Version dated: December 15, 2017

RH: disparate views on disparity.

Disparate views on disparity.

Thomas Guillerme[1,*], Natalie Cooper[2], ..., Philip Donoghue[3]

[1]*Imperial College London, Silwood Park Campus, Department of Life Sciences, Buckhurst Road, Ascot SL5 7PY, United Kingdom.*

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

Abstract

blablalbalba

(Keywords: disparity)

Introduction

Biodiversity is not a smooth gradient of forms, functions, ecologies, behaviours, etc., but it is rather variable and discontinuous. This is at the base of biological sciences and is linked to the concept of disparity (Wills, 2001; Hopkins and Gerber, 2017) with the observation that some groups of species are more similar or dissimilar than others. This diversity in morphologies is in fact proper in nature and decoupled from taxonomic diversity: some groups may exhibit low diversity of species by high diversity of shapes (Ruta et al., 2013; Hopkins, 2013, or the other way around). In palaeobiology, this concept, has been studied since the 90s describing (Gould, 1989, 1991; Briggs et al., 1992; Wills et al., 1994; Foote, 1994, 1996; Jernvall et al., 1996; Foote, 1997) and as been extended with modifications to macroevolutionary analysis in general since the 2000s (whether based on continuous traits; e.g. Harmon et al. 2003, 2008; or on geometric morphometrics; e.g. Claude 2008; Zelditch et al. 2012; Adams and Otárola-Castillo 2013; Adams et al. 2017).

Since then, morphological disparity has been used for testing a vast array of biological hypotheses such as: how did body plans evolve through time? (Wesley-Hunt, 2005); why does some morphologies exist and not other? Gerber (2017); what is the tempo and mode of morphological innovations through time? (Hughes et al., 2013); what is the effect of mass extinction on disparity? (Halliday and Goswami, 2016); how does competition between groups affect disparity? (Brusatte et al., 2008); how did morpho-ecological niches get filled through time? (Price et al., 2014); and many more.

This apparent versatility of morphological disparity analysis led to a phenomenal amount of papers based on disparity methods (Fig. 1)! However, "a lot of disparity studies are inductive fishing trips and are not often designed to test a specific hypothesis." Indeed, compared to the vast amount of publication using morphological disparity, only a handful highlight/solve problems and propose hypothesis testing.

What actually *is* disparity?

Prentice et al. (2011) define disparity as: "a term widely (albeit not always consistently) used to describe the range of forms in a group of organisms, or the difference among different body plans". Disparity can describe either the metric (Wills 2001; or index Hopkins and Gerber 2017) or the whole pipeline (Zelditch et al., 2012; Lloyd, 2016). To assess the usage of disparity in different published studies, we collected methodological data from the 500 first Google Scholar results for the key words "morphological disparity" per order of appearance (accessed on the 1st of November 2017). For the 230 relevant papers among the 500 matches, we collected the following methodological data: (1) What was the focal biological group? ; (2) What kind of data was measured (e.g. landmarks, discrete data, etc.)? ; (3) Was data collected on the full organism or not? ; (4) How was the morphospace explicitly defined (e.g. PCA, PCO, MDS, etc.)? ; (5) How was the disparity metric(s) explicitly defined? ; (6) Which statistical test was applied to test the disparity related hypothesis? ; (7) Was phylogeny taken into account or not? We used only the explicit definition of the

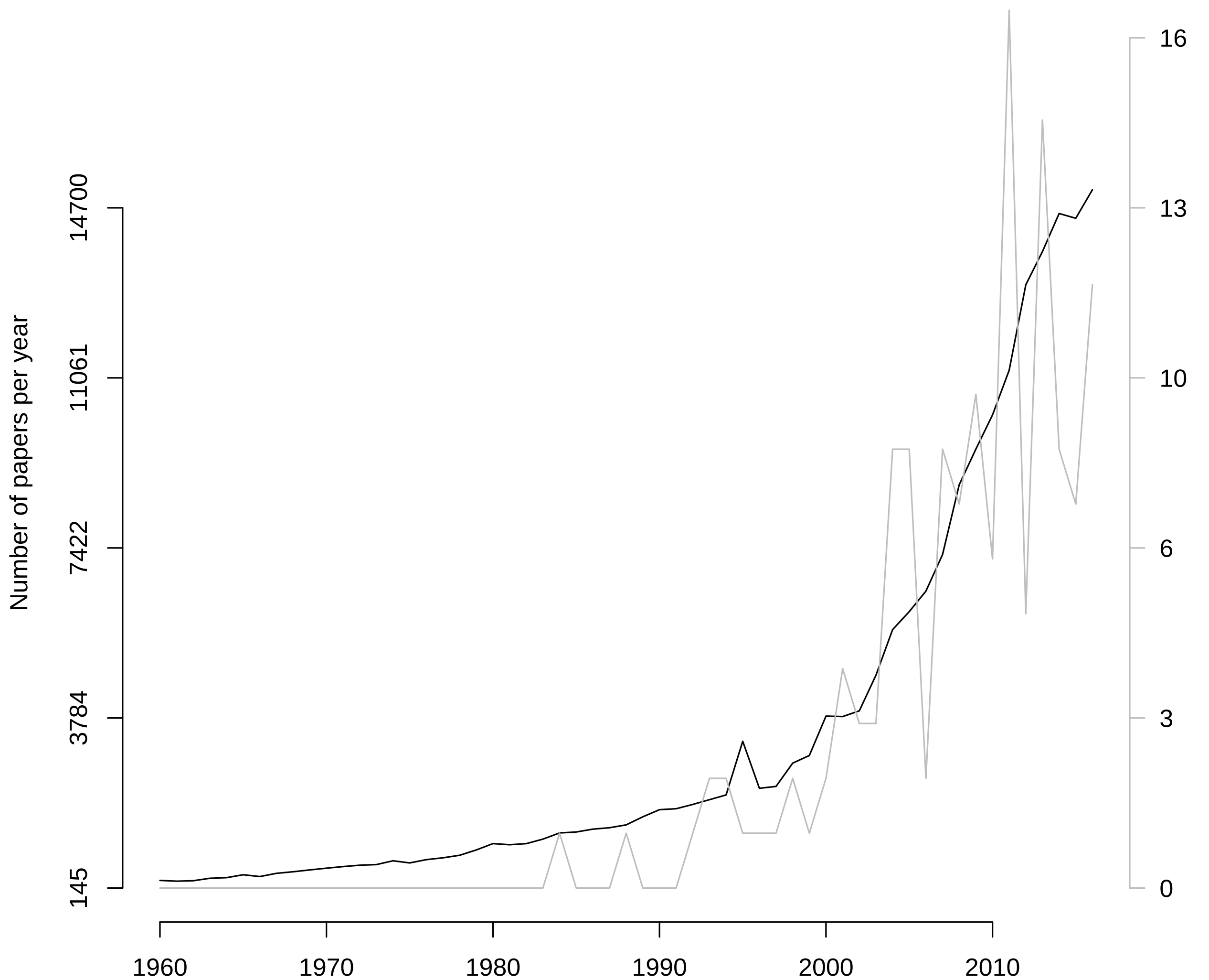

Figure 1: Number of papers on Google Scholar matching the search "morphological disparity" per year. In black, the match is in the paper and in grey, in the title. We collected the number of matches per year from 1960 to 2016 in Google Scholar for the terms "morphological disparity" both in the text (fuzzy matching) or in the title (exact matching). The data was collected on the 1st of November 2017.

morphospace and the disparity metric(s) in this search since a few number of papers had a vague definition of either or both (e.g. a disparity metric was measured but not described anywhere in the paper). The remaining 270 matches were disparity was mentioned but not measured felt in the following categories: papers out of topic, papers mentioning morphological disparity without measuring it, review papers, papers not accessible (either through a broken link or a paywall) or referenced citation without the paper (as a Google Scholar match). To reduce the amount of categories for the 230 recorded methods, we concatenated different methods in a smaller number of categories (see supplementary materials and `https://github.com/TGuillerme/Disparity_Working_Group/blob/master/Analysis/data_cleaning.Rmd`).

From the collected data, we can highlight the use of three main different disparity analysis with their associated data/morphospace/metric/tests and related to specific methodological implementations.

- **The "`Claddis`" approach:** this group of methods uses discrete morphological data (sometimes referred to as "Cladistic" characters) for the full organisms to build a PCO from the organism's pairwise distance as a morphospace. Disparity is then often measured as a variation of the ordinated matrix dimensions' variances or ranges (e.g. the sum of variance or/and the sum of ranges). Hypothesis are often tested using multivariate ANOVAs on the pairwise distance matrix or by simply comparing the confidence intervals overlap of the disparity from different groups.
- **The "`geomorph`" approach:** this group of method is based on landmark data (2D

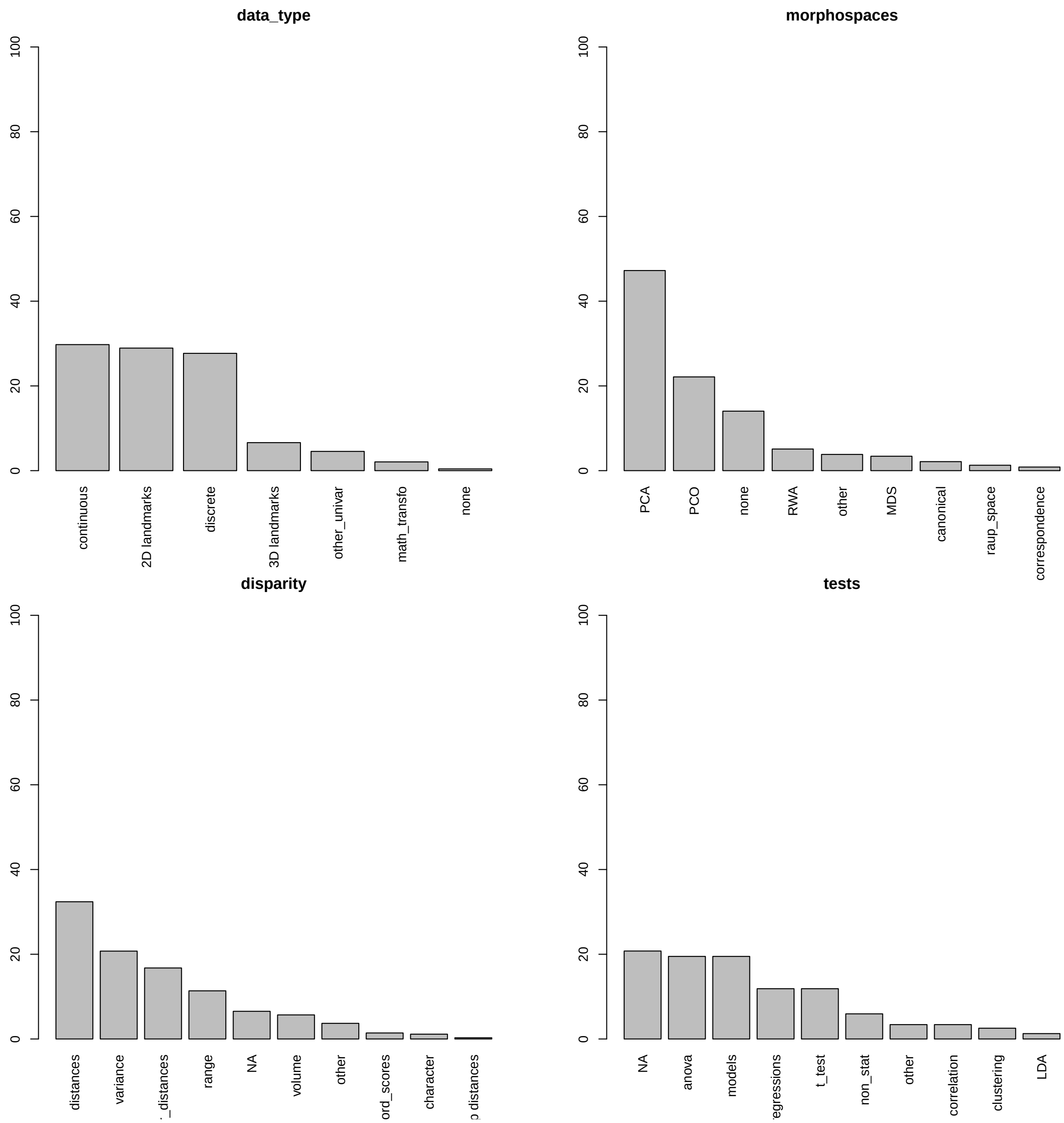

Figure 2: Disparity methods proportional usage: (1) Data type: which input data was used (...); (2) Morphospace: how was the morphospace obtained (...); (3) disparity: what type of disparity metric was calculated (...); (4) test: what type of test was applied (...).

or 3D) on parts of the organism studied usually the skull) and use a Procrustes transformation of the landmarks that are then directly ordinated using a PCA (but sometimes RWA). Disparity is often measured as a distance metric (e.g. the distance between the species and a point in the morphospace such as its centroid). Hypothesis are then tested using ANOVA type tests with usually no phylogenetic correction (although phylogeny is sometimes used to correct the morphospace).

- **The "`approach`" approach:** this method can directly use continuous or discrete data for the full organism without any ordination (but not necessarily), and will measure disparity as the average pairwise distance between species (whether euclidean or any other type of distance). Hypothesis of higher/lower disparity can then be measured using null evolutionary models.

The advent of these three approaches coincides with the explosion of disparity analysis (Fig. 1) and suggest a high popularity of disparity analysis. Unfortunately, however, few actual tests have been performed on whether any of these approaches *actually* allows to tackle the classic disparity questions (see Introduction). Are all these methods the right tools to answer these questions are they mere "inductive fishing trips" used to describe the data rather than testing any of the hypothesis related to the disparity questions?

Does data used in disparity analysis allows to test hypothesis?

In fact, the question of which data to analyse (Hetherington et al., 2015) and which part of the data to use has been analysed (Hopkins, 2017) has rarely been assessed. Can the data used in disparity studies actually test the phenomenon researchers are trying to capture?

For example, when studying discrete morphological characters (often following the "`Claddis`" approach), the data used is often recycled data used for phylogenetic analysis. These datasets are usually design to recover evolutionary history and differentiate clades and, often due to the preservations of the fossil record, are mainly based on hard tissue characters. In phylogenetics, this artefact have been shown to be leading to measurable topological recovery error (for soft *vs* hard characters; Sansom and Wills 2013; or dental *vs* cranial characters Sansom et al. 2017). Therefore, is the morphospace composed of "fossilisable" characters enough for answering evolutionary hypothesis? For example, are niche replacement hypothesis equivalent to fossilisable-character-based-niches replacement hypothesis?

Similarly, many studies in geometric morphometric are based solely on skull shape using landmark techniques combined with Procrustes sumperimposition (the "`geomorph`" approach; Zelditch et al., 2012; Adams et al., 2017). However, it has been shown that the whole skeleton displays different integrated modules with different specific evolutionary constraints leading to variable rates and modes of evolution (Goswami et al. 2014; and this even at the cranial level Goswami and Polly 2010). Is skull variation enough for measuring body plans variation through time for example?

In the case where disparity is measured a single continuous trait (or a collection of traits; following the "`geiger`" approach), for example in Price et al. (2014), beak and limb length as well as body mass are used to differentiate groups of birds and different zones in the morphospace can then be interpreted as different niches. Is this set of traits studied in isolation sufficient to test niche occupancy hypothesis?

Additionally, in many cases, the disparity metrics and hypothesis are often based on a transformation of the data (pairwise distances, PCA, PCO, MDS, RWA, etc.). Can these transformation readily reflect the raw data in our hypothesis? For examples a morphospace as a pairwise matrix reflects the dissimilarity between the data, a morphospace based on an ordination reflects the variability within the data, a morphospace based on both reflects the variability within the dissimilarity, are each method reflecting that?

Do the disparity metrics (or indices) really reflect variation in the studied phenomenon?

From the 230 papers analysed, we found 104 unique combination's of metrics to measure disparity (Fig. 2). Commonly, these metrics are used to approximate the volume occupancy within the morphospace (Wills, 2001; Donohue et al., 2013; Díaz et al., 2016).

However, it has been shown that different metrics reacts differently depending on the dataset

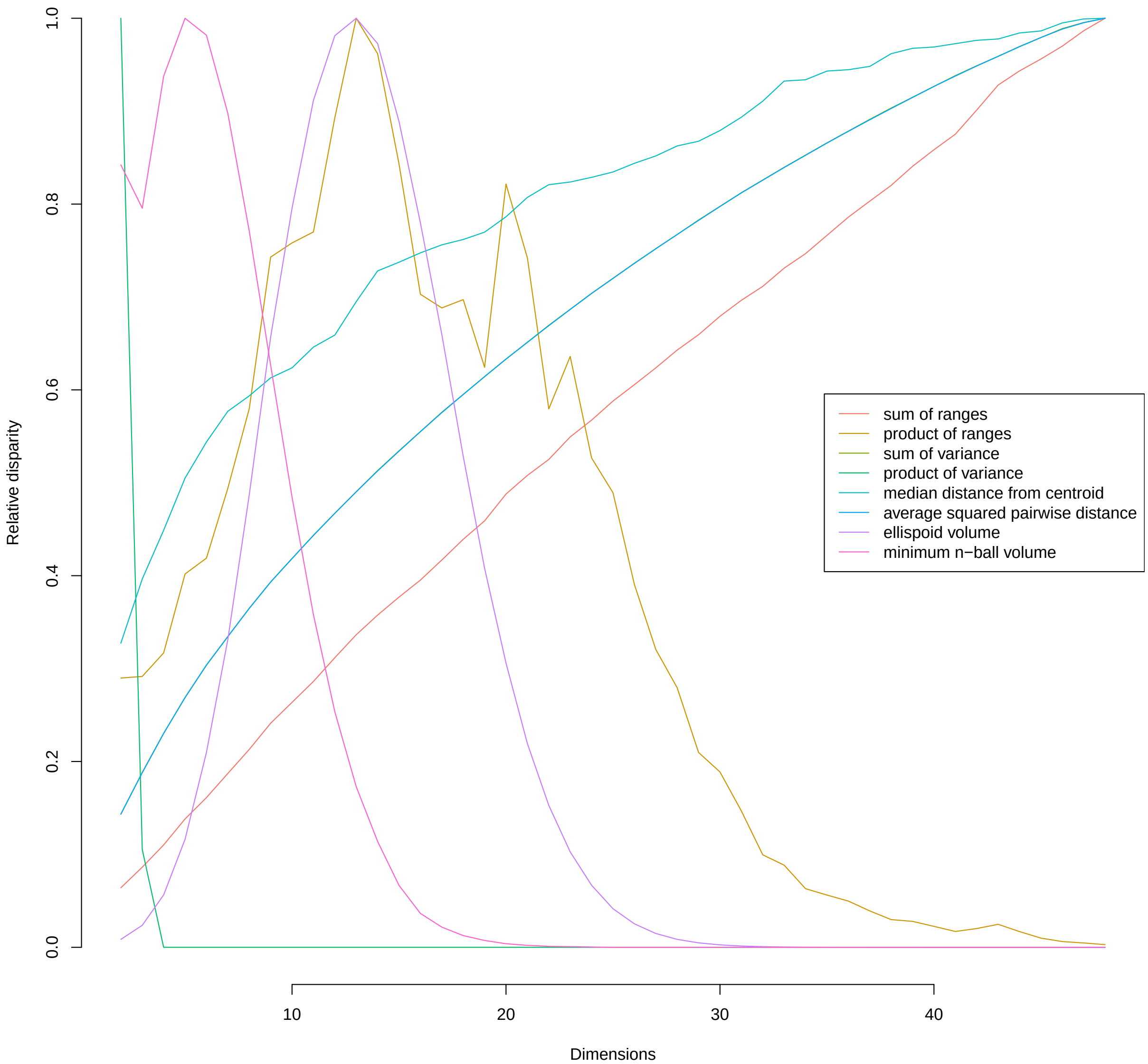

Figure 3: Effect of the number of dimensions on different disparity metrics (note that the sum of variance is equal to the average squared pairwise distance in this case).

Is disparity only relative?

Do the disparity metrics really reflect the variation in the hypothesis that are tested? E.g. Are the variance/range based metrics a description of the ordination or the observed biological phenomenon (i.e. they are not including covariance and phylogenetic signal) E.g. What does the average distance between species or from a point really convey? Is the distribution not more interesting? Do we need more like metrics measuring explicitly what is tested in the hypothesis (rather than approximate it). For example, for competitive replacement, should we not measure whether group A occupies the same place as group B through time and whether that is due to competition or simply because any other place in the space cannot be occupied (and why are such spaces unoccupied)?

Is the current statistical toolkit sufficient for testing disparity hypothesis?

Are the methods used to test hypothesis in disparity actually testing what we think theyâĂŹre testing (problem of multidimensional hypothesis) What does an anova or related variance based test real infirms/confirms regarding the disparity questions? How do we deal with bootstrapped data and pseudo-replications? Can we realistically use univariate gaussian models to test hypothesis based on multivariate, probably non-gaussian data?

What are we missing?

Learning lessons from other fields

In ecology, disparity bears strong parallels with β-diversity in ecology (a measure of ecological communities (dis)similarity): one biological observation described by a vast array of metrics (Baselga, 2010; Anderson et al., 2011; Donohue et al., 2016).

Conclusion

We need a more hypothesis driven way of analysing disparity. Disparity methods should be the tool for answering disparity hypothesis, not merely a description of the data.

Acknowledgments

The Royal Society. TG acknowledge support from European Research Council under the European Union's Seventh Framework Programme (FP/2007 âĂŞ 2013)/ERC Grant Agreement number 311092 awarded to Martin D. Brazeau.

*

References

Adams, D., M. Collyer, A. Kaliontzopoulou, and E. Sherratt. 2017. Geometric morphometric analyses of 2d/3d landmark data .

Adams, D. C. and E. Otárola-Castillo. 2013. geomorph: an r package for the collection and analysis of geometric morphometric shape data. Methods in Ecology and Evolution 4:393–399.

Anderson, M. J., T. O. Crist, J. M. Chase, M. Vellend, B. D. Inouye, A. L. Freestone, N. J. Sanders, H. V. Cornell, L. S. Comita, K. F. Davies, et al. 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology letters 14:19–28.

Baselga, A. 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography 19:134–143.

Briggs, D. E., R. A. Fortey, and M. A. Wills. 1992. Morphological disparity in the cambrian. Science 256:1670–1673.

Brusatte, S. L., M. J. Benton, M. Ruta, and G. T. Lloyd. 2008. Superiority, competition, and opportunism in the evolutionary radiation of dinosaurs. Science 321:1485–1488.

Claude, J. 2008. Morphometrics with R. Springer Science & Business Media.

Díaz, S., J. Kattge, J. H. Cornelissen, I. J. Wright, S. Lavorel, S. Dray, B. Reu, M. Kleyer, C. Wirth, I. C. Prentice, et al. 2016. The global spectrum of plant form and function. Nature 529:167.

Donohue, I., H. Hillebrand, J. M. Montoya, O. L. Petchey, S. L. Pimm, M. S. Fowler, K. Healy, A. L. Jackson, M. Lurgi, D. McClean, et al. 2016. Navigating the complexity of ecological stability. Ecology letters 19:1172–1185.

Donohue, I., O. L. Petchey, J. M. Montoya, A. L. Jackson, L. McNally, M. Viana, K. Healy, M. Lurgi, N. E. O'Connor, and M. C. Emmerson. 2013. On the dimensionality of ecological stability. Ecology Letters 16:421–429.

Foote, M. 1994. Morphological disparity in Ordovician-Devonian Crinoids and the early saturation of morphological space. Paleobiology 20:320–344.

Foote, M. 1996. Ecological controls on the evolutionary recovery of post-Paleozoic Crinoids. Science 274:1492–1495.

Foote, M. 1997. The evolution of morphological diversity. Annual Review of Ecology and Systematics 28:129–152.

Gerber, S. 2017. The geometry of morphospaces: lessons from the classic raup shell coiling model. Biological Reviews 92:1142–1155.

Goswami, A. and P. D. Polly. 2010. The influence of modularity on cranial morphological disparity in carnivora and primates (mammalia). PloS one 5:e9517.

Goswami, A., J. B. Smaers, C. Soligo, and P. D. Polly. 2014. The macroevolutionary consequences of phenotypic integration: from development to deep time. Philosophical Transactions of the Royal Society B: Biological Sciences 369.

Gould, s. J. 1989. Wonderful Life: The Burgess Shale and the Nature of History. W. W. Norton & Co., New York.

Gould, S. J. 1991. The disparity of the burgess shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace. Paleobiology 17:411–423.

Halliday, T. J. D. and A. Goswami. 2016. Eutherian morphological disparity across the end-cretaceous mass extinction. Biological Journal of the Linnean Society 118:152–168.

Harmon, L. J., J. A. Schulte, A. Larson, and J. B. Losos. 2003. Tempo and mode of evolutionary radiation in iguanian lizards. Science 301:961–964.

Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. Geiger: investigating evolutionary radiations. Bioinformatics 24:129–131.

Hetherington, A. J., E. Sherratt, M. Ruta, M. Wilkinson, B. Deline, and P. C. Donoghue. 2015. Do cladistic and morphometric data capture common patterns of morphological disparity? Palaeontology 58:393–399.

Hopkins, M. 2013. Decoupling of taxonomic diversity and morphological disparity during decline of the Cambrian trilobite family Pterocephaliidae. Journal of Evolutionary Biology 26:1665–1676.

Hopkins, M. 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 60:309–318.

Hopkins, M. J. and S. Gerber. 2017. Morphological Disparity Pages 1–12. Springer International Publishing, Cham.

Hughes, M., S. Gerber, and M. A. Wills. 2013. Clades reach highest morphological disparity early in their evolution. Proceedings of the National Academy of Sciences 110:13875–13879.

Jernvall, J., J. P. Hunter, M. Fortelius, et al. 1996. Molar tooth diversity, disparity, and ecology in cenozoic ungulate radiations. Science 274:1489–1492.

Lloyd, G. T. 2016. Estimating morphological diversity and tempo with discrete character-taxon matrices: implementation, challenges, progress, and future directions. Biological Journal of the Linnean Society 118:131–151.

Prentice, K. C., P. Ruta, and M. J. Benton. 2011. Evolution of morphological disparity in pterosaurs. Journal of Systematic Palaeontology 9:337–353.

Price, T. D., D. M. Hooper, C. D. Buchanan, U. S. Johansson, D. T. Tietze, P. Alström, U. Olsson, M. Ghosh-Harihar, F. Ishtiaq, S. K. Gupta, et al. 2014. Niche filling slows the diversification of himalayan songbirds. Nature 509:222–225.

Ruta, M., K. D. Angielczyk, J. Fröbisch, and M. J. Benton. 2013. Decoupling of morphological disparity and taxic diversity during the adaptive radiation of anomodont therapsids. Proceedings of the Royal Society of London B: Biological Sciences 280.

Sansom, R. S. and M. A. Wills. 2013. Fossilization causes organisms to appear erroneously primitive by distorting evolutionary trees. Scientific Reports 3:1–5.

Sansom, R. S., M. A. Wills, and T. Williams. 2017. Dental data perform relatively poorly in reconstructing mammal phylogenies: Morphological partitions evaluated with molecular benchmarks. Systematic biology Page syw116.

Wesley-Hunt, G. D. 2005. The morphological diversification of carnivores in North America. Paleobiology 31:35–55.

Wills, M. 2001. Morphological disparity: A primer. Pages 55–144 *in* Fossils, Phylogeny, and Form (J. M. Adrain, G. D. Edgecombe, and B. S. Lieberman, eds.) vol. 19 of *Topics in Geobiology*. Springer US.

Wills, M. A., D. E. G. Briggs, and R. A. Fortey. 1994. Disparity as an evolutionary index: A comparison of Cambrian and recent arthropods. Paleobiology 20:93–130.

Zelditch, M. L., D. L. Swiderski, and H. D. Sheets. 2012. Geometric morphometrics for biologists: a primer. Academic Press: Waltham.