Evaluating phenotypic variation, genetic variance, and trait divergence of aggregate traits in two sister species of cheilostome bryozoans

September 13, 2022

Sarah Leventhal[1], Sarah Jamison-Todd[2], Carl Simpson[1]

[1] University of Colorado Museum of Natural History and Department of Geological Sciences, University of Colorado, Boulder 265 UCB Boulder, CO 80309

[2] University of College London, Earth Sciences, 5 Gower Place, London, WC1E 6BS

email: sarah.leventhal@colorado.edu

Keywords: Bryozoans, division of labor, macroevolution, major transitions, trait variation

Abstract: The study of trait evolution in modular animals is more complicated than that in solitary animals, since a single genotype of a modular colony can express an enormous range of phenotypic variation. The trait variation in each modular colony can be summarized as a phenotypic distribution that corresponds to a particular genotype. We test whether such phenotypic distributions in modular colonies are evolvable in two sister species of the bryozoan

Stylopoma, grown and bred in a common garden breeding experiment. To test the assumption that a phenotypic variance-covariance matrix is comparable to the genetic variance-covariance matrix when calculating evolvabilities in modular animals, we use both parent-offspring covariance matrices and phenotypic variance-covariance matrices to compute our evolvability estimates. We find that the medians of phenotypic distributions are evolvable between generations of colonies, while measures of dispersion are largely not. Our result uncovers another unusual feature of these two species of *Stylopoma*: the major axis of variation in the ancestral species correlates with the pattern of morphological divergence between the two species, but that of the descendent species does not. This result indicates that rapid evolution of evolvability can occur after speciation, and that the pattern of species divergence does not always correlate with the major axis of variation in descendent species. More broadly, we find that the use of our parent-offspring covariance matrices and our phenotypic covariance matrices do not yield comparable evolvabilities, indicating that, at least for modular animals, approximating evolvabilities with use of a phenotypic covariance matrix is problematic for study of aggregate trait evolution.

Introduction

In a group of colonial animals called cheilostome bryozoans, colonies are composed of genetically identical individuals, termed zooids, that possess a wide range of phenotypic variation (McKinney and Jackson, 1991; Lidgard, 1990). This variation is expressed among zooids so that colonies represent a snapshot of the phenotypic variation that stems from a single genotype (Cheetham et al., 1995, 1993). At its most extreme, this variation within colonies is expressed as differentiated body types, called polymorphs, which serve different functions in the colony. How polymorphism in bryozoans evolved is unknown (Taylor, 2020), but it is impossible to resolve without first understanding whether phenotypic distributions within colonies are actually evolvable across colony generations.

Cheilostome bryozoans have a complex life cycle with distinct asexual and sexual modes of reproduction (Taylor, 2020; McKinney and Jackson, 1991). Asexual reproduction allows for a single colony to grow larger by increasing the number of zooids in the colony. The sexual stage of the cheilostome bryozoan life cycle leads to the formation of new colonies, via the release of gametes from colony members (Lidgard and Jeremy, 1989; McKinney and Jackson, 1991; Taylor, 2020). The larvae formed by sexual reproduction between colonies establish new, genetically-distinct colonies upon settling.

The phenotypic variation (V_P) exhibited within a colony is neither the result of trait similarity between asexually-budded zooids (Simpson et al., 2020) nor environmental change along axes of colony growth (Okamura, 1992). However, species of bryozoans have measurable zooid-level differences from each other (Jackson and Cheetham, 1990). Therefore, there must be a mechanism that allows for zooid morphologies to evolve. If phenotypic distributions of zooid-level traits within colonies stem from additive genetic variances (V_A) passed across the sexual stage of reproduction (Fig. 1), then zooid-level traits are evolvable as colony-level phenotypic

distributions. The evolvability of these distributions would support existing explanations for the evolution of division of labor in bryozoan colonies, because polymorph types are believed to be derived from preexisting body types, termed autozooids, within colonies (Taylor, 2020; McKinney and Jackson, 1991). If distributions can evolve, then they can presumably split and diverge over time, allowing for morphologically distinct zooids to evolve.

In this study we assessed trait distributions of parent-offspring colony pairs in two *Stylopoma* species to test the hypothesis that phenotypic distributions of zooid-level traits in colonies are evolvable. These same specimens were used in a series of pioneering studies on quantitative genetics of bryozoans during the 1990s (Cheetham et al., 1993, 1994, 1995). However, these studies focused on zooid-level traits and did not consider colony-level evolution. Here we use new measurements on these same specimens to quantify the evolvability (sensu Houle 1992) of colony-level traits, and therefore their evolutionary potential.

Our dataset also allows us to test a longstanding theory about phenotypic variation and its relationship to underlying additive genetic variation. Typically, evolvability can be estimated from a phenotypic variance-covariance matrix, which is generally considerd a sufficient proxy for an underlying additive genetic variance-covariance matrix (Cheverud, 1988; Roff, 1995; Sodini et al., 2018; Love et al., 2021). However, while this may hold true for solitary animals, its applicability to trait evolution in modular animals has, as far as we know, never been explicitly tested. This presents a problem, as the genetic underpinning of phenotypic variation in modular animals is not as straightforward as it is in solitary animals, since each clonal colony expresses a wide range of phenotypic variation (Taylor, 2020; McKinney and Jackson, 1991). Since evolvability is believed to be calculable by scaling either phenotypic variance-covariance matrices or additive genetic variance-covariance matrices (derived from a parent-offspring covariance matrix), we use both methods to calculate our evolvability values, thus testing whether phenotypic variance-covariance matrices are an adequate proxy of additive

genetic variation in bryozoans.

Methods

Specimens

We used colonies from two species of the bryozoan genus *Stylopoma*. These specimens were originally used in several studies on skeletal morphology and the assessment of morphological differences between species of *Stylopoma* (Jackson and Cheetham, 1990; Cheetham et al., 1993, 1994, 1995). Colonies were grown in the 1980s from a larval stage to adulthood in a common environment at the Smithsonian's field station in the San Blas Islands in Panama. The experimental design was based on Maturo's protocol (Maturo Jr, 1973) to isolate maternal colonies with a clean, flat substrate on which their larvae could settle. Once offspring colonies were established, they were reared in a common environment. The experimental design is intended to minimize the impact of environment on phenotypic variation in colonies, so that phenotypic differences between colonies can be attributed to genetic differences alone. Furthermore, the flat surfaces colonies use as substrates minimize distortion of zooid morphology for measurement collection from two-dimensional images. More details on the rearing of these specimens can be found in Jackson and Cheetham 1990; Cheetham et al. 1993; Simpson et al. 2020.

We used a total of 93 parent-offspring colony pairs of *Stylopoma species 1* (42 pairs) and *Stylopoma species 2* (53 pairs).

Quantifying aggregate colony-level traits

To digitize colonies, we first used a stackshot rail, a Canon EOS camera, and Helicon Remote software to generate z-stacks of colonies. Our camera was set to 4x magnification, and we

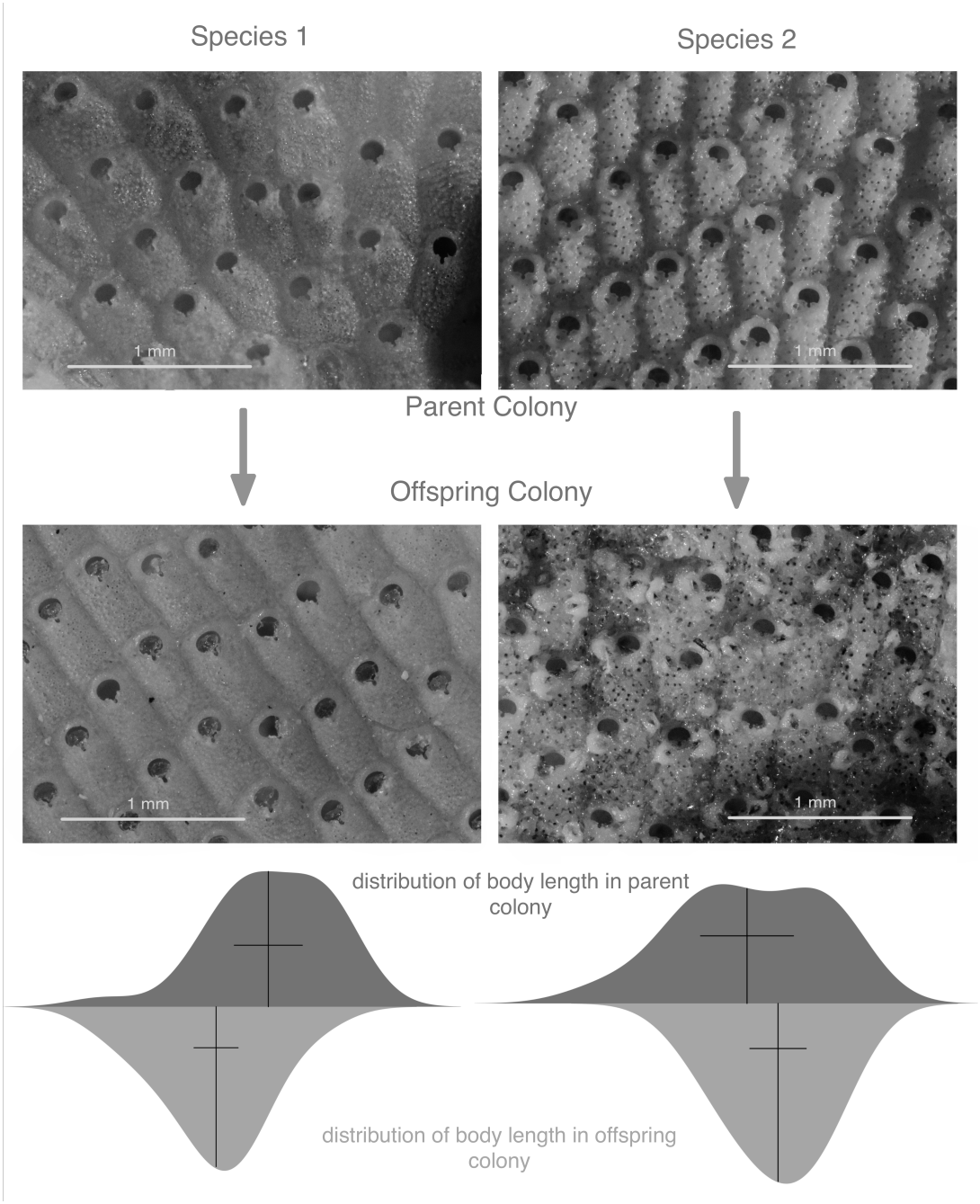

Figure 1: **An example of aggregate phenotypic similarity across a colony generation.** We show an example of two colony generations from two species of *Stylopoma*. The resulting distributions of zooid length values are displayed below each maternal and offspring colony. To quantify the distributions, we measure two aspects of these distributions: central tendency with the median (vertical black lines) and dispersion around the median with median absolute deviation (MAD, horizontal black lines). The analysis in this study measures parent-offspring covariation by comparing the medians and median absolute deviations of traits across colony generations.

set our stackshot rail to move in increments that have 10% focus overlap between photos. We used Helicon Focus software to stitch together our image stacks, yielding high-resolution focus-stacked photographs for measurement. We evaluated this technique for error by repeatedly measuring the same zooids in a colony photographed at different angles. We found that at angles of 20° to 30°, we had a difference of $\pm 20\mu m$ in measurements of autozooid body length (which is a $\sim$ 5% difference for this trait), which is only slightly greater than our measurement error on specimens lying flat.

We measured two summary statistics of the distributions of five autozooid traits in digitized parent and offspring colonies. To collect our measurements, we used ImageJ (Schneider et al., 2012). Within each colony, we selected thirty autozooids at random and measured four of their skeletal traits: zooid length, zooid width, orifice length, and orifice width. A fifth trait, frontal pore density, was quantified using a subset of five zooids per colony. We selected these traits because they are taxonomically significant in *Stylopoma*, as they allow for the morphological distinction of these two species (Simpson et al., 2020; Jackson and Cheetham, 1990). We did not consider avicularian morphology in this analysis, since the two species have different types of avicularia (vicarious in *S. species 1* and adventitious in *S. species 2*). The different structural properties of adventitious and vicarious avicularia make it difficult to compare homologous traits across species, so we focus exclusively on autozooidal traits.

To assess the measurement error for these traits, we measured each trait on a single zooid from each species ten times. The ranges of the measurements in each species were calculated and divided by the mean measured value in each species. Our error estimates were averaged between the two species, yielding single estimates of measurement error for each trait. The resulting error values were 3.8% for zooid length, 4.0% for zooid width, 4.1% for orifice length, 2.8% for orifice width, and 4.9% for frontal pore density.

The colony-level traits we selected are based on two summary statistics that quantify the

phenotypic distributions of constituent zooids. We first quantified the median phenotype (a measure of central tendency) of each trait in each colony. Then, we quantified dispersion among zooids within a colony by calculating the median absolute deviation (MAD) of each trait distribution. We chose these two summary statistics as proxies for the mean and variance because they are robust to small sample sizes (Whitley and Ball, 2001; Leys et al., 2013; Kashif et al., 2017) and if the data are normally distributed, these values are estimators of the mean and variance. Given the large number of measurements ($> 25,000$) required for this study, we prioritized measuring more colonies over sampling more zooids within a colony. The use of the median and MAD values ensures that central tendency and dispersion in the dataset are adequately captured. Of course, it is possible that evolution can occur at higher moments (skew and kurtosis) (Rice, 2004) of the distributions, but for this analysis we are focusing on first and second moments.

There are generally two types of colony-level traits: aggregate traits, which are summary statistics of traits in colony members, and emergent traits, which are only expressed at the colony-level (Grantham, 1995; Okasha, 2014; Lloyd and Gould, 1993). Evolutionary potential of these traits has rarely been assessed, with the exception of some studies (e.g., Jablonski 1986 and Jablonski and Hunt 2006). In bryozoans, emergent colony-level traits, such as the position and orientation of polymorphs relative to other zooids in colonies, are known to have evolutionary potential (Simpson et al., 2020). While the sorting of aggregate traits has been subject to some prior study (e.g., Lloyd and Gould 1993), the evolutionary potential of aggregate traits has, as far as we know, never been assessed. This methodology allows us to assess the evolutionary potential of aggregate traits in bryozoan colonies.

Measuring colony-level trait evolvability from the parent-offspring covariance matrix

We follow Schluter (Schluter, 1996) to calculate a parent-offspring covariance matrix using parent-offspring pairs, and we follow Hansen (Hansen et al., 2003; Hansen and Houle, 2008) to calculate evolvabilities. First, we calculated the covariances of median and MAD trait values across parent and offspring colony generations to form a parent-offspring covariance matrix (Fernandez and Miller, 1985; Rice, 2004; Falconer and Mackay, 1996). Since these colonies are sexually produced, we doubled this covariance matrix to estimate the additive genetic contribution of each colony-level trait (Falconer and Mackay, 1996). To deal with the natural asymmetry of a parent-offspring covariance matrix, we took the mean of entries across the diagonal (Schluter, 1996). We scaled entries in our parent-offspring covariance matrix by dividing them with the squared mean value for each trait (Hansen et al., 2011; Houle, 1992). For off-diagonal values, we use the product of the two trait means instead of a single trait squared mean. Scaling the parent-offspring covariance matrix in this way yields evolvabilities (Houle, 1992), which can be interpreted as a maximum phenotypic percent change per generation under direct selection.

Traits in our parent-offspring covariance matrices with diagonal values below zero were removed from our matrices, since negative values indicate that there is no additive genetic variance of those traits (Cheetham et al., 1995). In our case, the median trait values of orifice width were found to have no additive genetic variance for either species. For our MAD evolvabilities, each species has different traits with evolvabilities of zero (stemming from the corresponding parent-offspring covariance values falling below zero): for *Stylopoma species 1*, the MADs of pore density and orifice have additive genetic variances of zero, and for *Stylopoma species 2*, the MADs of body length and body width have additive genetic variances of zero. Due to the

relatively high number of traits with no additive genetic variance in either or both species, we concatenated the MAD and median traits for each species into a single matrix for analysis.

We calculated the standard error in our estimates of trait evolvability using the method outlined by Lynch and Marsh (Lynch and Walsh, 1998). To follow this method, we first found the standard deviation of our large sample variance for each trait and then we divided that standard deviation value by the trait mean squared (Lynch and Walsh, 1998; Garcia-Gonzalez et al., 2012) to propagate error through our evolvability calculation. The resulting values represent the standard error of each evolvability estimate.

Measuring colony-level trait evolvability from the phenotypic variance-covariance matrix

Our estimation of evolvability using a phenotypic variance-covariance matrix is more straightforward. We first pooled together our parent and offspring colony datasets, and generated a covariance matrix across all traits (Falconer and Mackay, 1996). This covariance matrix forms our phenotypic variance-covariance matrix. To calculate evolvability (Houle, 1992), we scaled these values by the mean trait values squared, as described in the above procedure. The resulting values represent our phenotypic variance-covariance-derived evolvabilities.

Our error estimates remain the same in this procedure as they do in the procedure above.

Comparison of evolvabilities derived from the parent-offspring covariance matrix and the phenotypic variance-covariance matrix

The evolvabilities derived from the phenotypic variance-covariance matrices were all above zero, while some of the evolvabilities we calculated from the parent-offspring covariance matrices had negative values, which we interpreted as variances of 0. This alone demonstrates that there are

inherent differences in the evolvabilities calculated from parent-offspring covariance matrices and from phenotypic variance-covariance matrices. However, to estimate how similar the nonzero evolvabilities are in each species, we filtered out traits with parent-offspring covariance matrix evolvabilities of zero in each species.

We compared the remaining species-specific evolvabilities using the random skewers method (Cheverud, 1996; Cheverud and Marroig, 2007) and Krzanowski Correlations (Krzanowski, 1993) to see if the matrix of nonzero evolvabilities calculated from our parent-offspring covariance matrices and our phenotypic variance-covariance matrices are similar.

Comparison of evolvabilities between species

We first compared the matrix of evolvabilities derived our phenotypic variance-covariance matrices across species using the random skewers method (Cheverud, 1996; Cheverud and Marroig, 2007) and Krzanowski Correlations (Krzanowski, 1993). This provides us with an estimate of how similar patterns of phenotypic variation are between species.

To compare our evolvabilities derived from our parent-offspring covariance matrices, we had to account for the differences in non-evolvable traits across species. Therefore, we removed traits with no additive genetic variance in one of species from the evolvability matrices of both species so that the matrices could be compared. The resulting dimensions of our two matrices correspond to the five colony-level aggregate traits we measured that have non-zero evolvability in both species: the evolvabilities of median zooid length, zooid width, orifice length, and frontal pore density, and the evolvability of the MAD of orifice width. We evaluated the similarity of our matrices using the random skewers method (Cheverud, 1996; Cheverud and Marroig, 2007) and Krzanowski Correlations (Krzanowski, 1993). All statistical tests were performed in RStudio 2021.09.0 (Team, 2015) using the EvolQG package (Melo et al., 2015).

Quantifying directions of evolutionary change

Since there are differences between our evolvability matrices depending on methodology, we use evolvabilities derived from our parent-offspring covariance matrices for the remainder of the analysis, as parent-offspring covariance provides a more accurate assessment of additive genetic variation than phenotypic variance-covariance matrices.

For this portion of the analysis, we consider each species's evolvability matrix separately, so that the underlying eigenstructure of each evolvability matrix can be compared to the pattern of trait divergence (for traits that are evolvable in each matrix). To calculate trait divergence between species, we first pooled the species together and calculated the pooled mean for each trait. Then we divided the mean trait values for each species by the pooled mean for each trait. We subtracted the scaled median and MAD trait means of *Stylopoma species 2* from those of *Stylopoma species 1* to estimate divergence for each trait. We used these trait divergences to create divergence vectors for our analysis (Hopkins, 2016).

For our trait divergence, we computed a vector of a single row, with each column corresponding to the dimension of a trait with nonzero evolvability in each species. We use this vector to compute a single angle of deviation from the first eigenvector (Schluter, 1996). For each species, the divergence vector contains 7 columns, one for each trait with nonzero evolvability (note that the traits with nonzero evolvability differ between species).

While we compute these angles to compare the major eigenvector structure of each species with the pattern of divergence, this method has been subject to criticism in the past for focusing exclusively on the major axis of variation in a species (Hunt, 2007). However, since the first eigenvectors of our evolvability matrices cover the majority of the variation in each species (51% in parent-offspring covariance evolvabilities for *Stylopoma species 1*, and 92% in parent-offspring covariance evolvabilities for *Stylopoma species 2*), we focus only on the first eigenvector of each evolvability matrix for this analysis.

Estimating trait-specific patterns of species divergence

Since these two species of *Stylopoma* are sisters (Jackson and Cheetham, 1994), we have the opportunity to understand how our measured evolvability values correspond to the pattern of trait divergence between these species. In our case, we want to evaluate the angle of deviation between the major axis of variation in each species and the dimensional axis of each trait. We therefore estimate the angle between the major axis of variation (the first eigenvector) of each species (Hansen and Houle, 2008) and a vector of each corresponding trait dimension.

For this part of the analysis, we use an expanded trait dimension vector, with an equal number of rows and columns. We use traits that have nonzero evolvabilities in each evolvability matrix. Each species has three traits with evolvabilities of zero, so we use the seven traits with nonzero values as our columns.

To estimate angles for trait-specific divergence, we computed the arccosine of the dot product first eigenvector for each species and each row of the corresponding dimension vector. The resulting angles represent the deviation between the major axis of variation in each evolvability matrix and the dimensional axis of each trait. This provides insight into how closely the major axis of variation in each species aligns with each trait dimension. We compared the angles (which we call θ_i, for traits 1 through i) calculated from the procedure above with the corresponding trait distance between species. For each comparison, we calculated correlations to assess the significance of the relationship.

To assess the direct relationship between trait evolvabilities and trait distance between the species, we took the mean evolvability value for each of the five nonzero traits across the two species and compared those mean evolvabilities with trait divergence between the species. We measured a correlation to evaluate the strength of this relationship.

Estimating null values

To estimate null values for comparison with our estimated values, we used a bootstrap procedure. We generated 1,000 randomized parent-offspring covariance matrices with replacement, and doubled these matrices and averaged them across the diagonals so that they would be symmetrical. To estimate null evolvabilities, we divided the diagonal values by the mean trait values squared, following the same methodology as we did for estimating evolvabilities from our measured parent-offspring covariance matrices (Schluter, 1996). We also calculated the variance of our bootstrapped evolvability estimates.

To assess the significance our calculated angles of deviation between trait-specific species divergence and the first eigenvector of our evolvability matrices, we generated a distribution of angles computed between 10,000 randomly-generated vector pairs. We computed the mean of this distribution to compare with our calculated angles.

Results and discussion

The impact of methodology on evolvability estimation: the variable contributions of additive genetic variance to phenotypic variation

Our evolvability estimates vary depending on whether they were calculated from the phenotypic variance-covariance matrix or the parent-offspring covariance matrix in each species (Tables 1 and 2).We find that there is a general overestimation of evolvability when values are derived from a phenotypic variance-covariance matrix. Using this method, traits that have no additive genetic variation across generations still have nonzero evolvabilities because there is phenotypic variation observed in colonies. However, for certain traits, that phenotypic variation does not correspond to underlying additive genetic variation.

Even when we restrict our comparison to traits that have nonzero evolvabilities, there are significant differences in evolvability values from the phenotypic variance-covariance matrix and the parent-offspring covariance matrix in *S. species 1* (random skewers $r = 0.38$, $p \sim 0.2$, Krzanowski Correlation $r = 0.35$). For *S. species 2*, direct comparison of nonzero values shows significant similarity, indicating that, at least for this particular species, the phenotypic variance-covariance matrix is a suitable estimate of evolvability for traits that have nonzero additive genetic variation (random skewers $r = 0.79$, $p \sim 0.01$, Krzanowski Correlation $r = 0.92$).

While some evolvability values can be adequately estimated from the phenotypic variance-covariance matrix, the use of this matrix as a proxy for the additive genetic variation fails for traits that are variant due to reasons other than underlying additive genetic variation. In modular animals, there is a large portion of phenotypic variation that is not underpinned by genetic variation (Taylor, 2020; McKinney and Jackson, 1991). Therefore, studies of modular animals that rely on phenotypic variance-covariance matrices for estimations of evolvability should be interpreted with caution, as evolvabilities may be overestimated in traits that have no underlying additive genetic variation.

Interestingly, the phenotypic variance-covariance matrices between species are similar (Random skewers $r = 0.85$, $p \sim 0.0005$; Krzanowski $r = 0.836$). We interpret this result to indicate that these two species of *Stylopoma* exhibit similar patterns of phenotypic variation despite possessing different patterns of underlying additive genetic variation. This result may reflect the common environment in which colonies were reared, since phenotypic variation in colonies may result from environmental plasticity rather than additive genetic variation (Bull, 1987).

Since many traits are not evolvable in either species, our result indicates that pleiotropy, epistasis, or genetic mutation may play a large role in modulating phenotypic expression in *Stylopoma* (Hansen et al., 2006). Evolvability refers to the propensity for change in trait variability (a trait's tendency to vary), and not simply an instance of expressed phenotypic

plasticity (Wagner and Altenberg, 1996). Accurate measures of evolvability therefore require robust population sampling. In modular animals, each colony exhibits a portion of a trait's potential variability, but the full measure of variability in a species can only be estimated from sampling many colonies. We find that the variability of traits across colonies is not the same as the pattern of phenotypic variation in a single colony, thereby showing that evolvability estimation from trait variation in colonies of modular animals alone is insufficient.

The partial evolvability of colony-level phenotypic distributions

Since we have two estimates of evolvability in each species, we focus on evolvabilities derived from our parent-offspring covariance matrices for the remainder of our analysis, since it is a more accurate representation of underlying additive genetic variation than the evolvabilities derived from phenotypic variance-covariance matrices.

In our parent-offspring covariance-derived evolvabilities, we find that median trait values of phenotypic distributions are largely evolvable, with more nuanced results for our MAD traits (Tables 1 and 2). While the mean evolvability of MAD values is higher than that of the medians in both species (0.0090 versus 0.0021, respectively, in *S. species 1*; 0.027 and 0.0048, respectively, in *S. species 2*), MAD evolvabilities also have higher error, likely due to the similarity in values between unrelated colonies. This error indicates that MAD evolvabilities may not be significantly above the null evolvability value. In contrast, most median evolvabilities are significantly above the null evolvability value even when accounting for error. This indicates that evolution of the central location of a colony-level phenotypic distribution in morphospace may evolve more readily than its dispersion around the median. In other words, aggregate traits within colonies primarily evolve according to translational shifts of the median.

The random skewers comparison of our parent-offspring evolvability matrices shows that, while on the margin of acceptance (random skewers $r = 0.64$, $p \sim 0.085$, Krzanowski Correla-

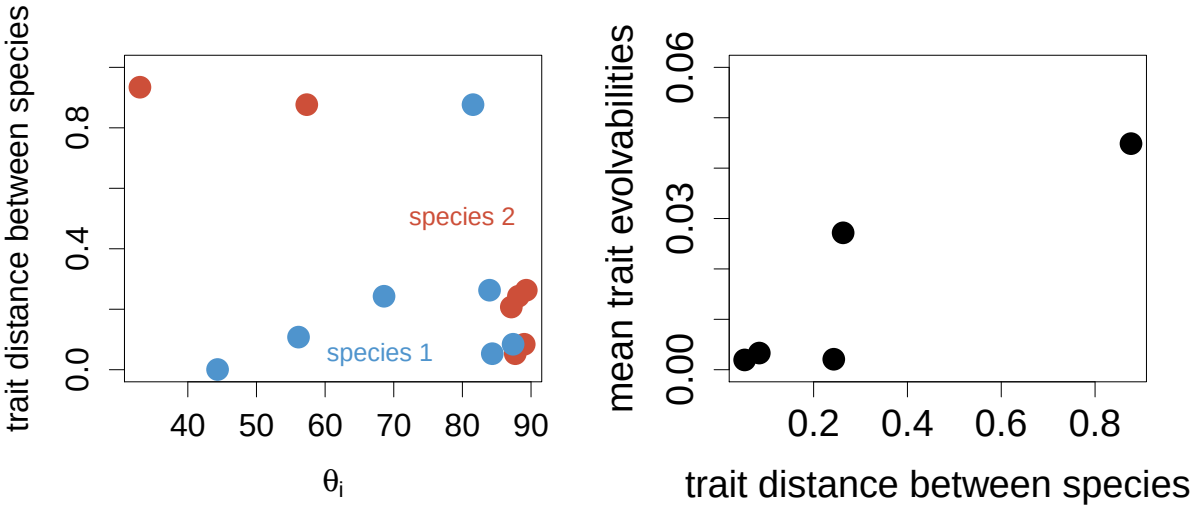

Figure 2: **Macroevolutionary impact of colony-level evolvability.** **A:** Angle between each component of the mean trait divergence vector and the first eigenvector of trait values in each species plotted against the trait divergence vector (r_{sp1} = 0.347, r_{sp2} = -0.939). **B:** Averaged trait evolvabilities (of the five traits with positive evolvabilities in both species) plotted against mean trait divergence (r = 0.897).

tion r = 0.73), each species's parent-offspring covariance matrix-derived evolvabilities may not respond in the same way to selection. This estimation of matrix correlation likely overestimates the degree of similarity between our matrices, since certain traits are evolvable in only one of the species, indicating that the major axes of variation in the two species are likely significantly different. Trait covariation in these two *Stylopoma* species therefore appears to have evolved since *S. species 1* and *S. species 2* diverged (Arnold et al., 2008).

The variable evolvability of MAD values is peculiar considering the large range of trait variation observable in *Stylopoma* colonies. We posit that phenotypic variation in colonies may be underpinned by pleiotropy or epistatic changes to gene expression (Wagner and Altenberg, 1996) rather than heritable variation. There are inherent developmental constraints to the range

of phenotypes that can be expressed by a genotype. Trait variation in colonies may represent the limits of each genotype's phenotypic range of expression, which is particular to each colony's genotype, rather than a clear aggregate trait that is inherited across colony generations. We believe that further investigation into how genotype-phenotype maps evolve in colonial animals like bryozoans can uncover how MAD traits may evolve without having measurable underlying additive genetic variation across colony generations.

The macroevolutionary impact of evolvable aggregate-level traits

The major axis of variation in *S. species 2* has a relatively small angle of deviation from the direction of trait divergence between species (41.1°). However, *S. species 1* has an angle of deviation that is relatively high (74.1° for *S. species 1*), reflecting the differences in each species's eigenstructure. Our computed angle for *S. species 1* is nearly identical to the mean of our null distribution (74.7°), indicative of a weak relationship between the major axis of variation in the species and the axis of species divergence.

In our comparison of trait-specific divergences and the angle of deviation between the first eigenvector for each species and each corresponding dimension of the divergence vector, we find that *S. species 2* has a negative correlation ($r \sim -0.939$), but *S. species 1* does not ($r \sim 0.347$) (Fig. 2A). We interpret this result as further indication of *S. species 1* having a different evolvability structure than *S. species 2*, which may have become unstable and shifted after speciation (Arnold et al., 2008; Doroszuk et al., 2008).

Together, these results indicate that speciation occurred according to the variance structure of the ancestral species (*S. species 2*), with that of the descendent species (Jackson and Cheetham, 1999) possibly diverging after speciation. In this particular example, speciation appears to depend on ancestral species rather than the descendent species, with the eigenvector structure of the ancestral species aligning more closely with the pattern of divergence than the descendent

species. Instability of additive genetic variation on macroevolutionary timescales has been observed previously (Roff, 2000; Arnold et al., 2008), and in the case of *Stylopoma*, it appears that the punctuation events marking speciation in the genus could be a result of changes in the additive genetic variance of traits. However, the direction of trait divergence suggests that *S. species 1* diverged from *S. species 2* before the structure of its genetic variation shifted. At the very least, this appears to be an example in which evolvabilities themselves evolve (Pigliucci, 2008)

Traits with higher evolvability tend to be more divergent between species (Fig. 2 B). Our measure of evolvability between colonies and trait divergence between species complements the previously observed positive correlation between trait variance in a population and trait divergence between species, termed evolution along genetic lines of least resistance (Schluter, 1996). In fossil species of ostracods (Hunt, 2007), in extant species of fiddler crabs (Hopkins et al., 2016), and in stickleback fish (Schluter, 1996), macroevolutionary change of organism-level traits occurs preferentially in directions of maximum phenotypic variance. In *Stylopoma*, we find that this pattern is also observable for colony-level aggregate traits, though this pattern is largely driven by *S. species 2*, as its major axis of variation aligns more closely with the direction of species divergence than does *S. species 1*.

How changes to evolvability may contribute to directions of speciation is a topic worthy of future investigation, since bryozoans appear to undergo morphological evolution on geologically rapid timescales (Di Martino and Liow, 2022). The evolution of aggregate traits in these animals appears to be controlled by shifting evolvabilities, which can contribute to observable differences in zooid-level morphology, as seen in *Stylopoma*. The key to understanding how modular animals like bryozoans express and maintain such a wide array of morphologies in colonies is the study of how their aggregate traits and underlying evolvabilities evolve.

Table 1: **Trait evolvabilities, with error and null estimates, for *Stylopoma species 1*.** Evolvability values for five trait medians and five trait MADs. ol=orifice length, ow=orifice width, pd=pore density, evol. = evolvability.

	parent-offspring covariance evol.	phenotypic variance-covariance evol.	standard sample error	Null evol.
median length	0.0019	0.0023	±0.000094	−0.000036 ± 0.00042
median width	0.0053	0.0043	±0.00020	−0.000095 ± 0.0018
median ol	0.0023	0.0037	±0.00069	0.0000063 ± 0.000063
median ow	0.00	0.022	±0.00038	0.0000068 ± 0.000072
median pd	0.017	0.016	±0.0029	0.00016 ± 0.00087
MAD length	0.0065	0.072	±0.0058	0.000066 ± 0.0019
MAD width	0.018	0.038	±0.0041	−0.00017 ± 0.0040
MAD ol	0.00	0.076	±0.033	0.00021 ± 0.000055
MAD ow	0.010	0.15	±0.035	0.00057 ± 0.000087
MAD pd	0.00	0.16	±0.050	−0.0058 ± 0.0048

Table 2: **Trait evolvabilities, with error and null estimates, for *Stylopoma species 2*.** Evolvability values for five trait medians and five trait MADs. ol=orifice length, ow=orifice width, pd=pore density, evol. = evolvability.

	parent-offspring covariance evol.	phenotypic variance-covariance evol.	standard sample error	Null evol.
median length	0.0021	0.0034	±0.00012	−0.000055 ± 0.0031
median width	0.0013	0.0031	±0.00019	−0.000025 ± .00052
median ol	0.0019	0.0026	±0.00074	0.0000095 ± 0.0000041
median ow	0.00	0.00093	±0.00028	0.000023 ± 0.0000034
median pd	0.073	0.041	±0.011	−0.00055 ± 0.00032
MAD length	0.00	0.064	±0.0050	−0.00067 ± 0.0014
MAD width	0.00	0.040	±0.0043	−0.00063 ± 0.0021
MAD ol	0.021	0.058	±0.037	−0.00015 ± 0.0000076
MAD ow	0.045	0.064	±0.032	−0.00069 ± 0.00015
MAD pd	0.15	0.20	±0.14	−0.0012 ± 0.000052

Competing Interests Statement

The authors declare that they have no competing interests.

Data Accessibility

All data needed to evaluate the conclusions in the paper are present in the paper and/or Supplementary Materials. Additional data related to this paper may be requested from the authors.

References and Notes

Arnold, S. J., Bürger, R., Hohenlohe, P. A., Ajie, B. C., and Jones, A. G. (2008). Understanding the evolution and stability of the g-matrix. *Evolution*, 62(10):2451–2461.

Bull, J. J. (1987). Evolution of phenotypic variance. *Evolution*, 41(2):303–315.

Cheetham, A. H., Jackson, J. B. C., and Hayek, L.-A. C. (1994). Quantitative genetics of bryozoan phenotypic evolution. ii. analysis of selection and random change in fossil species using reconstructed genetic parameters. *Evolution*, 48(2):360–375.

Cheetham, A. H., Jackson, J. B. C., and Hayek, L.-A. C. (1995). Quantitative genetics of bryozoan phenotypic evolution. iii. phenotypic plasticity and the maintenance of genetic variation. *Evolution*, 49(2):290–296.

Cheetham, A. H., Jackson, J. B. C., and Hayek, L. C. (1993). Quantitative genetics of bryozoan phenotypic evolution. i. rate tests for random change versus selection in differentiation of living species. *Evolution*, 47(5):1526–1538.

Cheverud, J. M. (1988). A comparison of genetic and phenotypic correlations. *Evolution*, 42(5):958–968.

Cheverud, J. M. (1996). Quantitative genetic analysis of cranial morphology in the cotton-top (saguinus oedipus) and saddle-back (s. fuscicollis) tamarins. *Journal of Evolutionary Biology*, 9(1):5–42.

Cheverud, J. M. and Marroig, G. (2007). Research article comparing covariance matrices: random skewers method compared to the common principal components model. *Genetics and Molecular Biology*, 30:461–469.

Di Martino, E. and Liow, L. H. (2022). Changing allometric relationships among fossil and recent populations in two colonial species. *Evolution*.

Doroszuk, A., Wojewodzic, M. W., Gort, G., and Kammenga, J. E. (2008). Rapid divergence of genetic variance-covariance matrix within a natural population. *The American Naturalist*, 171(3):291–304.

Falconer, D. and Mackay, T. (1996). *Introduction to quantitative genetics*, volume 3. Harlow, Essex, UK: Longmans Green.

Fernandez, G. and Miller, J. (1985). Estimation of heritability by parent-offspring regression. *Theoretical and Applied Genetics*, 70(6):650–654.

Garcia-Gonzalez, F., Simmons, L. W., Tomkins, J. L., Kotiaho, J. S., and Evans, J. P. (2012). Comparing evolvabilities: common errors surrounding the calculation and use of coefficients of additive genetic variation. *Evolution: International Journal of Organic Evolution*, 66(8):2341–2349.

Grantham, T. A. (1995). Hierarchical approaches to macroevolution: Recent work on species selection and the" effect hypothesis". *Annual Review of Ecology and Systematics*, 26(1):301–321.

Hansen, T. F. and Houle, D. (2008). Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology*, 21(5):1201–1219.

Hansen, T. F., Pélabon, C., Armbruster, W. S., and Carlson, M. L. (2003). Evolvability and genetic constraint in dalechampia blossoms: components of variance and measures of evolvability. *J Evol Biol*, 16(4):754–66.

Hansen, T. F., Pélabon, C., and Houle, D. (2011). Heritability is not evolvability. *Evolutionary Biology*, 38(3):258.

Hansen, T. F., ÁLvarez-Castro, J. M., Carter, A. J. R., Hermisson, J., and Wagner, G. P. (2006). Evolution of genetic architecture under directional selection. *Evolution*, 60(8):1523–1536.

Hopkins, M. J. (2016). Magnitude versus direction of change and the contribution of macroevolutionary trends to morphological disparity. *Biological Journal of the Linnean Society*, 118(1):116–130.

Hopkins, M. J., Haber, A., and Thurman, C. L. (2016). Constraints on geographic variation in fiddler crabs (ocypodidae: *Uca*) from the western atlantic. *Journal of Evolutionary Biology*, 29(8):1553–1568.

Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, 130(1):195–204.

Hunt, G. (2007). Evolutionary divergence in directions of high phenotypic variance in the ostracode genus *Poseidonamicus*. *Evolution*, 61(7):1560–1576.

Jablonski, D. (1986). Larval ecology and macroevolution in marine invertebrates. *Bulletin of marine science*, 39(2):565–587.

Jablonski, D. and Hunt, G. (2006). Larval ecology, geographic range, and species survivorship in cretaceous mollusks: organismic versus species-level explanations. *The American Naturalist*, 168(4):556–564.

Jackson, J. B. C. and Cheetham, A. H. (1990). Evolutionary significance of morphospecies: A test with cheilostome bryozoa. *Science*, 248(4955):579.

Jackson, J. B. C. and Cheetham, A. H. (1994). Phylogeny reconstruction and the tempo of speciation in cheilostome bryozoa. *Paleobiology*, 20(4):407–423.

Jackson, J. B. C. and Cheetham, A. H. (1999). Tempo and mode of speciation in the sea. *Trends in Ecology and Evolution*, 14(2):72–77.

Kashif, M., Aslam, M., Al-Marshadi, A. H., Jun, C., and Khan, M. I. (2017). Evaluation of modified non-normal process capability index and its bootstrap confidence intervals. *IEEE Access*, 5:12135–12142.

Krzanowski, W. J. (1993). Permutational tests for correlation matrices. *Statistics and Computing*, 3(1):37–44.

Leys, C., Ley, C., Klein, O., Bernard, P., and Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4):764–766.

Lidgard, S. (1990). Growth in encrusting cheilostome bryozoans: Ii. circum-atlantic distribution patterns. *Paleobiology*, 16(3):304–321.

Lidgard, S. and Jeremy, B. C. J. (1989). Growth in encrusting cheilostome bryozoans: I. evolutionary trends. *Paleobiology*, 15(3):255–282.

Lloyd, E. A. and Gould, S. J. (1993). Species selection on variability. *Proceedings of the National Academy of Sciences of the United States of America*, 90(2):595–599.

Love, A. C., Grabowski, M., Houle, D., Liow, L. H., Porto, A., Tsuboi, M., Voje, K. L., and Hunt, G. (2021). Evolvability in the fossil record. *Paleobiology*, pages 1–24.

Lynch, M. and Walsh, B. (1998). *Genetics and analysis of quantitative traits*. Sinauer Associates, Inc., Sunderland, MA.

Maturo Jr, F. (1973). Offspring variation from known maternal stocks of parasmittina nitida (verrill). *Living and Fossil Bryozoa, GP Larwood, Eds.(Academic Press, 1973)*, pages 577–584.

McKinney, F. and Jackson, J. (1991). *Bryozoan Evolution*. University of Chicago Press.

Melo, D., Garcia, G., Hubbe, A., Assis, A. P., and Marroig, G. (2015). Evolqg - an r package for evolutionary quantitative genetics [version 3; referees: 2 approved, 1 approved with reservations]. *F1000Research*, 4(925).

Okamura, B. (1992). Microhabitat variation and patterns of colony growth and feeding in a marine bryozoan. *Ecology*, 73(4):1502–1513.

Okasha, S. (2014). Emergent group traits, reproduction, and levels of selection. *Behavioral and Brain Sciences*, 37(3):268–269.

Pigliucci, M. (2008). Is evolvability evolvable? *Nature Reviews Genetics*, 9(1):75–82.

Rice, S. (2004). *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer.

Roff, D. (2000). The evolution of the g matrix: selection or drift? *Heredity*, 84(2):135–142.

Roff, D. A. (1995). The estimation of genetic correlations from phenotypic correlations: a test of cheverud's conjecture. *Heredity*, 74:481–490.

Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution*, 50(5):1766–1774.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). *NIH Image to ImageJ: 25 years of image analysis*.

Simpson, C., Herrera-Cubilla, A., and Jackson, J. B. C. (2020). How colonial animals evolve. *Science Advances*, 6(2).

Sodini, S. M., Kemper, K. E., Wray, N. R., and Trzaskowski, M. (2018). Comparison of genotypic and phenotypic correlations: Cheverud's conjecture in humans. *Genetics*, 209(3):941–948.

Taylor, P. (2020). *Bryozoan Paleobiology*. Wiley.

Team, R. (2015). Rstudio: integrated development for r. rstudio. *Inc., Boston, MA*, 700.

Wagner, G. P. and Altenberg, L. (1996). Perspective: Complex adaptations and the evolution of evolvability. *Evolution*, 50(3):967–976.

Whitley, E. and Ball, J. (2001). Statistics review 1: Presenting and summarising data. *Critical Care*, 6(1):66.