Morphometrics and evolution: the challenge of crossing rugged phenotypic landscapes with straight paths

P.D. Polly

Departments of Earth and Atmospheric Sciences, Biology, and Anthropology, Indiana University, Bloomington, IN 47405 USA

Geometric morphometrics is widely used to study underlying causal factors in phenotypic evolution and to reconstruct evolutionary history of phenotypes. However, non-linearities in the phenotypic landscape may exist such that analytical solutions derived from comparison of phenotypes in morphospace may have complex or contradictory relationships in the space of the underlying factors. Ancestral reconstruction of horn morphology based on two mammalian ungulates illustrates how biologically improbable results can arise from the mathematical properties of geometric morphometric morphospaces. Raup's shell coiling equations are used to illustrate the potential for contradictory conclusions to be drawn from ancestral reconstructions in parameter spaces (such as measurements of levels of gene expression or allele frequencies) versus shape spaces (such as morphospaces based on phenotypic analysis). These examples are generalizable to many real morphometric studies, suggesting that care should be taken when drawing conclusions about genetic, developmental, or environmental processes based on morphometric analyses. Dense sampling of shape space and the use of fully multivariate and, perhaps, nonlinear methods can help forestall potential problems.

Keywords: geometric morphometrics; mollusc shells; phenotypic evolution; phylogenetic comparative methods; semilandmarks.

Морфометрия и эволюция: проблема пересечения крутых фенотипических ландшафтов прямыми путями

П.Д. Полли

Отделения наук о Земле и атмосфере, биологии и антропологии, Индианский университет, Блумингтон, Индиана, 47405 США

Для изучения основополагающих факторов фенотипической эволюции и для реконструкции эволюционной истории фенотипов широко применяются методы геометрической морфометрии. Однако фенотипические ландшафты могут быть нелинейными настолько, что аналитические решения, полученные путем сравнения фенотипов в морфопространстве, будут иметь сложные или даже противоречивые взаимоотношения в пространстве факторов, определяющих эти фенотипы. Иллюстрацией того, как на основании математических свойств геометрических морфопространств получаются совершенно невероятные с точки зрения биологии результаты, служит реконструкция родословной морфологии рогов копытных млекопитающих. На модели Раупа, описывающей спиральность раковин, показано, что результаты реконструкции предковых форм в параметрических пространствах (таких как уровни экспрессии генов или частоты встречаемости аллелей) могут войти в противоречие с результатами реконструкции в пространствах форм (таких как фенотипические морфопространства). Приведенные примеры в полной мере относятся к морфометрическим исследованиям ныне существующих живых объектов, а значит, формулируя выводы о генетических, онтогенетических или экологических процессах на основании данных морфометрического анализа, надо соблюдать определенную осторожность. Плотное покрытие пространства форм и использование полностью многомерных и, возможно, нелинейных методов могут помочь предотвратить потенциальные проблемы.

Ключевые слова: геометрическая морфометрия; раковины моллюсков; эволюция фенотипов; сравнительные методы филогенетики; полуметки.

Belyaev's experiment with fox domestication exposed a remarkable phenotypic transformation. By selecting foxes for "tame" dog-like behavior, dog-like phenotypes emerged: piebald color patterns, drooping ears, and curled wagging tails (Belyaev, 1979; Belyaev et al., 1984). The existence of parallel traits in the two species are what Vavilov (1922) called "homologous series" in variation, shared genetic components in otherwise different genetic backgrounds. The linked transformations in these traits is an example of pleiotropy (Trut et al., 2013), the effect of one or a few genes on several disparate traits, and this particular set of connections extends so deeply in mammal phylogeny that parallel changes observed in other domestic species have been referred to as the "domestication syndrome" (Wilkins et al., 2014; Sánchez-Villagra et al., 2016). Pleiotropy and homologous series, along with other kinds of genetic and developmental factors, provide the biological underpinnings for what is sometimes called the phenotypic landscape, which describes the complex and frequently non-linear relationship between phenotypes and their underlying biological and environmental factors. The complexity of these underlying links have important implications for morphometrics because they may channel biological change along paths that are not the shortest or most direct from a purely morphometric perspective (Arnold et al., 2001; Klingenberg, Leamy, 2001; Hansen, 2006, 2008; Polly, 2008).

This paper explores the tension between geometric morphometric and biological models of phenotypic transformations. Morphometrics provide the analytical apparatus for measuring, ordinating, and comparing phenotypes – especially phenotypic shape – but its mechanics are purely mathematical. Evolution, development, and genetics are the processes by which one phenotype is transformed into another, and their mechanics are biological. When mathematical and biological paths between phenotypes differ, the phenotypic transformations predicted by morphometrics may contradict those predicted from biological processes. I first review the mathematical properties of geometric morphometric spaces to show how paths through them are defined. Then I review the concept of the phenotypic landscape and how a linear transformation in underlying genetic or developmental factors can produce a highly non-linear change in phenotypes (or vice versa). Then, using two hypothetical examples, I show how tracing paths through geometric morphometric morphospace – paths of the sort associated with ancestral shape reconstruction or other evolutionary modeling exercises – can produce results contradictory to the underlying biology of the change. The possibility of such contradictions has implications for how we use geometric morphometrics to study phylogeny or underlying genetic and developmental factors, so the paper ends with a discussion of practices that will help avoid misinterpretations.

Geometric morphometrics and morphospaces

Arguably the central concept of geometric morphometrics is the morphospace, a multivariate hypervolume in which each point represents distinct configuration of landmarks and which is scaled so that distances between objects equal the distances between their original shapes (Bookstein, 1991; Rohlf, 1993; Mitteroecker, Hutteger, 2009).

Fig. 1 illustrates the mathematical properties of a typical morphospace. To produce such a space, three dimensional

landmarks are applied to the object being analysed (six shells in this case). The landmarks are then registered to a common coordinate system using Procrustes superimposition, which rotates, translates, and rescales the points to minimize the sum of squared distances between the shapes (Rohlf, Slice, 1990). Principal component axes (PCs) are extracted using singular value decomposition of the covariance matrix of the superimposed coordinates (Dryden, Mardia, 1998). The PC axes, which are the eigenvectors of the covariance matrix, form the orthogonal axes of the multivariate morphospace. Each dimension of the space has an axis that represents a linear combination of landmark coordinates that covary with one another as the axis is traced from its negative to positive end. There is one axis for each degree of freedom in the data set, which is $2k-4$ for two dimensional landmarks or $3k-7$ for three dimensional landmarks, where k is the number of landmarks (Dryden, Mardia, 1998; Zelditch et al., 2004).

Each point in the morphospace corresponds to a unique configuration of landmarks (a unique shape). The addresses of objects on the axes are known as scores, which are simply the Cartesian coordinates of the position of a point along the several PCs. The scores on PC 1 for *Dalmanella* are illustrated in Fig. 1. Each axis of the space represents a spectrum of shapes that are described by a single linear combination of the landmark coordinates. The shape variation associated with the first two PC axes are illustrated with a series of shape models. These models were constructed along a line in which the score on all PCs except the one of interest are zero (scores of each shape model are reported underneath it). A position that is not precisely on the axis has a shape that corresponds to the additive combination of models on all of the PCs. For example, *Dalmanella*, which has scores 0.42 on PC 1 and 0.20 on PC 2 has a shape that is sum of the relevant shape models on PCs 1 and 2. The distance between two objects in the morphospace can be calculated as the Euclidean distance between their scores, as shown in Fig. 1.

Presuming that the morphospace was constructed from a covariance matrix as described above, that distance is identical to the so-called Procrustes distance, which is the summed distance between corresponding landmarks in the superimposed shapes. Thus, the scores completely describe the shape of each object and can be substituted for the Procrustes coordinates in subsequent statistical analyses (Dryden, Mardia, 1998). Because the scores on each axis are uncorrelated with scores on other axes and because the number of dimensions is equal to the degrees of freedom in the data set, the scores are more appropriate for many statistical analyses than the original Procrustes coordinates (Rohlf, 1993). The PC scores are therefore frequently referred to as shape variables.

Geometric morphometric morphospaces have two properties that are important for this discussion. First, and arguably most important, is that a vector in morphospace represents the shortest geometric transformation between one shape and another (Klingenberg, Leamy, 2001; Klingenberg, Monteiro, 2005). Second, any mathematical operation involving vectors or non-linear trajectories of change can be performed in the morphospace and converted into their corresponding shapes. Consequently, biologically interesting analyses such as predicting the effects of a selection vector on shape (Klingenberg, Leamy, 2001; Myers et al., 2006; Klingenberg et al., 2010;
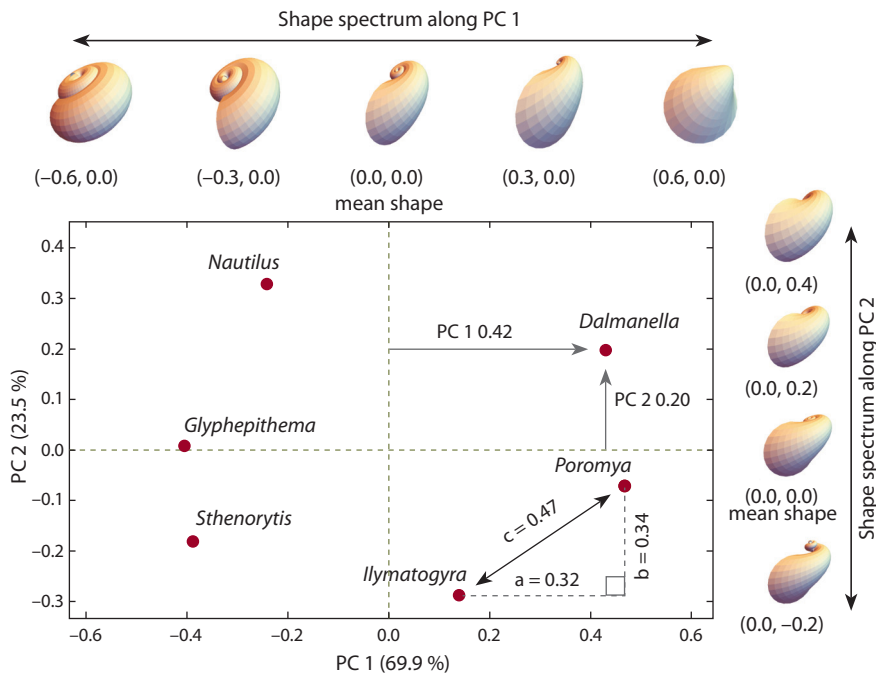
Fig. 1. The mathematics of a geometric morphometric morphospace. This morphospace was created with a principal components ordination of six shells, the three-dimensional landmarks of which were first Procrustes superimposed. See text for details.

Martínez-Abadías et al., 2011), comparing ontogenetic trajectories (Zelditch et al., 1992; Viðarsdóttir et al., 2002; Mitteroecker et al., 2005; Klingenberg, 2016), reconstructing an ancestral shape based on the topology of a phylogenetic tree (Rohlf, 2002; Klingenberg, Gidaszewski, 2010; Gómez-Robles et al., 2013; Mounier, Lahr, 2016), or simulating phenotypic evolution using a Brownian motion or Ornstein-Uhlenbeck model (Polly, 2004; Clavel et al., 2015; Polly et al., 2016) can, in principle, be accomplished effectively in a geometric morphometric morphospace by extrapolating vector paths or other more complex transects through it.

However, the fact that the shortest mathematical distance between two shapes is not necessarily the biologically simplest way to transform one phenotype into another has potentially profound consequences for how realistic the results of these operations will be (Bookstein, 2016).

Phenotypic landscapes and morphometrics
Phenotypic landscapes are representations of the genotype-phenotype map that describe transformations between phenotypes based on their relationship with underlying genetic, developmental, and environmental factors (Nijhout, 2002, 2007, 2008; Rice, 2002, 2004; Wolf, 2002). Phenotypic landscapes are not adaptive landscapes, although the two have clear relationships to one another. An adaptive landscape describes fitness as a function of underlying genetic factors or phenotypic traits (Fig. 2, *a*, *b*) (Wright, 1932; Simpson, 1944; Lande, 1976; Arnold et al., 2001). One could also imagine developmental or environmental factors on the axes, but the key identifying feature of an adaptive landscape is the height of the surface represents fitness. In contrast, the height of a phenotypic landscape's surface represents a trait's value as a function of the underlying factors that influence it (Fig. 2, *c*). Phenotypic landscapes provide the maps between the underlying factors and complex, multi-trait phenotypes, providing a conceptual bridge between morphometrics and the biological factors that produce phenotypic transformations (Polly, 2008).

A frequent goal of morphometric analysis is to deduce underlying biological factors or to test hypotheses about the effects of such factors on phenotypic change, whether they be genetic, developmental, or evolutionary (Bookstein, 1991; Klingenberg, 2002; Adams et al., 2004, 2013; Zelditch et al., 2004; Polly, Motz,

2017). For such studies the phenotypic landscape matters. If the relationship with the underlying factors is simple and linear, then a straight path in the phenotypic landscape would produce a straight path in morphospace and vice versa (Fig. 2, *c*). Even if a path is curved in one, it will be correspondingly curved in the other. This relationship is ideal for inferring underlying factors from morphometric analysis. But if the relationship between phenotypic traits and their underlying factors is complex and non-linear, then a straight path in a phenotypic landscape may produce a curved, convoluted, or broken path in the associated multivariate morphospace (Fig. 2, *d*). Inferring underlying factors from morphometric data would be challenging if the phenotypic landscape is complex. Furthermore, modeling phenotypic change in the two spaces could produce contradictory results, as discussed below.

Paths that go astray in morphospace: two hypothetical examples
The potential for paths through morphospace to defy biological principles is illustrated here with two examples of ancestral shape reconstruction.

In geometric morphometrics, the standard method for analyzing phylogenetic patterns of shape evolution is to project a phylogenetic tree into morphospace. Using the phylogenetic generalized least squares (PGLS) method (Martins, Hansen, 1997) or one of its equivalents and a Brownian motion model of evolution, the node shapes are the mean of the tip shapes weighted their branch distances from the node and the branches themselves form straight line between ancestor and descendant (Rohlf, 2002; Klingenberg, Gidaszewski, 2010). The branches form straight lines because, under Brownian motion, the expected outcome of random evolutionary change in a single lineage is equal to the starting trait value and the variance of the outcomes increases linearly with time (Felsenstein, 1973, 1988) (but see discussion of high dimensional random walks below). When applied to the problem of reconstructing node values and branch paths, this statistical expectation means that the most likely path between ancestor and descendant values is the shortest straight line between them.
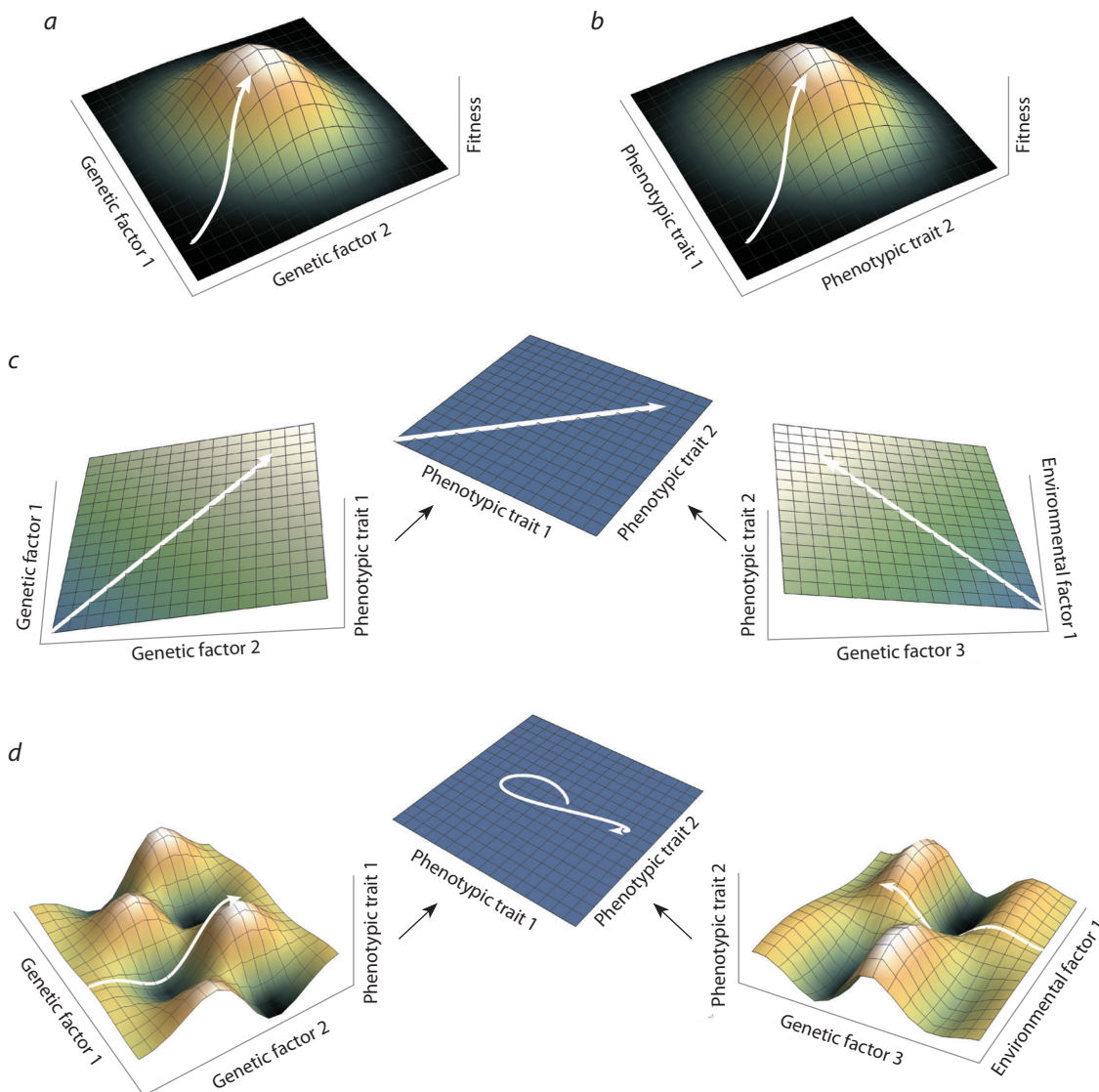
Fig. 2. Comparison of adaptive and phenotypic landscapes.

a, An adaptive landscape showing the relationship between two genetic factors and fitness. The white line traces a direct path to maximum fitness. *b*, An adaptive landscape showing relationship between two phenotypic traits and fitness. The white line traces a direct path to maximum fitness. *c*, A simple phenotypic landscape in which one phenotypic trait is linearly controlled by two genetic factors and a second trait is linearly controlled by a genetic and environmental factor. A linear path through these two landscapes produces a linear path in the associated morphospace. *d*, A complex, bivariate phenotypic landscape in which one phenotypic trait has a non-linear underlying relationship with two genetic factors and a second phenotypic trait has an underlying relationship with genetic and environmental factors. Straight paths across the two phenotypic trait landscapes produce a highly curved path in morphospace.

Antilocaprid horns

In the first example, an ancestral horn configuration is estimated from the profiles of two pronghorns (Fig. 3). *Hexobelomeryx* is an extinct antilocaprid from the Pliocene of North America that had a thick three-pointed horn just above the orbit. "*Nasomeryx*" is a fictional but biologically plausible animal with a forked horn at the tip of its snout, reminiscent of the nasal horn of the Miocene protoceratid *Synthetoceras*. The profile shapes of these animals, which are identical except for the horns, have been represented by outline curves composed of semilandmarks (landmarks that are spaced along a curve or surface rather than placed individually at homologous points). Semilandmark methods (Bookstein, 1997; MacLeod, 1999; Gunz et al., 2005; Gunz, Mitteroecker, 2013) are fre-

quently used to capture the topology of complex edges or surfaces, especially when surface features like the number of horns differ between taxa (Klingenberg, 2008; Polly, 2008; Gunz, Mitteroecker, 2013). Ancestral node reconstruction on this simple two-branch tree yields the unlikely hypothetical profile at the bottom, "*Yeahrightomeryx*". This animal, instead of having a single multipronged horn, has a large cone-shaped horn in the parietal region and a two-pronged horn at mid-snout.

This reconstruction is biologically improbable because of the cone-shape of the parietal horn, the splitting of a single horn into two, and the sliding of horns between one position and the other down the length of the forehead. One would not expect a broad cap-shaped horn structure, or a horn with a
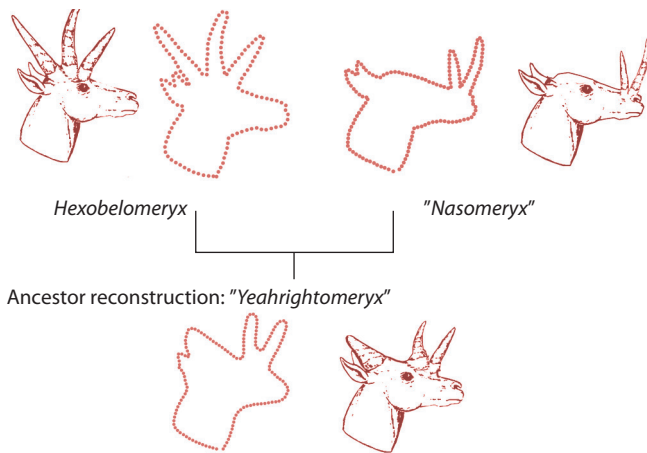
Fig. 3. Ancestral shape reconstruction based on the profile outline of
the Pliocene-aged North American antilocaprid *Hexobelomeryx* with
three horns above the eyes and a fictional but plausible antilocaprid
"*Nasomeryx*" with two horns on the snout. Ancestral reconstruction of
the profile produces the biologically implausible "*Yeahrightomeryx*" with
a broad cone-shaped horn on the back of the head and two horns in an
intermediate position high on the snout. (*Hexobelomeryx* reconstruction
by Barbara Webb, used with permission of the Florida Paleontological
Society, Inc.).

single base and multiple tips to split at the base into two horns,
or the change in horn position from forehead to nose to occur
by incremental movement from one location to the other.

Mathematically, however, the "*Yeahrightomeryx*" recon-
struction is completely predictable. In transforming a head
with three points into one with two, the shortest mathematical
path is simply to sink one of the horns into the domed parietal
area of the skull and to slide the other two points halfway from
one position to the other.

Admittedly this example is contrived. One can rightfully
argue that representing a profile as a simple outline is prob-
lematic, that other methods of semilandmark placement would
make a difference (such as sliding semilandmarks *ala* Gunz
et al., 2005), or that one should not use semilandmarks at all
and, instead, replace them with ordinary landmarks on clearly
homologous locations.

Regardless of these niceties, the principles illustrated in
this example hold for all morphometric studies. The points in
morphospace occupied by real objects represent real biological
shapes, but the moment that one interpolates into the empty
morphospace between objects the transformation is governed
by mathematical rules rather than biological ones. While this
fact is not problematic for many applications, attention should
be paid to whether lines in a mathematical morphospace map
cleanly onto the paths of biological processes if one's ques-
tion is about the relationship between phenotypic shape and
the biological factors that govern it. Biologically implausible
branch paths will be most likely when structures shift posi-
tions (heterotopy; Zelditch, Fink, 1996) or are gained and lost
completely, like the horns in this example.

A phenotypic landscape for shell coiling
In the second example, the consequences of non-linearities in
the phenotypic landscape for ancestor reconstruction is illus-

trated with a sample of shells generated using Raup's (1961,
1966) shell coiling equations. The basic structure of mollusc
and brachiopod shells is the logarithmic spiral that results
from growth of the animal, which secretes ever-enlarging
increments of shell as it gets bigger. The form of the spiral
can be modified with three parameters – W, T, and D – to
describe almost the full range of shell shapes found in nature.
Each shell form is generated by rotating a starting shape that
represents the outline of the aperture around a coiling axis.
W is the rate of whorl expansion with each complete revolu-
tion around the axis and ranges from 1 to infinity, T is the rate
of translation of the starting shape along the coiling axis and
can be any real number, and D is the proportional distance
of the starting shape from the coiling axis, ranging from 0
to 1. The differences in scaling mean that these parameters
define an affine trait space, which presents known com-
plexities for estimating trajectories (Mitteroecker, Hutteger,
2009). In cylindrical coordinates, the surface of the shell is
described by

$$r_\theta = r_0 W^{\theta/2\pi},$$
$$y_\theta = y_0 W^{\theta/2\pi} + r_c T(W^{\theta/2\pi} - 1),$$

where θ is the angle around the coiling axis, r is the distance
from the axis, y is the position along the axis, and r_0 and y_0
are the starting values of r and y on the corresponding point
on the generating shape (Raup, 1966).

While the shell coiling parameters and generating equations
are purely mathematical, they are analogous to developmental
genes and the processes that translate gene expression into
phenotypes. The logarithmic form of the generating equations
and the variation in scaling of the generating parameters create
a non-linear relationship between generating parameters and
shell shape like the ones illustrated in Fig. 2, *d*.

Ancestor reconstructions performed in parameter space
are markedly different from those performed in geometric
morphometric morphospace. Shell shapes were generated
that intentionally resemble six taxa, four living (*Nautilus*,
Poromya, *Glyphepithema*, and *Sthenorytis*) and two extinct
(*Dalmanella* and *Ilymatogyra*). They are shown along with
the parameters used to generate them and the phylogeny of
these taxa in Fig. 4, *a* (the same shells were used to construct
the morphospace of Fig. 1).

PGLS was first used to reconstruct ancestral values for
the generating parameters. These reconstructions are shown
in two ways, first in parameter space (Fig. 4, *b*) and then in
morphospace based on the corresponding outputs of the coiling
equations (Fig. 4, *c*). The nodes are all intermediate between
the tip taxa in parameter space, as is expected from a Brown-
ian motion model of evolution, but almost all the nodes fall
outside the range of tip shapes when the corresponding shell
shapes are projected into the morphospace. This reconstruc-
tion, which implies parallel acquisition of the tight coiling of
the gastropods and cephalopod on the left side of the mor-
phospace, would be appropriate if evolution has occurred by
Brownian change in the underlying parameters.

PGLS was then used to reconstruct ancestral node values
directly from the shell shapes (Fig. 4, *d*). These reconstruc-
tions all fall within the range of the tip shapes and there is
very little parallelism implied by the branch trajectories. The
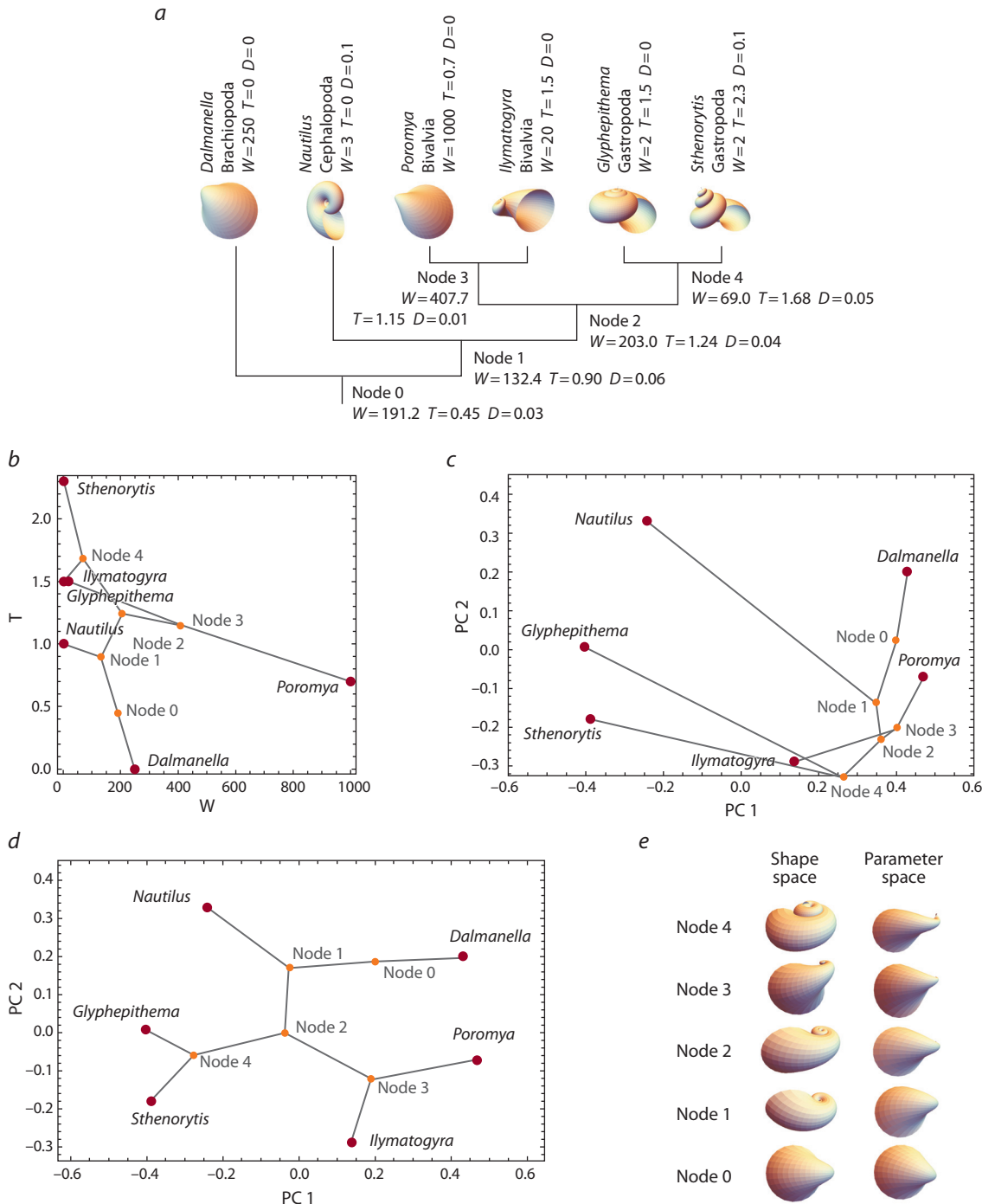shapes of these reconstructed ancestors are markedly differ-

Fig. 4. Ancestral reconstruction of six Raupian shells.

a, Phylogeny of the six shells, their generating parameters, and PGLS ancestral reconstructions. *b*, PGLS ancestral reconstruction in parameter space. *c*, Projection of the tree into geometric morphometric morphospace based ancestral reconstruction of generating parameters. *d*, Ancestral reconstruction based directly on shell shape. *e*, Comparison of the ancestral shapes reconstructed from shell shape and from generating parameters.

ent from those reconstructed directly from the parameters (Fig. 4, *e*), but would be appropriate if Brownian motion evolution has occurred in the phenotypes rather than their underlying parameters. This means that the modes of evolution reconstructed from parameter space and morphospace are expected to be different (e. g., Brownian motion in one and directional in the other).

Discussion

Real world examples. Hypothetical examples were used to illustrate the potential conflict between mathematically optimized and biologically optimized paths in order to have complete control of the relationship between phenotype and underlying parameters, but there are many real-world examples in which a similar conflict is likely to exist.

For example, mammalian molar tooth phenotypes are generated by cascades of developmental signalling that activates and inhibits cell growth within the developing tooth, growth which changes the spatial relationships between signalling centers and receptors (Jernvall, 1995; Jernvall, Thesleff, 2000). Up- or downregulation of individual signalling molecules in the cascade can produce non-linear changes in seemingly unrelated parts of the phenotype, such as the number of cusps, the presence of lophs, and the number of teeth (Jernvall et al., 2000; Kangas et al., 2004). In fact, incremental changes in the level of expression of activator and inhibitor molecules can produce not only non-linear trajectories in morphospace, but large jumps from one part of morphospace to another as cusps are gained and lost (Salazar-Ciudad, Jernvall, 2004; Harjunmaa et al., 2014). The phenotypic landscape of teeth would create non-linear maps between trajectories in morphospace and the space of the underlying molecular parameters much like in the shell example (Polly, 2008, 2015). Indeed, the real developmental processes involved in shell coiling probably operate much like the signalling cascades that control cell proliferation and secretion in teeth and thus produce phenomena like the ones illustrated in the Raup shell coiling example above (Rice, 1998; Marin et al., 2012).

Non-intermediate phenotypes in hybrid animals that result from overdominance effects are another example. If the genetics of phenotypic traits were purely additive, then the phenotype of a hybrid would lie on a line between the parent phenotypes in morphospace. However, the observed phenotypes of hybrids often lie away from a simple transect, as they do in the mandible shape of both mice (Klingenberg et al., 2001) and common shrews (Polly et al., 2013). This phenomenon has been linked to differences between the additive and dominance effects of quantitative trait loci (QTLs) on mandible shape (Klingenberg et al., 2001), which create a non-linear genotype-phenotype map analogous to the ones described above. Like with tooth morphology, the consequence is that the pattern of change in the underlying parameters, especially if it were to be categorized as corresponding to one mode of evolution or another, would be different from the pattern in morphspace.

Non-linear trajectories in morphospace are expected in some cases. As described above, the mathematical ordering of morphospace is based on minimizing Procrustes distances and ordinating shapes with PCA. The former defines the scaling between axes that determines what is straight and what is not in the multivariate morphospace and the later defines the coordinate system and, thus, the perspective from which we see trajectories in low dimensional plots like Fig. 1. The former contributes to why linear patterns in underlying parameters may correspond to non-linear trajectories in morphospace (or vice versa), the latter affects our ability to recover the morphospace trajectory. If a phenotypic trajectory is non-linear or if it is not perfectly aligned with the PC axes, its appearance in just one or two dimensions may be misleading. It may also be difficult or impossible to capture with univariate methods, such as performing a regression or ANOVA on just one PC.

A gradational series of highly multivariate shapes can artificially create the appearance of a strongly non-linear trajectory in low dimensional projection (i.e., a two-dimensional PC plot). A gradational series plotted on the first two axes of PC space will usually form a curve or horseshoe because the

first PC axis, which describes the axis of greatest variance, tends to separate the first and last members of the series because they are the most different in shape and therefore account for the greatest amount of variation in the data set. Similarly, the second PC separates the middle of the series from its two ends, because that contrast accounts for the greatest amount of shape difference not already described by the first PC. Successive PCs separate successive subdivisions of the series. Bookstein (2012) described this phenomenon in detail for random walk time series and presented statistical expectations for the amount of curvature observed on the first few principle components of morphospace. Examples of real-world gradients in shape that are subject to this phenomenon are ontogenetic series of hominid skulls (Mitteroecker et al., 2005) or anterior-posterior series of vertebrae (Head, Polly, 2015). Phylogenetic trees are examples of branched time series, therefore PC ordinations of species are also subject to the same artefactual effects even though the curvature may be much less noticeable (Polly et al., 2013).

Curved or irregular trajectories can also be caused by the non-linearities of the phenotypic landscape. Raup's shell coiling equations introduce such non-linearities. Linear change in *W* produces a broadly curvilinear path in shape space (Polly, Motz, 2017). Changes in parameters controlling the generating shape can produce sigmoid or even complex stair-stepped trajectories in shape space depending on how they interact with the ranks of semilandmarks used to represent the shape (Polly, Motz, 2017).

Recommendations

Consider the possibility of non-linear relationships. Because of the phenomena described in this paper, care should be taken in making inferences about the factors underlying shape based on analyses conducted in morphospace. In ideal situations the phenotypic landscape is understood so that patterns in morphospace can be translated into expected patterns in the underlying factors (Rice, 2002; Polly, 2008; Bookstein, 2016). However, the phenotypic landscape is seldom known for the complex multivariate phenotypes that are analysed with geometric morphometrics. Caution should therefore be exercised about making inferences about underlying developmental, genetic, or environmental factors from the morphometric analysis of phenotypes alone.

This caution should extend to interpretation of evolutionary patterns. For example, evolutionary model fitting may find that patterns of evolution in morphospace are best explained by Brownian motion, but that would not necessarily imply that evolution in the parameters of the underlying developmental and genetic factors are random. Conversely, a random pattern of change in the underlying factors could produce a non-random pattern of evolution in the phenotypes, as in the example with the snail coiling presented above.

Always use fully multivariate analytical methods. Even though PCs are aligned with major axes of variation and the scores of the sample being analysed are uncorrelated from one PC to another, the underlying factors of interest may not be aligned with the PCs. The alignment of the PCs is completely sample dependent – adding or subtracting an additional taxon or enlarging a sample to increase statistical power will reorient the PC axes, but it will not change the underlying relationship

between phenotypic shape and its causal factors (see example based on shell coiling in Polly, Motz, 2017). This means that fully multivariate statistical techniques like multivariate MANOVA or multivariate regression must be used to evaluate the relationship between shape and putative causal factors like environmental differences, locomotor function, or gene expression patterns (Polly et al., 2013; Goolsby, 2015; Uyeda et al., 2015; Bookstein, 2016; Polly, Motz, 2017). Despite widespread practice to the contrary, univariate tests should never be performed on individual PCs, regardless of whether they are statistical tests for differences between groups or fitting evolutionary models.

Avoid linear extrapolations over large shape distances. The potential problems illustrated above will be exacerbated when trajectories are extended over long distances in morphospace. Densely sampled series will be more likely to trace the paths created by non-linear phenotypic landscapes than will sparse samples, thus minimizing the amount of extrapolation that is necessary to reconstruct ancestral shapes or to predict evolutionary outcomes. For phylogenetic problems, consider including fossil taxa to minimize the evolutionary time between nodes and tips. Fossil members of a clade will have shorter branch lengths than its living members and thus helpfully constrain node and branch reconstructions toward the biologically plausible (Polly, 2001; Finarelli, Flynn, 2006; Slater et al., 2012; Gómez-Robles et al., 2013; Slater, Harmon, 2013).

Conclusion

When Belyaev started his experiment with fox domestication he knew very little about the genotype-phenotype map. Selection for tameness produced phenotypic changes similar to those found in domestic dogs, thus demonstrating that pleiotropy between these traits exists in canids and thus may explain the parallel acquisition of similar features in a range of domestic species that are rare in their wild relatives. The foxes are an example of how trajectories in phenotypic space may differ from their associated trajectories in genetic space.

The potential for the mathematics of morphometric spaces to produce shape trajectories that are biologically counterintuitive is very real. The two hypothetical examples described above illustrate two of the main ways in which this can happen: (1) the trajectory may intersect with shapes that are biologically unlikely or even impossible; and (2) the patterns of change in morphospace may have a complex, non-linear relationship with change in underlying genetic, developmental, or environmental factors. While reconstructions of phylogenetic nodes and branches were emphasized here, many other operations that are commonly employed in morphometrics produce trajectories that are subject to the same issues. Estimating selection vectors, evolutionary response vectors, or ontogenetic trajectories are examples mentioned above. Even calculating a mean shape involves a mathematical extrapolation to a new point in morphospace. In principle, all of these operations could produce counterintuitive or contradictory results as readily as the phylogenetic examples shown here.

Care should therefore be taken in making inferences about causal factors or evolutionary processes, especially if the phenotypes being analysed are very different, when the underlying phenotypic landscape is unknown. Potential problems can be minimized with dense sampling, which may involve adding

fossil taxa, with fully multivariate statistical methods that are able to fit models with trajectories that cross morphospace axes, and by considering non-linear model fitting when appropriate. The strongest inferences will be those framed as hypotheses based on based *a priori* evidence that tested for consistency with shape variation rather than inductive identification of causal factors from the shape patterns along. For example, rather than conjecturing that the sequence of shapes along PC 1 in Fig. 1 represents the phenotypic effect a hitherto unrecognized developmental regulatory gene (or whatever), one is on safer grounds predicting the effects on shape of a known regulatory gene and testing how well it explains the observed variation in shape. Most importantly, it should be recognized that trajectory shapes in morphospace may be quite different in the space of the underlying factors. For example, the mode of evolution reconstructed in morphospace (e. g., Brownian motion) may be different in the associated factor space (e. g., directional), which means that inferring whether evolutionary patterns are random may depend on which space is being analysed.

But these cautions are not intended to discourage the use of geometric morphometrics to study evolution. Evolutionary modelling, phylogenetic reconstruction, and statistical testing of factors that may affect evolution of the phenotype can be accomplished productively with morphometrics, as many studies demonstrate. Furthermore, the principle of parsimony dictates that theoretically informed predictions derived from morphospace, such as ancestral shape reconstructions, should be preferred in the absence of evidence to the contrary (e. g., Polly, 2001; Finarelli, Flynn, 2006; Gómez-Robles et al., 2013). Furthermore, fitness is determined by the phenotype, not its underlying factors. Therefore one might expect phenotypic trajectories to incrementally follow paths of selection, whereas trajectories in the underlying factor space may be the ones that trace complex, non-linear paths in response to selection on the phenotype. Some studies of phylogenetic changes in phenotypic evolution in the context of selection for functional performance have recovered predictable change in geometric morphometric morphospace (Polly et al., 2016). One might therefore expect drift and selection processes to differ in the trajectories they trace: drift, which is a process that involves random sampling of the underlying genetic pool, might produce linear transformations in the underlying factors with non-linear responses in the phenotypes, but selection, which is systematic sampling of phenotypic variation, might produce the opposite in situations where the phenotypic landscape is non-linear.

Acknowledgements

Conflict of interest

The author declares that there are no conflicts of interest.

References

Adams D.C., Rohlf F.J., Slice D.E. Geometric morphometrics: ten years of progress following the 'revolution'. Italian J. Zool. 2004;71:5-16. DOI 10.1080/11250000409356545.

Adams D.C., Rohlf F.J., Slice D.E. A field comes of age: geometric morphometrics in the 21st century. Hystrix. 2013;24:7-14. DOI 10.4404/hystrix-24.1-6283.

Arnold S.J., Pfrender M.E., Jones A.G. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Genetica. 2001;112-113:9-32. DOI 10.1023/A:1013373907708.

Belyaev D.K. Destabilizing selection as a factor in domestication. J. Heredity. 1979;70:301-308.

Belyaev D.K., Plyusnina I.Z., Trut L.N. Domestication in the Silver fox (*Vulpes fulvus* DESM): changes in physiological boundaries of the sensitive period of primary socialization. Appl. Anim. Behav. Sci. 1984;13:359-370. DOI 10.1016/0168-1591(85)90015-2.

Bookstein F.L. Morphometric Tools for Landmark Data. Cambridge, UK: Cambr. Univ. Press, 1991.

Bookstein F.L. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med. Image Anal. 1997;1:225-243.

Bookstein F.L. Random walk as a null model for high-dimensional morphometrics of fossil series: geometrical considerations. Paleobiology. 2012;39:52-74. DOI 10.1016/S1361-8415(97)85012-8.

Bookstein F.L. The inappropriate symmetries of multivariate statistical analysis in geometric morphometrics. Evol. Biol. 2016;43:277-313. DOI 10.1007/s11692-016-9382-7.

Clavel J., Escarguel G., Merceron G. mvMORPH: an R package for fitting multivariate evolutionary models to morphometric data. Meth. Ecol. Evol. 2015;6:1311-1319. DOI 10.1111/2041-210X.12420.

Dryden I.L., Mardia K.V. Statistical Shape Analysis. Chichester, UK: John Wiley & Sons, 1998.

Felsenstein J. Maximum-likelihood estimation of evolutionary trees from continuous characters. Am. J. Hum. Genet. 1973;25:471-492.

Felsenstein J. Phylogenies and quantitative characters. Ann. Rev. Ecol. Syst. 1988;19:445-471. DOI 10.1146/annurev.es.19.110188.002305.

Finarelli J.A., Flynn J.J. Ancestral state reconstruction of body size in the Caniformia (Carnivora, Mammalia): the effects of incorporating data from the fossil record. Syst. Biol. 2006;55:301-313. DOI 10.1080/10635150500541698.

Goolsby E.W. Phylogenetic comparative methods for evaluating the evolutionary history of function-valued traits. Syst. Biol. 2015;64: 568-578. DOI 10.1093/sysbio/syv012.

Gómez-Robles A., Bermúdez de Castro J.M., Arsuaga J.-L., Carbonell E., Polly P.D. No known hominin species matches the expected dental morphology of the last common ancestor of Neanderthals and modern humans. Proc. Natl. Acad. Sci. USA. 2013;110:18196-18201. DOI 10.1073/pnas.1302653110.

Gunz P., Mitteroecker P. Semilandmarks: a method for quantifying curves and surfaces. Hystrix. 2013;24:103-109. DOI 10.4404/hystrix-24.1-6292.

Gunz P., Mitteroecker P., Bookstein F.L. Semilandmarks in three dimensions. Modern Morphometrics in Physical Anthropology. Ed. D.E. Slice. New York, NY: Kluwer, 2005;73-93.

Hansen T.F. The evolution of genetic architecture. Ann. Rev. Ecol. Evol. Syst. 2006;37:123-157. DOI 10.1146/annurev.ecolsys.37.091305.110224.

Hansen T.F. Macroevolutionary quantitative genetics? A comment on Polly (2008). Evol. Biol. 2008;35:182-185. DOI 10.1007/s11692-008-9027-6.

Harjunmaa E., Seidel K., Häkkinen T., Renvoisé E., Corfe I.J., Kallonen A., Zhang Z.Q., Evans A.R., Mikkola M.L., Salazar-Ciudad I., Klein O.D., Jernvall J. Replaying evolutionary transitions from the dental fossil record. Nature. 2014;512:44-48. DOI 10.1038/nature13613.

Head J.J., Polly P.D. Evolution of the snake body form reveals homoplasy in amniote *Hox* gene function. Nature. 2015;520:86-89. DOI 10.1038/nature14042.

Jernvall J. Mammalian molar cusp patterns. Acta Zool. Fenn. 1995; 198:1-61.

Jernvall J., Thesleff I. Reiterative signaling and patterning during mammalian tooth morphogenesis. Mech. Develop. 2000;92:19-29. DOI 10.1016/S0925-4773(99)00322-6.

Jernvall J., Keränen S.V., Thesleff I. Evolutionary modification of development in mammalian teeth: quantifying gene expression patterns and topography. Proc. Natl. Acad. Sci. USA. 2000;97:14444-14448. DOI 10.1073/pnas.97.26.14444.

Kangas A.T., Evans A.R., Thesleff I., Jernvall J. Nonindependence of mammalian dental characters. Nature. 2004;432:211-214. DOI 10.1038/nature02927.

Klingenberg C.P. Morphometrics and the role of the phenotype in studies of the evolution of developmental mechanisms. Gene. 2002;287: 3-10. DOI 10.1016/S0378-1119(01)00867-8.

Klingenberg C.P. Novelty and "homology-free" morphometrics: What's in a name? Evol. Biol. 2008;35:186-190. DOI 10.1007/s11692-008-9029-4.

Klingenberg C.P. Size, shape, and form: concepts of allometry in geometric morphometrics. Dev. Genes Evol. 2016;226:113-137. DOI 10.1007/s00427-016-0539-2.

Klingenberg C.P., Debat V., Roff D.A. Quantitative genetics of shape in cricket wings: developmental integration in a functional structure. Evolution. 2010;64:2935-2951. DOI 10.1111/j.1558-5646.2010.01030.x.

Klingenberg C.P., Gidaszewski N.A. Testing and quantifying phylogenetic signals and homoplasy in morphometric data. Syst. Biol. 2010;59:245-261. DOI 10.1093/sysbio/syp106.

Klingenberg C.P., Leamy L.J. Quantitative genetics of geometric shape in the mouse mandible. Evolution. 2001;55:2342-2352. DOI 10.1554/0014-3820(2001)055[2342:QGOGSI]2.0.CO;2.

Klingenberg C.P., Leamy L.J., Routman E.J., Cheverud J.M. Genetic architecture of mandible shape in mice: effects of quantitative trait loci analyzed by geometric morphometrics. Genetics. 2001;157: 785-802.

Klingenberg C.P., Monteiro L.R. Distances and directions in multidimensional shape spaces: implications for morphometric applications. Syst. Biol. 2005;54:678-688. DOI 10.1080/10635150590947258.

Lande R. Natural selection and random genetic drift in phenotypic evolution. Evolution. 1976;30:314-334. DOI 10.2307/2407703.

MacLeod N. Generalizing and extending the eigenshape method of shape space visualization and analysis. Paleobiol. 1999;25:107-138. DOI 10.1666/0094-8373(1999)025<0107:GAETEM>2.3.CO;2.

Marin F., Le Roy N., Marie B. The formation and mineralization of mollusk shell. Front. Biosci. 2012;4:1099-1125.

Martínez-Abadías N., Esparza M., Sjøvold T., González-José R., Santos M., Hernández M., Klingenberg C.P. Pervasive genetic integration directs the evolution of human skull shape. Evolution. 2011;66:1010-1023. DOI 10.1111/j.1558-5646.2011.01496.x.

Martins E.P., Hansen T.F. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. Am. Nat. 1997;149:646-667. DOI 10.1086/286013.

Mitteroecker P., Hutteger S.M. The concept of morphospaces in evolutionary and developmental biology: mathematics and metaphors. Biol. Theory. 2009;4:54-67. DOI 10.1162/biot.2009.4.1.54.

Mitteroecker P., Gunz P., Bookstein F.L. Heterochrony and geometric morphometrics: a comparison of cranial growth in *Pan paniscus* versus *Pan troglodytes*. Evol. Develop. 2005;7:244-258. DOI 10.1111/j.1525-142X.2005.05027.x.

Mounier A., Lahr M.M. Virtual ancestor reconstruction: revealing the ancestor of modern humans and Neandertals. J. Hum. Evol. 2016; 91:57-72. DOI 10.1016/j.jhevol.2015.11.002.

Myers E.M., Janzen F.J., Adams D.C., Tucker J.K. Quantitative genetics of plastron shape in slider turtles (*Trachemys scripta*). Evolution. 2006;60:563-572. DOI 10.1554/05-633.1.

Nijhout H.F. The nature of robustness in development. BioEssays. 2002;24:553-563. DOI 10.1002/bies.10093.

Nijhout H.F. Complex traits: genetics, development, and evolution. Integrating Evolution and Development: From Theory to Practice. Eds. R. Sansom, R.N. Brandon. Cambridge, MA: MIT Press, 2007; 93-112.

Nijhout H.F. Developmental phenotypic landscapes. Evol. Biol. 2008; 35:100-103. DOI 10.1007/s11692-008-9024-9.

Polly P.D. Paleontology and the comparative method: Ancestral node reconstructions versus observed node values. Am. Nat. 2001;157:596-609. DOI 10.1086/320622.

Polly P.D. On the simulation of the evolution of morphological shape: multivariate shape under selection and drift. Palaeont. Electr. 2004; 7.2.7A:1-28.

Polly P.D. Developmental dynamics and G-matrices: Can morphometric spaces be used to model evolution and development? Evol. Biol. 2008;35:83-96. DOI 10.1007/s11692-008-9020-0.

Polly P.D. Gene networks, occlusal clocks, and functional patches: new understanding of pattern and process in the evolution of the dentition. Odontol. 2015;103:117-125. DOI 10.1007/s10266-015-0208-3.

Polly P.D., Motz G.J. Patterns and processes in morphospace: geometric morphometrics of three-dimensional objects. Paleont. Soc. Pap. 2017;22:71-99.

Polly P.D., Lawing A.M., Fabre A.-C., Goswami A. Phylogenetic principal components analysis and geometric morphometrics. Hystrix. 2013;24:1-9. DOI 10.4404/hystrix-24.1-6383.

Polly P.D., Stayton C.T., Dumont E.R., Pierce S.E., Rayfield E.J., Angielczyk K.D. Combining geometric morphometrics and finite element analysis with evolutionary modeling: towards a synthesis. J. Vert. Paleont. 2016;36(4):p.e1111225. DOI 10.1080/02724634.2016.1111225.

Raup D.M. The geometry of coiling in gastropods. Proc. Natl. Acad. Sci. USA. 1961;47:602-609.

Raup D.M. Geometric analysis of shell coiling: general problems. J. Paleont. 1966;40:1178-1190.

Rice S.H. The bio-geometry of mollusc shells. Paleobiol. 1998;24:133-149. DOI 10.1017/S0094837300020017.

Rice S.H. A general population genetic theory for the evolution of developmental interactions. Proc. Natl. Acad. Sci. USA. 2002;99:15518-15523. DOI 10.1073/pnas.202620999.

Rice S.H. Evolutionary Theory: Mathematical and Conceptual Foundations. Sunderland, MA: Sinauer Associates, Inc, 2004.

Rohlf F.J. Relative warp analysis and an example of its application to mosquito wings. Contributions to Morphometrics. Eds. L.F. Marcus, E. Bello, A. Garcia-Valdecasas. Madrid, Spain: Museum Nacional de Ciencias Naturales, 1993;131-159.

Rohlf F.J. Geometric morphometrics and phylogeny. Morphology, Shape, and Phylogenetics. Eds. N. Macleod, P. Forey. Abingdon, United Kingdom: Taylor and Francis, 2002;175-193.

Rohlf F.J., Slice D. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst. Zool. 1990;39:40-59. DOI 10.2307/2992207.

Salazar-Ciudad I., Jernvall J. How different types of pattern formation mechanisms affect the evolution of form and development. Evol. Develop. 2004;6:6-16. DOI 10.1111/j.1525-142X.2004.04002.x.

Sánchez-Villagra M.R., Geiger M., Schneider R.A. The taming of the neural crest: a developmental perspective on the origins of morphological covariation in domesticated mammals. R. Soc. Open Sci. 2016;3:160107. DOI 10.1098/rsos.160107.

Simpson G.G. Tempo and Mode in Evolution. New York, NY: Columbia Univ. Press, 1944.

Slater G.J., Harmon L.J. Unifying fossils and phylogenies for comparative analyses of diversification and trait evolution. Meth. Ecol. Evol. 2013;4:699-702. DOI 10.1111/2041-210X.12091.

Slater G.J., Harmon L.J., Alfaro M.E. Integrating fossils with molecular phylogenies improves inference of trait evolution. Evolution. 2012; 66:3931-3944. DOI 10.1111/j.1558-5646.2012.01723.x.

Trut L.N., Herbeck Y.E., Kharlamova A.V., Gulevich R.G., Kukekova A.V. Fox domestication: molecular mechanisms involved in selection for behavior. Russ. J. Genet. Appl. Res. 2013;3:419-425. DOI 10.1134/S2079059713060117.

Uyeda J.C., Caetano D.S., Pennell M.W. Comparative analysis of principal components can be misleading. Syst. Biol. 2015;64:677-689. DOI 10.1093/sysbio/syv019.

Vavilov N.I. The law of homologous series in variation. J. Genet. 1922; 12:47-89. DOI 10.1007/BF02983073.

Viðarsdóttir U.S., O'Higgins P., Stringer C. A geometric morphometric study of regional differences in the ontogeny of the modern human facial skeleton. J. Anat. 2002;201:211-229. DOI 10.1046/j.1469-7580.2002.00092.x.

Wilkins A.S., Wrangham R.W., Fitch W.T. The "domestication syndrome" in mammals: A unified explanation based on neural crest cell behavior and genetics. Genetics. 2014;197:795-808. DOI 10.1534/genetics.114.165423.

Wolf J.B. The geometry of phenotypic evolution in developmental hyperspace. Proc. Natl. Acad. Sci. USA. 2002;99:15849-15851. DOI 10.1073/pnas.012686699.

Wright S. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. 6th Int. Congr. Genet. 1932;356-366.

Zelditch M.L., Fink W.L. Heterochrony and heterotopy: stability and innovation in the evolution of form. Paleobiol. 1996;22:241-254. DOI 10.1017/S0094837300016195.

Zelditch M.L., Bookstein F.L., Lundrigan B.L. Ontogeny of integrated skull growth in the cotton rat *Sigmodon fulviventer*. Evolution. 1992;46:1164-1180. DOI 10.2307/2409763.

Zelditch M.L., Swiderski D.L., Sheets H.D., Fink W.L. Geometric Morphometrics for Biologists: a Primer. Amsterdam, The Netherlands: Elsevier Acad. Press, 2004.