Trait-fitness associations via fecundity and competition in a two-million-year-long fossil record

Lee Hsiang Liow[1,2,*], Arthur Porto[3], Emanuela Di Martino[1,4*]

[1]Natural History Museum, University of Oslo, Oslo, Norway
[2]Centre for Planetary Habitability, Department of Geosciences, University of Oslo, Oslo, Norway
[3]Florida Museum of Natural History, *University of Florida, Gainesville, FL, USA*
[4]Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Università di Catania, Corso Italia 57, 95129, Catania, Italy
[*]Corresponding authors
Orcids
Liow**:** https://orcid.org/0000-0002-3732-6069
Porto: https://orcid.org/0000-0002-9210-8750
Di Martino: https://orcid.org/0000-0002-3892-4036

Running title

Fossil trait-fitness associations

Abstract

The evolution of phenotypic traits is usually studied on generational times or across species on million-year timescales. We bridge this conceptual gap by using high density sampling of a species lineage, *Microporella agonistes* (Bryozoa, Cheilostomatida), over 2 million years of its evolutionary history, to ask if trait-fitness associations are consistent with evolutionary trait models often applied to phenotypic time series. We use average fecundity and competitive outcome as two different fitness components, where competitive outcome is a proxy for partial survival. Examining three quantitative traits in multivariate analyses, we present evidence that some traits experienced substantial selective pressures, in part controlled by past environments. A complex interplay of resource competition with an altering set of competitors and past temperatures, has contributed to the changing patterns of phenotypes within the focal species. A comparison with congeneric species living in the same regional community suggests that size traits are more temporally variable and less constrained than shape traits. Our analyses also show that while controls on phenotypes are complex and varied in time, ecological and evolutionary processes that unfold on shorter time scales are not inconsistent with macroevolutionary patterns observed on longer timescales.

Keywords: Selection, phenotypic evolution, cheilostome bryozoans, automated phenotyping, resource competition

Introduction

Phenotypic traits are the result of evolution, including stochastic processes and varied constraints. Realized trait distributions in empirical populations can be used to estimate the relative contributions of selection, constraints and drift, and conditions under which they play larger or smaller roles in the outcomes of evolution. Most of the empirical case studies of population-level trait evolution, from which we have gained insights on the form and rate of evolution, are based on contemporary populations monitored over time intervals that are short (Hendry and Kinnison 1999; Stroud et al. 2023) relative to the available time for evolution of a species lineage. Spatially separated populations/communities are often used as substitutes for temporally separated populations/communities subject to different environments (Damgaard 2019). However, insights from data from temporally separated populations from the fossil record (Hunt 2007; Hunt et al. 2015; Voje et al. 2020) offer a more direct understanding on constraints (Firmat et al. 2014; Di Martino and Liow 2022) and long-term changes in adaptive landscapes (Hunt and Roy 2006; Brombacher et al. 2017).

Phenotypic data from temporally separated populations can be modeled with variations on the Ornstein-Uhlenbeck (OU) process that characterize evolutionary optima and their accompanying stochastic variation (Hunt et al. 2008; Voje 2020), just like they have been applied in comparative phylogenetic contexts to model adaptation and selection (Butler and King 2004). A better fit to an OU model is often interpreted as support for adaptation and selection, while a better fit to a purely stochastic model, like Brownian motion, discredits selection for the trait in question (Butler and King 2004). An alternative approach to studying fossil time series involves the measurement of fitness components alongside target morphological traits. This approach enables a more direct investigation of the relative contributions of selection versus constraints to trait evolution, as well as the identification of sources and reasons for conflicting signals at shorter versus longer timescales (e.g. Voje et al. 2023), which can be potentially couched in terms of micro- versus macroevolution.

However, proxies of fitness components can be rarely estimated from the fossil record, hence fitness is seldom studied in fossil populations. Rare examples include Kurtén’s pioneering cross-sectional population investigation on *Ursus spelaeus* (a Pleistocene cave bear), where he used age-at-mortality as a fitness component to study selection on tooth shape (Kurtén 1958). Likewise, Van Valen used age-at-mortality to study selection intensity of Miocene horses and rightly pointed out that few types of organisms are amenable to the quantification of selection using the fossil record (Van Valen 1963).

Cheilostome bryozoans offer a unique opportunity in evolutionary research. These marine colonial invertebrates have a high preservation potential because of their life-habit (they often encrust hard substrates) and their calcium carbonate skeletons. Earlier work has suggested that their skeletal morphologies can be robustly used for species-level identification, even in the absence of soft tissues (Jackson and Cheetham 1990), and molecular studies have corroborated this finding (Orr et al. 2021, 2022). Many cheilostome species brood their larvae in skeletal structures called ovicells (Fig. 1). As each ovicell harbors a single embryo that develops into a larva, the density of ovicells within a colony of genetically identical feeding zooids (autozooids) can be used as a proxy for fecundity, a fitness component (Jackson and Wertheimer 1985; Di Martino and Liow 2021*a*). The size of ovicells has also been used as a proxy for larval (offspring) size or maternal investment (Herrera et al. 1996), where larger larvae may have a survival advantage (Marshall et al. 2006). Note that in this contribution, we use ovicell size as a trait potentially under selection, rather than as a fitness component (proxy for larval size and hence larval survival), although we do examine trade-offs involving ovicell size.

Mortality or survival is a commonly studied fitness component in contemporary populations of solitary metazoans (Crone 2001). In the case of colonial metazoans, such as corals, experiments have shown that partial mortality can lead to reduced allocation of resources towards reproduction (Graham and van Woesik 2013). Cheilostome bryozoans, many of which are sessile and encrusting as adults, compete for space on substrates. Such competition often results in one colony partially or completely overgrowing another, which can cause death or partial death of the overgrown colony (Buss 1979; Buss and Jackson 1979). Partial mortality is believed to have detrimental effects on the fitness of bryozoan colonies (Buss 1986; Turner and Todd 1994). Autozooid relative size is a relatively good predictor of winning competitive overgrowths (Liow et al. 2017), although other non-size traits also contribute to competitive outcomes (Liow et al. 2019). To the benefit of paleoecological studies, overgrowth competition, analogous to those studied in contemporary bryozoan communities, is also preserved in fossil assemblages. Overgrowth can be quantified with relative ease, in the sense that the colony overgrowing another and that being overgrown can be identified as such using the geometry and direction of growth for inference (Taylor 2016).

It is commonly accepted that organismal traits, and especially life-history traits, like fecundity, offspring size and mortality often trade-off under different selective pressures (Stearns 1989; Schluter et al. 1997). For example, bryozoans that were experimentally halved (i.e. partial mortality) produced fewer, and smaller larvae more variable in size (Marshall and Keough 2004), plausibly to channel resources to survival. While large autozooids are in general advantageous for winning overgrowth competition, colonies with larger autozooids are sometimes less fecund (Di Martino and Liow 2021*a*). Autozooid size has also been shown experimentally to be in part controlled by the temperature at which ontogeny occurs (O'Dea and Okamura 2000). Trade-offs can also be conditional on environmental conditions (Lancaster et al. 2017) and are incompletely understood, even in well-studied groups like birds (Santos and Nakagawa 2012). Unsurprisingly, much less is known about such trade-offs in most marine species including bryozoans (Di Martino and Liow 2021*a*).

Although colonies with larger autozooids have a competitive advantage within a community setting (Liow et al. 2017), those with smaller autozooids have higher fecundity in certain species (Di Martino and Liow 2021*a*). On the other hand, on a global, multi-species, macroevolutionary scale, autozooid size in cheilostome bryozoans seems to be constrained, even though there is a detectable tendency for within-lineage evolution of larger autozooid sizes (Liow and Taylor 2019). These apparently contradictory findings highlight the inadequacy of simply scaling up within-lineage observations to explain across-lineage evolution (Kingsolver and Pfenning 2004; Belk et al. 2020), sometimes referred to as the continuity of micro- and macroevolution. While we do not believe that invoking "higher-level" processes is necessary to explain macroevolution, we argue that within-species, among-lineage and macroevolutionary patterns and processes should all be considered to establish plausible evolutionary scenarios (Liow et al. 2023), and to gain a comprehensive understanding of evolutionary history and processes. Ideally, disparate levels of analyses should be done on the same organismal group, such that natural history knowledge can help shed light on the findings.

To contribute to the pursuit of understanding the sources of selection, constraints and drift within species, as well as their implications for macroevolutionary patterns and processes, we build on the insights and methods from an earlier study (Di Martino and Liow 2021*a*) on the cheilostome *Antarctothoa tongima* (Ryland and Gordon 1977). In *A. tongima,* multivariate analyses detected negative trait-fitness associations for autozooid sizes, but no particular relationship between fecundity and ovicell size. On the other hand, in terms of autozooid shape, colonies with longer zooids (interpreted as having higher feeding efficiency) were more fecund. Here, we explore a different cheilostome species, *Microporella agonistes* Gordon 1984, and a substantially larger and richer dataset, for more than 2 million years of its history, up to the Recent, in part compiled using computer-vision approaches (Di Martino et al. 2023). Using *Microporella,* we ask questions on the nature of phenotypic evolution and trait-fitness associations, using measures of fecundity and overgrowth (partial mortality) as proxies for fitness. While our primary focus is on *M. agonistes* (Fig. 1), we also compare congeners (Fig. S1A–C) that overlap with this target species in both time and space to study interspecific similarities and differences. We stress that, although we adopt a quantitative framework similar to Lande and Arnold (1983), the slopes or curvatures of the trait-fitness relationships should not be interpreted as selection gradients per se, as the populations under study, even the contemporary ones, are not “instantaneous” snapshots, but rather temporally- (and spatially-) averaged populations.

Our study addresses the following questions: (1) What forms of trait-fitness associations are seen in different traits and are these detectably constant or changing through time? (2) Is there a detectable trade-off between fecundity and overgrowth, or fecundity and offspring size? (3) Both fecundity and overgrowth are interpretable as fitness components but how do the estimated trait-fitness associations differ when using either component as a fitness proxy? (4) Are there detectable autozooid size advantages for overgrowth, and to what extent is size controlled by the abiotic environment? (5) Do closely related lineages show similar patterns of trait evolution or forms of trait-fitness associations in the given time intervals?

We discuss the implications of our findings, considering the interplay between processes occurring at shorter time scales, such as ecological interactions and reproduction, and patterns that manifest over longer timescales, such as the long-term evolution of trait means.

Material and methods

Material

Microporella agonistes is a commonly observed encruster of hard substrates currently having a widespread distribution off the coast of New Zealand. It has been reported from several sites including Kermadec Ridge, Cook Strait, Marlborough Sounds, Tasman Bay, Westland, Fiordland, Puysegur Bank, and the western approaches to Foveaux Strait, at a broad depth range (i.e. 26–549 m) (Gordon 1989). *Microporella agonistes* is also a commonly observed component of the Pleistocene bryozoan fauna of the Wanganui Basin in the North Island of New Zealand (Liow et al. 2016).

We examined and imaged 985 fossil, subrecent and Recent colonies of *M. agonistes* encrusting mollusk shells, mainly bivalves and subordinately gastropods (see “Imaging” below). The fossil colonies were collected during three field expeditions conducted in January 2014, March 2017, and February 2020, from Pleistocene formations of the Wanganui Basin in New Zealand (see Table S1), where six of these temporal intervals had more than 50 colonies sampled and measured. The formations sampled are shell-beds from transgressive system tracks (Rust and Gordon 2011), part of a cyclic depositional sequence about 2 km thick, with a well-established, high-resolution chronostratigraphy spanning the last 2.5 Myr (Carter and Naish 1998; Naish et al. 1998, 2005; Abbott et al. 2005). Subrecent colonies (i.e. those with some remains of soft tissues) of an unknown age (but confidently between the last fossil formation and the contemporary populations, see “Whanganui core” in Table S1) were available from a piston core (Gillespie et al. 1998). Contemporary colonies, like the subrecent ones, are analogous to fossil samples, consisting of bryozoan encrusted mollusk shells dredged from Cook Strait (2012) and South Taranaki Bight (2017), New Zealand with average depths of 120 m. We treat the colonies studied from separate formations, piston core, and Recent populations as comparable “populations”, and for brevity we use the term population or formation to refer to the data corresponding to the rows in Table S1.

We also imaged colonies of three congeners, namely *Microporella discors* Uttley and Bullivant, 1972, *M. intermedia* Livingstone, 1929 and *M. speculum* Brown 1952 (Fig. S1A–C), which co-exist with the focal species (see Table S1 for sample sizes), using the same protocol (see “Imaging” below). *Microporella rusti* Di Martino, Taylor, Gordon, Liow 2017 also co-exists with *M. agonistes* but it is rarely sampled (Di Martino and Liow 2022) and was hence excluded here.

Imaging

Colonies were digitized using a Hitachi TM4000plus Tabletop scanning electron microscope (SEM). Prior to scanning electron microscopy, contemporary colonies were treated overnight in a dilute solution of commercial bleach to remove soft tissues and to expose the skeleton. Many of the colonies, whether fossil or Recent, were often only partially preserved or retained on the substrate, but these partial colonies (fragments) are assumed to be unbiased with respect to the traits and spatial interactions we measure. To avoid duplicating measurements, we ensured that multiple SEM images taken of the same colony did not overlap. Image magnification was standardized at 50x, such that each image covers approximately the same area (c. 4.5 mm^2). Images were taken ensuring that the regions of the colony from which traits were measured were not tilted with respect to the frontal plane. Different colonies of the same species on the same substrate were distinguished by their geometry and direction of growth.

Spatial interactions

We followed an established protocol (Liow et al. 2016, 2017) for observing inter- and intraspecific spatial interactions. Here, we focus on win-lose overgrowths (Fig. S1D), where the growing edge of one colony is observed to cover one or more orifices of zooids in the second colony, but we also documented other spatial interactions, including reciprocal overgrowths, stand-offs and fouling (Taylor 2016). In addition, we also tabulated the direction of approach for win-lose and reciprocal overgrowths, where “frontal” means the two colonies grew towards each other, flank means one colony approached the other laterally, and “rear” means one colony approached the proximal end of the other (Di Martino et al. 2020). Only in frontal approaches do both interactors play active roles in the competitive interaction. As part of our standard protocol, we also noted the relative average autozooid size of the combatants (Liow et al. 2017). If the relative autozooid size of the two combatants could not be distinguished by visual comparison, they were scored as “equal”. Stand-offs and reciprocal overgrowths necessarily happen syn-vivo (i.e. occurring when both partners are in life). These have high proportional representation in our data hence we assume that instances of overgrowth after death contribute noise but not signal to overgrowth interaction data (McKinney 1995; Taylor 2016). For analyses in this paper, we use only the interactions where at least one partner was a *Microporella* species (e.g. Fig. S1D). Out of the 8438 interactions scored for *Microporella*, 6521 were collected for the purpose of the current study, while the remaining 1917 were previously published in Liow et al. (2016).

Data extraction using Computer Vision

The phenotypic data needed to estimate fecundity are the observed number of modules (zooids) and the number of fertile polymorphs (i.e. modules bearing a brood chamber called ovicell) in each SEM image, which often encompass part of a colony (see Fig. 1). Each colony can be formed by several to tens or even hundreds of zooids. The three focal traits are the area and shape (maximum length divided by maximum width) of feeding modules (autozooids), and area of ovicells (Fig. 1). Our total dataset comprises over 400,000 structures, a task that would be inhibitive to measure without automation. Given an input SEM image, we use a Swin Transformer-based Mask R-CNN for instance segmentation (Liu et al. 2021) trained on a dataset of 935 annotated cheilostome bryozoan images (Di Martino et al. 2023). After segmentation, measurements are based on the predicted masks using contour functions from the *opencv 4.6.0* python library (Bradski 2000). The command-line was used for high-throughput computer vision inference. For more comprehensive details, we refer to DeepBryo (Di Martino et al. 2023), a deep-learning web-application specifically designed to perform these tasks. The accuracy of both object classification and measurement for diverse structures ($r^2 > 0.89$) is satisfactorily high for DeepBryo, given the trade-off for the amount of data we can gain, compared with manual measurements (Di Martino et al. 2023).

Characterizing phenotypic evolution

Our phenotypic time series are short and is well-known that short time series tend to best fit random walks, even when the underlying generating model differs. Given this caveat, we fit five commonly invoked models, namely an unbiased random walk (URW), a biased random walk (GRW), stasis (S), strict stasis (SS) and Ornstein-Uhlenbeck (OU) to each measured *Microporella agonistes* trait individually using paleoTS (Hunt 2015) and evoTS (Voje 2023). Data best-fitting an OU model is typically interpreted as demonstrating stabilizing selection and those best fitting a biased random walk interpreted as directional selection (Hunt 2006). The URW reflects pure drift (Butler and King 2004). The stasis model is implemented as white noise around a constant mean, and strict stasis a special case of stasis with zero variance around the long-term mean (Hunt et al. 2015), both meaning to capture a part of the "punctuated equilibrium" concept commonly applied to the fossil record (Eldredge and Gould 1972).

Multivariate trait-fitness modeling: binomial model for fecundity

We estimate a component of fitness as average fecundity of colonies, equivalent to genetic individuals in solitary organisms, using counts of gravid female polymorphs (i.e. zooids with ovicells that hold embryos or larvae) per polymorph within the colony (Fig. 1). This measure of fitness we term fecundity is used as the response variable in a binomial generalized linear model (glm). The predictor (focal) traits, namely autozooid area, ovicell area and autozooid shape (shown in Fig. 1), are each natural logged and standardized to their global mean, following (Di Martino and Liow 2021*a*). Note that autozooid areas have limited predictive power for ovicell areas ($r^2 = 0.38$ see Di Martino and Liow 2022), and the variance inflation factor (VIF) is approximately 1 for each predictor trait in our dataset.

Initially, we model trait-fitness relationships without distinguishing among the populations from which the data originate, using linear and quadratic combinations of the traits to allow forms of trait-fitness associations that resemble stabilizing or directional selection. Subsequently, we test whether including population identity (i.e. the population from a specific formation or a particular Recent location) as a factor improves the fit of the best models, after removing the populations with less than 20 colonies measured. As including population identity as a factor improves the model fit, we proceed to estimate trait-fitness relationships within each population separately. As before, we fit various linear and quadratic combinations of the traits. We present both model averaged fits and those of the best models in the main text and SI results, respectively. While there are some trait outliers in the dataset, we chose not to remove them for two reasons. First, removing outliers made the logged data non-normally distributed. Second, we believe that the outlier data represent genuine empirical values we are interested in analyzing. To explore the sensitivity of our results to the vagaries of sampling, we fit and plot the same models to 30 random draws (without replacement) of 75% of the available data in each time interval.

Modeling spatial interactions

We use interaction data where *Microporella* is observed to have overgrown or was overgrown by another cheilostome species. We estimate the binomial probability of *M. agonistes* winning the combat, first regardless of the relative average autozooid size of the competitor, and then the binomial probability of *M. agonistes* winning the combat when it is the larger combatant. As a more stringent alternative, we also used a subset of the data comprising only frontal encounters, as there is a much greater chance that both colonies are investing metabolically to tackle the head-on encounter that involves survival. This is because encrusting cheilostomes respond to spatial encounters by growing in their distal end (i.e. growing edge), while rear encounters do not elicit an active overgrowth response.

Trait-fitness modeling: binomial model for partial survival

The colonies from which phenotypic data and spatial interactions are quantified do not entirely overlap. The following analyses were performed exclusively with colonies where both phenotypic data and overgrowth outcomes were available. We use frontal overgrowth outcomes as a proxy for partial mortality/survival (an alternative fitness component to fecundity). In this case, a win is scored as for partial survival, while a loss is indicated as 0 for partial death. These outcomes are used as the response variable in a binomial glm, with autozooid area, ovicell area and autozooid shape as predictor variables, just like in the fecundity above. Due to the limited data available for each population, we initially examined trait-fitness relationships without distinguishing among populations. Subsequently, we analyzed selected populations (each with at least 20 measured colonies) with richer data to gain further insights.

Paleoenvironment links to phenotypes

Cheilostome autozooid size has been shown to be influenced by the temperature at the time of autozooid formation within a colony (O’Dea and Okamura 2000). This size-temperature association is also detectable across multi-species communities, where smaller zooids are linked to higher temperatures (Okamura and Bishop 1988; Kuklinski and Taylor 2008). Here, we are interested in the association between traits and paleotemperature as approximated by δ^{18}O, a widely used proxy for seawater temperature in paleontological studies. Using each trait in turn as a response variable, we compared linear models that combined the other two traits as well as the mean and standard deviation of the δ^{18}O values estimated using the data available in each time interval (data from Lisiecki and Raymo 2005), approximated with their known duration (see Table S1). For the Recent populations, we obtained δ^{18}O data spanning 1988–2016 for the southern hemisphere latitude encompassing New Zealand (25–90°S) from NASA data repository (https://data.giss.nasa.gov/018data) to ensure an equivalent dataset (see SI). We present fits obtained from the best models selected using AIC criteria.

Analyses for congeners

Data for *Microporella discors*, *M. intermedia* and *M. speculum* were analyzed in the same fashion as for the focal *M. agonistes*.

Data and code

The data and R code necessary to replicate our results can be found at Dryad https://doi.org/10.5061/dryad.3ffbg79rj. The images used to generate the data can be found at https://zenodo.org/records/10795305. The DeepBryo source code can be found at https://github.com/agporto/DeepBryo.

Results

Phenotypic traits through time

There is substantial temporal variation in each of the three quantitative traits and fecundity over the observed evolutionary history of *Microporella agonistes*, most notably in autozooid size (Fig. 2). Best-fitting models among the five studied models for trait means are the unbiased random walk (URW), generalized random walk indicating directional evolution (GRW) and strict stasis (SS) for log autozooid size, log ovicell size and autozooid shape, respectively (Fig. S2). Notice that the mean-standardized standard deviations are strikingly different for each trait, with autozooid areas having considerably lower values compared to autozooid shapes (Fig. S2), as shapes within colonies are very plastic due to epigenetic factors. Because Figs. 2 and S2 suggest that temporal trait means are different, we tested for inter-population differences (Table S2), finding that autozooid size is in general more variable temporally. For each trait, variances are generally comparable across populations with a few exceptions (Table S2). However, there is no detectable relationship between population variance and temporal duration represented by the data for any of the traits, i.e. the amount of time represented did not inflate variance (Fig. S3).

Trait-fitness associations via fecundity

When considering all temporal populations of *Microporella agonistes* as a unified “super-population”, the best-fitting multivariate model for trait-fitness associations includes quadratic terms for all three traits, introducing curvature to the relationships (Table S3, Fig. S4). According to this model, colonies with autozooids at both size extremes are predicted to have higher fitness, as are colonies with larger and smaller ovicells, even though the 95% CIs are wider at these extremes. In contrast, colonies with average shapes have the highest fitness. However, including formation/population identity as a factor significantly improves the log likelihood values of equivalent models (Table S4), hence we proceed to analyze each population separately, as the patterns in Fig. S4 may result from combining heterogenous populations. Three of the seven populations analyzed show a similar concave prediction of autozooid size fitness relationships, while others show a convex pattern (Fig. 3), although not all of these fits yield statistically significant coefficients (Fig. S5). It is also worth noting that the highest density for autozooid size (Fig. 3) do not align at the same relative size (x-axis values), suggesting that the optimal autozooid size is different in different temporal populations. The ovicell size fitness relationships in most populations are perceptibly linear (Fig. 3 and Fig. S5). These are in contrast to autozooid shape which seems to have a “universal” optimum (Fig. 3 and S5). In summary, trait-fitness associations in *M. agonistes* are time-varying for autozooid size but are relatively stable for ovicell size and autozooid shape. Subsampled versions of these data show similar results where those populations with smaller sample sizes show more sampling noise (Fig. S6). There is no to little trade-off between fecundity and ovicell area (indicative in middle panels of Fig. 3, see also Fig. S7).

Trait-fitness associations via overgrowth

Microporella agonistes is not, on average, a consistent winner in spatial combats (Fig. 4A, see also Liow et al. (2016) for examples of consistent winners). However, it tends to win when it is the larger partner in such combats (Fig. 4B). Using overgrowth outcomes (win or lose) as a proxy for fitness in 284 colonies of *M. agonistes* for which we also have phenotypic measurements, we find that the best model of multivariate trait-fitness association is the null model, although competing models have very similar loglikelihood values (Table S5). Model averaged predictions look quite “flat” as suggested by the (best) null model (Fig. S8).

Using a two-sample t-test, there is no relationship between autozooid size (t = 0.23, df = 305, p-value = 0.80, Fig. 5A) or fecundity (t = 1.53, df = 309, p-value = 0.13, Fig. 5B) and the probability of winning, also when populations are examined separately (Table S6). These results do not necessarily contradict with Fig. 4 as this shows the probability of winning an overgrowth regardless of the opposing species when *M. agonistes* is comparatively larger, while Fig. 5 compares size differences in conspecifics.

Relationship between phenotypic traits and the environment

Autozooid size is more strongly associated with δ^{18}O than ovicell size or autozooid shape (Figs. 6 and S9, Tables S7 and S8), although there is still much unexplained variation in the models. A one-unit change in δ^{18}O is predicted to result in 0.19 ± 0.02 log unit of change in autozooid size, and 0.13± 0.04 log unit of change in ovicell size and -0.01 ± 0.005 change in autozooid shape, respectively for each of the best models. Consequently, while we detect relative size advantages in overgrowth interactions, it is likely that long-term mean size evolution is in part controlled by both paleotemperature and its variability.

Traits in closely related species

Three *Microporella* species, other than *M. agonistes*, are observed in substantial numbers (Table S1). When disregarding any heterogeneity among temporal populations, the overall trait-fitness association for autozooid shape are similar among the four species (compare Fig. S4 and Fig. S10**)**, exhibiting an optimal value, i.e. a concavity in the quadratic function. However, autozooid area has a different estimated curve for each species. When comparing temporal populations, the pattern for autozooid area deviates more significantly compared to ovicell area and shape (compare Fig. 3 and Fig. S11A–C).

When other species are treated to the same models as *M. agonistes* (Fig. 6) to investigate the temperature control of traits, best linear models for autozooid size when accounting for the other two traits always includes δ^{18}O, with a significant or close to significant positive slope, even if slight (Figs. S12 and S13), indicating that higher temperatures are associated with smaller autozooids. This is not the case for ovicell size and is inconsistent for shape (Figs. S12 and S13).

Discussion

Phenotypes change over the lifetime of an organism, with development, growth, and via interactions with the environment. Ecologists and evolutionary biologists studying natural, contemporary systems have access to repeated observations of the same individuals from a population, sometimes across consecutive generations, to capture information on fitness and phenotypes (e.g. Petrullo et al. 2023). Phenotypic data from individuals of different populations and/or related species can shed light on the relationships between adaptation on shorter versus longer time scales, represented by the evolutionary time that has elapsed among species (e.g. De Lisle et al. 2022). A complementary but more direct approach to accessing past phenotypes is through fossil remains. Here, we use the rich fossil record of a lineage to study its morphology, including intra-individual, intra-population, inter-population, temporal and interspecies variation. Our empirical system allows us to combine charting phenotypic change using standard evolutionary models of trait evolution, characterizing trait-fitness associations and quantifying the influence of factors associated with phenotypic differences, including paleoclimate and competition on the same species lineage.

In general, we find that trait-fitness associations are temporally variable for some traits (autozooid size) and quite constant for others (shape). We did not detect the trade-offs which might be expected (between fecundity and overgrowth; fecundity and offspring size). Competition (survival) and fecundity are different fitness components and do not show similar trait-fitness associations. While overgrowth outcomes are highly predictable at community level, there is no clear selective advantage at the species level. On the other hand, temperature predicts autozooid size to some extent for the species we studied. Given the somewhat complex empirical results we have presented, we discuss each of the traits we studied in turn, in the context of fitness, the evolutionary community and temporal backdrop. We then explore the implications of our findings in the light of the current literature discussing micro- versus macroevolution, contrast between contemporary species versus fossil species approaches, before turning to discussion of caveats necessary for interpreting our results.

Size is plastic and likely subject to both biotic and abiotic controls

The three quantitative traits we studied are best fit to different phenomenological models of trait evolution. Autozooid size is a trait as important ecologically and evolutionarily as body size in solitary organisms (Peters 1983), but perhaps more analogous to leaf size in plants (e.g. Reich et al. 1998) as autozooids are clonally replicated within a genetic individual (the colony) and are tasked with feeding. Autozooid size is variable within a colony (see e.g. Di Martino and Liow 2021*b*) and shows fluctuating colony means through the evolutionary lifetime of *Microporella agonistes* (Fig. 1). With the caveat that our time series is short (Hunt et al. 2015), the best-fit univariate phenomenological model is an unbiased random walk and hence consistent with either fluctuating evolutionary optima or stochastic temporal changes. These temporal fluctuations in average autozooid size are also consistent with changing multivariate trait-fitness associations, estimated from time-interval to time-interval (Fig. 3

and S5). There are time-intervals during which there is a detectable single optimal size, while in others, selection seems to be disruptive where both larger and smaller autozooids are favored. Autozooid size may not be a trait that is good at tracking selective peaks: when there are changes in (multiple) selective pressures, populations may lag behind in autozooid size changes. While what these selective pressures are is unclear, there is strong independent evidence that larger relative autozooid sizes increase competitive abilities (Fig. 4, and see Liow et al. 2016, 2017) among a changing species competition pool (Rust and Gordon 2011). Yet, when either fecundity (Figs. 3, S4, S5) or competitive outcomes (Fig. S8) were used as a fitness component, there were indications that *both* larger and smaller zooids are at a selective advantage, which is puzzling. Simultaneously, there is no detectable advantage of being a larger autozooid-sized *M. agonistes* colony with a population of conspecifics (Fig. 4). There is also independent evidence showing that sea-water temperature contributes to the control of autozooid size (this study and see also O’Dea and Okamura 2000; Di Martino and Liow 2021*a*). Via δ^{18}O, temperature is the only environmental variable accessible to us for the Pleistocene. Temperature is statistically associated with autozooid size (Figs. 6 and S9) but explains a rather small part of the phenotypic variation (Tables S7 and S8). Ideally, to firmly establish a causal relationship between temperature and autozooid size, formal time-series analyses should be implemented (Reitan and Liow 2019), but this was not possible because of the short length of the phenotypic time series. In summary, while autozooid size is one of the best studied phenotypic traits in cheilostome bryozoans, we are far from being able to accurately predict autozooid size variation as multiple selective pressures and constraints seem to be interacting in complex ways.

Ovicell size and autozooid shape each show different evolutionary patterns

Ovicell size reflects maternal investment into offspring where the size of the ovicell is correlated with the size of the larva (Herrera et al. 1996). As in autozooid size, ovicell size fluctuates through the evolutionary history of *M. agonistes* (Fig. 2). The best-fit univariate phenomenological model is a biased random walk and hence consistent with either directional changing evolutionary optima or stochastic temporal changes that are directional. Ovicell sizes closer to the recent are somewhat larger than those from earlier in the Pleistocene (Figs. 2, S2 and Table S2), however, within any time-interval, there is barely any indication of a relationship between ovicell area and fitness (Figs. 3 and S5), consistent with findings from an earlier study of the same genus (Di Martino and Liow 2021*b*). Likewise, there is no convincing pattern for multivariate trait-fitness association between competition and ovicell size (Fig. S8). Our third multivariate trait was autozooid shape. An earlier study of a different cheilostome species showed that colonies with longer autozooids were favored (Di Martino and Liow 2021*a*). In *Microporella*, however, colonies with “average” shaped autozooids consistently have the highest fecundity through time (Fig. 3), and this pattern of “stabilizing selection” is consistent with the best phenomenological univariate model being strict stasis (Fig. S2), where the variation in the data can be attributed to sampling noise. Shape, on the other hand, has no detectable consequence for competitive outcomes (Fig. S8).

Do shorter term processes scale up to macroevolutionary patterns?

There is an ongoing discussion on whether (and when) microevolutionary processes can predict macroevolutionary patterns (Rolland et al. 2023). While conceptual and theoretical research can contribute to this research, empirical studies are important in helping us think about this question. In contrast to the study of a different species of cheilostome bryozoans from the same community (Di Martino and Liow 2021*a*), where shorter-term trait-fitness associations were interpreted as not scaling up to macroevolutionary patterns, the current study suggests that they could. With orders of magnitude more measurements generated in part with deep-learning tools, and more data in terms of the number of time points for time series analyses, a more nuanced picture of trait evolution on shorter (tens of thousands of years, represented by trait-fitness associations within time intervals) to longer (hundreds of thousands of years, represented by phenomenological evolutionary models) time scales is emerging. However, what seems key over the lifetime of a genetic individual, such as a relatively larger autozooid size to contribute positively to competition for resources and survival, does not indefinitely increase, as autozooid size is kept in check by environmental constraints (temperature) and trade-offs (smaller zooids sometimes are associated with more fecund colonies), and conspecifics with larger zooids do not win competitions more frequently, suggesting additional predictors of such combats (see e.g. Liow et al. 2019). These complexities within the study system, where we have only scratched the surface, suggest that shorter-term processes may often have limited predictive power for longer-term outcomes (Rolland et al. 2023). Nevertheless, there are no deep conflicts among plausible explanations over different time scales, as previously suggested in the past (Erwin 2000) with variation accountable for by sampling noise, environment, multivariate constraints and likely changing selective pressures.

Stasis versus sources of variation

Eldgredge and Gould's widely cited work on phenotypic stasis in the fossil record (Eldredge and Gould 1972) has sparked considerable research and debate (Pennell et al. 2014) with some of the key conflicts arising due to a mismatch between the interpretation of links between data and conceptual frameworks. The current study adds data and insights to the "Punc-Eq" battleground. The documented long-term temporal trait variation in our focal species is clear and hardly exceptional among cheilostome bryozoans (Liow et al. 2017; Voje et al. 2020; Di Martino and Liow 2021*a*) and other organisms (Hunt et al. 2008; Brombacher et al. 2017; Voje 2020). We do agree broadly that "stasis is data" (Gould 1991), finding that one of the traits we investigate (autozooid shape in *M. agonistes*) best fits a model of little change and lots of sampling noise. Yet "best-fit" is conditioned on the suite of models applied, the length of the time series and the amount of available data per time interval, where a fossil time series best fitting a stasis model can still exhibit significant temporal variability (Voje 2018). Deciphering the causes of this variability, whether due to poor sampling, mixing of temporal populations within sedimentary layers, or other factors or combinations of factors, is challenging but with multiple lines of investigation and high-throughput phenotyping as an aid, may be attainable, as shown by our study system.

Beyond the focal species: contrasting congenerics with distant relatives

The *Microporella* species in New Zealand, including those presented here, are likely to be monophyletic (Orr et al. 2022). Given their phenotypic similarity and close phylogenetic relationships, we expected these congenerics to respond more similarly to their past shared environment, compared to a distinct genus, *Antarctothoa* that lived in the same community over 2 million years (Di Martino and Liow 2021*a*). While average autozooid sizes are consistently associated with temperature across all the *Microporella* examined, trait-fitness associations differed among them, also in time (Fig. S11). In contrast, unlike *Antarctothoa*, where there is selection for skinner zooids, all *Microporella* species we studied show evidence for an optimal average autozooid shape ("stabilizing" selection). In general, our results (Figs. 3, 5, S11) here are consistent with the general lack of a relationship or trade-off between fecundity and ovicell size, a proxy for maternal investment, at least among marine invertebrates (Marshall and Keough 2007; Di Martino and Liow 2021*b*).

Caveats, challenges and opportunities

Cheilostome bryozoans are great model systems for evolutionary studies for the reasons discussed in the introduction. In addition, their clonal nature allows us to partially account for phenotypic plasticity that is due solely to environmental variation (Di Martino and Liow 2021*b*, 2022). But the clonal/colonial/sessile-as-adult lifestyle likely faces different constraints because of the plasticity of the modular units, the possibility of survival after partial mortality, and asexual growth as a means of population growth and spread (Sebens 2002; Gorné and Díaz 2019). While no study system is perfect for answering general questions, diverse empirical systems can lend support to increasing our understanding of common processes, including selection, constraints and the processes that link shorter versus longer timescale patterns. A brief list of the caveats of the study system and approaches is warranted. First, the trait-fitness associations we studied are reminiscent of those observed in selection studies (Lande and Arnold 1983; Hendry and Kinnison 1999), yet the data presented within each temporal population represent multiple generations that have experienced a plethora of environmental pressures. The amount of time potentially represented in each of our temporal sample does not detectably inflate phenotypic variance (Fig. S3, see also Hunt 2004). And while the shape of the multivariate trait-fitness relationships (Fig. 3) may not be selection gradients, they do represent broad-sense selection and can indicate changing optima. Second, even though we have greatly increased the amount of phenotypic data available for analyses by applying automated phenotyping tools, some combinations of species and time are still noticeably poor in data in this study, even in such a rich fossil system. This is in part due the well-known pattern in community ecology, some species are naturally more abundant than others, but also due to remaining bottlenecks in collecting, processing, and analyzing fossil material. Third, a key fitness component may be missing from our analyses, namely colony growth rate, a parameter that is almost impossible to estimate in fossils and challenging to estimate in natural populations. Growth rates for only a handful of cheilostome bryozoans have been compiled (Smith 2014). Fourth, while we have performed separate analyses for the two fitness components that we identified (fecundity versus competition as a proxy for colony survival), it might have been ideal to combine these via population growth (Gómez-Llano et al. 2023) but population dynamics is a challenge to estimate well (Reitan et al. 2022). Fifth, we have estimated partial mortality using discrete outcomes of competition, yet such spatial competition could involve very varied numbers of zooids in both parties. Last, we have assumed that in the colonies we examined each ovicell is used only once for the production of a single larva, but this might not hold true in all the samples we had available.

Concluding remarks

Generation to generation survival and reproduction must contribute to the existence of a species, but linking these to longer-term population dynamics on the geological time scales is challenging (Reitan et al. 2022). Selection agents can be obscure even for contemporary populations (MacColl 2011) and stabilizing selection, even on generational timescales, may be rarer than we might have realized (Stroud et al. 2023). Non-model calcifying marine organisms, such as cheilostome bryozoans, have a rich fossil record that allows us to peer into past using ecological and evolutionary parameters that are equivalent to and comparable with those commonly studied among contemporary species. Colonial organisms in addition allow us to quantify environmental variation (Di Martino and Liow 2022), yet there are only a handful of selection studies based on extant colonial organisms (Kenyon et al. 2020; Burgess and Bueno 2021) leaving us to base much of our premises and assumptions in this study on those from solitary organisms. Cheilostome bryozoans are one of our best bets for solidifying the mechanistic and phenomenological links between micro- and macroevolution (Rolland et al. 2023) and require more attention from evolutionary biologists.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant no.724324 to LHL), the Research Council of Norway (grant no. 314499 to EDM) and NVIDIA Corporation (Nvidia Hardware Grant to AP). We thank S. Rust, K.L. Voje, P.D. Taylor, A. Beu, D.P. Gordon, M.H. Ramsfjell, J. Boeve, C. Schack, E. Enevoldesen, and D.R. Yanakopulos for their field contributions in the fossil collection seasons of 2014, 2017, and 2020 and GNS for arranging permits. We are also grateful to M.H. Ramsfjell for helping with Scanning Electron Microscopy, and S. Nodder and D.P. Gordon for the piston core samples of the Whanganui Core. NIWA collected the samples in the South Taranaki Bight as part of the project SUSS17305 “Sediment tolerance and mortality thresholds of offshore benthos” funded by the New Zealand Sustainable Seas National Science Challenge (2017), and those from Cook Strait as part of the commercial projects “TTR11301 Benthos” and “TTR11301

Deepwater" with Trans-Tasman Resources (TTR) Ltd, in 2012. S. Singhal, P.D. Polly and an anonymous reviewer have provided constructive comments that helped with the revisions of the original ms.

Statement of Authorship

Author All authors acquired funding for the study. LHL developed the concept of the study and performed the statistical analyses. EDM identified and curated the bryozoan material and organized the image compilation. AP wrote the pipeline for phenotyping images and automated the data acquisition from the images. LHL wrote the first draft of the ms and all authors contributed to subsequent editing.

Data and Code Accessibility

The data and R code necessary are deposited at https://doi.org/10.5061/dryad.3ffbg79rj. The images used to generate the data can downloaded at https://zenodo.org/records/10795305.

References

Abbott, S. T., T. R. Naish, R. M. Carter, and B. J. Pillans. 2005. Sequence stratigraphy of the Nukumaruan Stratotype (Pliocene-Pleistocene, c. 2.08-1.63 Ma), Wanganui Basin, New Zealand. Journal of the Royal Society of New Zealand 35:123–150.

Belk, M. C., S. J. Ingley, and J. B. Johnson. 2020. Life history divergence in live-bearing fishes in response to predation: is there a microevolution to macroevolution barrier? Diversity 12:179.

Bradski, G. 2000. The openCV library.

Brombacher, A., P. A. Wilson, I. Bailey, and T. H. G. Ezard. 2017. The breakdown of static and evolutionary allometries during climatic upheaval. The American Naturalist 193:350–362.

Brown, D. A. 1952. The Tertiary cheilostomatous Polyzoa of New Zealand. Trustees of the British Museum (Natural History), London.

Burgess, S. C., and M. Bueno. 2021. When does growth rate influence fitness in a colonial marine invertebrate? Marine Biology 168.

Buss, L. 1979. Bryozoan overgrowth interactions—the interdependence of competition for space and food. Nature 281:475–477.

Buss, L. W. 1986. Competition and community organization on hard surfaces in the sea. Pages 517–536 *in* J. Diamond and T. J. Case, eds. Community Ecology. Princeton University Press.

Buss, L. W., and J. B. C. Jackson. 1979. Competitive networks: nontransitive competitive relationships in cryptic coral reef environments. The American Naturalist 113:223–234.

Butler, M. A., and A. A. King. 2004. Phylogenetic comparative analysis: A modeling approach for adaptive evolution. The American Naturalist 164:683–695.

Carter, R. M., and T. R. Naish. 1998. A review of Wanganui Basin, New Zealand: global reference section for shallow marine, Plio-Pleistocene (2.5-0 Ma) cyclostratigraphy. Sedimentary Geology 122:37–52.

Crone, E. 2001. Is survivorship a better fitness surrogate than fecundity? Evolution 55:2611–2614.

Damgaard, C. 2019. A critique of the space-for-time substitution practice in community ecology. Trends in Ecology & Evolution 34:416–421.

De Lisle, S. P., M. I. Mäenpää, and E. I. Svensson. 2022. Phenotypic plasticity is aligned with phenological adaptation on both micro- and macroevolutionary timescales. Ecology Letters 25:790–801.

Di Martino, E., B. Berning, D. P. Gordon, P. Kuklinski, L. H. Liow, M. H. Ramsfjell, H. L. Ribeiro, et al. 2023. DeepBryo: A web app for AI-assisted morphometric characterization of cheilostome bryozoans. Limnology and Oceanography: Methods 21:542–551.

Di Martino, E., and L. H. Liow. 2021*a*. Trait-fitness associations do not predict within-species phenotypic evolution over 2 million years. Proceedings of the Royal Society B-Biological Sciences 288:20202047.

———. 2021*b*. Larger offspring associated with lower temperatures across species of *Microporella*, a widespread colonial invertebrate. Marine Ecology Progress Series 662:1–13.

———. 2022. Changing allometric relationships among fossil and Recent populations in two colonial species. Evolution 76:2424–2435.

Di Martino, E., L. H. Liow, T. Perkins, R. W. Portell, and P. D. Taylor. 2020. Sneaking up on ‘enemies’: alleviating inherent disadvantages in competitive outcomes in a nearly 3-million-year-old palaeocommunity from Florida, USA. Lethaia 53:553–562.

Di Martino, E., P. D. Taylor, D. P. Gordon, and L. H. Liow. 2017. New bryozoan species from the Pleistocene of the Wanganui Basin, North Island, New Zealand. European Journal of Taxonomy 345:1–15.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. Pages 82–115 *in* T. J. M. Schopf, ed. Models in Paleobiology. W.H. Freeman & Company.

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

Firmat, C., I. Lozano-Fernández, J. Agustí, G. H. Bolstad, G. Cuenca-Bescós, T. F. Hansen, and C. Pélabon. 2014. Walk the line: 600000 years of molar evolution constrained by allometry in the fossil rodent *Mimomys savini*. Philosophical Transactions of the Royal Society B: Biological Sciences 369:20140057.

Gillespie, J. L., C. S. Nelson, and S. D. Nodder. 1998. Post-glacial sea-level control and sequence stratigraphy of carbonate-terrigenous sediments, Wanganui shelf, New Zealand. Sedimentary Geology 122:245–266.

Gómez-Llano, M., W. A. Boys, T. Ping, S. P. Tye, and A. M. Siepielski. 2023. Interactions between fitness components across the life cycle constrain competitor coexistence. Journal of Animal Ecology 92:2297–2308.

Gordon, D. 1984. The marine fauna of New Zealand: Bryozoa: Gymnolaemata from the Kermadec Ridge. Memoirs of the New Zealand Oceanographic Institute 91:1–198.

———. 1989. The marine fauna of New Zealand: Bryozoa: Gymnolaemata (Cheilostomata Ascophorina) from the western South Island continental shelf and slope. Memoirs of the New Zealand Oceanographic Institute 97:1–158.

Gorné, L. D., and S. Díaz. 2019. Meta-analysis shows that rapid phenotypic change in Angiosperms in response to environmental change is followed by stasis. The American Naturalist 194:840–853.

Gould, S. J. 1991. Opus 200. Natural History 100:12–18.

Graham, J. E., and R. van Woesik. 2013. The effects of partial mortality on the fecundity of three common Caribbean corals. Marine Biology 160:2561–2565.

Hendry, A. P., and M. T. Kinnison. 1999. The pace of modern life: measuring rates of contemporary microevolution. Evolution 53:1637–1653.

Herrera, A., J. B. C. Jackson, D. J. Hughes, J. Jara, and H. Ramos. 1996. Life-history variation in three coexisting cheilostome bryozoan species of the genus *Stylopoma* in Panama. Marine Biology 126:461–469.

Hunt, G. 2004. Phenotypic variation in fossil samples: modeling the consequences of time-averaging. Paleobiology 30:426–443.

———. 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.

Hunt, G. 2015. PaleoTS https://CRAN.R-project.org/package=paleoTS.

Hunt, G., M. A. Bell, and M. P. Travis. 2008. Evolution toward a new adaptive optimum: Phenotypic evolution in a fossil stickleback lineage. Evolution 62:700–710.

Hunt, G., M. J. Hopkins, and S. Lidgard. 2015. Simple versus complex models of trait evolution and stasis as a response to environmental change. Proceedings of the National Academy of Sciences 112:4885–4890.

Hunt, G., and K. Roy. 2006. Climate change, body size evolution, and Cope's Rule in deep-sea ostracodes. Proceedings of the National Academy of Sciences of the United States of America 103:1347–1352.

Jackson, J. B. C., and A. H. Cheetham. 1990. Evolutionary significance of morphospecies - a test with cheilostome Bryozoa. Science 248:579–583.

Jackson, J. B., and S. P. Wertheimer. 1985. Patterns of reproduction in five common species of Jamaican reef-associated bryozoans. Pages 161–168 *in* Bryozoa, Ordovician to Recent: Papers presented at the 6th International Conference on Bryozoa, Vienna, 1983. Fredensborg, Olsen & Olsen.

Kenyon, T. M., C. Doropoulos, S. Dove, G. E. Webb, S. P. Newman, C. W. H. Sim, M. Arzan, et al. 2020. The effects of rubble mobilisation on coral fragment survival, partial mortality and growth. Journal of Experimental Marine Biology and Ecology 533:151467.

Kingsolver, J. G., and D. W. Pfenning. 2004. Individual-level selection as a cause of Cope's Rule of phyletic size increase. Evolution 58:1608–1612.

Kuklinski, P., and P. D. Taylor. 2008. Are bryozoans adapted for living in the Arctic? Virginia Museum of Natural History, Special Publication 101–110.

Kurtén, B. 1958. The life and death a Pleistocene cave bear, a study in paleoecology. Acta Zoologica Fennica 95:1–59.

Lancaster, L. T., G. Morrison, and R. N. Fitt. 2017. Life history trade-offs, the intensity of competition, and coexistence in novel and evolving communities under climate change. Philosophical Transactions of the Royal Society B: Biological Sciences 372:20160046.

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

Liow, L. H., E. Di Martino, G. Krzeminska, M. Ramsfjell, S. Rust, P. D. Taylor, and K. L. Voje. 2017. Relative size predicts competitive outcome through 2 million years. Ecology Letters 20:981–988.

Liow, L. H., E. Di Martino, K. L. Voje, S. Rust, and P. D. Taylor. 2016. Interspecific interactions through 2 million years: are competitive outcomes predictable? Proceedings of the Royal Society B-Biological Sciences 283:20160981.

Liow, L. H., T. Reitan, K. L. Voje, P. D. Taylor, and E. Di Martino. 2019. Size, weapons, and armor as predictors of competitive outcomes in fossil and contemporary marine communities. Ecological Monographs 89:e01354.

Liow, L. H., A. Porto, and E. Di Martino. 2024. Data from: “Trait-fitness associations via fecundity and competition in a two-million-year-long fossil record” The American Naturalist.

Liow, L. H., and P. D. Taylor. 2019. Cope’s Rule in a modular organism: Directional evolution without an overarching macroevolutionary trend. Evolution 73:1863–1872.

Liow, L. H., J. Uyeda, and G. Hunt. 2023. Cross-disciplinary information for understanding macroevolution. Trends in Ecology & Evolution 38:250–260.

Lisiecki, L. E., and M. E. Raymo. 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. Paleoceanography 20:PA1003.

Liu, Z., Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, et al. 2021. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. Pages 9992–10002 *in*2021 IEEE/CVF International Conference on Computer Vision (ICCV).

Livingstone, A. A. 1929. Papers from Dr Th. Mortensen’s Pacific Expedition 1914-16. XLIX. Bryozoa Cheilostomata from New Zealand. Videnskabelige Meddelelser fra Dansk naturhistorisk Forening i Kjøbenhavn 87:45–104.

MacColl, A. D. C. 2011. The ecological causes of evolution. Trends in Ecology & Evolution 26:514–522.

Marshall, D. J., C. N. Cook, and R. B. Emlet. 2006. Offspring size effects mediate competitive interactions in a colonial marine invertebrate. Ecology 87:214–225.

Marshall, D. J., and M. J. Keough. 2004. When the going gets rough: effect of maternal size manipulation on larval quality. Marine Ecology Progress Series 272:301–305.

Marshall, D. J., and M. J. Keough. 2007. The evolutionary ecology of offspring size in marine invertebrates. Pages 1–60 *in* Advances in Marine Biology (Vol. 53). Academic Press.

McKinney, F. K. 1995. Taphonomic effects and preserved overgrowth relationships among encrusting marine organisms. Palaiois 10:279–282.

Naish, T. R., S. T. Abbott, V. Alloway, A. G. Beu, R. M. Carter, A. R. Edwards, T. D. Journeaux, et al. 1998. Astronomical calibration of a southern hemisphere Plio-Pleistocene reference section, Wanganui Basin, New Zealand. Quaternary Science Reviews 17:695–710.

Naish, T. R., B. Field, H. Zhu, A. Melhuish, R. J. Carter, S. T. Abbott, S. Edwards, et al. 2005. Integrated outcrop, drill core, borehole and seismic stratigraphic architecture of a cyclothemic, shallow-marine depositional system, Wanganui Basin, New Zealand. Journal of the Royal Society of New Zealand 35:91–122.

O’Dea, A., and B. Okamura. 2000. Intracolony variation in zooid size in cheilostome bryozoans as a new technique for investigating palaeoseasonality. Palaeogeography Palaeoclimatology Palaeoecology 162:319–332.

Okamura, B., and J. D. D. Bishop. 1988. Zooid size in cheilostome bryozoans as an indicator of relative palaeotemperature. Palaeogeography Palaeoclimatology Palaeoecology 66:145–152.

Orr, R. J. S., E. Di Martino, D. P. Gordon, M. H. Ramsfjell, H. L. Mello, A. M. Smith, and L. H. Liow. 2021. A broadly resolved molecular phylogeny of New Zealand cheilostome bryozoans as a framework for hypotheses of morphological evolution. Molecular Phylogenetics and Evolution 161:107172.

Orr, R. J. S., E. D. Martino, M. H. Ramsfjell, D. P. Gordon, B. Berning, I. Chowdhury, S. Craig, et al. 2022. Paleozoic origins of cheilostome bryozoans and their parental care inferred by a new genome-skimmed phylogeny. Science Advances 8:eabm7452.

Pennell, M. W., L. J. Harmon, and J. C. Uyeda. 2014. Is there room for punctuated equilibrium in macroevolution? Trends in Ecology & Evolution 29:23–32.

Peters, R. H. 1983. The Ecological Implications of Body Size. Cambridge University Press.

Petrullo, L., S. Boutin, J. E. Lane, A. G. McAdam, and B. Dantzer. 2023. Environment mismatch errors enhance lifetime fitness in wild red squirrels. Science 379:269–272.

Reich, P. B., D. S. Ellsworth, and M. B. Walters. 1998. Leaf structure (specific leaf area) modulates photosynthesis–nitrogen relations: evidence from within and across species and functional groups. Functional Ecology 12:948–958.

Reitan, T., T. H. Ergon, and L. H. Liow. 2022. Relative species abundance and population densities of the past: developing multispecies occupancy models for fossil data. Paleobiology 49:23–38.

Reitan, T., and L. H. Liow. 2019. layeranalyzer: Inferring correlative and causal connections from time series data in r. Methods in Ecology and Evolution 10:2183–2188.

Rolland, J., L. F. Henao-Diaz, M. Doebeli, R. Germain, L. J. Harmon, L. L. Knowles, L. H. Liow, et al. 2023. Conceptual and empirical bridges between micro- and macroevolution. Nature Ecology & Evolution 7:1181–1193.

Rust, S., and D. Gordon. 2011. Plio-Pleistocene bryozoan faunas of the Wanganui Basin, New Zealand: stratigraphic distribution and diversity. New Zealand Journal of Geology and Geophysics 54:151–165.

Ryland, J. S., and D. P. Gordon. 1977. Some New Zealand and British species of *Hippothoa* (Bryozoa: Cheilostomata). Journal of the Royal Society of New Zealand 7:17–49.

Santos, E. S. A., and S. Nakagawa. 2012. The costs of parental care: a meta-analysis of the trade-off between parental effort and survival in birds. Journal of Evolutionary Biology 25:1911–1917.

Schluter, D., T. D. Price, L. Rowe, and P. R. Grant. 1997. Conflicting selection pressures and life history trade-offs. Proceedings of the Royal Society of London. Series B: Biological Sciences 246:11–17.

Sebens, K. P. 2002. Energetic constraints, size gradients, and size limits in benthic marine invertebrates. Integrative and Comparative Biology 42:853–861.

Smith, A. M. 2014. Growth and calcification of marine bryozoans in a changing ocean. The Biological Bulletin 226:203–210.

Stearns, S. C. 1989. Trade-offs in life-history evolution. Functional Ecology 3:259–268.

Stroud, J. T., M. P. Moore, R. B. Langerhans, and J. B. Losos. 2023. Fluctuating selection maintains distinct species phenotypes in an ecological community in the wild. Proceedings of the National Academy of Sciences 120:e2222071120.

Taylor, P. D. 2016. Competition between encrusters on marine hard substrates and its fossil record. Palaeontology 59:481–497.

Turner, S. j., and C. D. Todd. 1994. Competition for space in encrusting bryozoan assemblages: the influence of encounter angle, site and year. Journal of the Marine Biological Association of the United Kingdom 74:603–622.

Uttley, G. H., and J. S. Bullivant. 1972. Biological results of the Chatham Islands 1954 Expedition. Part 7. Bryozoa Cheilostomata. New Zealand Oceanographic Institute Memoir 57:1–61.

Van Valen, L. 1963. Selection in natural populations: *Merychippus primus*, a fossil horse. Nature 197:1181–1183.

Voje, K. L. 2018. Assessing adequacy of models of phyletic evolution in the fossil record. Methods in Ecology and Evolution 9:2402–2413.

———. 2020. Testing eco-evolutionary predictions using fossil data: Phyletic evolution following ecological opportunity. Evolution 74:188–200.

———. 2023. Fitting and evaluating univariate and multivariate models of within-lineage evolution. Paleobiology 49:747–764.

Voje, K. L., E. Di Martino, and A. Porto. 2020. Revisiting a landmark study system: no evidence for a punctuated mode of evolution in *Metrarabdotos*. The American Naturalist 195:899–917.

Voje, K. L., M. Grabowski, A. Holstad, A. Porto, M. Tsuboi, and G. H. Bolstad. 2023. Does lack of evolvability constrain adaptation? If so, on what timescales. Evolvability: A Unifying Concept in Evolutionary Biology? The MIT Press.

Figure 1: Study species and traits measured. A scanning electron micrograph shows part of a *Microporella agonistes* colony (Recent, Cook Strait KAH1204.04, shell 1, colony 23; SEM number: edm8185) and the traits measured. Outlined are the perimeter of an autozooid (the space enclosed is the first focal trait autozooid area), and the perimeter of an ovicell (the space enclosed is the second focal trait ovicell area). Scale bar is 500 μm. Autozooid length (ZL) and autozooid width (ZW) are indicated on the stylized figure of a single autozooid. The third focal trait autozooid shape is the ratio ZL/ZW.

Figure 2: Trait and fecundity variation of *Microporella agonistes.* A. through D. show natural logged mean autozooid area (μm^2), logged mean ovicell area (μm^2), mean autozooid shape (length divided by width, dimensionless), and mean colony fecundity (approximated by ovicell per polymorph, i.e. density). Colony means (blue dots) are jittered for visibility. In A. to C., population means and 1 standard error are shown, while in D., means and 95% quantiles are marked. The numbers on the x axes are tags for the formations/populations studied (see Table S1). Panel background colors delineate populations within each million-year, with olive green (furthest right panels) indicating the Recent and Whanganui core populations. Some colonies do not have ovicells preserved (see Table S1).

Figure 3: Model-averaged trait-fitness associations for temporal populations of *M. agonistes.* Populations are ordered from top (oldest) to bottom (youngest) labeled with their names (Shakespeare Cliff Basal Shellbed is shortened) and ages. Columns from left to right show fecundity (ovicell density) versus mean-standardized log autozooid area measured in μm^2 (Az.), mean-standardized log ovicell area measured in μm^2 (Ov.), and mean-standardized autozooid shape (Sh.) is unitless. Grey dots are colony averages and solid colored lines are the model averaged predictions of the multivariate trait-fitness relationships within each population. For the ease of comparison, y-axes (ovicell density) are the same throughout, and x-axes the same for each trait.

Figure 4: Outcomes of overgrowth competition for *Microporella agonistes*. A. The binomial probability of wins, regardless of the relative autozooid size of *M. agonistes* in the frontal overgrowth competition for populations. B. The binomial probability of wins, given that *M. agonistes* is the larger combatant (i.e. this is a subset of A where *M. agonistes* has larger autozooids). Numbers are sample sizes and vertical lines are 95% CI. Horizontal grey dashes at 0.5 show the null expectation. The x-axes are in millions of years ago (mya).

Figure 5: Winners are not predicted by autozooid size and fecundity. Panels show boxplots of A. autozooid size and B. fecundity for losers (L) and winners (W), see Fig. S1D for an example. T-tests confirm that the two groups are not different, also when analyzed population by population. C. shows that there is no trade-off between ovicell area and fecundity. The data are in grey (all populations) and black lines show individual predictions from 10 populations for which there are > 20 colonies for which data are available. See SI for details.

Figure 6: Relationship between paleoclimate and *Microporella agonistes* traits. Grey dots are average trait values for colonies given mean δ^{18}O value for each of the formations plotted with a jitter for visibility. Black dots are mean traits for each time interval. The black line is the prediction when other variables in the model are held constant and the grey its 95% CI for the best models. In this figure, the Recent is excluded as δ^{18}O values are extreme compared to the rest of the Pleistocene (but see Fig. S9 and Table S8 for plot and estimates including data from the Recent).