Chapter 3: one- and two-dimensional representations of foraminifera body size

Body size is one of the most commonly measured traits in ecology and evolution because it indicates ecological success and covaries with environmental and life-history traits. Foraminiferal body size is generally estimated by measuring the shell diameter or area from two-dimensional images. However, these one- and two-dimensional measurements could present a sub-optimal representation of body size in an evolving lineage when shell shapes vary through time. I measured test diameter and area of over 500 individuals of the species *Globoconella puncticulata* using two-dimensional images, and compared the results to measurements of test volume of the same individuals as measured by a recently developed high-throughput method for analysing fully volumetric three-dimensional morphometrics. I found that both test diameter and area covary strongly with test volume over the entire range of test sizes studied, implying that even in these evolving lineages, test diameter and area can provide consistent proxies for body size. The method presented here can be readily repeated on other species of planktonic foraminifera, validating their potential to study the driving mechanisms of changing body size in deep time.

Chapter 4: static and evolutionary allometries during climatic upheaval

The influence of within-species variation and covariation on evolutionary patterns is well established for generational and macroevolutionary processes, most prominently through genetic lines of least resistance (Bégin et al., 2003; Hunt, 2007; Lande, 1979; Lande and Arnold, 1983). However, it is not known whether intraspecific phenotypic variation also directs microevolutionary trajectories into the long term when a species is subject to varying environmental conditions. I present a high-resolution record of size and shape changes among 12,633 individual planktonic foraminifera over a 0.5-Myr time interval spanning the late Pliocene to earliest Pleistocene intensification of Northern Hemisphere glaciation.

I found that within climate intervals marked by constant environmental variability the within-population allometries predict evolutionary change from one time-step to the next, and that the within-phase among-population (i.e. evolutionary) allometries match their corresponding static (within-population) allometries. However, the evolutionary allometry across the three climate phases deviates significantly from the static and phase-specific evolutionary allometries in the extinct-going species. These results imply that allometric constraints can be overcome by environmental perturbations.

Chapter 5: synergistic climate forcing of foraminifera evolution

Changes in biodiversity are often linked to climate change, usually represented by single parameters such as global temperature (Erwin, 2009; Mayhew et al., 2008). However, climate consists of many interacting variables and species likely respond to the entire climate system as opposed to individual variables (Harnik et al., 2012). In this chapter, I use the ecological and morphological trait measurements of the two species of planktonic foraminifera generated in Chapter 3 and compare them to local and regional records of sea surface temperature, productivity, nutrient availability and global atmospheric CO_2. The results show that phenotypic and ecological changes are indeed driven by the interactions of multiple environmental factors. No dominant climatic driver could be identified and temperature alone explains remarkably little variance, implying that successful reconstructions of palaeobiological dynamics will require multiple ecological and evolutionary metrics.

Chapter 6: evolution across space and through time

To predict biotic responses to future climate change, a common approach is to extrapolate species' responses to current environmental conditions into the future. These projections assume that the species' adaptive potential is similar across its geographical range, and that environmental variability among contemporary populations is a relevant predictive tool. Using covariations between ecological and morphological traits in 29,438 individuals of two planktonic foraminifera from six Atlantic populations, I show that within-population temporal dynamics had negligible predictive power for the among-population changes. These results imply that each population responds to site-

specific environmental and biotic settings, complicating extrapolation to other locations. As within-population variation does not predict among-population dynamics over time, diversity should be tracked through both space and time as one can be a poor substitute for the other.

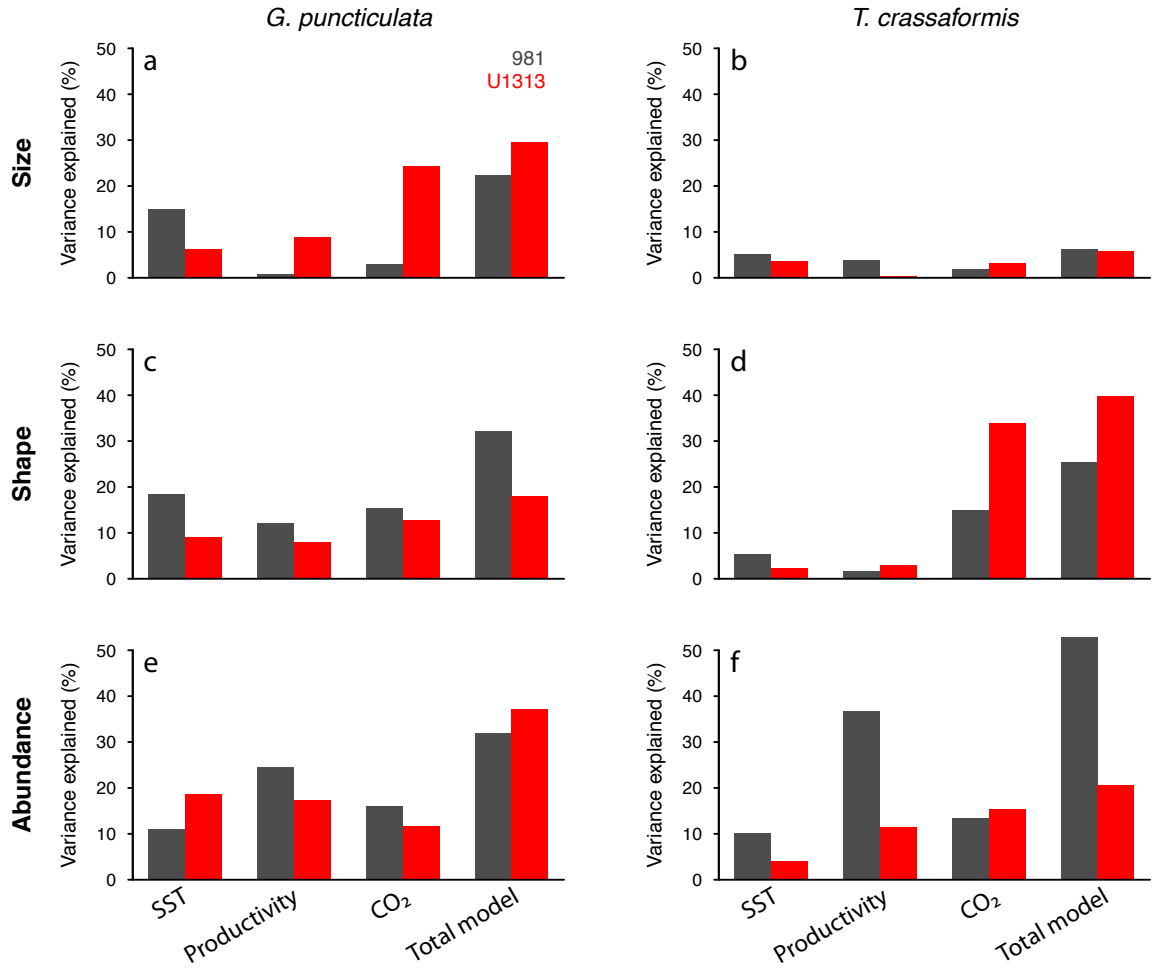

Figure 7.1 Variance explained in size (a,b), shape (c,d) and abundance (e,f) of *Globoconella puncticulata* and *Truncorotalia crassaformis* from North Atlantic Sites 981 (grey) and U1313 (red) by the available environmental parameters and their interactions ('Total model').

Future work

The fossil record of planktonic foraminifera is ideally suited to study the driving mechanisms of evolutionary processes. Its high resolution and continuous deposition of full-body fossils enable reconstructions of evolutionary time series at both micro- and macroevolutionary time scales, and the possibility to generate

climate records from the same sediment samples allows for direct comparisons between biotic and abiotic change. The automated analysis and measurement setup described in this thesis makes it possible to generate the large datasets required for long-term and high-resolution evolutionary time series.

The work presented here focusses on only two species, over a relatively short time interval in their existence. To increase our understanding of foraminifera evolution the study interval could be expanded to include species' entire biostratigraphic ranges. This will enable studies of the consistency of allometries and other trait covariations within and across species, and test whether evolutionary change is relatively constant, or speeds up/slows down during different key stages in its existence such as during speciation and preceding extinction. It would also show if a species' response to climate is stable across its range, or whether it varies through time. It would also be beneficial to include more species in the studied time interval from various ranges and depth habitats. I show that two study species respond differently to environmental change even though they inhabit the same range and depth habitat, suggesting even more varying responses in species with different ecological preferences.

The traits described in this thesis represent very simple representations of overall shell size and shape. However, due to the preservation of full-body fossils in the foraminiferal fossil record many more traits can be extracted from foraminiferal shells, such as the size and shape of individual chambers, size and shape of the aperture, and various landmarks around the shell's perimeter. Studying the variation in and covariation among these traits will increase our understanding of trait covariation in evolving lineages through time.

The results presented in Chapter 6 warn against the use of spatial variation to predict biotic responses through time, which is particularly problematic for predictions of present-day biodiversity to projected future climate change. One way to improve model predictions could be to study trait covariances among contemporary populations. The degree of overlap among site-specific covariation patterns, as well as biotic interactions could provide an indication of the similarity between populations' responses, and therefore help validate the model's results. Additionally, any biotic response to environmental change should be tested among populations. Preliminary work on two populations in the core (Site U1313) and edge (Site 981) of the study species' ranges show that response to environmental change varies between sites (Figure 7.1), further consolidating the results of Chapter 5 that environmental variability is not a constant predictor of

biotic change across space. Therefore, to assess whether climatic variability is a relevant predictive tool in contemporary populations, multiple-year time series of biotic response to environmental variability are required.

Additionally, despite the growing body of work on past climatic reconstructions most evolutionary studies still compare biotic records to single environmental variables such as the benthic oxygen isotope stack. However, the variance in biodiversity explained by this parameter is generally relatively low (10-35% (Lazarus et al., 2014; Mayhew et al., 2008)) and could likely be improved by including multiple variables. For example, Lazarus et al. (2014) show that diatom diversity is also strongly correlated to $\delta^{13}C$ and atmospheric CO_2, implying that a multivariate model combining all studied factors would do better at explaining biotic change. More Cenozoic reconstructions of environmental parameters such as local sea surface temperature, productivity and nutrient availability (in the marine realm), and proxies for the hydrological cycle and aridity such as dust deposition and charcoal records are becoming available, enabling direct comparisons between local environmental and biotic change.

To improve understanding of evolutionary change as a result of environmental variability, many more time intervals characterized by a wide range of environmental settings and rates of climate change should be considered. For example, comparing biotic responses during Paleogene transient climate events such as the Paleocene Eocene Thermal Optimum or the Middle Eocene Climatic Optimum to late Neogene glacial-interglacial climate cyclicity will enable studies of species' adaptation potential to a changing environment. Together, the results will expand the knowledge of the processes shaping life on Earth.

Bibliography

Agrawal, A.F., Stinchcombe, J.R., 2009. How much do genetic covariances alter the rate of adaptation? Proceedings of the Royal Society B: Biological Sciences 276, 1183-1191.

Alizon, S., Kucera, M., Jansen, V.A., 2008. Competition between cryptic species explains variations in rates of lineage evolution. Proceedings of the National Academy of Sciences of the United States of America 105, 12382-12386.

Allmon, W.D., Martin, R.E., 2014. Seafood through time revisited: the Phanerozoic increase in marine trophic resources and its macroevolutionary consequences. Paleobiology 40, 256-287.

Anderson, O., Spindler, M., Bé, A., Hemleben, C., 1979. Trophic activity of planktonic foraminifera. Journal of the Marine Biological Association of the United Kingdom 59, 791-799.

Araújo, M.B., New, M., 2007. Ensemble forecasting of species distributions. Trends in ecology & evolution 22, 42-47.

Arnold, A.J., 1983. Phyletic evolution in the Globorotalia crassaformis (Galloway and Wissler) lineage: a preliminary report. Paleobiology 9, 390-397.

Arnold, S.J., 1992. Constraints on phenotypic evolution. American Naturalist 140, S85-S107.

Arnold, S.J., Bürger, R., Hohenlohe, P.A., Ajie, B.C., Jones, A.G., 2008. Understanding the evolution and stability of the G-matrix. Evolution 62, 2451-2461.

Arnold, S.J., Pfrender, M.E., Jones, A.G., 2001. The adaptive landscape as a conceptual bridge between micro-and macroevolution. Genetica 112, 9-32.

Aze, T., Ezard, T.H., Purvis, A., Coxall, H.K., Stewart, D.R., Wade, B.S., Pearson, P.N., 2011. A phylogeny of Cenozoic macroperforate planktonic foraminifera from fossil data. Biological reviews of the Cambridge Philosophical Society 86, 900-927.

Bailey, I., Hole, G.M., Foster, G.L., Wilson, P.A., Storey, C.D., Trueman, C.N., Raymo, M.E., 2013. An alternative suggestion for the Pliocene onset of major northern hemisphere glaciation based on the geochemical provenance of North Atlantic Ocean ice-rafted debris. Quaternary Science Reviews 75, 181-194.

Barnosky, A.D., 2001. Distinguishing the effects of the Red Queen and Court Jester on Miocene mammal evolution in the northern Rocky Mountains. Journal of Vertebrate Paleontology 21, 172-185.

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., 2011. Has the Earth/'s sixth mass extinction already arrived? Nature 471, 51-57.

Bartoli, G., Sarnthein, M., Weinelt, M., Erlenkeuser, H., Garbe-Schönberg, D., Lea, D.W., 2005. Final closure of Panama and the onset of northern hemisphere glaciation. Earth and Planetary Science Letters 237, 33-44.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.

Bégin, M., Roff, D.A., Phillips, P., 2003. The constancy of the G matrix through species divergence and the effects of quantitative genetic constraints on phenotypic evolution: a case study in crickets. Evolution 57, 1107-1120.

Beldade, P., Koops, K., Brakefield, P.M., 2002. Developmental constraints versus flexibility in morphological evolution. Nature 416, 844-847.

Bell, D.A., Jung, S.J.A., Kroon, D., Lourens, L.J., Hodell, D.A., 2014. Local and regional trends in Plio-Pleistocene d18O records from benthic foraminifera. Geochemistry, Geophysics, Geosystems, 3304 - 3321.

Bell, D.B., Jung, S.J., Kroon, D., Hodell, D.A., Lourens, L.J., Raymo, M.E., 2015. Atlantic deep-water response to the Early Pliocene shoaling of the Central American seaway. Scientific reports 5.

Bell, M.A., Travis, M.P., Blouw, D.M., 2006. Inferring natural selection in a fossil threespine stickleback. Paleobiology 32, 562-577.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., Courchamp, F., 2012. Impacts of climate change on the future of biodiversity. Ecology letters 15, 365-377.

Benton, M.J., 2009. The Red Queen and the Court Jester: Species Diversity and the Role of Biotic and Abiotic Factors Through Time. Science 323, 728-732.

Berggren, W.-A., Hilgen, F., Langereis, C., Kent, D.V., Obradovich, J., Raffi, I., Raymo, M.E., Shackleton, N., 1995a. Late Neogene chronology: new perspectives in high-resolution stratigraphy. Geological Society of America Bulletin 107, 1272-1287.

Berggren, W.A., Kent, D.V., Swisher, C.C., Aubry, M.-P., 1995b. A revised Cenozoic geochronology and chronostratigraphy. Geochronology, Time Scales, and Global Stratigraphic Correlation, SPEM special publication 54, 129-212.

Bergmann, C., 1848. Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse.

Bickert, T., Curry, W.B., Wefer, G., 1997. Late Pliocene to Holocene (2.6 - 0 Ma) Western Equatorial Atlantic deep-water circulation: inferences from benthic stable isotopes. Proceedings of the Ocean Drilling Program, Scientific Results 154, 239-254.

Bijma, J., Erez, J., Hemleben, C., 1990. Lunar and semi-lunar reproductive cycles in some spinose planktonic foraminifers. Journal of foraminiferal research 20, 117-127.

Biolzi, M., 1991. Morphometric analyses of the Late Neogene planktonic foraminiferal lineage *Neogloboquadrina dutertrei*. Marine Micropaleontology 18, 129-142.

Blois, J.L., Hadly, E.A., 2009. Mammalian response to Cenozoic climatic change. Annual Review of Earth and Planetary Sciences 37, 181-208.

Blow, W.H., 1969. Late Middle Eocene to Recent planktonic foraminiferal biostratigraphy, Proceedings of the first international conference on planktonic microfossils. EJ Brill Leiden, pp. 199-422.

Bolli, H.M., Saunders, J.B., Perch-Nielsen, K., 1989. Planktic Foraminifera, Calcareous Nannofossils and Calpionellids. Cambridge University Press, New York.

Bolton, C.T., Lawrence, K.T., Gibbs, S.J., Wilson, P.A., Herbert, T.D., 2011. Biotic and geochemical evidence for a global latitudinal shift in ocean biogeochemistry and export productivity during the late Pliocene. Earth and Planetary Science Letters 308, 200-210.

Bolton, C.T., Wilson, P.A., Bailey, I., Friedrich, O., Beer, C.J., Becker, J., Baranwal, S., Schiebel, R., 2010. Millennial-scale climate variability in the subpolar North Atlantic Ocean during the late Pliocene. Paleoceanography 25.

Brassell, S.C., Eglinton, G., Marlowe, I.T., Pflaumann, U., Sarnthein, M., 1986. Molecular stratigraphy: a new tool for climatic assessment. Nature 320.

Brombacher, A., Wilson, P.A., Bailey, I., Ezard, T.H.G., The breakdown of static and evolutionary allometries during climatic upheaval. In press at American Naturalist.

Brombacher, A., Wilson, P.A., Bailey, I., Ezard, T.H.G., Temperature is a poor proxy for synergistic climate forcing of plankton evolution. Submitted to PNAS.

Brombacher, A., Wilson, P.A., Ezard, T.H., 2017. Calibration of the repeatability of foraminiferal test size and shape measures with recommendations for future use. Marine Micropaleontology 133, 21-27.

Brook, B.W., Sodhi, N.S., Bradshaw, C.J., 2008. Synergies among extinction drivers under global change. Trends Ecol Evol 23, 453-460.

Budd, A.F., Johnson, K.G., Stemann, T.A., 1996. Plio-Pleistocene turnover and extinctions in the Caribbean reef coral fauna. Evolution and Environment in Tropical America. University of Chicago Press, Chicago, 168-204.

Caromel, A.G., Schmidt, D.N., Fletcher, I., Rayfield, E.J., 2016. Morphological change during the ontogeny of the planktic foraminifera. Journal of Micropalaeontology 35, 2-19.

Caromel, A.G.M., Schmidt, D.N., Phillips, J.C., Rayfield, E.J., 2014. Hydrodynamic constraints on the evolution and ecology of planktic foraminifera. Marine Micropaleontology 106, 69-78.

Ceballos, G., Ehrlich, P.R., Barnosky, A.D., García, A., Pringle, R.M., Palmer, T.M., 2015. Accelerated modern human–induced species losses: Entering the sixth mass extinction. Science advances 1, e1400253.

Chaisson, W., 2003. Vicarious living: Pliocene menardellids between an isthmus and an ice sheet. Geology 31, 1085-1088.

Channell, J.E.T., Kanamatsu, T., Sato, T., Stein, R., Alvarez Zarikian, C.A., Malone, M.J., Expedition303/306Scientists, 2006a. Site U1313. Proc. IODP, 303/306: College Station TX (Integrated Ocean Drilling Program Management International, Inc.). 303/306, 1-124.

Channell, J.E.T., Kanamatsu, T., Sato, T., Stein, R., Alvarez Zarikian, C.A., Malone, M.J., Expedition 303/306 Scientists, 2006b. Site U1313. Proceedings of the Integrated Ocean Drilling Program 303/306.

Chapman, M.R., Funnell, B.M., Weaver, P.P.E., 1998. Isolation, extinction and migration within Late Pliocene populations of the planktonic foraminiferal lineage

Globorotalia (*Globoconella*) in the North Atlantic. Marine Micropaleontology 33, 203-222.

Cheverud, J.M., 1984. Quantitative genetics and developmental constraints on evolution by selection. Journal of Theoretical Biology 110, 155-171.

Cheverud, J.M., 1996. Developmental integration and the evolution of pleiotropy. American Zoologist 36, 44-50.

Clavel, J., Morlon, H., 2017. Accelerated body size evolution during cold climatic periods in the Cenozoic. Proceedings of the National Academy of Sciences 114, 4183-4188.

Cohen, J., 1988. Statistical power analysis for the behavioural sciences. Hillside. NJ: Lawrence Earlbaum Associates.

Conover, R., 1982. Interrelations between microzooplankton and other plankton organisms, Annales de l'Institut océanographique, pp. 31-46.

Cope, E.D., 1887. The origin of the fittest: essays on evolution. D. Appleton.

Coxall, H.K., Pearson, P.N., Shackleton, N.J., Hall, M.A., 2000. Hantkeninid depth adaptation: an evolving life strategy in a changing ocean. Geology 28, 87-90.

Coxall, H.K., Wilson, P.A., Pälike, H., Lear, C.H., Backman, J., 2005. Rapid stepwise onset of Antarctic glaciation and deeper calcite compensation in the Pacific Ocean. Nature 433, 53-57.

Curry, W.B., Shackleton, N.J., Richter, C., Expedition 154 Scientists, 1995. Site 925. Proceedings of the Ocean Drilling Program, Initial Reports 154.

Darling, E.S., Côté, I.M., 2008. Quantifying the evidence for ecological synergies. Ecology letters 11, 1278-1286.

Darling, K.F., Wade, C.M., 2008. The genetic diversity of planktic foraminifera and the global distribution of ribosomal RNA genotypes. Marine Micropaleontology 67, 216-238.

Darling, K.F., Wade, C.M., Kroon, D., Brown, A.J.L., 1997. Planktic foraminiferal molecular evolution and their polyphyletic origins from benthic taxa. Marine Micropaleontology 30, 251-266.

Darling, K.F., Wade, C.M., Kroon, D., Leigh Brown, A.J., Bijma, J., 1999. The diversity and distribution of modern planktic foraminiferal small subunit ribosomal RNA genotypes nd their potential as tracers of present and past ocean circulation. Paleoceanography 14, 3 - 12.

Davis, A.J., Jenkinson, L.S., Lawton, J.H., Shorrocks, B., Wood, S., 1998. Making mistakes when predicting shifts in species range in response to global warming. Nature 391, 783-786.

de Vargas, C., Norris, R.D., Zaninetti, L., Gibb, S.W., Pawlowski, J., 1999. Molecular evidence of crytic speciation in planktonic foraminifers and their relation to oceanic provinces. PNAS 96, 2864 - 2868.

DeConto, R.M., Pollard, D., Wilson, P.A., Pälike, H., Lear, C.H., Pagani, M., 2008. Thresholds for Cenozoic bipolar glaciation. Nature 455, 652-656.

Edgar, K.M., Bohaty, S.M., Gibbs, S.J., Sexton, P.F., Norris, R.D., Wilson, P.A., 2013. Symbiont 'bleaching' in planktic foraminifera during the Middle Eocene Climatic Optimum. Geology 41, 15-18.

Ehrmann, W., 1998. Implications of late Eocene to early Miocene clay mineral assemblages in McMurdo Sound (Ross Sea, Antarctica) on paleoclimate and ice dynamics. Palaeogeography, Palaeoclimatology, Palaeoecology 139, 213-231.

Emiliani, C., 1954. Depth habitats of some species of pelagic foraminifera as indicated by oxygen isotope ratios. American Journal of Science 252, 149-158.

Emiliani, C., 1955. Pleistocene temperature. The Journal of Geology 63.

Erez, J., Almogi-Labin, A., Avraham, S., 1991. On the life history of planktonic foraminifera: lunar reproduction cycle in Globigerinoides sacculifer (Brady). Paleoceanography 6, 295-306.

Erwin, D.H., 2000. Macroevolution is more than the repeated rounds of microevolution. Evolution & Development 2, 78-84.

Erwin, D.H., 2009. Climate as a driver of evolutionary change. Current biology : CB 19, R575-583.

Estes, S., Arnold, S.J., 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. The American Naturalist 169, 227-244.

Ezard, T.H., Purvis, A., 2016. Environmental changes define ecological limits to species richness and reveal the mode of macroevolutionary competition. Ecology letters 19, 899-906.

Ezard, T.H.G., Aze, T., Pearson, P.N., Purvis, A., 2011. Interplay Between Changing Climate and Species' Ecology Drives Macroevolutionary Dynamics. Science 332, 349-351.

Fenton, I.S., Pearson, P.N., Jones, T.D., Purvis, A., 2016. Environmental Predictors of Diversity in Recent Planktonic Foraminifera as Recorded in Marine Sediments. PloS one 11, e0165522.

Figueirido, B., Janis, C.M., Pérez-Claros, J.A., De Renzi, M., Palmqvist, P., 2012. Cenozoic climate change influences mammalian evolutionary dynamics. Proceedings of the National Academy of Sciences 109, 722-727.

Firmat, C., Lozano-Fernández, I., Agustí, J., Bolstad, G.H., Cuenca-Bescós, G., Hansen, T.F., Pélabon, C., 2014. Walk the line: 600000 years of molar evolution constrained by allometry in the fossil rodent *Mimomys savini*. Phil. Trans. R. Soc. B 369, 20140057.

Foote, M., Sepkoski, J.J., 1999. Absolute measures of the completeness of the fossil record. Nature 398, 415-417.

Foster, J.B., 1964. Evolution of mammals on islands. Nature 202, 234-235.

Franz, S.-O., Tiedemann, R., 2002. Stable isotope data of Hole 154-925D (Tab. A2), PANGAEA, doi: 10.1594/PANGAEA. 67527 Reference (s): Franz, Sven-Oliver (1999): Pliozäne Zeitreihen zur Rekonstruktion der Tiefenwasserzirkulation und der siliziklastischen Amazonasfracht im äquatorialen Westatlantik (Ceara Schwelle, ODP Leg 154). GEOMAR Report, GEOMAR Research Center for Marine Geosciences, Christian Albrechts University in Kiel 84, 183.

Friedrich, O., Wilson, P.A., Bolton, C.T., Beer, C.J., Schiebel, R., 2013. Late Pliocene to early Pleistocene changes in the North Atlantic Current and suborbital-scale sea-surface temperature variability. Paleoceanography 28, 274-282.

Fritz, S.A., Schnitzler, J., Eronen, J.T., Hof, C., Böhning-Gaese, K., Graham, C.H., 2013. Diversity in time and space: wanted dead and alive. Trends in Ecology & Evolution 28, 509-516.

Fuchs, H.L., Franks, P.J., 2010. Plankton community properties determined by nutrients and size-selective feeding. Marine Ecology Progress Series 413, 1-15.

Fukami, T., Wardle, D.A., 2005. Long-term ecological dynamics: reciprocal insights from natural and anthropogenic gradients. Proceedings of the Royal Society of London B: Biological Sciences 272, 2105-2115.

Futuyma, D.J., 2010. Evolutionary constraint and ecological consequences. Evolution 64, 1865-1884.

Garcia, R.A., Cabeza, M., Rahbek, C., Araujo, M.B., 2014. Multiple dimensions of climate change and their implications for biodiversity. Science 344, 1247579.

Gavrilets, S., Losos, J.B., 2009. Adaptive radiation: contrasting theory with data. Science 323, 732-737.

Gilchrist, G.W., 1995. Specialists and generalists in changing environments. I. Fitness landscapes of thermal sensitivity. The American Naturalist 146, 252-270.

Gingerich, P.D., 2001. Rates of evolution on the time scale of the evolutionary process. Genetica 112, 127-144.

Godbold, J.A., Solan, M., 2013. Long-term effects of warming and ocean acidification are modified by seasonal variation in species responses and environmental conditions. Phil. Trans. R. Soc. B 368, 20130186.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecological modelling 135, 147-186.

Haber, A., 2016. Phenotypic Covariation and Morphological Diversification in the Ruminant Skull. The American Naturalist 187, 576-591.

Hansen, T.F., Houle, D., 2008. Measuring and comparing evolvability and constraint in multivariate characters. J Evol Biol 21, 1201-1219.

Hansen, T.F., Pelabon, C., Armbruster, W.S., Carlson, M.L., 2003. Evolvability and genetic constraint in Dalechampia blossoms: components of variance and measures of evolvability. Journal of evolutionary biology 16, 754-766.

Hargreaves, A.L., Samis, K.E., Eckert, C.G., 2014. Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. The American Naturalist 183, 157-173.

Harnik, P.G., Lotze, H.K., Anderson, S.C., Finkel, Z.V., Finnegan, S., Lindberg, D.R., Liow, L.H., Lockwood, R., McClain, C.R., McGuire, J.L., O'Dea, A., Pandolfi, J.M., Simpson, C., Tittensor, D.P., 2012. Extinctions in ancient and modern seas. Trends in Ecology & Evolution 27, 608-617.

Hart, M., Hylton, M., Oxford, M., Price, G., Hudson, W., Smart, C., 2003. The search for the origin of the planktic Foraminifera. Journal of the Geological Society 160, 341-343.

Haug, G.H., Sigman, D.M., Tiedemann, R., Pedersen, T.F., Sarnthein, M., 1999. Onset of permanent stratification in the subarctic Pacific Ocean. Nature 401.

Haug, G.H., Tiedemann, R., 1998. Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. Nature 393.

Haug, G.H., Tiedemann, R., Zahn, R., Ravelo, A.C., 2001. Role of Panama uplift on oceanic freshwater balance. Geology 29, 207-210.

He, F., Hubbell, S.P., 2011. Species-area relationships always overestimate extinction rates from habitat loss. Nature 473, 368-371.

Hecht, A.D., 1976a. An ecologic model for test size variation in recent planktonic foraminifera: applications to the fossil record. Journal of Foraminiferal Research 6, 295-311.

Hecht, A.D., 1976b. An ecological model for test size variation in recent planktonic foraminifera: applications to the fossil record. Journal of Foraminiferal Research 6, 295-311.

Hemleben, C., Spindler, M., Anderson, O.R., 1989. Modern planktonic foraminifera.

Hennissen, J.A.I., Head, M.J., De Schepper, S., Groeneveld, J., 2014. Palynological evidence for a southward shift of the North Atlantic Current at ~2.6 Ma during the intensification of late Cenozoic Northern Hemisphere glaciation. Paleoceanography 29, 564-580.

Herbert, T.D., Ng, G., Cleaveland Peterson, L., 2015. Evolution of Mediterranean sea surface temperatures 3.5–1.5 Ma: Regional and hemispheric influences. Earth and Planetary Science Letters 409, 307-318.

Herbert, T.D., Peterson, L.C., Lawrence, K.T., Liu, Z., 2010. Tropical ocean temperatures over the past 3.5 million years. Science 328, 1530-1534.

Hillenbrand, C.-D., Cortese, G., 2006. Polar stratification: A critical view from the Southern Ocean. Palaeogeography, Palaeoclimatology, Palaeoecology 242, 240-252.

Holt, R.D., 1990. The microevolutionary consequences of climate change. Trends in Ecology & Evolution 5, 311-315.

Hsiang, A.Y., Elder, L.E., Hull, P.M., 2016. Towards a morphological metric of assemblage dynamics in the fossil record: a test case using planktonic foraminifera. Phil. Trans. R. Soc. B 371, 20150227.

Hull, P.M., Norris, R.D., 2009. Evidence for abrupt speciation in a classic case of gradual evolution. Proc Natl Acad Sci U S A 106, 21224-21229.

Hunt, G., 2006. Fitting and comparing models of phyletic evolution: random walks and beyond. Paleobiology 32, 578-601.

Hunt, G., 2007. Evolutionary divergence in directions of high phenotypic variance in the ostracode genus *Poseidonamicus*. Evolution 61, 1560-1576.

Huntley, B., Berry, P.M., Cramer, W., McDonald, A.P., 1995. Special paper: modelling present and potential future ranges of some European higher plants using climate response surfaces. Journal of Biogeography, 967-1001.

Huxley, J., 1932. Problems of relative growth. Methuen, London.

Irigoien, X., Huisman, J., Harris, R.P., 2004. Global biodiversity patterns of marine phytoplankton and zooplankton. Nature 429, 863-867.

Jackson, J.B., Johnson, K.G., 2009. Life in the last few million years.

James, F.C., 1970. Geographic size variation in birds and its relationship to climate. Ecology 51, 365-390.

Jansen, E., Fronval, T., Rack, F., Channell, J.E., 2000. Pliocene-Pleistocene ice rafting history and cyclicity in the Nordic Seas during the last 3.5 Myr. Paleoceanography 15, 709-721.

Jansen, E., Raymo, M.E., Blum, P., Expedition 162 Scientists, 1996. Sites 980/981. Proceedings of the Ocean Drilling Program, Initial Reports 162.

Jaramillo, C., Rueda, M.J., Mora, G., 2006. Cenozoic plant diversity in the Neotropics. Science 311, 1893-1896.

Jenkins, D.G., 1992. The paleogeography, evolution and extinction of Late Miocene-Pleistocene planktonic foraminifera from the southwest Pacific. Centenary of Japanese Micopaleontology. Terra Scientific Publishing Company, Tokyo, 27-35.

Jonkers, L., Reynolds, C.E., Richey, J., Hall, I.R., 2015. Lunar periodicity in the shell flux of planktonic foraminifera in the Gulf of Mexico. Biogeosciences 12, 3061-3070.

Karas, C., Nürnberg, D., Bahr, A., Groeneveld, J., Herrle, J.O., Tiedemann, R., 2017a. Pliocene oceanic seaways and global climate. Scientific reports 7, 39842.

Karas, C., Nürnberg, D., Tiedemann, R., Bahr, A., Groeneveld, J., Herrle, J., deMenocal, P., 2017b. Global cooling and the role of the Pliocene constrictions of tropical seaways. Scientific Reports 7.

Kawagata, S., Hayward, B.W., Gupta, A.K., 2006. Benthic foraminiferal extinctions linked to late Pliocene–Pleistocene deep-sea circulation changes in the northern Indian Ocean (ODP Sites 722 and 758). Marine Micropaleontology 58, 219-242.

Keigwin, L., 1982. Isotopic paleoceanography of the Caribbean and East Pacific: role of Panama uplift in late Neogene time. Science 217, 350-353.

Kennett, J.P., Srinivasan, M.S., 1983. Neogene planktonic foraminifera. A phylogenetic atlas. Hutchinson Ross Publishing Company, Stroudsburg, Pennsylvania.

Kitchener, A.C., Dugmore, A.J., 2000. Biogeographical change in the tiger, Panthera tigris. Animal Conservation 3, 113-124.

Knappertsbusch, M., 2000. Morphologic evolution of the coccolithophorid *Calcidiscus leptoporus* from the Early Miocene to Recent. Journal of Paleontology 74, 712-730.

Knappertsbusch, M., 2007. Morphological variability of *Globorotalia menardii* (planktonic foraminifera) in two DSDP cores from the Caribbean Sea and the Eastern Equatorial Pacific. Carnets de Géologie/Notebooks on Geology, Brest, Article 4, 1-34.

Knappertsbusch, M., Binggeli, D., Herzig, A., Schmutz, L., Stapfer, S., Schneider, C., Eisenecker, J., Widmer, L., 2009. AMOR - A new system for automated imaging of microfossils for morphometric analyses. Palaeontologia Electronica 12, 1-20.

Knuth, D.E., 1969. The Art of Computer Programming vol. II Seminumerical Algorithms. Addison-Wesley, Boston, USA.

Kucera, M., Malmgren, B., 1998. Differences between evolution of mean form and evolution of new morphotypes: an example from Late Cretaceous planktic foraminifera. Paleobiology 24, 49-63.

Lande, R., 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33, 402-416.

Lande, R., 1980. Genetic variation and phenotypic evolution during allopatric speciation. American Naturalist 116, 463-479.

Lande, R., Arnold, S.J., 1983. The measurement of selection on correlated characters. Evolution 37, 1210-1226.

Lang, D.C., Bailey, I., Wilson, P.A., Beer, C.J., Bolton, C.T., Friedrich, O., Newsam, C., Spencer, M.R., Gutjahr, M., Foster, G.L., Cooper, M.J., Milton, J.A., 2014. The transition on North America from the warm humid Pliocene to the glaciated Quaternary traced by eolian dust deposition at a benchmark North Atlantic Ocean drill site. Quaternary Science Reviews 93, 125-141.

Lang, D.C., Bailey, I., Wilson, P.A., Chalk, T.B., Foster, G.L., Gutjahr, M., 2016. Incursions of southern-sourced water into the deep North Atlantic during late Pliocene glacial intensification. Nature Geoscience 9, 375-379.

Lawrence, K.T., Liu, Z., Herbert, T.D., 2006. Evolution of the eastern tropical Pacific through Plio-Pleistocene glaciation. Science 312, 79-83.

Lawrence, K.T., Sigman, D., Herbert, T., Riihimaki, C., Bolton, C., Martinez-Garcia, A., Rosell-Mele, A., Haug, G., 2013. Time-transgressive North Atlantic productivity changes upon Northern Hemisphere glaciation. Paleoceanography 28, 740-751.

Lazarus, D., Barron, J., Renaudie, J., Diver, P., Turke, A., 2014. Cenozoic planktonic marine diatom diversity and correlation to climate change. PLOS one 9.

Liebrand, D., Lourens, L., Hodell, D., De Boer, B., Van de Wal, R., Pälike, H., 2011. Antarctic ice sheet and oceanographic response to eccentricity forcing during the early Miocene. Climate of the Past 7, 869-880.

Lisiecki, L.E., Raymo, M.E., 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. Paleoceanography 20.

Lohmann, G., Malmgren, B.A., 1983. Equatorward migration of *Globorotalia truncatulinoides* ecophenotypes through the late Pleistocene: gradual evolution or ocean change? Paleobiology 9, 414-421.

Lombard, F., Labeyrie, L., Michel, E., Bopp, L., Cortijo, E., Retailleau, S., Howa, H., Jorissen, F., 2011. Modelling planktic foraminifer growth and distribution using an ecophysiological multi-species approach. Biogeosciences 8, 853-873.

Lombard, F., Labeyrie, L., Michel, E., Spero, H.J., Lea, D.W., 2009. Modelling the temperature dependent growth rates of planktic foraminifera. Marine Micropaleontology 70, 1-7.

Lončarić, N., Brummer, G.-J.A., Kroon, D., 2005. Lunar cycles and seasonal variations in deposition fluxes of planktic foraminiferal shell carbonate to the deep South Atlantic (central Walvis Ridge). Deep Sea Research Part I: Oceanographic Research Papers 52, 1178-1188.

Lunt, D.J., Valdes, P.J., Haywood, A., Rutt, I.C., 2008. Closure of the Panama Seaway during the Pliocene: implications for climate and Northern Hemisphere glaciation. Climate Dynamics 30, 1-18.

Lynch, M., 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. American Naturalist 136, 727-741.

Maclean, I.M., Wilson, R.J., 2011. Recent ecological responses to climate change support predictions of high extinction risk. Proceedings of the National Academy of Sciences 108, 12337-12342.

MacLeod, N., Carter, J.L., 1984. A method for obtaining consistent specimen orientations for use in microfossil biometric studies. micropaleontology, 306-310.

Malmgren, B.A., Berggren, W.A., Lohmann, G., 1983. Evidence for punctuated gradualism in the Late Neogene Globorotalia tumida lineage of planktonic foraminifera. Paleobiology, 377-389.

Malmgren, B.A., Kennett, J.P., 1981. Phyletic gradualism in a Late Cenozoic planktonic foraminiferal lineage; DSDP Site 284, southwest Pacific. Paleobiology 7, 230-240.

Marroig, G., Cheverud, J.M., 2005. Size as a line of least evolutionary resistance: diet and adaptive morphological radiation in New World monkeys. Evolution 59, 1128-1142.

Martínez-Botí, M., Foster, G., Chalk, T., Rohling, E., Sexton, P., Lunt, D., Pancost, R., Badger, M., Schmidt, D., 2015. Plio-Pleistocene climate sensitivity evaluated using high-resolution CO2 records. Nature 518, 49-54.

Marx, F.G., Uhen, M.D., 2010. Climate, critters, and cetaceans: Cenozoic drivers of the evolution of modern whales. Science 327, 993-996.

Mayhew, P.J., Jenkins, G.B., Benton, T.G., 2008. A long-term association between global temperature and biodiversity, origination and extinction in the fossil record. Proceedings of the Royal Society of London B: Biological Sciences 275, 47-53.

Mc Intyre, K., Ravelo, A., Delaney, M., 1999. North Atlantic intermediate waters in the late Pliocene to early Pleistocene. Paleoceanography 14, 324-335.

McClymont, E.L., Sosdian, S.M., Rosell-Melé, A., Rosenthal, Y., 2013. Pleistocene sea-surface temperature evolution: Early cooling, delayed glacial intensification, and implications for the mid-Pleistocene climate transition. Earth-Science Reviews 123, 173-193.

Molnar, P., 2008. Closing of the Central American Seaway and the Ice Age: A critical review. Paleoceanography 23.

Molnar, P., 2017. Comment (2) on “Formation of the Isthmus of Panama” by O’Dea et al. Science Advances 3, e1602320.

Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J., Valencia, V., Ayala, C., Pérez-Angel, L., Rodriguez-Parra, L., Ramirez, V., 2015. Middle Miocene closure of the Central American seaway. Science 348, 226-229.

Moore, C., Mills, M.M., Milne, A., Langlois, R., Achterberg, E.P., Lochte, K., Geider, R.J., La Roche, J., 2006. Iron limits primary productivity during spring bloom development in the central North Atlantic. Global Change Biology 12, 626-634.

Mudelsee, M., Raymo, M.E., 2005. Slow dynamics of the Northern Hemisphere glaciation. Paleoceanography 20.

Naafs, B.D.A., Hefter, J., Acton, G., Haug, G.H., Martínez-Garcia, A., Pancost, R., Stein, R., 2012. Strengthening of North American dust sources during the late Pliocene (2.7Ma). Earth and Planetary Science Letters 317-318, 8-19.

Naafs, B.D.A., Stein, R., Hefter, J., Khélifi, N., De Schepper, S., Haug, G.H., 2010. Late Pliocene changes in the North Atlantic Current. Earth and Planetary Science Letters 298, 434-442.

NASA, 2016. http://neo.sci.gsfc.nasa.gov/view.php?datasetId=MYD28M.

Norris, R.D., 1991a. Biased extinction and evolutionary trends. Paleobiology 17, 388-399.

Norris, R.D., 1991b. Parallel evolution in the keel structure of planktonic foraminifera. Journal of Foraminiferal Research 21, 319-331.

Norris, R.D., 1996. Symbiosis as an evolutionary innovation in the radiation of Paleocene planktic foraminifera. Paleobiology 22, 461-480.

Norris, R.D., Hull, P.M., 2012. The temporal dimension of marine speciation. Evolutionary Ecology 26, 393-415.

Nürnberg, D., Bijma, J., Hemleben, C., 1996. Assessing the reliability of magnesium in foraminiferal calcite as a proxy for water mass temperatures. Geochimica et Cosmochimica Acta 60, 803-814.

Nyman, T., Linder, H.P., Pena, C., Malm, T., Wahlberg, N., 2012. Climate-driven diversity dynamics in plants and plant-feeding insects. Ecology letters 15, 889-898.

O'Dea, A., Lessios, H.A., Coates, A.G., Eytan, R.I., Restrepo-Moreno, S.A., Cione, A.L., Collins, L.S., de Queiroz, A., Farris, D.W., Norris, R.D., 2016. Formation of the Isthmus of Panama. Science Advances 2, e1600883.

Pagani, M., Huber, M., Liu, Z., Bohaty, S.M., Henderiks, J., Sijp, W., Krishnan, S., DeConto, R.M., 2011. The role of carbon dioxide during the onset of Antarctic glaciation. Science 334, 1261-1264.

Pearson, P.N., Coxall, H.K., Stouge, S., 2014. Origin of the Eocene planktonic foraminifer *Hantkenina* by gradual evolution. Palaeontology 57, 243-267.

Pearson, P.N., Ezard, T.H.G., 2014. Evolution and speciation in the Eocene planktonic foraminifer *Turborotalia*. Paleobiology 40, 130-143.

Pearson, P.N., McMillan, I.K., Wade, B.S., Dunkley Jones, T., Coxall, H.K., Bown, P.R., Lear, C.H., 2008. Extinction and environmental change across the Eocene-Oligocene boundary in Tanzania. Geology 36, 179-182.

Pélabon, C., Firmat, C., Bolstad, G.H., Voje, K.L., Houle, D., Cassara, J., Rouzic, A.L., Hansen, T.F., 2014. Evolution of morphological allometry. Annals of the New York Academy of Sciences 1320, 58-75.

Pereira, H.M., Leadley, P.W., Proença, V., Alkemade, R., Scharlemann, J.P., Fernandez-Manjarrés, J.F., Araújo, M.B., Balvanera, P., Biggs, R., Cheung, W.W., 2010. Scenarios for global biodiversity in the 21st century. Science 330, 1496-1501.

Peters, R.H., 1983. The ecological implications of body size. Cambridge University Press, Cambridge.

Pfuhl, H.A., Shackleton, N.J., 2004. Changes in coiling direction, habitat depth and abundance in two menardellid species. Marine Micropaleontology 50, 3-20.

Phillimore, A.B., Hadfield, J.D., Jones, O.R., Smithers, R.J., 2010. Differences in spawning date between populations of common frog reveal local adaptation. Proceedings of the National Academy of Sciences 107, 8292-8297.

Puttick, M.N., Thomas, G.H., Benton, M.J., 2014. High rates of evolution preceded the origin of birds. Evolution 68, 1497-1510.

Quental, T.B., Marshall, C.R., 2010. Diversity dynamics: molecular phylogenies need the fossil record. Trends in Ecology & Evolution 25, 434-441.

R Core Team, 2013. R: A language and environment for statistical computing, R Foundation for Statiastical Computing, Vienna, Austria.

Raffi, S., Stanley, S.M., Marasti, R., 1985. Biogeographic patterns and Plio-Pleistocene extinction of Bivalvia in the Mediterranean and southern North Sea. Paleobiology 11, 368-388.

Raup, D.M., 1966. Geometric analysis of shell coiling: general problems. Journal of Paleontology 40, 1178-1190.

Raymo, M., Ruddiman, W., Shackleton, N., Oppo, D., 1990. Evolution of Atlantic-Pacific δ 13 C gradients over the last 2.5 my. Earth and Planetary Science Letters 97, 353-368.

Renaud, S., Auffray, J.C., Michaux, J., 2006. Conserved phenotypic variation patterns, evolution along lines of least resistance, and departure due to selection in fossil rodents. Evolution 60, 1701-1717.

Renaud, S., Schmidt, D.N., 2003. Habitat tracking as a response of the planktic foraminifer Globorotalia truncatulinoides to environmental fluctuations during the last 140 kyr. Marine Micropaleontology 49, 97-122.

Rosenthal, Y., Boyle, E.A., Slowey, N., 1997. Temperature control on the incorporation of magnesium, strontium, fluorine, and cadmium into benthic foraminiferal shells from Little Bahama Bank: Prospects for thermocline paleoceanography. Geochimica et Cosmochimica Acta 61, 3633-3643.

Ruddiman, W., Kutzbach, J., 1989. Forcing of late Cenozoic northern hemisphere climate by plateau uplift in southern Asia and the American West. Journal of Geophysical Research: Atmospheres 94, 18409-18427.

Ruddiman, W.F., Kidd, R.B., Thomas, E., Expedition 94 Scientists, 1987. Site 606. Initial Reports of the Deep Sea Drilling Project 94.

Ruddiman, W.F., Sarnthein, M., Baldauf, J.G., Expedition 108 Scientists, 1988. Site 659. Proceedings of the Ocean Drilling Program, Initial Reports 108.

Ruegg, K.C., Hijmans, R.J., Moritz, C., 2006. Climate change and the origin of migratory pathways in the Swainson's thrush, Catharus ustulatus. Journal of Biogeography 33, 1172-1182.

Ruggieri, E., Herbert, T., Lawrence, K.T., Lawrence, C.E., 2009. Change point method for detecting regime shifts in paleoclimatic time series: Application toδ18O time series of the Plio-Pleistocene. Paleoceanography 24.

Rutherford, S., D'Hondt, S., Prell, W., 1999. Environmental controls on the geographic distribution of zooplankton diversity. Nature 400, 749-753.

Schiebel, R., 2002. Planktic foraminiferal sedimentation and the marine calcite budget. Global Biogeochemical Cycles 16.

Schluter, D., 1996. Adaptive radiation along genetic lines of least resistence. Evolution 50, 1766 - 1774.

Schmidt, D.N., Rayfield, E.J., Cocking, A., Marone, F., 2013. Linking evolution and development: Synchrotron Radiation X-ray tomographic microscopy of planktic foraminifers. Palaeontology 56, 741-749.

Schmidt, D.N., Renaud, S., Bollmann, J., 2003. Response of planktic foraminiferal size to late Quaternary climate change. Paleoceanography 18, n/a-n/a.

Schmidt, D.N., Renaud, S., Bollmann, J., Schiebel, R., Thierstein, H.R., 2004a. Size distribution of Holocene planktic foraminifer assemblages: biogeography, ecology and adaptation. Marine Micropaleontology 50, 319-338.

Schmidt, D.N., Thierstein, H.R., Bollmann, J., Schiebel, R., 2004b. Abiotic forcing of plankton evolution in the Cenozoic. Science 303, 207-210.

Scott, G.H., Kennett, J.P., Wilson, K.J., Hayward, B.W., 2007. *Globorotalia puncticulata*: Population divergence, dispersal and extinction related to Pliocene–Quaternary water masses. Marine Micropaleontology 62, 235-253.

Shackleton, N.J., Backman, J., Zimmerman, H., Kent, D.V., Hall, M.A., Roberts, D.G., Schnitker, D., Baldauf, J.G., Desprairies, A., Horighausen, R., Huddlestun, P., Keene, J.B., Kaltenback, A.J., Krumsiek, K.A.O., Morton, A.C., Murray, J.W., Westberg-Smith, J., 1984. Oxygen isotope calibration of the onset of ice-rafting and history of glaciation in the North Atlantic region. Nature 307, 620-623.

Sigman, D.M., Jaccard, S.L., Haug, G.H., 2004. Polar stratification in a cold climate. Nature 428, 59-63.

Sikes, E.L., Keigwin, L.D., Curry, W.B., 1991. Pliocene paleoceanography: circulation and oceanographic changes associated with the 2.4 Ma glacial event. Paleoceanography 6, 245 - 257.

Silverman, B.W., 1986. Density estimation for statistics and data analysis. CRC press.

Simpson, G.G., 1944. Tempo and mode in evolution. Columbia University Press, New York, N.Y.

Simpson, G.G., 1953. The Baldwin effect. Evolution 7, 110-117.

Sodhi, N.S., Bickford, D., Diesmos, A.C., Lee, T.M., Koh, L.P., Brook, B.W., Sekercioglu, C.H., Bradshaw, C.J., 2008. Measuring the meltdown: drivers of global amphibian extinction and decline. PloS one 3, e1636.

Stanley, S.M., 1986. Anatomy of a regional mass extinction: Plio-Pleistocene decimation of the western Atlantic bivalve fauna. Palaios, 17-36.

Steeman, M.E., Hebsgaard, M.B., Fordyce, R.E., Ho, S.Y., Rabosky, D.L., Nielsen, R., Rahbek, C., Glenner, H., Sorensen, M.V., Willerslev, E., 2009. Radiation of extant cetaceans driven by restructuring of the oceans. Systematic biology 58, 573-585.

Stenseth, N.C., Maynard Smith, J., 1984. Coevolution in ecosystems: Red Queen evolution or stasis? Evolution, 870-880.

Stewart, D.R.M., 2003. Evolution of Neogene globorotaliid foraminifera and Miocene climate change, Department of Earth Sciences. PhD Thesis. University of Bristol, Department of Earth Sciences, p. 269.

Stickley, C.E., St John, K., Koç, N., Jordan, R.W., Passchier, S., Pearce, R.B., Kearns, L.E., 2009. Evidence for middle Eocene Arctic sea ice from diatoms and ice-rafted debris. Nature 460, 376.

Stigall, A.L., Saupe, E., 2013. Analysing links between biogeography, niche stability and speciation: the impact of complex feedbacks on macroevolutionary patterns. Palaeontology 56, 1225-1238.

Svenning, J.C., 2003. Deterministic Plio-Pleistocene extinctions in the European cool-temperate tree flora. Ecology letters 6, 646-653.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F., De Siqueira, M.F., Grainger, A., Hannah, L., 2004. Extinction risk from climate change. Nature 427, 145-148.

Thuiller, W., 2003. BIOMOD–optimizing predictions of species distributions and projecting potential future shifts under global change. Global change biology 9, 1353-1362.

Thuiller, W., 2004. Patterns and uncertainties of species' range shifts under climate change. Global Change Biology 10, 2020-2027.

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD–a platform for ensemble forecasting of species distributions. Ecography 32, 369-373.

Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T., Prentice, I.C., 2005. Climate change threats to plant diversity in Europe. Proceedings of the National Academy of Sciences of the united States of America 102, 8245-8250.

Tiedemann, R., Sarnthein, M., Shackleton, N.J., 1994. Astronomic timescale for the Pliocene Atlantic d18O and dust flux records of Ocean Drilling Program site 659. Paleoceanography 9, 619 - 638.

Tonni, E.P., Alberdi, M.T., Prado, J., Bargo, M.S., Cione, A.L., 1992. Changes of mammal assemblages in the pampean region (Argentina) and their relation with the Plio-Pleistocene boundary. Palaeogeography, Palaeoclimatology, Palaeoecology 95, 179-194.

Van der Putten, W.H., Macel, M., Visser, M.E., 2010. Predicting species distribution and abundance responses to climate change: why it is essential to

include biotic interactions across trophic levels. Philosophical Transactions of the Royal Society B: Biological Sciences 365, 2025-2034.

Van Valen, L., 1973a. A new evolutionary law. Evolutionary Theory 1, 1-30.

Van Valen, L., 1973b. Pattern and the balance of nature. Evolutionary Theory 1, e49.

Vance, D., Burton, K., 1999. Neodymium isotopes in planktonic foraminifera: a record of the response of continental weathering and ocean circulation rates to climate change. Earth and Planetary Science Letters 173, 365-379.

Voje, K.L., Hansen, T.F., Egset, C.K., Bolstad, G.H., Pelabon, C., 2014. Allometric constraints and the evolution of allometry. Evolution 68, 866-885.

Vrba, E.S., 1995. The fossil record of African antelopes (Mammalia, Bovidae) in relation to human evolution and paleoclimate.

Wade, B.S., 2004. Planktonic foraminiferal biostratigraphy and mechanisms in the extinction of Morozovella in the late middle Eocene. Marine Micropaleontology 51, 23-38.

Wade, B.S., Olsson, R.K., 2009. Investigation of pre-extinction dwarfing in Cenozoic planktonic foraminifera. Palaeogeography, Palaeoclimatology, Palaeoecology 284, 39-46.

Wade, B.S., Pearson, P.N., 2008. Planktonic foraminiferal turnover, diversity fluctuations and geochemical signals across the Eocene/Oligocene boundary in Tanzania. Marine Micropaleontology 68, 244-255.

Wade, B.S., Pearson, P.N., Berggren, W.A., Pälike, H., 2011. Review and revision of Cenozoic tropical planktonic foraminiferal biostratigraphy and calibration to the geomagnetic polarity and astronomical time scale. Earth-Science Reviews 104, 111-142.

Wade, B.S., Poole, C.R., Boyd, J.L., 2016. Giantism in Oligocene planktonic foraminifera Paragloborotalia opima: Morphometric constraints from the equatorial Pacific Ocean. Newsletters on Stratigraphy 49, 421-444.

Wake, D.B., Vredenburg, V.T., 2008. Are we in the midst of the sixth mass extinction? A view from the world of amphibians. Proceedings of the National Academy of Sciences 105, 11466-11473.

Wei, K.-Y., 1994a. Allometric heterochrony in the Pliocene-Pleistocene planktic foraminiferal clade *Globoconella*. Paleobiology 20, 66-84.

Wei, K.-Y., 1994b. Stratophenetic tracing of phylogeny using SIMCA pattern recognition technique: a case study of the Late Neogene planktic foraminifera *Globoconella* clade. Paleobiology 20, 52-65.

Wei, K.-Y., Kennett, J., 1986. Taxonomic evolution of Neogene planktonic foraminifera and paleoceanographic relations. Paleoceanography 1, 67-84.

Wei, K.-Y., Kennett, J.P., 1988. Phyletic gradualism and punctuated equilibrium in the late Neogene planktonic foraminiferal clade *Globoconella*. Paleobiology 14, 345-363.

Woodard, S.C., Rosenthal, Y., Miller, K.G., Wright, J.D., Chiu, B.K., Lawrence, K.T., 2014. Antarctic role in Northern Hemisphere glaciation. Science 346, 847-851.

Wright, S., 1931. Evolution in Mendelian populations. Genetics 16, 97-159.

Wright, S., 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. na.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686-693.

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.

Zachos, J.C., Kroon, D., Blum, P., Expedition 208 Scientists, 2004. Site 1264. Proceedings of the Ocean Drilling Program, Initial Reports 208.