RH: disparate disparity analyses.

**Disparities in the analysis of morphological disparity**

Thomas Guillerme, Natalie Cooper, Stephen L. Brusatte, Katie E. Davis, Andrew L Jackson, Sylvain Gerber, Anjali Goswami, Kevin Healy, Melanie Hopkins, Marc EH Jones, Graeme T. Lloyd, Joseph E. O’Reilly, Abi Pate, Mark N Puttick, Emiliy Rayfield, Erin E. Saupe, Emma Sherratt, Graham Slater, Vera Weisbecker, Gavin H Thomas and Philip Donoghue  
 *School of Biological Sciences, University of Queensland, St. Lucia, Queensland, Australia.*  
*Department of Life Sciences, Natural History Museum, London, Cromwell Road, London, SW7 5BD, UK.*

*3School of GeoSciences, University of Edinburgh, Grant Institute, Edinburgh EH9 3FE, UK*

4 Department of Biology, University of York, YO10 5DD

5 Department of Zoology, School of Natural Sciences, Trinity College.

6 Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, 57 rue Cuvier CP39, 75005 Paris, France.

7 Ryan Institute, School of Natural Sciences, National University of Ireland Galway, Ireland.

10 *School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK*  
11*School of Earth Sciences, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK*

12 *Milner Centre for Evolution, University of Bath, BA2 7AY*

13 *Department of Earth Sciences, University of Oxford, S Parks Road, Oxford OX1 3AN, UK*

14 *School of Biological Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia*

+ These authors contributed equally to the manuscript.

\***Corresponding author:** *guillert@tcd.ie*

**NOTE TO ALL THE CO-AUTHORS:**

**-The word limit for this is 5000 words so we can’t add any large sections.**

**-There is no reference limit per se but please don’t add too many. I will sort out the references later so don’t correct the formatting of them here.**

**-Add your affiliation in the author list above**. The affiliations count in the word count so please 1) keep it short (max 1 per person!) and 2) try to use one that’s already there. Also add any initials I’ve missed in your names.

- If in doubt, check formatting here: <https://royalsociety.org/journals/authors/author-guidelines/>

Thanks!

**Abstract (200 words max)**

Statistical nalyses of morphological disparity have been used to characterise and investigate the evolution of variation in the anatomy, function, and ecology of organisms since the 1980s. While a diversity of methods have been employed, it is unclear whether they provide equivalent insights. Here we review the most commonly used approaches for characterising and analysing morphological disparity, all of which have associated limitations that, if ignored, can lead to misinterpretation. We provide best practice guidelines for disparity analyses, while noting that there can be no “one-size-fits-all” approach. The available tools should always be used in the context of a specific biological question that will determine data and method selection at every stage of the analysis.

**Keywords**: multidimensionality, palaeobiology, ecology, morphology, disparity, variance/variation

# 1. Introduction

Clades of organisms are characterised by variation in both numbers of species and range of phenotypes through time. At the extremes, clades may be exceptionally species-rich and phenotypically diverse (e.g. cichlids and molluscs), species-rich but phenotypically conservative (e.g. bacteria and nematodes), species-poor but phenotypically diverse (e.g. Afrotheria), or depauperate in both species and phenotypic diversity (e.g. lungfish). These phenomena suggest that taxonomic and phenotypic diversity are not inextricably linked, raising important questions about how phenotypic diversity arises, such as: How does diversity evolve? Are some morphologies more common than others? Can anatomy evolve in all “directions” or are some anatomies impossible to achieve? (Mike Foote 1997, 1995) What role does ecology play in structuring morphological diversity?. Analyses of species diversity have a venerable history, but those of phenotypic diversity (hereafter *disparity*) are a comparatively more recent phenomenon. Originally defined as “multidimensional morphological dissimilarity at a macroevolutionary scale” (Runnegar 1987, [Gould 2000](https://paperpile.com/c/sTGYvp/Uns3)), this concept of disparity emerged from attempts by palaeobiologists to characterise the evolutionary origin of animal bodyplans and from attempts by comparative developmental biologists to provide causal explanations for their emergence. However, analyses of “disparity” have since expanded into comparative biology as a means of capturing the effect of intrinsic and extrinsic causal agents in morphological evolution. Typically, methods to capture disparity are based on multidimensional spaces where each dimension represents an aspect of morphological variation (a trait) and biological observations (taxa) can be placed in this space based on their trait values. Such multidimensional spaces (or morphospaces) can then be used to tackle a diverse array of questions that can be grouped into four main (non-mutually exclusive) classes (Fig. 1):

(1) *Descriptive disparity.* The pioneering studies of disparity characterised the shapes of organisms and how they differed among groups [(Foote 1995; Wills, Briggs, and Fortey 1994)](https://paperpile.com/c/sTGYvp/fTJ3+eZ3F). These studies consist of describing multidimensional patterns in the diversity of morphological traits, addressing questions such as: why are some morphological trait combinations more common than others and what are the biological (or mathematical) properties of the resulting morphospace? [(Foote 1995; Raup 1961; Gerber 2017)](https://paperpile.com/c/sTGYvp/fTJ3+I0Ic+QVvv).

(2) *Disparity-through-time.* This approach investigates how the morphologies of organisms have changed on a temporal scale, focussing on the disparity of taxa in particular time intervals or slices. This approach has been used widely in palaeobiology to answer a range of macroevolutionary questions, such as: how does disparity accumulate over the history of a clade [(Guillerme and Cooper 2018; Wright 2017)](https://paperpile.com/c/sTGYvp/ekU4+s33b), or how does disparity change leading up to and across mass extinction events [(Friedman 2010)](https://paperpile.com/c/sTGYvp/EETc)?

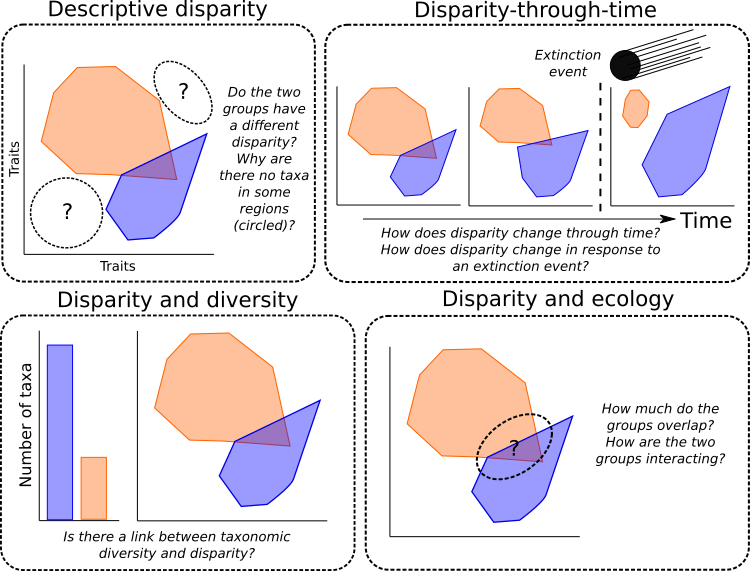
(3) *Disparity and taxonomic diversity.* Morphological disparity provides another perspective on biodiversity; high morphological disparity represents a high diversity of morphologies (i.e. shapes or body plans) and is, presumably, associated with high levels of ecological and functional diversity (though see below). This makes disparity an informative complement to diversity measures based on species richness alone. Indeed, most studies that have investigated disparity and taxonomic diversity support an effective decoupling [(e.g., Fortey, Briggs, and Wills 1996; Ruta et al. 2013; M. J. Hopkins 2013; Moyne and Neige 2007)](https://paperpile.com/c/sTGYvp/2tbJ+geAO+hea5+aVVj). This approach has been used to investigate whether some groups are more successful than others in their exploration of new evolutionary strategies [(Losos 2011)](https://paperpile.com/c/sTGYvp/dJHu).

(4) *Disparity as a proxy for ecology.* The disparity of a group can be used as a proxy for either the functional role it plays within an ecosystem or its ecological niche. This approach assumes that groups with high disparity are also likely to be functionally and ecologically diverse, and that groups found in similar regions of shape space will have similar functional and ecological roles [(Friedman 2010; Pierce, Angielczyk, and Rayfield 2008; P. S. L. Anderson et al. 2011)](https://paperpile.com/c/sTGYvp/EETc+tSIy+qjj9). The links between form and function, however, are not always clear cut. Traits can be linked to multiple functions and multiple functions can be linked to a single trait [(Wainwright et al. 2005)](https://paperpile.com/c/sTGYvp/Ejzr). This approach has been used to investigate hypotheses of competitive replacement [(Stephen L. Brusatte et al. 2008)](https://paperpile.com/c/sTGYvp/EeC8) and changes in ecosystem function during and after mass extinctions [(Friedman 2010)](https://paperpile.com/c/sTGYvp/EETc). It is particularly common in palaeobiology where it is not possible to directly observe the ecological or functional characteristics of extinct species [(Wainwright et al. 2005)](https://paperpile.com/c/sTGYvp/Ejzr).

(5) Disparity and evo-devo. Morphological disparity blalbalbal

Fundamental insights into evolutionary biology have been elicited from these four types of analysis. One of the most important insights is the discovery that morphological disparity is often greatest early in the evolutionary history of clades [(Hughes, Gerber, and Wills 2013; Foote 1997;](https://paperpile.com/c/sTGYvp/xxh5+yqPw) [Erwin 2007)](https://paperpile.com/c/sTGYvp/EPJ2), indicating that the capacity for evolutionary innovation wanes with clade age, which some have argued reflects the evolutionary assembly of gene regulatory networks that constrain later fundamental change [(Hughes, Gerber, and Wills 2013; Erwin 2011)](https://paperpile.com/c/sTGYvp/xxh5+Z6l6). However, this example highlights one of the most challenging problems confronting researchers who are attempting, increasingly, to obtain general insights from multiple independent studies: can the insights gained from studies using the diversity of methods, approaches and data types employed be considered equivalent?

In attempting to answer this question, we review current methods and highlight their limitations, as part of a more general attempt to propose best practice guidelines for studies of disparity . We first discuss the appropriate data required for characterising disparity, then review various challenging aspects of these approaches. Throughout, it is important to remember these tools should always be used in the context of a specific scientific question, as this will drive data and methodological choices at every stage of the process.

**

*Figure disparities [figure\_disparities.pdf]: The four main types of disparity analysis.* ***Descriptive disparity*** *focuses on describing the features of morphospace occupation;* ***disparity-through-time*** *investigates the evolution of the morphospace through time including the effect of extinction events;* ***disparity and taxonomic diversity*** *compares different measures of biodiversity ;* ***disparity as a proxy for ecology*** *uses disparity as a proxy for the ecological or functional role of a group. These categories are not independent and many disparity studies will cover more than one.*

# 2. Data and disparity

Disparity analyses are based on traits, but traits can be characterised in a number of ways: (1) discrete “cladistic” characters, e.g. coding the absence or presence of features or a discrete characteristic of a trait [(Close et al. 2015)](https://paperpile.com/c/sTGYvp/PbSx); (2) continuous measurements of features (e.g., lengths) [(P. S. L. Anderson et al. 2011)](https://paperpile.com/c/sTGYvp/qjj9); or (3) more mathematical descriptors from geometric morphometric landmark data (e.g. Procrustes coordinates) [(Cooney et al. 2017)](https://paperpile.com/c/sTGYvp/RjqE), and Fourier coefficients [(Foote 1989;](https://paperpile.com/c/sTGYvp/2Neu) [Spriggs et al. 2018)](https://paperpile.com/c/sTGYvp/ZEDR)) (Fig. 2). None of these approaches are superior, but they may be more or less well-suited to characterising traits under comparison and to the question being asked of those traits (Hetherington et al 2015; Hopkins 2017).

For example, if investigating variation of bat wing shapes, both homologous landmarks and continuous measurements of bones may be appropriate to capture patterns of wing variation. However, if the question focuses on comparing wings between bats and birds, different measurements might be more appropriate depending on the specific question: for example if the focus is on wing function, i.e. whether the aerodynamic properties of wings vary within bats or between bats and birds, the traits collected should reflect these aerodynamic properties (e.g. wingspan, aspect ratio, etc.). However, if the focus is on convergence between different bats and birds, it would be preferable to use traits that have facilitated flight in both groups (e.g. digit length, integumentary system, etc.). Where there is any doubt about the appropriate traits to choose, it may be preferable to use different kinds of data for the same feature to determine whether they capture the same pattern of disparity.

The points above assume that researchers are collecting their own data for disparity analyses, but often this is not the case. Discrete character data are commonly recycled from phylogenetic studies (e.g.[Foote 1989; Deline et al. 2018](https://paperpile.com/c/sTGYvp/2Neu+0y4V)). This is an efficient approach to character sampling, but it may artifactually increase disparity between phylogenetically distinct groups because phylogenetic characters are often collected to discriminate among groups [(Foote 1995)](https://paperpile.com/c/sTGYvp/fTJ3). This needs to be considered when interpreting results, especially as synapomorphies will naturally lead to an apparent shift or increase in disparity when new clades appear. Furthermore, many datasets are limited to subsets of anatomy that are at least implicit samples of overall anatomy, but explicit tests of this assumption have shown that different aspects of morphology can exhibit different patterns of disparity [(Melanie J. Hopkins 2017)](https://paperpile.com/c/sTGYvp/xLdm). This effect of anatomical part on disparity patterns can be especially challenging when working on datasets where the available data has non-random missing anatomical parts, such as the absence of soft tissue in the fossil record [(Deline et al. 2018)](https://paperpile.com/c/sTGYvp/0y4V).

Trait data suffer from the same shortcomings as most biological datasets – data can be missing, non-overlapping, hierarchical, inapplicable, ambiguous, polymorphic, correlated, or there may be an insufficient sample size [(Brazeau, Guillerme, and Smith 2017](https://paperpile.com/c/sTGYvp/Yrbg); [Palci and Lee 2018)](https://paperpile.com/c/sTGYvp/yO2t). Biological phenomena such as allometry and sexual dimorphism may also influence trait data. More mundanely, data collection is constrained by the time and money available, making collating a “perfect” dataset difficult. Ultimately, disparity analyses characterise the data, and subsamples of the universe of possible data may not have the power to uncover holistic patterns of disparity. Therefore, trait data should be collected with the question in mind, or the question asked should be tailored to the limits of the data available.

3.

*Figure data [figure\_data.pdf]: Major routes to obtain morphological data for disparity analyses. Data can be collected as discrete trait observations (e.g. presence or absence data) or as continuous data. Continuous data can be collected by various methods including linear measurements and landmark coordinates or contours (curves). These measurements can then be mathematically transformed (logarithm transforms, scaling, Procrustes superimposition, elliptic Fourier transforms, etc.). Regardless of the method, data collection produces a trait matrix where the observed traits constitute columns and the studied elements (generally taxa or OTUs) the rows.*

# Disparity methods

Once suitable trait data have been collected, the design of the disparity analysis itself needs to be considered. Study design encompasses several key aspects including (a) the difficulty of dealing with multidimensional data; (b) the variety of ordination (dimension reduction) techniques available and their limitations; (c) the metrics used to summarise the relative disparity of groups; (d) the methods used for hypothesis testing within the disparity analysis framework; and (e) ancestral state estimations in disparity-through-time analyses. We consider these aspects in turn below.

## (a) Multidimensional mayhem

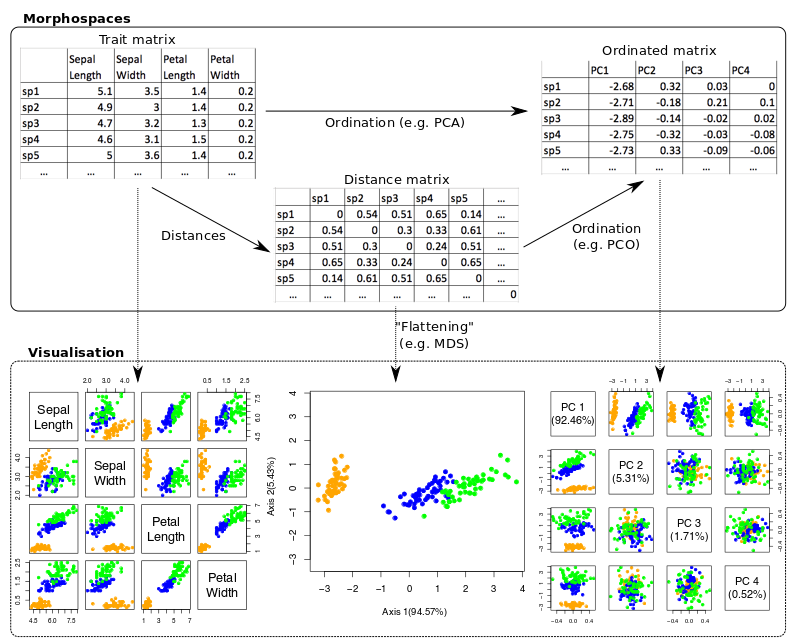
Disparity analyses often use ordination techniques for dimension reduction, making it easier for researchers to comprehend patterns in just two or three spatial dimensions at a time - or, indeed, just the first few axes of a principal components analysis (PCA) or principal coordinates analysis (PCO, PCoA). Unfortunately, dimension reduction can introduce errors of interpretation when these principal axes are unrepresentative of the data more generally [(Bookstein 1997, 2015, [b] 2017, [a] 2017)](https://paperpile.com/c/sTGYvp/1SD2+sN5d+xaUx+o4w7)[CITE Vera and Thomas paper]. Visual interpretations of multidimensional data can be particularly misleading, not least since multidimensional spaces are not necessarily Euclidean even when analysing morphometric data [(Deline et al. 2018; Gerber 2017)](https://paperpile.com/c/sTGYvp/0y4V+QVvv). Categorical data are a good deal more problematic, since the characters themselves are invariably non-equivalent, non-independent, and mostly non-Euclidean. Such non-Euclidean spaces often have non-intuitive properties, for example, straight lines viewed in bivariate plots of some dimensions are not actually straight and, even less intuitively, distances are non-metric [(Gerber 2014)](https://paperpile.com/c/sTGYvp/SJbC).

## (b) To ordinate or not to ordinate - that is the (multidimensional) question

Dimension reduction (or ordination) is a powerful tool in disparity analyses, helping to elucidate the underlying structure within datasets. However, like most other aspects of disparity analyses, reducing dimensionality can be tricky. Ordination techniques come in many flavours depending on the data and the desired morphospace properties. Quantitative (continuous) data can be reduced using PCA, and qualitative data (or mixed data types) can be reduced using PCoA, which is equivalent to metric multidimensional scaling (MDS) or non-metric multidimensional scaling (NMDS) (see Legendre and Legendre 2012,chapter 9, for a detailed explanation of ordination methods and properties).

Ordination is advantageous for plotting and visualising data, and can reveal properties of the morphospace not intuitively captured by disparity metrics (see Metrics section below [metrics]). Additionally, after ordinating the data, it is possible to reduce further the number of dimensions by focussing on just a subset of axes (e.g. selecting only those axes that describe the majority of the variation) or by fitting the variation to a prescribed number of axes. In the case of geometric morphometric data, ordination is particularly useful as it conserves the mathematical properties of the data while efficiently reducing the dimensions [(Legendre and Legendre 2012)](https://paperpile.com/c/sTGYvp/oFiP). This has clear advantages for interpreting the results. For example, the axes will represent gradients of biological variation (e.g. elongation and flattening of the beak in birds [(Cooney et al. 2017](https://paperpile.com/c/sTGYvp/RjqE)). However, there are a number of potential problems with these transformations of the data. In the case of ordination in a geometric morphometric context, subsampling axes from the ordination can lead to misinterpretation of the results. Furthermore, interpreting biological variation along the axes is always a *post-hoc* procedure and may have little relation to the overall question; for example, if the first few ordination axes represent the elongation of the beak in birds, but the question is about wing disparity.

However, In many cases, ordination might not be necessary. For example, if a metric (e.g. sum or ranges of variance) used to characterise disparity can use all of the data (see Metrics section below [metrics]), it is not necessary to calculate it on ordinated data [e.g.[(Close et al. 2015)](https://paperpile.com/c/sTGYvp/PbSx)]. Additionally, in some cases, reducing the dimensionality of a dataset can render its interpretation more problematic. For example, when the analysed data is non-Euclidean (e.g. discrete morphological characters), interpreting the resulting ordinated space can be challenging, even if applying an ordination technique is straightforward, [(Gerber 2014)](https://paperpile.com/c/sTGYvp/SJbC). This is problematic when comparing the position of groups in multidimensional space, as the distances might not be linear. Finally, *post-hoc* interpretations of the gradient of variation on the ordination axes may be biologically meaningless or simply impossible. Although some gradients are easy to detect or interpret (e.g. the elongation and depth of mandibles in fishes on first and second PC axes, respectively; [[Hill et al. 2018]](https://paperpile.com/c/sTGYvp/3JPy)), some are not (e.g. [Weisbecker et al. 2019)](https://paperpile.com/c/sTGYvp/TZzO). For example, with discrete morphological data, a gradient between the species that have many characters in state 1 and the ones that have more in state 0, has no biological meaning if these are binary alternate states. For all of these reasons, multidimensional data analyses should not be ordinated automatically, and careful consideration should be given to whether the aim of the study can be achieved without ordination (Lloyd 2016, 2018).



*Figure methods: Morphospaces: different mathematical representations of a morphospace. A trait matrix can be transformed into a distance matrix (e.g. in Close et al., 2015) or an ordinated matrix (e.g. in Brusatte et al. 2008). In each case, all three matrices represent the morphospace for the observations at hand. Visualisation: different ways to represent the morphospace in 2D. Visualisations can use either trait plots (directly from the trait matrix); “flattening” ordination of the data (e.g. using multidimensional scaling forcing all the variance to be contained in two dimensions); or ordination axis plots (directly from the ordinated matrix).*

## (c) Disparity metrics

Most disparity datasets are multidimensional and, consequently, a large component of any disparity analysis involves considering how to extract a meaningful (i.e. interpretable) summary of disparity. This is usually achieved with a disparity metric or index [(Melanie J. Hopkins and Gerber 2017)](https://paperpile.com/c/sTGYvp/vTHS). As with any summary of multidimensional data, disparity metrics will reflect only some aspects of the morphological variation, never its whole complexity [metrics paper]. It is therefore often beneficial to use more than one metric to summarise different aspects of variation, guided by the aim of the study.

When considering only one dimension, disparity metrics can be used to reflect the spread of the distribution (e.g. the range, quantiles or variance) or its central tendency (i.e., mean, median or mode). Among these metrics, some will have more attractive properties than others, such as sensitivity to outliers. Range and mean are highly sensitive, whereas quantiles, variance and median are less so, making them more or less appropriate for different questions. For example, if we want to characterise the extent of morphospace occupied by a group (e.g. does group A occupy as much space as group B?), metrics related to the spread of the group in the morphospace are most appropriate (e.g. volume [(Díaz et al. 2016)](https://paperpile.com/c/sTGYvp/47fI), distance from the centroid [(Melanie J. Hopkins and Gerber 2017; Finlay and Cooper 2015)](https://paperpile.com/c/sTGYvp/vTHS+yyNa), variance and range [(S. L. Brusatte et al. 2008)](https://paperpile.com/c/sTGYvp/tGyd)). Conversely, if we wish to describe the “position” of a group in a morphospace (e.g. does group A occupy the same space as group B?), metrics related to the distance between the elements within a group and a fixed point in the morphospace are most appropriate (e.g.[(Guillerme et al., n.d.)](https://paperpile.com/c/sTGYvp/u1KE)). Finally, if we aim to characterise the density of morphospace occupation (e.g. is group A more closely packed than group B?) metrics related to the pairwise distances between elements will be most appropriate (e.g. nearest neighbour distance, pairwise distances, etc. [(Close et al. 2015)](https://paperpile.com/c/sTGYvp/PbSx)).

In addition to considering what properties of disparity metrics should capture, it is important to also consider the mathematical properties of the metrics and their associated caveats [(Wills 2001; Ciampaglio, Kemp, and McShea 2001)](https://paperpile.com/c/sTGYvp/nFf7+ROH8). For example, measuring the full sum of the variance in each dimension of the space does not require we add the covariance between the axes in an ordinated space using a PCA. However, this is not true of other mathematical spaces or when not all dimensions or elements are considered, even in a PCA [Legendre & Legendre 2012].

Furthermore, multidimensional spaces also have some counter-intuitive properties that need to be considered such as the “curse of dimensionality” [(Bellman 1966)](https://paperpile.com/c/sTGYvp/Qsl3). In spaces with some axis of variance lower than 1, product-based metrics used as proxies of volumes (e.g. product of ranges, hypervolume, hypercube, etc.) can tend towards zero fairly quickly for spaces with even a modest number of dimensions [(Bellman 1966)](https://paperpile.com/c/sTGYvp/Qsl3); Donoho 2000). Some other types of metrics are also extremely sensitive to outliers and can be biased easily by sample size, for example range [(Butler et al. 2012)](https://paperpile.com/c/sTGYvp/aSSL) or convex hull based metrics [(Butler et al. 2012; Jackson et al. 2011)](https://paperpile.com/c/sTGYvp/aSSL+PwyQ).

## (d) Testing hypotheses on the evolution of disparity

No matter which disparity metrics have been calculated, the research question must be framed in an appropriate statistical context. The multidimensional statistical toolkit for ecology and evolution has been greatly expanded in recent years [(Adams and Collyer 2018)](https://paperpile.com/c/sTGYvp/ZnDd), but some of these advances have yet to be implemented in disparity analyses. Instead, hypothesis testing has mostly been confined to a small set of well-established methods. One commonly used test is the non-parametric permutation analysis of variance [(M. J. Anderson and Walsh 2013; M. J. Anderson 2001)](https://paperpile.com/c/sTGYvp/3hy2+SC6L), an analysis of variance (ANOVA) of the pairwise distances between different groups. Although statistically valid, this test is not always directly related to the hypothesis under question. For example, PERMANOVA tests whether two groups share the same variance/covariance in a “distance-space”. This is not the same as testing whether the two groups overlap in morphospace. Statistical tests should be employed that are tailored to the question at hand, rather than simply following common practices.

It is also important to consider which data should be subjected to a statistical test. For example, in morphological disparity analysis, especially for palaeobiological questions, data are often bootstrapped. This has two advantages: (i) when the disparity metric is unidimensional (e.g. the sum of variances), bootstrapping the data generates a distribution of the metric that can be analysed using the vast statistical toolkit available for comparing distributions; (ii) when data are scarce, bootstrapping the data allows users to introduce variance, rendering the test less sensitive to outliers. However, bootstrapped data are pseudoreplicates and thus non-independent. This violates the assumptions of most parametric statistical tests. Furthermore, the number of bootstrap pseudoreplicates will inevitably increase the Type I error rate. These factors are often ignored in disparity analyses.

## (e) Phylogenetic autocorrelation

As with all comparative datasets, the data used in disparity analyses are not independent because close relatives will tend to have more similar morphologies than more distant relatives [(Harvey and Pagel 1998)](https://paperpile.com/c/sTGYvp/WXik). Thus, for many disparity analyses, phylogenetic relationships should be taken into account. However, it has been noted that some popular phylogenetic correction methods (like pPCA) can be inappropriate if incorrect assumptions are made about the data (Uyeda et al). [(Adams and Collyer 2018)](https://paperpile.com/c/sTGYvp/ZnDd) provide a thorough review of multivariate phylogenetic comparative methods, and so we do not consider them further here.

## (f) Ancestral state estimation in disparity-through-time analyses

Disparity-through-time analyses often use ancestral state estimation to extract disparity estimates for non-sampled taxa and/or nodes of a phylogeny. Ancestral state estimation can be performed at two points in the disparity analysis pipeline: either (1) pre-ordination, i.e. the estimation is done before transformation of the data (e.g. ordination, or distance matrix construction) and is simply based on the original data; or (2) post-ordination, i.e. the estimation is done after transformation of the data by estimating the ancestral states using the transformed matrix (e.g. the ordination scores) [(Lloyd 2018)](https://paperpile.com/c/sTGYvp/53SJ).

Pre-ordination ancestral state estimation will change the way the ordination space is defined – i.e. the relationship between the points are not yet estimated – and requires longer computational times. However, once the morphospace is defined its properties will not change. Post-ordination ancestral state estimation will not change the empirical ordination space and is faster to compute, but it will add elements in the space, whose estimated positions can be problematic for statistical tests and evolutionary inferences down the line [(Lloyd 2018)](https://paperpile.com/c/sTGYvp/53SJ).

No ancestral state estimation method is without drawbacks and above all else are highly dependent on the data and method used. In general, using ancestral state estimation can help with recovering patterns of changes in disparity but should not be used simply to generate extra data points to increase statistical power. In fact, these extra points are not independent and can also have problematic side effects, especially when testing for the influence of mass extinctions on disparity.

# 4. Disparity analyses for the future

Morphological disparity analyses are widely employed in evolutionary palaeobiology, and they are based on a diversity of methods and data. There is no “one-size-fits-all” pipeline for morphological disparity analyses. As with any multidimensional analysis, there are many variables that have to be considered when deciding which data to use and how to analyse it, stemming from the explicit hypotheses being tested. Many of the problems in morphological disparity analysis arise from “blind” application of established methodological pipelines without consideration of the biological question being addressed. We advocate the bespoke assembly of analytic approaches and, given the computational efficiency of these methods, an experimental approach that explores the impact of competing approaches, such as choice of distance measure, ordination method and ancestral state estimation method on disparity analysis results. Many of the methods employed in disparity analysis are used more widely in other fields, including genomics and ecology, which also encompass analyses of multidimensional datasets [(Donohue et al. 2013; Canter et al. 2018; Saupe et al. 2015)](https://paperpile.com/c/sTGYvp/krNU+60H0+cV3v). Innovations in morphological disparity analyses likely await discovery in their respective literatures.

While studies of morphological disparity would benefit from advances in multidimensional analysis in other fields, the concept of a morphospace could reciprocally benefit other disciplines. For example, the multidimensional analysis of Diaz [(Díaz et al. 2016)](https://paperpile.com/c/sTGYvp/47fI), which analysed patterns of form and function in plants, is essentially an ecomorphospace; isotopic analyses of organisms [(Jackson et al. 2011)](https://paperpile.com/c/sTGYvp/PwyQ) can be represented as an isotope-space; ecosystem functioning in Donohue et al. (2013) as an ecosystem-space, etc. These generalisations could also be exported for any set of traits (e.g. acousto-spaces for acoustic traits or glotto-spaces for linguistic traits). Cognate approaches have been adopted recently in the analysis of single cell comparative transcriptome data [(Sebé-Pedrós et al. 2018)](https://paperpile.com/c/sTGYvp/856K) where interpretation of the resulting transcriptome-spaces would be improved by heeding the concerns we highlight concerning morphospaces.

Although disparity analyses are now simple to implement in freely available softwares [(e.g., Guillerme 2018; Adams and Otárola-Castillo 2013; Bouxin 2005; Harmon et al. 2008; Lloyd 2016; Navarro 2003; Dixon 2003)](https://paperpile.com/c/sTGYvp/xDqf+J2G1+9JdS+9Zoi+bCsU+EmTR+2KmX), it is crucial to remember that they are multidimensional analyses and multidimensional analyses are complex. We assert that future morphological analyses will benefit by emphasising the methodological decisions made, rather than simply using disparity analysis because *we can*.

# 5. Author contributions

TG, NC and PD proposed this review; TG and NC led the writing supported by PD and GT. All authors edited drafts and approved the final version.

# 6. Acknowledgments

This article results from discussion at the Royal Society International Science Seminar on ‘Reconciling disparate views on disparity’ held at Chicheley Hall, 9-10th January, 2018.

-AG was funded by European Research Council Starting grant 637171 ADaPTiVE.

-ALJ was funded by an Irish Research Council Laureate Award IRCLA/2017/186. ES was funded by a University of Adelaide Research Fellowship;

-EES was funded by a Leverhulme Trust Research Project Grant (DGR01020).

-PD was funded by NERC (NE/P013678/1; NE/N002067/1) and BBSRC (BB/N000919/1);

-SB was funded by European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 756226, ERC Starting Grant: PalM) and a Leverhulme Trust Research Project Grant (RPG-2017-167);

-TG was funded by ARC DP170103227 and FT180100634 awarded to VW;

Donoho, D. L. (2000). High-dimensional data analysis: The curses and blessings of dimensionality. *AMS math challenges lecture*, *1*(2000), 32.

References

[Adams, Dean C., and Michael L. Collyer. 2018. “Multivariate Phylogenetic Comparative Methods: Evaluations, Comparisons, and Recommendations.” *Systematic Biology* 67 (1): 14–31.](http://paperpile.com/b/sTGYvp/ZnDd)

[Adams, Dean C., and Erik Otárola-Castillo. 2013. “Geomorph: Anrpackage for the Collection and Analysis of Geometric Morphometric Shape Data.” *Methods in Ecology and Evolution / British Ecological Society* 4 (4): 393–99.](http://paperpile.com/b/sTGYvp/J2G1)

[Anderson, Marti J. 2001. “A New Method for Non-Parametric Multivariate Analysis of Variance.” *Austral Ecology* 26 (1): 32–46.](http://paperpile.com/b/sTGYvp/SC6L)

[Anderson, Marti J., and Daniel C. I. Walsh. 2013. “PERMANOVA, ANOSIM, and the Mantel Test in the Face of Heterogeneous Dispersions: What Null Hypothesis Are You Testing?” *Ecological Monographs* 83 (4): 557–74.](http://paperpile.com/b/sTGYvp/3hy2)

[Anderson, Philip S. L., Matt Friedman, Martin D. Brazeau, and Emily J. Rayfield. 2011. “Initial Radiation of Jaws Demonstrated Stability despite Faunal and Environmental Change.” *Nature* 476 (7359): 206–9.](http://paperpile.com/b/sTGYvp/qjj9)

[Bellman, R. 1966. “Dynamic Programming.” *Science* 153 (3731): 34–37.](http://paperpile.com/b/sTGYvp/Qsl3)

[Bookstein, Fred L. 1997. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press.](http://paperpile.com/b/sTGYvp/1SD2)

[———. 2015. “The Relation between Geometric Morphometrics and Functional Morphology, as Explored by Procrustes Interpretation of Individual Shape Measures Pertinent to Function.” *Anatomical Record*  298 (1): 314–27.](http://paperpile.com/b/sTGYvp/sN5d)

[———. 2017a. “A Newly Noticed Formula Enforces Fundamental Limits on Geometric Morphometric Analyses.” *Evolutionary Biology* 44 (4): 522–41.](http://paperpile.com/b/sTGYvp/o4w7)

[———. 2017b. “A Method of Factor Analysis for Shape Coordinates.” *American Journal of Physical Anthropology* 164 (2): 221–45.](http://paperpile.com/b/sTGYvp/xaUx)

[Bouxin, Guy. 2005. “Ginkgo, a Multivariate Analysis Package.” *Journal of Vegetation Science: Official Organ of the International Association for Vegetation Science* 16 (3): 355–59.](http://paperpile.com/b/sTGYvp/9JdS)

[Brazeau, Martin D., Thomas Guillerme, and Martin R. Smith. 2017. “Morphological Phylogenetic Analysis with Inapplicable Data.” https://doi.org/](http://paperpile.com/b/sTGYvp/Yrbg)[10.1101/209775](http://dx.doi.org/10.1101/209775)[.](http://paperpile.com/b/sTGYvp/Yrbg)

[Brusatte, S. L., M. J. Benton, M. Ruta, and G. T. Lloyd. 2008. “The First 50 Myr of Dinosaur Evolution: Macroevolutionary Pattern and Morphological Disparity.” *Biology Letters* 4 (6): 733–36.](http://paperpile.com/b/sTGYvp/tGyd)

[Brusatte, Stephen L., Michael J. Benton, Marcello Ruta, and Graeme T. Lloyd. 2008. “Superiority, Competition, and Opportunism in the Evolutionary Radiation of Dinosaurs.” *Science* 321 (5895): 1485–88.](http://paperpile.com/b/sTGYvp/EeC8)

[Butler, Richard J., Stephen L. Brusatte, Brian Andres, and Roger B. J. Benson. 2012. “How Do Geological Sampling Biases Affect Studies of Morphological Evolution in Deep Time? A Case Study of Pterosaur (Reptilia: Archosauria) Disparity.” *Evolution; International Journal of Organic Evolution* 66 (1): 147–62.](http://paperpile.com/b/sTGYvp/aSSL)

[Canter, Erin J., Catalina Cuellar-Gempeler, Abigail I. Pastore, Thomas E. Miller, and Olivia U. Mason. 2018. “Predator Identity More than Predator Richness Structures Aquatic Microbial Assemblages in Sarracenia Purpurea Leaves.” *Ecology* 99 (3): 652–60.](http://paperpile.com/b/sTGYvp/60H0)

[Ciampaglio, Charles N., Matthieu Kemp, and Daniel W. McShea. 2001. “Detecting Changes in Morphospace Occupation Patterns in the Fossil Record: Characterization and Analysis of Measures of Disparity.” *Paleobiology* 27 (4): 695–715.](http://paperpile.com/b/sTGYvp/ROH8)

[Close, Roger A., Matt Friedman, Graeme T. Lloyd, and Roger B. J. Benson. 2015. “Evidence for a Mid-Jurassic Adaptive Radiation in Mammals.” *Current Biology: CB* 25 (16): 2137–42.](http://paperpile.com/b/sTGYvp/PbSx)

[Cooney, Christopher R., Jen A. Bright, Elliot J. R. Capp, Angela M. Chira, Emma C. Hughes, Christopher J. A. Moody, Lara O. Nouri, Zoë K. Varley, and Gavin H. Thomas. 2017. “Mega-Evolutionary Dynamics of the Adaptive Radiation of Birds.” *Nature* 542 (7641): 344–47.](http://paperpile.com/b/sTGYvp/RjqE)

[Deline, Bradley, Jennifer M. Greenwood, James W. Clark, Mark N. Puttick, Kevin J. Peterson, and Philip C. J. Donoghue. 2018. “Evolution of Metazoan Morphological Disparity.” *Proceedings of the National Academy of Sciences of the United States of America* 115 (38): E8909–18.](http://paperpile.com/b/sTGYvp/0y4V)

[Díaz, Sandra, Jens Kattge, Johannes H. C. Cornelissen, Ian J. Wright, Sandra Lavorel, Stéphane Dray, Björn Reu, et al. 2016. “The Global Spectrum of Plant Form and Function.” *Nature* 529 (7585): 167–71.](http://paperpile.com/b/sTGYvp/47fI)

[Dixon, Philip. 2003. “VEGAN, a Package of R Functions for Community Ecology.” *Journal of Vegetation Science: Official Organ of the International Association for Vegetation Science* 14 (6): 927.](http://paperpile.com/b/sTGYvp/2KmX)

[Donohue, Ian, Owen L. Petchey, José M. Montoya, Andrew L. Jackson, Luke McNally, Mafalda Viana, Kevin Healy, Miguel Lurgi, Nessa E. O’Connor, and Mark C. Emmerson. 2013. “On the Dimensionality of Ecological Stability.” *Ecology Letters* 16 (4): 421–29.](http://paperpile.com/b/sTGYvp/krNU)

[Erwin, Douglas H. 2007. “DISPARITY: MORPHOLOGICAL PATTERN AND DEVELOPMENTAL CONTEXT.” *Palaeontology*. https://doi.org/](http://paperpile.com/b/sTGYvp/EPJ2)[10.1111/j.1475-4983.2006.00614.x](http://dx.doi.org/10.1111/j.1475-4983.2006.00614.x)[.](http://paperpile.com/b/sTGYvp/EPJ2)

[———. 2011. “Evolutionary Uniformitarianism.” *Developmental Biology* 357 (1): 27–34.](http://paperpile.com/b/sTGYvp/Z6l6)

[Finlay, Sive, and Natalie Cooper. 2015. “Morphological Diversity in Tenrecs (Afrosoricida, Tenrecidae): Comparing Tenrec Skull Diversity to Their Closest Relatives.” *PeerJ* 3 (April): e927.](http://paperpile.com/b/sTGYvp/yyNa)

[Foote, Mike. 1989. “Perimeter-Based Fourier Analysis: A New Morphometric Method Applied to the Trilobite Cranidium.” *Journal of Paleontology*. https://doi.org/](http://paperpile.com/b/sTGYvp/2Neu)[10.1017/s0022336000036556](http://dx.doi.org/10.1017/s0022336000036556)[.](http://paperpile.com/b/sTGYvp/2Neu)

[———. 1995. “Morphological Diversification of Paleozoic Crinoids.” *Paleobiology* 21 (03): 273–99.](http://paperpile.com/b/sTGYvp/fTJ3)

[———. 1997. “THE EVOLUTION OF MORPHOLOGICAL DIVERSITY.” *Annual Review of Ecology and Systematics* 28 (1): 129–52.](http://paperpile.com/b/sTGYvp/yqPw)

[Fortey, R. A., D. E. G. Briggs, and M. A. Wills. 1996. “The Cambrian Evolutionary ‘explosion’: Decoupling Cladogenesis from Morphological Disparity.” *Biological Journal of the Linnean Society. Linnean Society of London* 57 (1): 13–33.](http://paperpile.com/b/sTGYvp/2tbJ)

[Friedman, Matt. 2010. “Explosive Morphological Diversification of Spiny-Finned Teleost Fishes in the Aftermath of the End-Cretaceous Extinction.” *Proceedings. Biological Sciences / The Royal Society* 277 (1688): 1675–83.](http://paperpile.com/b/sTGYvp/EETc)

[Gerber, Sylvain. 2014. “Not All Roads Can Be Taken: Development Induces Anisotropic Accessibility in Morphospace.” *Evolution & Development* 16 (6): 373–81.](http://paperpile.com/b/sTGYvp/SJbC)

[———. 2017. “The Geometry of Morphospaces: Lessons from the Classic Raup Shell Coiling Model.” *Biological Reviews of the Cambridge Philosophical Society* 92 (2): 1142–55.](http://paperpile.com/b/sTGYvp/QVvv)

[Gould, Stephen Jay. 2000. *Wonderful Life: The Burgess Shale and the Nature of History*. Random House.](http://paperpile.com/b/sTGYvp/Uns3)

[Guillerme, Thomas. 2018. “dispRity : A Modular R Package for Measuring Disparity.” *Methods in Ecology and Evolution / British Ecological Society* 9 (7): 1755–63.](http://paperpile.com/b/sTGYvp/xDqf)

[Guillerme, Thomas, and Natalie Cooper. 2018. “Time for a Rethink: Time Sub-Sampling Methods in Disparity-through-Time Analyses.” *Palaeontology* 61 (4): 481–93.](http://paperpile.com/b/sTGYvp/ekU4)

[Guillerme, Thomas, Mark N. Puttick, Ariel E. Marcy, and Vera Weisbecker. n.d. “Shifting Spaces: Which Disparity or Dissimilarity Metrics Best Summarise Occupancy in Multidimensional Spaces?” https://doi.org/](http://paperpile.com/b/sTGYvp/u1KE)[10.1101/801571](http://dx.doi.org/10.1101/801571)[.](http://paperpile.com/b/sTGYvp/u1KE)

[Harmon, Luke J., Jason T. Weir, Chad D. Brock, Richard E. Glor, and Wendell Challenger. 2008. “GEIGER: Investigating Evolutionary Radiations.” *Bioinformatics*  24 (1): 129–31.](http://paperpile.com/b/sTGYvp/9Zoi)

[Harvey, Paul H., and Mark D. Pagel. 1998. *The Comparative Method in Evolutionary Biology*. Oxford University Press, USA.](http://paperpile.com/b/sTGYvp/WXik)

[Hill, Jennifer J., Mark N. Puttick, Thomas L. Stubbs, Emily J. Rayfield, and Philip C. J. Donoghue. 2018. “Evolution of Jaw Disparity in Fishes.” *Palaeontology*. https://doi.org/](http://paperpile.com/b/sTGYvp/3JPy)[10.1111/pala.12371](http://dx.doi.org/10.1111/pala.12371)[.](http://paperpile.com/b/sTGYvp/3JPy)

[Hopkins, Melanie J. 2017. “How Well Does a Part Represent the Whole? A Comparison of Cranidial Shape Evolution with Exoskeletal Character Evolution in the Trilobite Family Pterocephaliidae.” *Palaeontology* 60 (3): 309–18.](http://paperpile.com/b/sTGYvp/xLdm)

[Hopkins, Melanie J., and Sylvain Gerber. 2017. “Morphological Disparity.” In *Evolutionary Developmental Biology*, 1–12.](http://paperpile.com/b/sTGYvp/vTHS)

[Hopkins, M. J. 2013. “Decoupling of Taxonomic Diversity and Morphological Disparity during Decline of the Cambrian Trilobite Family Pterocephaliidae.” *Journal of Evolutionary Biology* 26 (8): 1665–76.](http://paperpile.com/b/sTGYvp/hea5)

[Hughes, Martin, Sylvain Gerber, and Matthew Albion Wills. 2013. “Clades Reach Highest Morphological Disparity Early in Their Evolution.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (34): 13875–79.](http://paperpile.com/b/sTGYvp/xxh5)

[Jackson, Andrew L., Richard Inger, Andrew C. Parnell, and Stuart Bearhop. 2011. “Comparing Isotopic Niche Widths among and within Communities: SIBER - Stable Isotope Bayesian Ellipses in R.” *The Journal of Animal Ecology* 80 (3): 595–602.](http://paperpile.com/b/sTGYvp/PwyQ)

[Legendre, P., and Loic F. J. Legendre. 2012. *Numerical Ecology*. Elsevier.](http://paperpile.com/b/sTGYvp/oFiP)

[Lloyd, Graeme T. 2016. “Estimating Morphological Diversity and Tempo with Discrete Character-Taxon Matrices: Implementation, Challenges, Progress, and Future Directions.” *Biological Journal of the Linnean Society. Linnean Society of London* 118 (1): 131–51.](http://paperpile.com/b/sTGYvp/bCsU)

[———. 2018. “Journeys through Discrete-Character Morphospace: Synthesizing Phylogeny, Tempo, and Disparity.” *Palaeontology* 61 (5): 637–45.](http://paperpile.com/b/sTGYvp/53SJ)

[Losos, Jonathan B. 2011. *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*. Univ of California Press.](http://paperpile.com/b/sTGYvp/dJHu)

[Moyne, Sébastien, and Pascal Neige. 2007. “The Space-Time Relationship of Taxonomic Diversity and Morphological Disparity in the Middle Jurassic Ammonite Radiation.” *Palaeogeography, Palaeoclimatology, Palaeoecology* 248 (1-2): 82–95.](http://paperpile.com/b/sTGYvp/aVVj)

[Navarro, Nicolas. 2003. “MDA: A MATLAB-Based Program for Morphospace-Disparity Analysis.” *Computers & Geosciences* 29 (5): 655–64.](http://paperpile.com/b/sTGYvp/EmTR)

[Palci, Alessandro, and Michael S. Y. Lee. 2018. “Geometric Morphometrics, Homology and Cladistics: Review and Recommendations.” *Cladistics: The International Journal of the Willi Hennig Society*. https://doi.org/](http://paperpile.com/b/sTGYvp/yO2t)[10.1111/cla.12340](http://dx.doi.org/10.1111/cla.12340)[.](http://paperpile.com/b/sTGYvp/yO2t)

[Pierce, Stephanie E., Kenneth D. Angielczyk, and Emily J. Rayfield. 2008. “Patterns of Morphospace Occupation and Mechanical Performance in Extant Crocodilian Skulls: A Combined Geometric Morphometric and Finite Element Modeling Approach.” *Journal of Morphology* 269 (7): 840–64.](http://paperpile.com/b/sTGYvp/tSIy)

[Raup, D. M. 1961. “THE GEOMETRY OF COILING IN GASTROPODS.” *Proceedings of the National Academy of Sciences of the United States of America* 47 (4): 602–9.](http://paperpile.com/b/sTGYvp/I0Ic)

[Ruta, Marcello, Kenneth D. Angielczyk, Jörg Fröbisch, and Michael J. Benton. 2013. “Decoupling of Morphological Disparity and Taxic Diversity during the Adaptive Radiation of Anomodont Therapsids.” *Proceedings. Biological Sciences / The Royal Society* 280 (1768): 20131071.](http://paperpile.com/b/sTGYvp/geAO)

[Saupe, Erin E., Huijie Qiao, Jonathan R. Hendricks, Roger W. Portell, Stephen J. Hunter, Jorge Soberón, and Bruce S. Lieberman. 2015. “Niche Breadth and Geographic Range Size as Determinants of Species Survival on Geological Time Scales.” *Global Ecology and Biogeography*. https://doi.org/](http://paperpile.com/b/sTGYvp/cV3v)[10.1111/geb.12333](http://dx.doi.org/10.1111/geb.12333)[.](http://paperpile.com/b/sTGYvp/cV3v)

[Sebé-Pedrós, Arnau, Elad Chomsky, Kevin Pang, David Lara-Astiaso, Federico Gaiti, Zohar Mukamel, Ido Amit, Andreas Hejnol, Bernard M. Degnan, and Amos Tanay. 2018. “Early Metazoan Cell Type Diversity and the Evolution of Multicellular Gene Regulation.” *Nature Ecology & Evolution* 2 (7): 1176–88.](http://paperpile.com/b/sTGYvp/856K)

[Spriggs, Elizabeth L., Samuel B. Schmerler, Erika J. Edwards, and Michael J. Donoghue. 2018. “Leaf Form Evolution in Viburnum Parallels Variation within Individual Plants.” *The American Naturalist* 191 (2): 235–49.](http://paperpile.com/b/sTGYvp/ZEDR)

[Wainwright, Peter C., Michael E. Alfaro, Daniel I. Bolnick, and C. Darrin Hulsey. 2005. “Many-to-One Mapping of Form to Function: A General Principle in Organismal Design?” *Integrative and Comparative Biology* 45 (2): 256–62.](http://paperpile.com/b/sTGYvp/Ejzr)

[Weisbecker, Vera, Thomas Guillerme, Cruise Speck, Emma Sherratt, Hyab Mehari Abraha, Alana C. Sharp, Claire E. Terhune, Simon Collins, Stephen Johnston, and Olga Panagiotopoulou. 2019. “Individual Variation of the Masticatory System Dominates 3D Skull Shape in the Herbivory-Adapted Marsupial Wombats.” *Frontiers in Zoology* 16 (November): 41.](http://paperpile.com/b/sTGYvp/TZzO)

[Wills, Matthew A. 2001. “Morphological Disparity: A Primer.” In *Topics in Geobiology*, 55–144.](http://paperpile.com/b/sTGYvp/nFf7)

[Wills, Matthew A., Derek E. G. Briggs, and Richard A. Fortey. 1994. “Disparity as an Evolutionary Index: A Comparison of Cambrian and Recent Arthropods.” *Paleobiology* 20 (02): 93–130.](http://paperpile.com/b/sTGYvp/eZ3F)

[Wright, David F. 2017. “Phenotypic Innovation and Adaptive Constraints in the Evolutionary Radiation of Palaeozoic Crinoids.” *Scientific Reports* 7 (1): 13745.](http://paperpile.com/b/sTGYvp/s33b)