Combining Geometric Morphometrics and Finite Element Analysis with Evolutionary Modeling: Towards a Synthesis

Authors: Polly, P. David, Stayton, C. Tristan, Dumont, Elizabeth R., Pierce, Stephanie E., Rayfield, Emily J., et al.

COMBINING GEOMETRIC MORPHOMETRICS AND FINITE ELEMENT ANALYSIS WITH EVOLUTIONARY MODELING: TOWARDS A SYNTHESIS

P. DAVID POLLY,[1] C. TRISTAN STAYTON,[2] ELIZABETH R. DUMONT,[3] STEPHANIE E. PIERCE,[4]
EMILY J. RAYFIELD,[5] and KENNETH D. ANGIELCZYK[*,6]

[1]Departments of Geological Sciences, Biology, and Anthropology, Indiana University, 1001 E. 10th Street, Bloomington, Indiana
47405, U.S.A., pdpolly@indiana.edu;
[2]Department of Biology, Bucknell University, 001 Dent Drive, Lewisburg, Pennsylvania 17837, U.S.A., tstayton@bucknell.edu;
[3]Biology Department, University of Massachusetts, 611 North Pleasant Street, Amherst, Massachusetts 01003, U.S.A.,
bdumont@bio.umass.edu;
[4]Museum of Comparative Zoology and Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street,
Cambridge, Massachusetts 02138, U.S.A., spierce@oeb.harvard.edu;
[5]School of Earth Sciences, University of Bristol, Wills Memorial Building, Queen's Road, Bristol BS8 1RJ, U.K.,
e.rayfield@bristol.ac.uk;
[6]Integrative Research Center, Field Museum of Natural History, 1400 South Lake Shore Drive, Chicago, Illinois 60605, U.S.A.,
kangielczyk@fieldmuseum.org

ABSTRACT—Geometric morphometrics (GM) and finite element analysis (FEA) are increasingly common techniques for the study of form and function. We show how principles of quantitative evolution in continuous phenotypic traits can link the two techniques, allowing hypotheses about the relative importance of different functions to be tested in a phylogenetic and evolutionary framework. Finite element analysis is used to derive quantitative surfaces that describe the comparative performance of different morphologies in a morphospace derived from GM. The combination of two or more performance surfaces describes a quantitative adaptive landscape that can be used to predict the direction morphological evolution would take if a combination of functions was selected for. Predicted paths of evolution also can be derived for hypotheses about the relative importance of multiple functions, which can be tested against evolutionary pathways that are documented by phylogenies or fossil sequences. Magnitudes of evolutionary trade-offs between functions can be estimated using maximum likelihood. We apply these methods to an earlier study of carapace strength and hydrodynamic efficiency in emydid turtles. We find that strength and hydrodynamic efficiency explain about 45% of the variance in shell shape; drift and other unidentified functional factors are necessary to explain the remaining variance. Measurement of the proportional trade-off between shell strength and hydrodynamic efficiency shows that throughout the Cenozoic aquatic turtles generally sacrificed strength for streamlining and terrestrial species favored stronger shells; this suggests that the selective regime operating on small to mid-sized emydids has remained relatively static.

SUPPLEMENTAL DATA—Supplemental materials are available for this article for free at www.tandfonline.com/UJVP

INTRODUCTION

The role of functional performance in the evolution of morphological form remains a central research question in vertebrate paleontology. Today, two methods dominate the quantitative study of form and function in vertebrate

paleontology: geometric morphometrics (GM) and finite element analysis (FEA). Although both methods use digital representations of morphological shape (Fig. 1), the questions they address are fundamentally different. Geometric morphometrics quantifies differences in morphological shape, including static differences between individuals, sexes, or species, as well as transformational differences between ontogenetic stages, between stratigraphic units, or along branches of a phylogenetic tree (e.g., Zelditch et al., 2004, 2012). Combined with multivariate statistics and phylogenetics, GM can be used to analyze the relationship between shape and a variety of evolutionary, developmental, ecological, and functional factors (e.g., Adams et al., 2013). An extensive body of mathematical theory gives the approach much of its power but also sets broad limits on the ways in which it can be used (e.g., Bookstein, 1991; Dryden and Mardia, 1998).

In contrast, FEA is a numerical technique used to predict the performance of complex structures (e.g., Clough, 1990). It has

Color versions of one or more of the figures in this article can be found online at www.tandfonline.com/ujvp.
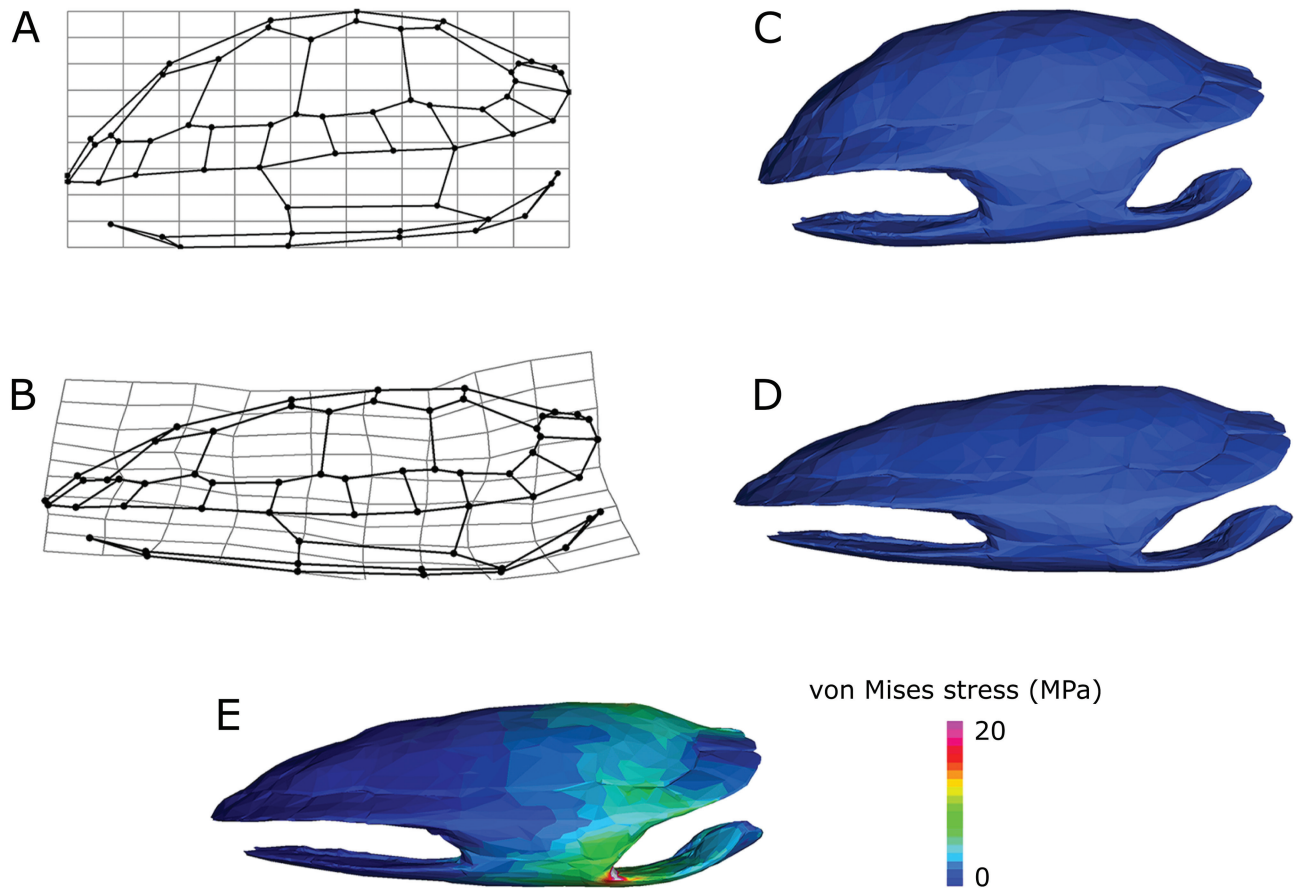
FIGURE 1. Illustrations of steps in the process of generating new FE models from landmark data and an existing FE model. Anterior is to the right for all figures. **A**, landmark data for a 'base' FE model (shown in **C**). Black lines illustrate features of the corresponding shell; gray lines show a grid associated with the landmark coordinates; **B**, landmark data from another specimen of interest. Again, black lines indicate shell features; gray lines show the deformation required to transform the 'base' model coordinates into the new coordinates, as interpolated by a thin-plate spline function. Note that in this case, the two-dimensional visualization is for representational purposes only and does not fully describe the shape differences between the turtle shells in three dimensions. See Klingenberg (2013) for more information on visualizations in geometric morphometrics; **C**, the 'base' FE model; **D**, a new FE model generated by applying the thin-plate spline interpolation function to the coordinates of all nodes in the 'base' model; **E**, stresses resulting from the application of an anterior midline load to the new FE model. Warmer colors indicate high stresses; cooler colors indicate low stresses.

recently become popular in evolutionary morphology to assess the mechanical behavior of anatomical structures by quantifying the topographical distribution of stresses (the amount of force per unit area experienced by tissues) and strains (the physical displacement of the tissues) within a morphological structure as it is deformed by an external force (e.g., Dumont et al., 2005). Finite element analysis is often used in paleontology to compare the abilities of extinct taxa to withstand loads induced by functions such as chewing in a specified manner, adopting a particular limb posture, being bitten by a particular predator, or supporting an estimated body mass (e.g., Rayfield, 2007). Both FEA and GM are valuable to paleontology because they allow the form and function of long extinct species to be analyzed and functional hypotheses to be tested through digital manipulation of morphology, materials, or applied loads and constraints.

In this review, we show how GM and FEA can be combined within a framework of quantitative evolutionary theory to test hypotheses about the role of functional factors in the evolution of morphological form. Such a quantitative synthesis was called for by D'Arcy Thompson in his book *On Growth and Form* (1917), in which he argued that evolutionary transformations in the shape of

organisms can be described with mathematical expressions based on the physical laws of the forces acting upon them. Thompson famously illustrated evolutionary transformations by deforming grids to show how the shape of one organism could be modified to produce the shape of another. His artistically constructed (rather than quantitatively derived) grid deformations were the inspiration behind the development of GM in the 1980s, which used the mathematics of thin-plate spline deformation to produce deformation grids quantitatively (e.g., Bookstein, 1989; also see below). However, Thompson considered not only the deformation of shape but also the structural efficiency and mechanical forces related to those deformations. He approached morphology from an engineering standpoint, noting how the trabecular organization in bones tends follow lines of loading stress (in a manner similar to Wolff, 1892) and how the axial skeleton resembles a cantilevered bridge that supports the weight of the body. Thompson argued that mechanics, not evolution, was the primary determinant of organismal form, stating that an animal's skeleton "... is to a very large extent determined by mechanical considerations, and tends to manifest itself as a diagram, or reflected image, of mechanical stress" (Thompson, 1917:712). Thompson's emphasis

on stress and strain is paralleled in FEA. In this article, we combine GM's strength for analyzing evolutionary transformations, FEA's strength for analyzing the structural performance of morphology, and quantitative evolutionary theory in order to test competing functional hypotheses about the evolution of form. Although our conception of the form–function relationship is fundamentally different from Thompson's, the goals of our methodological synthesis resemble his.

Our purpose differs from other recent syntheses of GM and FEA, which have focused on joining the two techniques at the level of the morphological structure. Some important advances have been made in using GM to measure shape deformation (strain) in FEA models after forces have been applied and to isolate the components of shape that are most directly related to that deformation (Gröning et al., 2011; O'Higgins et al., 2011; O'Higgins and Milne, 2013). Other syntheses have used GM to measure shape differences between taxa or evolutionary transformations in shape as a context for comparing strain patterns measured with FEA (Pierce et al., 2008, 2009; Stayton, 2011; Tseng, 2013). Nevertheless, there has been debate over the conceptual and theoretical connections between mechanical deformation, shape deformations, differences in strain patterns, and differences in shape (Weber et al., 2011; Bookstein, 2013a; O'Higgins and Milne, 2013), and a true combination of GM and FEA remains elusive.

Here we focus on the evolution of form, and we synthesize GM and FEA in the context of evolutionary theory in order to draw on quantitative expressions of function, performance, fitness, selection, and phylogenetic change to test the putative roles of functional factors in the evolution of morphological diversity. Morphological evolution is the net result of selective and random (drift) processes acting on phenotypes and is channeled by constraints imposed by developmental, mechanical, genetic, and physical processes (Wright, 1968; Raff, 1996). Function plays a role in selection because the fitness of a population is in part determined by the overall functional performance of its morphologies in particular environmental contexts (Arnold, 1983; Endler, 1986; McGill et al., 2006). Many aspects of functional performance are simultaneously at play on complex phenotypes, often involving trade-offs between competing and sometimes opposing influences on the phenotype's performance in one environment or another.

Building on the work of Pierce et al. (2008, 2009), Stayton (2011), Young et al. (2011), and Dumont et al. (2014), we show how FEA can be used to develop phenotype–performance functions that serve as hypotheses for how morphology is expected to respond to selection in a range of functional or environmental situations (Fig. 2). These FEA-derived performance functions can be translated into adaptive landscape models (Simpson, 1944; Arnold, 2003). Summed together, performance surfaces based on several functional factors describe a variable adaptive landscape whose slope gradients point in the direction(s) that phenotypic evolution would take if selection favors functional factors that affect fitness. These adaptive landscapes can then be used to test the extent to which the actual morphological evolution of a clade, measured with a combination of GM and phylogenetic comparative methods, fits the expectations of the performance functions. Of special interest to paleontology, the relative contribution of competing functional factors to evolution within a clade can be estimated from the morphologies of extant and fossil specimens using maximum likelihood once the performance surfaces have been quantified. When there is a functional trade-off between two or more functional properties (Levens, 1962; Wainwright, 2007), we show how the shifting balance between the opposing performance functions can be reconstructed by finding the proportion that best fits the realized morphologies of each member of a clade and mapping the changes onto a phylogenetic tree.

Our goal in this article is not to present a detailed introduction to FEA or GM but to develop the theoretical and methodological basis for combining them to study the evolution of the form–function complex. Toward this end we briefly review the principles of the two methods and how they fit into an evolutionary framework. Elaborating on the work of Stayton (2011), we then present a worked example using the evolution of carapace morphology in the New World pond and box turtles (Emydidae), a radiation that has involved multiple transitions between aquatic and terrestrial habits. Low-domed shells with low drag coefficients are thought to perform better in regard to agility in the water and swimming efficiency, but high-domed shells that are resistant to crushing are thought to better resist predation and other physical attacks. We model the expected performance effects of these two opposing factors on carapace evolution, find the proportional balance between them that best explains the realized morphologies of both aquatic and terrestrial emydids, estimate changes in that balance that have occurred during the phylogenetic history of the clade, and ask whether these performance factors alone are sufficient to explain the clade's morphological diversity or whether factors such as drift, phylogenetic inertia, and genetic covariance also were important. Finally, although we focus on using the methods we present to synthesize GM and FEA data, it is important to note that in principle they can be applied to any type of continuous physiological, morphological, or performance variables that might bear a relationship to shape (as quantified by GM).

Institutional Abbreviations—FMNH, Field Museum of Natural History, Chicago, Illinois; **FSM**, Florida Museum of Natural History, Gainesville, Florida; **UNSM**, University of Nebraska State Museum of Natural History, Lincoln, Nebraska.

REVIEW OF METHODS

Geometric Morphometrics

Geometric morphometrics is a group of methods for the analysis of shape, all of which use Cartesian geometric coordinates of anatomical structures (i.e., landmarks) rather than linear dimensions or other measurements (Fig. 1A, B). These methods focus on homologous or analogous points rather than pairs of points (as in endpoints of linear measurements) or ratios, which has advantages for cleanly partitioning the mathematical effects of size and for visualizing results as graphical transformations of the actual shape of the object. Landmarks or semi-landmarks (iterative points placed along a curve or across a surface) represent discrete anatomical loci on the specimens of interest; mathematically, they are the points of correspondence between specimens (Dryden and Mardia, 1998). Coordinates for 2D landmarks can be captured from digital photographs (Zelditch et al., 2004). Obtaining coordinates for 3D landmarks requires specialized equipment, such as a Microscribe 3D point digitizer, a laser or optical surface scanner, an X-ray computed tomography (CT) or magnetic resonance imaging volumetric scanner, or the ability to combine coordinates from multiple 2D images. Photogrammetry can also be used to reconstruct 3D surfaces from a series of digital photos taken from different angles (Falkingham, 2012; Olsen and Westneat, 2015).

Landmark coordinates have no natural orientation or scale, which means that they must be registered to make the coordinates of one specimen comparable to others (Kendall, 1977). The most common method, and our focus, is generalized Procrustes superimposition in which each set of landmarks is rescaled, aligned with other sets at their geometric centers (centroids), and rotated until the sum of squared distances between them is minimized (Gower, 1975; Rohlf and Slice, 1990). Several variants of Procrustes superimposition exist, which differ in the way that size is calculated for rescaling, the point used to rotate the object and whether shapes are fit to one another or to the
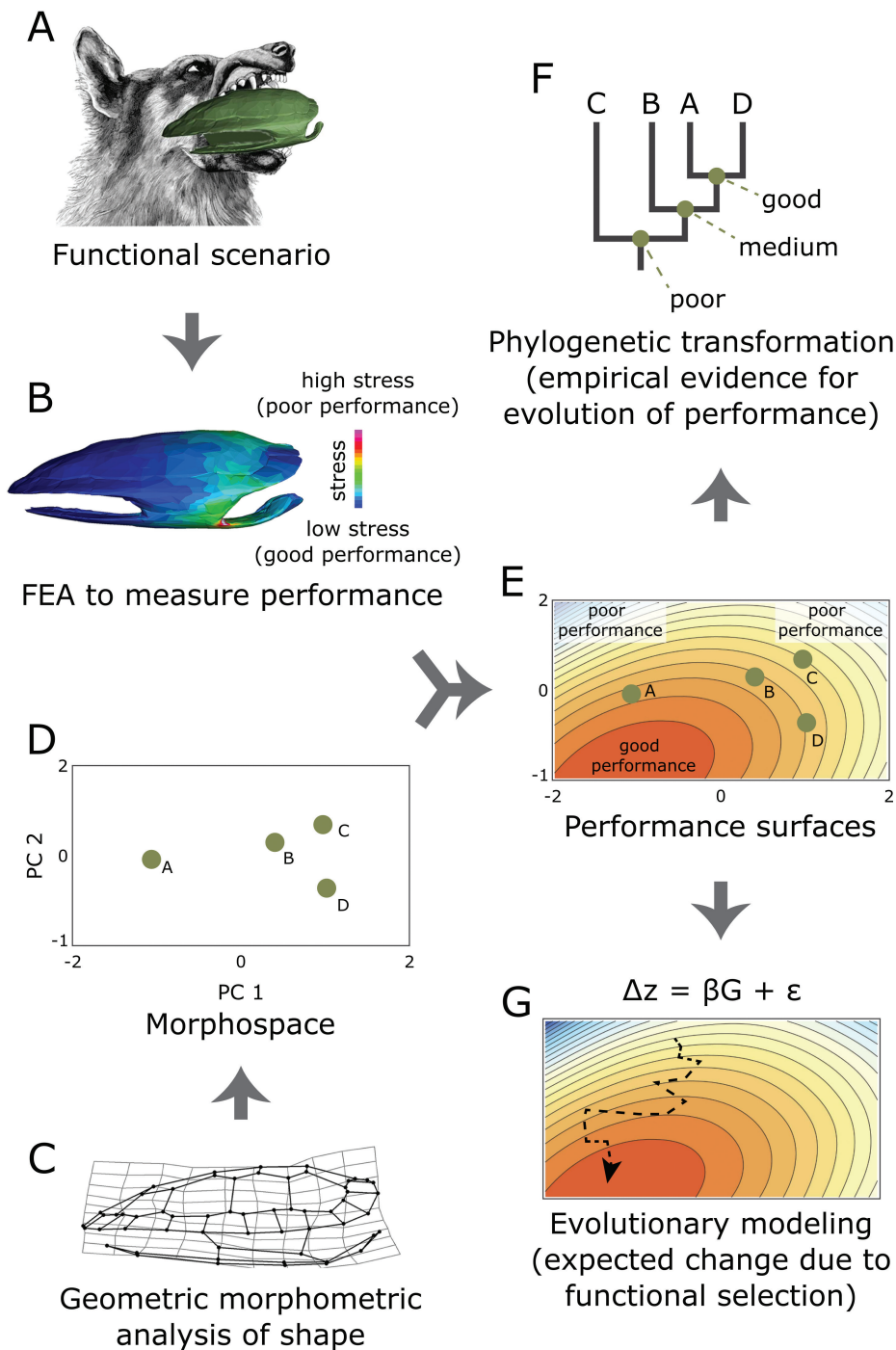
FIGURE 2. Schematic overview of using FEA and GM to study the evolution of morphology. **A**, a functional scenario of interest is identified, such as the resistance of streamlined turtle shells to predatory bites; **B**, FEA is used to model the performance of an organism's morphology in a given functional scenario. Here the stress induced on a turtle shell by a predator's bite is shown. The model is repeated across all taxa in the analysis and a performance index is derived from the stress or strain patterns; **C**, GM is used to characterize the shape of the morphology. In this case landmarks are used to quantify the shape of a turtle shell; **D**, multivariate ordination, such as PCA, is used to create a morphospace that portrays aspects of shape variation among many specimens or taxa; **E**, multivariate regression or polynomial surface fitting is used to estimate a performance surface by fitting the performance indices of the taxa to their spacing in morphospace; **F**, performance indices, trade-off weights, selection intensity, or other parameters can be mapped onto phylogenetic trees to estimate evolutionary changes from empirical data; **G**, performance surfaces for one or more functional scenarios can be combined into an adaptive landscape with which the expected evolutionary outcomes of selection can be simulated.

sample mean (see reviews in Zelditch et al., 2004; Slice, 2005). The removal of size, orientation, and translation reduce the degrees of freedom of the Procrustes aligned coordinates (loss of 4 degrees of freedom for 2D landmarks and 7 for 3D landmarks). The reduced dimensionality constrains variation such that shapes are distributed in a non-Euclidean mathematical space with the form of a hyperdimensional sphere or hemisphere (Kendall, 1984, 1985; Dryden and Mardia, 1998; Slice, 2001). Because of the non-Euclidean geometry of shape space, Procrustes coordinates are usually projected to a Euclidean tangent space, although in practice this is often unnecessary for biological shapes because developmental and functional integration typically constrains shape variation sufficiently that the non-Euclidean curvatures of shape space are inconsequential (Rohlf and Slice, 1990; Rohlf, 1999; Slice, 2001).

The Procrustes superimposed coordinates are the entry point for many multivariate analyses of shape. The coordinates themselves may be used as shape variables, but their covariances and reduced degrees of freedom must be taken into account when calculating *p*-values or other statistics. More commonly the coordinates are converted into one of two types of shape variables so that they have the proper number of degrees of freedom. The first method is to factor the coordinates into partial warp and uniform component

scores using a thin-plate spline decomposition (Bookstein, 1989, 1991; Rohlf, 1995; Zelditch et al., 2004). Each dimension of warp scores captures shape variation on different spatial scales. These scores have the correct number of degrees of freedom, but the different scales are still intercorrelated. The second and more common method of producing shape scores is to project the Procrustes coordinates onto their principal component (PC) axes (eigenvectors; Rohlf, 1993; Dryden and Mardia, 1998). Each dimension of PC shape represents a module of correlated variation in landmark coordinates that is orthogonal to other dimensions. Scores of the objects on the PC axes are shape variables that have both the proper number of degrees of freedom and are uncorrelated. Note that principal component decomposition of the partial warp and uniform component scores, known as relative warps analysis, produces identical shape variables so long as all the partial and uniform components are weighted equally (Rohlf, 1993). Standard multivariate statistical procedures such as multivariate regression and multivariate analysis of variance can be performed on the shape variables, but nonparametric permutation-based tests are recommended over traditional tests based on parametric F-distributions or Wilks' lambda because shape variation seldom meets the required assumptions of normality and sample sizes are seldom balanced (Manly, 1991; Zelditch et al., 2004; Mitteroecker and Gunz, 2009; Kowalewski and Novack-Gottshall, 2010). Nonparametric Procrustes analysis of variance tests based on multivariate distance between objects rather than on shape variables are also available (Klingenberg and McIntyre, 1998; Klingenberg, 2015).

The thin-plate spline can be used as a device for visualizing shape differences, decomposing shape variation by spatial scale, and (of particular interest for this study) morphing one digital object into the form of another using landmarks as tie points (Fig. 1; Bookstein, 1989). A spline function is fit to the displacements of landmarks caused by transforming a reference shape into a target shape and is used to interpolate the displacement of points that lie between the object's landmarks. The spline function is frequently used to deform a square grid placed on the reference object into a D'Arcy Thompson-like diagram of the shape transformation. The thin-plate spline also can be used to decompose shape differences into geometric components. The uniform (or affine) component describes shape differences that affect a target specimen equally across the whole specimen, whereas the nonuniform component describes localized shape differences (e.g., Bookstein, 1989, 1991, 1996b). Thin-plate spline visualizations are usually used with two-dimensional data, but they also can be applied to three-dimensional shapes (Bookstein, 1996b; Rohlf and Bookstein, 2003). Importantly, thin-plate splines can be used to transform larger grids or blocks of coordinate data based on a selected set of homologous landmarks (Bookstein, 1991; Zelditch et al., 2004; Wiley et al., 2005). The spline functions are fit to either two- or three-dimensional landmarks that have been placed on more complete digital objects, such as a photograph or 3D scan. The spline functions are then used to interpolate all the data points from the reference to the target object. Stayton (2009) showed how this procedure could be used to morph a reference FEA model into the forms of several target specimens, alleviating the need for an independent original scan of each specimen.

Important for our own analysis, principal component analysis (PCA) of Procrustes coordinates also produces a multivariate shape space, or morphospace (Fig. 2; Rohlf, 1993; Mardia and Dryden, 1998; Mitteroecker and Huttegger, 2009). Each point in morphospace represents a unique shape and each axis of the space represents an aspect of shape variation that is mathematically independent of variation on other axes. The distribution of real objects in morphospace is based on their shape similarity,

the distances between them being identical to their Procrustes shape distances (the summed distances between corresponding landmarks) in the non-Euclidean geometry of shape space. Note that when shapes belonging to a time series, such as an ontogenetic series, evolutionary lineage, or serially repeated structures, are projected into PC space, the trajectory connecting them can be highly nonlinear (Mitteroecker et al., 2004; Adams and Collyer, 2009; Collyer and Adams, 2013), forming predictable curves that describe increasingly smaller subdivisions of the entire series (Bookstein, 2013b). Multivariate regression of shape onto continuous independent factors such as body size, age, stress for a given load, or fitness yields prediction equations that can be visualized as lines, contour maps, surfaces, or hyperdimensional surfaces within the morphospace. In this article we visualize the relationship between shell shape and its functional properties by using such surfaces, whose contours show the pathways that shape evolution is expected to take if it is influenced by those functional factors.

Readers interested in Procrustes-based shape analysis can refer to recent reviews of the history of the technique (Bookstein, 1996a; 1998; Adams et al., 2004; Slice, 2005), associated analytical methodologies (Rohlf and Marcus, 1993; O'Higgins, 2000; Mitteroecker and Gunz, 2009; Klingenberg, 2013; Bookstein, 2014), and practical applications (Slice, 2007; Klingenberg, 2010; Lawing and Polly, 2010; Adams et al., 2013). Several book-length introductions to methods and theory are also available (Dryden and Mardia, 1998; Claude, 2008; Weber and Bookstein, 2011; Zelditch et al., 2004, 2012). Procrustes-based analysis can also be carried out on curves and surfaces using semi-landmarks (Bookstein, 1991, 1997; Gunz et al., 2005). Different methods exist for superimposing semi-landmarks, based on minimizing distances between individuals and the reference form (Sampson et al., 1996) or minimizing the bending energy of the thin-plate spline describing the difference between a specimen and the reference form (Green, 1996; Bookstein, 1997), with minor implications for the statistical properties of the superimposed data (see review in Zelditch et al., 2012).

Geometric morphometrics has become a standard tool in vertebrate paleontology. Recent examples of its application include studies of developmental patterns and mechanisms (Böhmer et al., 2015; Bhullar et al., 2015; Head and Polly, 2015), morphological integration (Goswami et al., 2015), intraspecific variation and sexual dimorphism (Cullen et al., 2014; Drake et al., 2015), taxonomy and phylogenetics (Abdala et al., 2014; Geraads, 2014; Benoit et al., 2015; Sansalone et al., 2015), functional morphology (Fabre et al., 2014; Martín-Serra et al., 2014; Fearon and Varricchio, 2015; Piras et al., 2015), responses to climate change (Meachen et al., 2014; O'Keefe et al., 2014), paleoecology (K. E. Jones et al., 2014; Mallon and Anderson, 2014; Dieleman et al., 2015; Meloro et al., 2015), and ichnology (Castanera et al., 2015; Ledoux and Boudadi-Maligne, 2015), as well as advances in morphometric methodology (Arbour and Brown, 2014; Hetherington et al., 2015).

Finite Element Analysis

Finite element analysis is a computational technique that invokes the mathematical principles of the finite element method to predict the behavior of a structure with defined material properties in response to user-determined loads and constraints. The technique is commonplace in engineering analysis (see review of its early development in Clough, 1990) and has been used extensively in orthopedic medicine and implant studies for a number of decades (Zienkiewicz et al., 1983). Since the early 2000s, FEA has gained traction as a technique for studying the mechanical behavior and performance of extinct and extant organisms, most commonly vertebrates, with particular emphasis on studies of

the vertebrate skull (Rayfield et al., 2001; Rayfield, 2007; Strait et al., 2007; Wroe et al., 2007; Wroe, 2008; Dumont et al., 2009, 2014; Tseng, 2009; Tseng and Wang, 2010; Fitton et al., 2012; Porro et al., 2013) or teeth (Spears and Macho, 1998; Macho and Spears, 1999; Anderson et al., 2011; Anderson and Rayfield, 2012; see D. Jones et al. [2012a]; D. Jones et al. [2012b]; Murdock et al. [2014]; and Martínez-Pérez et al. [2014] for examples using FEA on conodont elements). Recently, FEA of the postcranial skeleton has been gaining attention (e.g., Schwarz-Wings et al., 2009; Piras et al., 2012; Rega et al., 2012; Brassey et al., 2013; O'Higgins and Milne, 2013). Typically these analyses assume a linear elastic behavior for bone and model loading under static rather than dynamic conditions (but see Wang et al., 2012). Because a number of detailed biologically focused FEA reviews exist (e.g., Richmond et al., 2005; Rayfield, 2007; Panagiotopoulou, 2009), we only briefly describe the method here.

Like GM, FEA begins by capturing the geometry of the anatomical structure of interest, in either two or three dimensions (Fig. 1C, D). Two-dimensional FE models are usually built from outlines of photographs (e.g., Rayfield, 2005; Neenan et al., 2014; Serrano-Fochs et al., 2015), whereas 3D FE models are commonly assembled using X-ray CT (X-ray micro-computed tomography, synchotron radiation micro-computed tomography) or magnetic resonance imaging (e.g., Curtis et al., 2013; Button et al., 2014; Gill et al., 2014; Fortuny et al., 2015; Tseng and Flynn, 2015). Models with simple geometry can also be built using computer-aided design software (e.g., Anderson et al., 2011). Three-dimensional imaging has the advantage of visualizing both the external and internal geometry of biological structures, including internal cavities and thicknesses of cortical or cancellous bone layers, all of which can affect the mechanical behavior of a structure. Once a model has been generated, the digital geometric area or volume is discretized to a finite number of entities with a simple geometry, called elements, all of which are joined at their apices by points called nodes. This constitutes the finite element 'mesh.' Meshes can be composed of different numbers and shapes of elements. An increase in the numbers of elements generates more precise geometric models, but this does not necessarily increase the accuracy of the solution and requires greater computing power. Thus, repeated refinement of the mesh is necessary to create the most precise, accurate, and efficient FEA (Dumont et al., 2005; Bright, 2014).

Once a mesh is generated, elements are assigned material properties, such as the Young's modulus of elasticity and Poisson's ratio of compressibility, which can vary in magnitude or orientation to reflect heterogeneous or anisotropic properties of the source material, respectively. For extant taxa, it is possible to measure the material properties of the bone of interest using ex vivo experimentation (e.g., Krauss et al., 2009; Rhee et al., 2009; Zapata et al., 2010; Magwene and Socha, 2013), but bone of extinct taxa cannot be directly measured. The material properties of fossil bone can be optimized based on the extant phylogenetic bracket (Witmer, 1995) or by comparing bone histological morphology with a selection of comparative material of known properties (Rayfield et al., 2001). Boundary conditions, including constraints against free body motion and loads of defined magnitudes and orientations, are then applied to the FE model to mimic the function to be simulated (e.g., bilateral biting, midstance of a stride). Loads can reflect an external force such as an impact or intrinsic loads such as applied muscle or joint reaction forces. Realistic loads in extant taxa can be estimated via in vivo experimentation (i.e., bite force, ground reaction force), measurement of muscle origins, insertions, and physiological cross-sectional area through cadaveric dissection or, more recently, digital segmentation of contrast enhanced X-ray CT (e.g., Cox et al., 2011). Forces in extinct taxa are unknown, but the extant phylogenetic bracket can be applied to optimize muscle reconstructions, and in some cases muscle scars may be present or the

physical dimensions of the muscle attachment area can be measured to estimate muscle volume and cross-sectional area (e.g., 'dry-skull method'; Thomason, 1991). Whether or not absolute values of loads (muscle, bite, joint force) for FE models are needed depends on the question being asked. For instance, in large comparative studies, either the FE model or the loading parameters may be corrected for size if evolutionary or ontogenetic changes in shape alone also are of interest (see Dumont et al., 2009).

Finite element analyses compute the deflection of nodes within FE models in response to the applied boundary conditions and return values of nodal strains, element stress, and strain energy. These values, or permutations of them, are often used as quantitative indicators of the mechanical performance of structures. Many comparative FE analyses of skeletal structures have assumed that natural selection acts to minimize these values. Low stress and/or strain are considered to point to adaptations for structural strength, whereas low values of strain energy are thought to be indicative of adaptations for energy efficiency (e.g., Tanner et al., 2008; Wroe, 2008; Strait et al., 2010; Dumont et al., 2011). These assumptions are beginning to be challenged by analyses that address specific hypotheses of adaptation using data from relatively large clades with well-documented phylogenies (Dumont et al., 2014). For example, is there a signal of selection for low stress among species that feed on harder foodstuffs? The answer is not always 'yes.' A great deal more population-based and clade-level research is required to understand the implications of improved mechanical performance on fitness and evolutionary success. Indeed, we have yet to understand which metrics generated by FEA are the most suitable predictors of functional prowess. The choice of metric will certainly depend on the questions being asked.

Finite element models can be subject to validation tests by comparing experimentally recorded bone strains to strains computed by an FE model subject to similar loading and boundary conditions (see Bright [2014] for a review). Such analyses compare how accurately computational models can replicate experimental recorded strains. Experimental strains may be recorded in vivo (e.g., Strait et al., 2005; Ross et al., 2011) or ex vivo (see Kupczik et al., 2007; Bright and Rayfield, 2011) using strain gauges, digital speckle interferometry, or photoelastic materials (e.g., Anderson and Rayfield, 2012; Gröning et al., 2012). The challenge is to replicate experimentally derived boundary conditions in silico. For analyses of ex vivo strain, experimental loading conditions (i.e., where the experimental sample is fixed and how it is loaded) can be controlled and replicated in the FE model. In vivo loading parameters may be experimentally measured; for example, via electromyography of muscle activation patterns or by kinematic analysis. Typically these studies focus on the crania of mammals (mostly primates, but also see Bright and Rayfield, 2011), although archosaurs have also been subject to validation tests (Rayfield, 2011; Porro et al., 2013). The more accurately FE models capture material properties of bone, or specimen geometry, the better the prediction of experimentally recorded strain. However, validation studies are able to provide useful information on how well a composite material such as bone, with complex geometry, is approximated by a computerized discretization of digital rendering of hard and soft tissue anatomy. For studies on fossils, where material properties cannot be directly measured and even internal and external geometry can be difficult to assess, validation studies provide boundaries within which sensible questions of functional and mechanical performance may be asked. Also important to assessing the viability of FE models is sensitivity analysis, assessing the degree to which variation in input parameters or model geometry influence model results (e.g., Fitton et al., 2012, 2015). Other tests of model validity include comparing experimentally recorded bite forces to those predicted by FE modeling (Curtis et al., 2010) or

comparing jaw kinematics derived from multibody dynamic analysis (e.g., Shi et al., 2012).

Previous Work on Combining GM and FEA

On their own, GM and FEA are powerful tools for examining form and function of the vertebrate musculoskeletal system, but broader morphological, developmental, and evolutionary questions can be addressed by integrating the two approaches (e.g., O'Higgins et al., 2011, 2012). In its simplest form, GM can be used to 'clean' the geometric data upon which an FEA is performed. For example, landmarks and thin-plate splines (TPS) can be used to reconstruct missing features, retrodeform taphonomically altered fossil specimens, and fill in holes to create water-tight FE models (e.g., Angielczyk and Sheets, 2007; Gunz et al., 2009; Tallman et al., 2014). Geometric morphometric analyses also can form the basis for constructing FE models of realized and hypothetical morphologies within a given morphospace or to select the specimen that is closest to the mean shape in a population (e.g., Pierce et al., 2008, 2009; Cox et al., 2011). More elaborately, GM methods can be harnessed to warp one digital shape into another to generate new meshes for FEA (e.g., Sigal et al., 2008; Stayton, 2011; Parr et al., 2012; Piras et al., 2012; Tseng, 2013). With this approach, comparatively inexpensive landmark data are used to warp an FE mesh built from one specimen (a base model) into different target shapes that represent either real taxa or hypothetical intermediate morphologies whose deformation and strains can then be studied.

Thin-plate spline warping of FE meshes into new morphologies significantly reduces model building time and allows larger-scale comparative studies to be performed. However, a major caveat when warping FE meshes based on CT reconstructions is that the internal architecture of the bone (e.g., cortical thickness, trabecular organization) may not be preserved; thus, the morphology of the target specimen, and analyses based on it, may not be realistic (O'Higgins et al., 2012). Consider a skull with landmark points on its outer surface. As the surface is warped, the cortex is stretched to fit and the correct cortical thickness may not be maintained. Relatively little work has been done to validate whether or not target specimen FEAs based on TPS warping perform similarly to those created from CT reconstructions of the actual target morphology (although see Grassi et al., 2011), and development of protocols to mitigate against such potentially error-inducing effects is an important area for further research. Depending on the questions being asked, a possible solution is to create 2D surface FE models or solid 3D models, particularly when investigating the mechanical behavior of size/shape across ontogeny or through evolution (O'Higgins et al., 2012). Nevertheless, although a combined approach represents an exciting new development and holds great promise for future comparative studies, a number of challenges still remain and this technique should be used with caution (Adams et al., 2013).

Beyond data generation and manipulation, GM has been used to examine strain deformation within and across FEAs and, in particular, to investigate how sensitive FE models are to input parameters (e.g., material properties, load position). Such sensitivity analyses are done by collecting a series of easily identifiable landmarks on an unloaded FE model and on a variety of loaded and deformed FE models after simulation. One common approach is then to compare absolute strain response between the models at equivalent anatomical landmarks, using strain plots or PCA with strain as the variable (O'Higgins et al., 2011). Alternatively, the landmarks themselves can be analyzed via GM to assess global deformation across the whole object (e.g., Cox et al., 2011; Gröning et al., 2011, 2012). However, it is unclear how representations of shape differences (Procrustes distances) from GM relate to the strains described by FEA, particularly when shape changes between landmarks are interpolated (Bright, 2014), because no mathematical theory relating the two currently exists (Weber et al., 2011; Bookstein, 2013a; Adams et al., 2013). Thus, at present, assessment of global deformation using GM is best restricted to sensitivity analyses within one base FE model, rather than cross-species or cross-model studies.

Geometric morphometrics and finite element analysis have been used to explore broad macroevolutionary questions, such as the link between ecomorphology and functional performance. Using traditional statistical modeling techniques (e.g., linear regression, analysis of variance) the relationships between species distribution in morphospace (e.g., PC coordinates) and functional performance (e.g., von Mises stress) can be determined (Pierce et al., 2008, 2009). In this way, evolutionary hypotheses can be formulated about the underlying selective pressures that have shaped the patterns of morphological variation. Following on from this, maximum-likelihood methods can be used to assess which model of evolution (e.g., Brownian motion, directional change, stasis, Ornstein-Uhlenbeck model) best explains the distribution of shape and function on a phylogenetic tree (Young et al., 2011). An integration of FEA and shape transformation has also been used to investigate the evolution of performance optima within a given morphospace (Stayton, 2011; Dumont et al., 2014). In this context, FE models can be built for a range of shapes corresponding to a set of evenly spaced points throughout a morphospace. Performance values (e.g., von Mises stress) from the corresponding models are then assigned to each point in the set, and an interpolation function is used to construct a continuous surface showing expected performance for every shape in morphospace. The method we develop here draws strongly on such performance surfaces by using them as the basis for developing evolutionary models of selection.

Evolutionary Modeling

The theory of quantitative trait evolution is well developed and its principles can be used to model expected changes in phenotypes, including shape, under many different evolutionary scenarios. Growing from the work of Fisher (1930), Haldane (1924), Wright (1932, 1968), and Simpson (1944, 1953), this theory explains how the mean phenotype of a population is expected to respond to selection or random drift processes based on the variance and heritability of the traits, population size, genetic or developmental correlations with other traits, and fitness (Lande, 1976; Lande and Arnold, 1983; Arnold et al., 2001). Central to this theory is Simpson's concept of the adaptive landscape, a three-dimensional (but expandable to any number of dimensions) surface in which the horizontal x and y axes represent phenotypes and the landscape's height in the z axis represents a population's fitness or success in reproducing itself (Fig. 3; Wright, 1932; Simpson, 1944; see recent review in Svensson and Calsbeek, 2012). Selective change in the phenotype is the direct consequence of fitness, so a population's mean phenotype is expected to move upslope on the adaptive landscape's surface toward a local peak; the shape of the peak thus defines the magnitude and direction of selection. The rate and direction of phenotypic change are a consequence not only of selection but also the heritability of the phenotype, the covariances among the traits composing it, and drift, which is a random component of change that arises by chance sampling of offspring phenotypes from the parental pool and thus depends on size of the reproductive population (Lande, 1976).
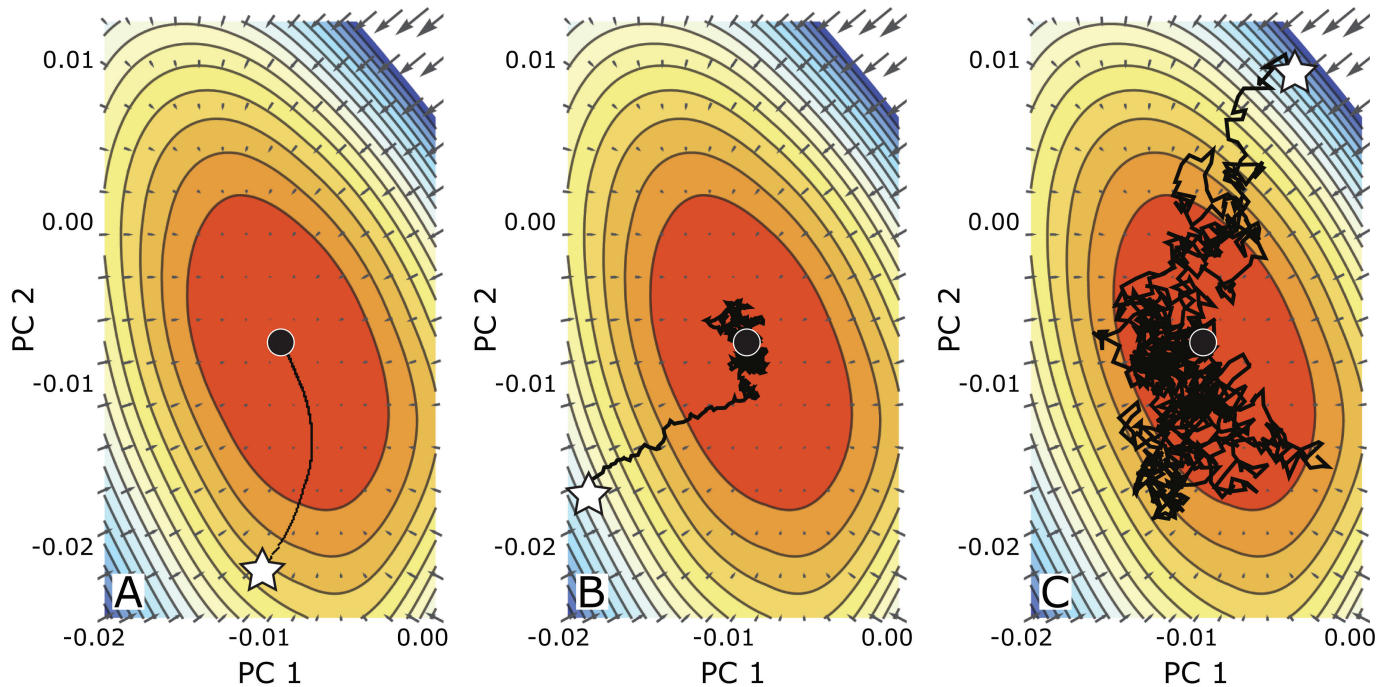
FIGURE 3. Evolution on an adaptive peak for a bivariate quantitative morphological trait, defined here as the first two principal components of turtle shell shape (*x* and *y* axes). Orange is the area of highest fitness and blue is the area of lowest fitness. Arrows show the direction and strength of selection defined by the gradient of the landscape's surface (Eq. (2)). Evolutionary pathways are shown in black. Stars mark the starting points; black dots mark the peaks. Evolutionary change occurs by selection plus drift (Eq. (1)). **A**, when the drift component is zero, selection moves the phenotype directly to the peak following the contours defined by the selection gradients; **B**, when drift is small relative to the selection gradients, the phenotype climbs erratically to the peak and wanders around it in close proximity; **C**, when drift is equal in magnitude to selection, the phenotype wanders widely around the peak.

Evolutionary change in the average phenotype of a species ($\Delta \bar{z}$) on an adaptive landscape (Arnold et al., 2001) can be described as

$$\Delta \bar{z} = G\beta + G/N , \qquad (1)$$

where *G* is the genetic covariance matrix, which is the heritable component of the phenotypic variances of two or more traits; *N* is the reproductive population size (*G/N* is the amount of change by genetic drift); and β is the vector of selection gradients defined by the slope of the adaptive landscape surface $\overline{W}$ evaluated at the trait mean:

$$\beta = \frac{\partial \ln \overline{W}}{\partial \bar{z}} . \qquad (2)$$

If *N* or β is large, then drift becomes a negligible component of change and the second term of Eq. (1) can be ignored (Fig. 3).

Evolutionary patterns that emerge on a topographically complex and changing adaptive landscape can be varied and multifaceted, but they can often be classified into one of three stereotypical modes of evolution: stabilizing processes (including stasis or Ornstein-Uhlenbeck processes), directional evolution, and Brownian motion (random change; Bookstein, 1987; McKinney, 1990; Gingerich, 1993; Hunt, 2006). The classic adaptive peak model of evolution is an example of a stabilizing process because the selection gradients on the slopes point upward, dropping to zero at the top of the peak (Fig. 3). The adaptive peak thus carries the mean phenotypes in populations from their positions on the slopes to a stable point at the peak. Once the peak has been attained, evolutionary change ceases until external

forces provide opportunities for selection to come into play once again. The paths often follow curved trajectories that are defined by the direction of the selection gradient vectors, which are in turn defined by the curvature of the adaptive landscape (Fig. 3; Supplementary Data Movie 1). Directional selection is either a short-lived phenomenon that occurs while the phenotype is being selected toward the peak or it is driven by an adaptive peak whose optimum is changing in a uniform direction. Brownian motion corresponds to random change, which can be generated by several distinct processes. Neutral drift, also known as genetic drift, is a Brownian motion process that occurs by chance sampling of one generation from the previous and is a function of population size (small populations are more prone to drift than large ones). Randomly changing directional selection can also produce a Brownian motion pattern, a process known as selective drift (Kimura, 1954). These two processes can be distinguished by estimating the rate of evolutionary change. The rate of true drift should be small, equal to the heritable component of the population-level phenotypic variance divided by the breeding population size, whereas selective drift may be much faster (Lande, 1976; Polly, 2004). For animated examples of phenotypes evolving under each of these three modes using geometric morphometrics, see Polly (2004).

If evolution occurs according to one of these three simple modes, the outcomes are statistically predictable if parameters such as the rate of change at each step in the evolutionary process, the phenotypic variances of the populations, the heritability, and the shape of the adaptive surface are known. Under Brownian motion, for example, the average outcome of a single evolving lineage will have the same phenotypic mean as the ancestor and a variance equal to the squared step rate (step variance) times the number of temporal steps (Raup, 1977;

McKinney, 1990; Berg, 1993). On a phylogenetic tree, the same outcomes apply to each branch and the variance among the tip taxa is a function of the squared step rate, the time since common ancestry, and the covariances expected from the topology of the tree (Felsenstein, 1988). These properties can be used to estimate the rate of trait evolution and the ancestral trait values at nodes within trees, or they can be used to remove confounding effects of phylogeny from statistical analyses of traits and independent factors (Felsenstein, 1988; Martins and Hansen, 1997). Each mode of evolution has known statistical properties, so the same techniques can be applied to situations other than Brownian motion. In fact, rates, modes, and other parameters can be estimated from data on a phylogeny or within a single lineage by

finding the values that maximize the likelihood under each mode and then selecting the best mode using the Akaike information criterion or similar model selection criteria (Butler and King, 2004; Hunt, 2006; Slater et al., 2011).

Adaptive landscapes are often used to represent the fitness effects of a single selective factor, but they can also be used to represent the net outcome of many selective factors that affect the phenotype (Arnold, 1983, 2003). Complex traits, such as the skeletal elements frequently preserved in the fossil record, perform in more than one functional context and their net performance is a trade-off between competing demands (Fig. 4; Wainwright, 2007). Trade-offs may involve fundamentally different functions, such as locomotion and prey capture, or the same
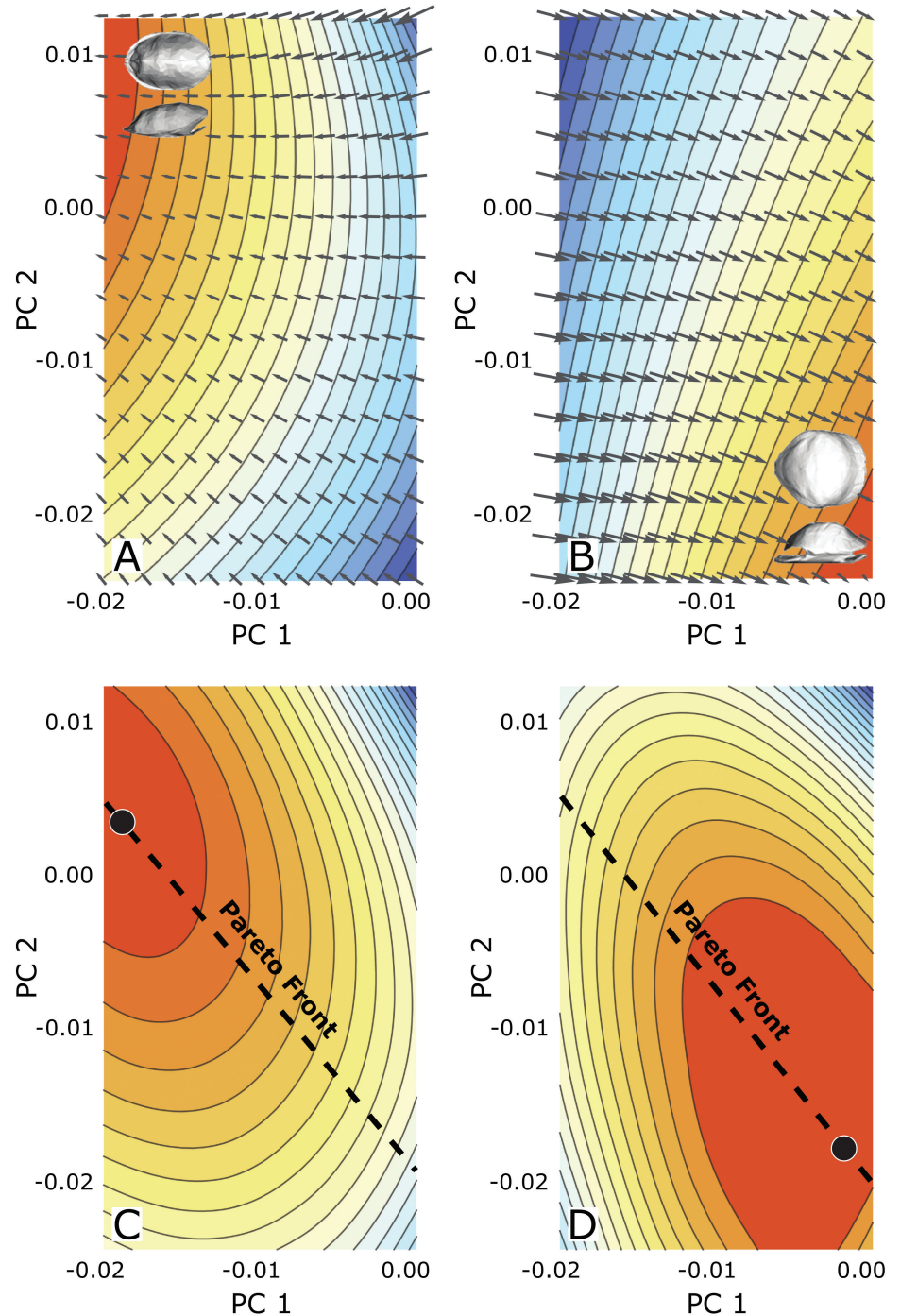
FIGURE 4. Performance surfaces defined by **A**, shell strength (resistance to crushing) and **B**, shell cross-sectional area, a component of hydrodynamic efficiency. Orange is the area of strongest performance (greatest strength and smallest cross-sectional area, respectively) and blue is the area of weakest performance. Arrows show selection gradients associated with optimization of each function. Insets show optimal shell morphologies; **C**, the adaptive landscape that results from combining the two surfaces, with selection for hydrodynamic performance weighted 0.2 relative to selection for shell strength and **D**, with hydrodynamic performance weighted 1.2 relative to selection for shell strength. The selective optimum (black dot) always lies along the Pareto front. Note that the adaptive peak in Fig. 2 is the combination of the two performance surfaces weighted equally.

function in heterogeneous environments, such as locomotion on land and in water, each of which can be quantified as a performance gradient. If performance affects fitness, then each performance gradient will contribute to the net selection gradient (Arnold, 1983). Moreover, if the contours of each performance gradient are known, they can be combined into a single adaptive landscape that describes the net result of selection on the phenotype by all the factors (Levens, 1962; Arnold, 1983, 2003). Depending on the topography of the constituent gradients and the multivariate complexity of the phenotype, the net adaptive landscape may be a simple peak or it may have a ridged or rugged multidimensional topography, including contours of equal performance for different phenotypes, and may even contain holes (Arnold, 2003; Gavrilets, 2004; Wainwright, 2007; Shoval et al., 2012).

Important for our purposes, the performance advantage of a particular phenotype is context dependent, which means that it can change as environments and circumstances change (Simpson, 1944; Lande, 1986; Wade and Kalisz, 1990). Here we will distinguish between the environmental heterogeneity experienced by individuals as they move through space and time, the net effects of which determine the contours of the adaptive landscape of a population or species, and the changing environments that occur over geological time or that are encountered by a lineage as it evolves into a new environmental niche. Both types of environmental change affect the balance of performance factors and, consequently, change the net adaptive landscape and therefore evolutionary trajectories.

SYNTHESIZING GM AND FEA IN AN EVOLUTIONARY CONTEXT: AN EXAMPLE USING TURTLES

A synthesis of GM, FEA, and evolutionary modeling has great potential as a means to explore the shifting balance between aspects of performance in changing environments. To demonstrate this potential, we will use these techniques to measure the changing evolutionary balance between hydrodynamic performance of turtle shells in aquatic environments and the strength of the shell against crushing in any environment (Stayton, 2011). Our analysis focuses on the Emydidae, one of the major extant clades of testudinoid turtles, whose members are commonly known as the New World pond and box turtles (although two emydid species, *Emys orbicularis* and *Emys trinacris*, occur in Europe, western Asia, and North Africa). Emydidae is one of the most speciose extant clades of turtles (Turtle Taxonomy Working Group, 2014) and it consists of two main subclades, the more species-rich but morphologically and ecologically conservative Deirochelyinae and the morphologically disparate but taxonomically depauperate Emydinae (e.g., Gaffney and Meylan, 1988; Stephens and Wiens, 2003, 2008; Wiens et al., 2010). Emydids are of particular interest in the current context because they include species specialized for both aquatic and terrestrial lifestyles, and various aspects of their shell shapes are known to covary with habitat preference (e.g., Claude et al., 2003; Claude, 2006; Angielczyk et al., 2011; Stayton, 2011). In particular, Stayton (2011) showed that there is a broad trade-off between hydrodynamic efficiency and shell strength in emydids, and even within individual emydid species there is fine-scale variation in shape related to strength and streamlining that likely has significant functional and evolutionary implications (Rivera and Stayton, 2011, 2013; Vega and Stayton, 2011; Fish and Stayton, 2014).

Our general approach is diagrammatically illustrated in Fig. 2. We use FEA and three-dimensional geometry to estimate performance factors associated with resistance of the shell to crushing and hydrodynamic performance in the water (Fig. 4; Stayton, 2011). The relationships of shape to hydrodynamic cross-sectional area and shell strength can be modeled as performance surfaces, each of which has a selective effect that varies

depending on habitat, predation, and related factors. The two performance surfaces thus jointly contribute to the overall adaptive landscape insofar as the performance of shell shape in the local environment contributes to the total fitness of emydid populations. Stayton (2011) showed that the evolution of shell shape is subject to a trade-off between optimization for reduced hydrodynamic drag and optimization for greater resistance to deformation during predatory attacks, because increasing one property tends to decrease the other. The net fitness of shell shape is therefore a trade-off between how much a species relies on efficient movement in water and how frequently it is subject to crushing attacks. Functional performance of shell shape can be determined empirically because it arises from the structural and mechanical properties of this trade-off, but its effect on fitness depends on ecology. We use GM and maximum likelihood to estimate the changing balance between these competing factors on a phylogenetic tree, and we ask whether the two performance factors are sufficient to explain the diversification of emydid shell form.

Methods

Fifty-nine landmarks were used to quantify shell shape in 60 specimens of 47 species of emydid turtle (Supplementary Data Table 1; part of a larger data set of 385 specimens of 238 species from all hard-shelled turtle families). All landmarks represent triple junctions between scutes (enlarged scales found on the surface of most turtle shells) or maxima of curvature along intersections between two scutes. Included in this data set was a specimen of *Glyptemys muhlenbergii* (FSM 85274) that had been CT scanned at the University of Texas at Austin's high-resolution X-ray CT facility. The scan consisted of 1,215 slices, each 0.1 mm thick with a field of reconstruction of 90 mm and 0.1 mm between slices. The CT data were used to build a 41,615 element, 10,629-node FE model of the turtle's shell (see Stayton [2009] for additional details of model construction). A set of four restraints, permitting rotation but not translation, were placed on the carapace of the model, close to the bridge. Eight evenly-spaced load cases were assigned to the carapace of the model, each consisting of a 200 N load directed normal to the surface of the shell. All elements were assigned a Young's modulus of 10 GPa and a Poisson's ratio of 0.3 (typical values for turtle shell bone; Stayton, 2009; Magwene and Socha, 2013).

A GM relative warps analysis was conducted on the landmark data to ordinate all emydid turtle specimens in a multivariate shape space (Supplementary Data Fig. 1). Performance surfaces were constructed for the region of morphospace occupied by the emydid species, along the first two PCs, as follows. A set of 42 theoretical shapes, located at all combinations of six values along the first PC axis and seven along the second PC axis, with average values for emydids along all other PC axes, was generated. These values were chosen to cover the region of shape space occupied by emydid turtles, as well as some distance beyond (examples of the theoretical shapes and their positions in morphospace are shown in Supplementary Data Fig. 2). A thin-plate spline interpolation function was then generated for the transformation of the CT-scanned *Glyptemys muhlenbergii* landmarks to the landmark coordinates of each of the 42 theoretical shapes. These interpolation functions were applied to the coordinates of each of the nodes in the *G. muhlenbergii* FE model, resulting in 42 new FE models, each with the same number of elements, loads and constraints as the base FE model (see Stayton [2009] for details on the warping procedure). All models were scaled to the same surface area, to ensure that differences in mechanical performance were entirely due to shape (Dumont et al., 2009). Each of the 42 new FE models was analyzed using a linear elastic model; von Mises stresses were extracted for each element for each load case and then averaged to determine mean stress for

each model. Because all loads were at the same location and of the same magnitude, higher stresses indicate weaker shells. Performance surfaces were constructed for each of these functional parameters by fitting a third-order polynomial surface to the averaged von Mises stresses (z axis) as arranged on the first two dimensions of the morphospace (x, y axes). Finally, to determine hydrodynamic performance, the frontal area for each FE model was calculated by constructing a minimum convex hull around the set of all nodes projected into the y, z plane (i.e., a frontal view of the shell) and then calculating the area inside that hull.

We used maximum likelihood to find the proportional balance between the two performance surfaces that best explains the shape of emydid shells. Arnold (1983, 2003) showed how the phenotype–fitness relationship that defines an adaptive landscape $\overline{W}$ can be decomposed into separate selective factors (β_i), each of which contributes separately to net fitness. Furthermore, each selective factor can be broken into components that link phenotype to performance (β_{Fi}) and performance to fitness (β_{Wi}). Shell strength and hydrodynamics are thus both performance factors whose relevance to fitness is determined by the ecology of the species, including the combined importance of being able to resist predatory attacks and to maneuver efficiently in aquatic environments. When two performance factors act simultaneously on the phenotype, their expected net effect can be described by substituting the constituent factors into Eq. (1):

$$\Delta\overline{z} = G(\beta_{F1}\beta_{W1} + \beta_{F2}\beta_{W2}) + \varepsilon, \qquad (3)$$

where ε represents other unconsidered functional factors that affect the fitness of the phenotype (in the current example, factors such as the internal volume of the shell or its self-righting capability). Note that for fossils and many living taxa, we do not know G, the additive genetic covariance matrix of the phenotype. We can, however, estimate P, the phenotypic covariance matrix, from morphometric analysis, although care should be taken to use a sample of adequate size (Cheverud, 1982; Polly, 2004; Goswami and Polly, 2010). Furthermore, we know that $P = G + E$ (Lande, 1976), where E is the nongenetic component of the phenotype. Note that though P is related to G, it is not identical. This means that when G is unknown, P can only be substituted by considering E to be part of the net error term (Cheverud, 1988). In fact, we can make a further simplification by transferring all of the phenotypic covariances to the error term, allowing us to focus purely on the effects of the performance surfaces F_i, whose slopes and curvatures define the direction and magnitude of β_{Fi}.

With those simplifications, the adaptive landscape surface can be described as the summation of the two performance surfaces weighted by their relative fitness w,

$$\overline{W} = w_1F_1 + w_2F_2 + \varepsilon. \qquad (4)$$

Equation (4) is the key to estimating the trade-off between performance in multiple functions: hydrodynamics and shell strength in our example. Because selection is expected to maximize fitness, we expect to find the phenotype $\overline{z}$ at the highest point on $\overline{W}$ (Figs. 3, 4). We know F_i from the FEA performance data, but we do not know w_i, which is determined by the ecology of the species. Hydrodynamic performance is likely to make a stronger contribution to fitness in aquatic species, whereas shell strength is likely to contribute to fitness in all species but will vary by their exposure to predator attack, their ability to avoid attack, and other factors. However, if we assume that selection has maximized fitness, then we can estimate the values of w_i by finding the combination that results in the peak of $\overline{W}$ being as close as possible to the observed shell shape. The most likely values of w_i are the ones that maximize the fitness of the phenotype

(its height on $\overline{W}$ relative to the height of the peak) given shell shape and the two fitness surfaces, the log likelihood of which is

$$l(w_1, w_2 \mid \overline{z}, F_1, F_2) = \ln \frac{w_1F_1[\overline{z}] + w_2F_2[\overline{z}]}{Max[w_1F_1 + w_2F_2]}, \qquad (5)$$

where F_i is the multivariate function for the height of performance surface i at a particular location in the morphospace, $\overline{z}$ is the position of a particular shell in the morphospace, and $Max[w_1F_1 + F_2F_2]$ is the optimum, or peak, of the combined performance surfaces. For a given value of w_i, the log likelihood will be zero at the peak and negative everywhere else; maximizing this function for a given shell shape ($\overline{z}$) thus finds the w_i for which $\overline{z}$ is nearest the height of the peak. Note that as both w_i values decrease, the adaptive peak becomes lower and broader. In fact, if $w1 = w2 = 0$, then $\overline{W}$ flattens into a plane and all phenotypes become equally likely under Eq. (5). Thus, if both w_i parameters are left free to vary, then maximum likelihood will always converge on $w_i = 0$. This problem can be avoided by holding one weight constant and using likelihood to estimate the proportional weight of the other parameter. Here we set the weight for shell strength function to 1 ($w_1 = 1$) because shell strength is likely to be important for most turtles (softshell turtles, the pancake tortoise, and the leatherback sea turtle are potential exceptions), regardless of their habitat, and thus makes a useful basis of comparison with hydrodynamic efficiency, which probably varies widely in importance among species. Equation (5) is written for the mean shell shape of a single species, but it can be applied to a group by summing the log likelihoods across all $\overline{z}$ to find the trade-off weight that collectively maximizes the fitness of the group as a whole.

Note that the peak values of $\overline{W}$ all fall along a single line or curve no matter what values w_i takes on (Fig. 4; Supplementary Data Movie 1). Such a line or curve, known as the Pareto front, arises in any situation in which optimization is a trade-off between two contributing factors (Shoval et al., 2012). The Pareto front represents a series of points along which no improvement can be made in the performance of one function without a decrease in performance in another function. Such a front can be important for interpreting the relationship between form and function because it describes the morphologies that result from selection for both shell strength and cross-sectional area. Shell shapes that lie away from the Pareto front have not been optimized for both factors and thus represent departures from any hypothesis that shell strength and cross-sectional area are the predominant factors governing evolution of emydid morphology.

Trade-off weights (w_2) were mapped onto a phylogenetic tree using the generalized linear model method of Martins and Hansen (1997) with a Brownian motion model, which gives results identical to those of the squared-change parsimony method of Maddison (1991) and the one-rate maximum likelihood method of Mooers and Schluter (1997). Phylogeny and original branch lengths are from Wiens et al. (2010), the same phylogeny used by Stayton (2011). Branch lengths were transformed to time since divergence using the method of Pagel (1992). The age of the base of the tree was set at 40 million years ago using the estimated age of the last common ancestor of crown Emydidae as calibrated by Joyce et al. (2013) and Warnock et al. (2015), based on the oldest occurrence of *Chrysemys antiqua* in the Chadronian North American Land Mammal Age (NALMA) of South Dakota (Hutchison, 1996; note that this calibration is younger than that used by Lourenço et al., 2012).

Our analyses were limited to the first two principal components of shell morphospace, which account for 26% of the total variance. Note that the morphometric analysis from which this morphospace was derived considered a larger clade than just emydids; readers interested in the details of the morphospace

construction and patterns of shape variation in the full clade are referred to the original paper (Stayton, 2011). Ideally, our analyses could (and should) be carried out on the full multidimensional morphospace, but the analytical and conceptual complexity increases enormously with three or more dimensions because the performance surfaces become multidimensional manifolds that are not easily illustrated or described (Arnold, 2003). Two dimensions serve better for demonstrating the principle, which is our intention here. Our results should be interpreted with this limitation in mind. Customized permutation and Monte Carlo tests were used to assess statistical significance as needed (Manly, 1991; Kowalewski and Novack-Gottshall, 2010). Each test is described along with the corresponding results below.

Evolution of Functional Trade-offs in Emydidae

Here we apply the quantitative evolutionary framework described above to test hypotheses about the role of functional morphology in the evolution of form that cannot be tested using either GM or FEA in isolation. Although we use the roles of shell strength and hydrodynamic performance in the evolution of shell shape in emydid turtles as an example, the following questions are of a form that can be generalized to almost any clade.

Can Shell Shape be Explained by the Two Functional Performance Surfaces?—If strength and cross-sectional area are the primary factors influencing the evolution of shell shape in emydids, then selection should optimize the morphology of all taxa to lie exactly along the Pareto front, regardless of the precise weighting of the trade-off between the two factors, because that front represents the peak of the adaptive landscape derived from them. Clearly not all of the morphologies lie along this line (Fig. 5A), although some are quite close, which means that the two factors do not completely explain emydid shell evolution. This result is not unexpected because it would be surprising if just two factors could explain completely the evolution of a complex morphological structure in a diverse clade. Nevertheless, the value of a quantitative evolutionary framework is that the relative importance of the factors can be assessed further.

Is Selection for Smaller Cross-sectional Area Stronger in Aquatic Species?—This question can be addressed simply by estimating the weight of the hydrodynamic performance gradient (w_2) for each species and calculating the mean for aquatic and terrestrial species. The weight w_2 is closely related to the position of each species relative to the Pareto front. Close inspection of Fig. 5 and Supplementary Data Movie 1 shows that the contours of the combined performance surface are complex, but the terrestrial species fall closer to the end where $w_2 = 0.6$ (weak hydrodynamic contribution) than to the end where $w_2 = 1.2$ (strong hydrodynamic contribution; Fig. 5C–E). Estimated w_2 weights for all taxa can be found in Appendix 1. Importantly, w_2 is literally an estimate of the strength of selection for smaller cross-sectional area relative to the strength of selection for shell strength.

On average, selection for smaller cross-sectional area is stronger in aquatic species ($w_2 = 1.00$) than terrestrial species ($w_2 = 0.949$; Appendix 1), and these numbers imply that selection for strong shells and small cross-sectional areas are almost perfectly balanced in aquatic species. The significance of the difference in selection between aquatic and terrestrial species can be assessed by a permutation test in which the habitat category is randomly permuted among the species and the difference between means recalculated. By repeating the permutation many times, the observed difference can be tested against the distribution of differences that arise by chance. A test with 10,000 permutations indicates that the difference is not significant ($p = 0.149$).

The difference between aquatic and terrestrial turtles fails the significance test because *Glyptemys insculpta* and *Glyptemys*

muhlenbergii are classified as terrestrial but have high w_2 values (the 'T' taxa with PC2 values less than −0.01 in Fig. 5A). If these two species are dropped, the permutation test is significant ($p = 0.014$). In some ways, these species are 'the exception that proves the rule' because they are perhaps better described as amphibious than strictly terrestrial or aquatic. *Glyptemys inculpta* typically spends late spring and summer in fields, forests, and wetland environments within a few hundred meters of slow moving streams, whereas they spend more time in aquatic habitats in spring, fall and during hibernation (Ernst et al., 1994). *Glyptemys muhlenbergii* is a habitat specialist found in spring-fed sphagnum bogs, marshy meadows, and swamps. Slow-moving brooks with muddy substrates, shallow pools, and a mixture of wet and dry substrates are important habitat features (Ernst et al., 1994). Although these habitat preferences would be consistent with relaxed selective pressure for a streamlined shape compared to highly aquatic taxa such as many of the *Trachemys* or *Pseudemys* species in our analysis, continued interaction with aquatic environments by the *Glyptemys* species may mean that they still experience more selection than highly terrestrial species such as *Terrapene ornata*, which favors grasslands and spends very little time in standing or flowing water (Ernst et al., 1994). Alternatively, there may be other factors that favor a shell with a small cross-sectional area in *Glyptemys*, such as the need to navigate in dense wetland vegetation or leaf litter, climbing ability, or even constraints imposed by the unusually small body size of *G. muhlenbergii* (see Angielczyk and Feldman [2013] for a discussion of the effects of size and ontogeny on plastron shape in *Glyptemys*).

Does the Trade-off Between the Two Functional Factors Constrain Variation in Shell Shape?—Regardless of whether the balance of selection differs between aquatic and terrestrial species, selection for shell strength balanced against its hydrodynamic performance could still play an important role in emydid evolution. Our likelihood framework allows this question to be tested by determining whether the real shell morphologies are more likely under the functional trade-off model than are random shell morphologies. In other words, we can ask whether the real shell morphologies are closer to functional peaks than expected by chance.

The likelihoods of w_2 for each species (Appendix 1) can be thought of as measures of the height of the species on the combined performance surface at its optimal trade-off value. Summing the likelihoods provides a measure of the likelihood of shell shape for all emydids under a model in which shell strength and hydrodynamic performance are both contributing selective factors. If all shells lie along the Pareto front, the summed likelihood would be 0; for the real shell data, the summed likelihood was −1.403.

We tested whether this value was greater than expected by chance using a Monte Carlo test in which we randomly generate points within the morphospace defined by the emydid turtles, one for each species ($N = 47$), and estimate their combined optimal likelihood by maximizing Eq. (5) and summing across the species. Repeating this a large number of times approximates the distribution of likelihoods that would arise by chance, against which the real likelihood can be tested. Only five out of 1,000 Monte Carlo simulations had likelihoods equal to or greater than −1.403, which means that the probability of this pattern arising by chance is $p = 0.005$. Therefore, emydid shell shapes are significantly closer to the optimal trade-off than expected by chance, implying that shell strength and streamlining do constrain variation within the clade.

How Much Variation in Shell Morphology Can Be Explained by the Two Functional Factors?—Having confirmed that shell shape in emydids is constrained by the trade-off between strength and hydrodynamic efficiency, how strongly has the trade-off influenced morphology? One way of answering this
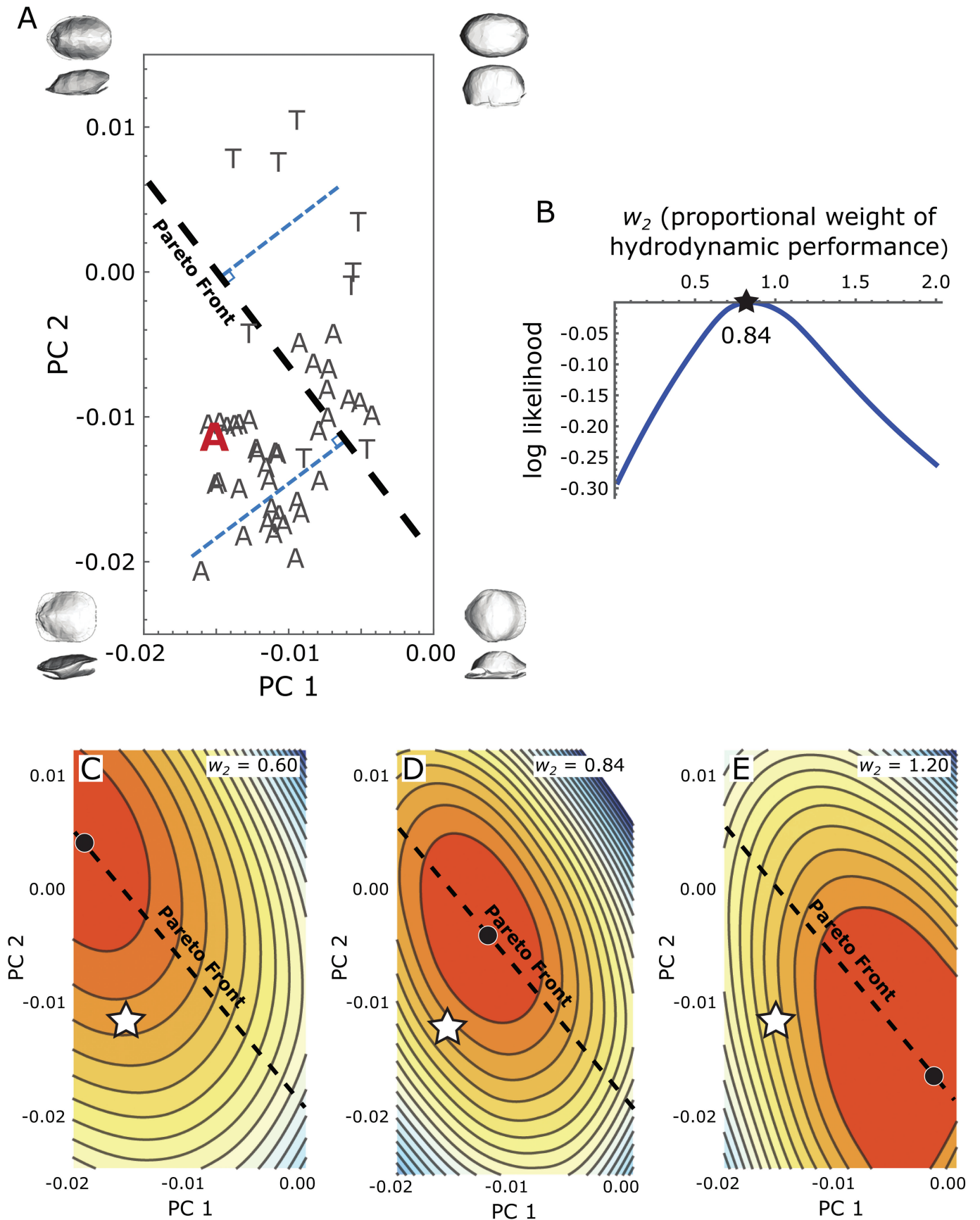
question is to measure the proportion of variance in shell shape that is correlated with the Pareto front (R^2), which represents the expected position of shells in morphospace if they were optimized by the two functional factors.

The functional trade-off as measured by the Pareto front explains 54.6% of the variance in shell shape ($R^2 = 0.564$), which is only marginally significant ($p = 0.0662$). R^2 was estimated as $1 - \text{RSS}_{\text{Pareto}}/\text{SS}_{\text{Total}}$, where $\text{RSS}_{\text{Pareto}}$ is the residual sum of squares around the Pareto front line and SS_{Total} is the total sum of squares of the PC1 and PC2 scores. Note that residual distances are calculated orthogonal to the Pareto front, similar to an orthogonal or reduced major axis regression, because variation around the line occurs along both the x and y axes. The p-values were derived from a Monte Carlo distribution of R^2 for 10,000 sets of random points in morphospace, each with $N = 47$. This result suggests that an unknown factor or factors is systematically causing deviation of shell shape from the optimal trade-off value between strength and cross-sectional area.

Note, however, that because of the curvature of the combined performance surfaces, points that are seemingly far from the center of the peak can still be near the top and thus in a favorable position in terms of functional optimization (cf. Supplementary Data Movie 1). This nonlinearity between the location of the Pareto front and the contours of the performance surface complicates interpretation of this particular result.

What Factors Prevent Shell Shape from Being Optimized by the Two Performance Factors?—The results so far show that shell shape is optimized for a trade-off between strength and hydrodynamic efficiency more than expected by chance but that the variance explained by the trade-off is less than expected by chance. Furthermore, terrestrial and aquatic taxa do not differ significantly in the extent to which they are optimized for hydrodynamic performance (primarily because of the hydrodynamic shapes of the two terrestrial *Glyptemys* species). Possible explanations for the departure include genetic drift, phylogenetic inertia, the influence of genetic covariances, and the operation of additional unknown selective factors. We consider each of these possibilities in turn.

In principle, drift is unlikely to explain the departure of shell morphology from the strength–hydrodynamic trade-off because the variation away from the optimum is not random. Above we showed that less variance was explained by the Pareto front than expected for random shell shapes because more variation in the data set is orthogonal to the front than expected by chance. Drift is an example of a chance process and will produce a fairly circular distribution of phenotypes (Fig. 3C).

Note, however, that drift is more than fast enough to explain some of the observed divergence among emydid shell shapes, though probably not all of it. The rate of drift per generation is a function of the genetic variance G and population size N (Eq. (1)). The genetic variance is the product of the phenotypic variance P and heritability h^2. The amount of phenotypic variance that will accumulate by drift alone in the absence of selection can be estimated as h^2Pt/N, where t is the time in generations (Lande, 1976). We do not know the true values any of these parameters for emydids, but we can make reasonable estimates. The relevant population size parameter of interest is the average size of the breeding population of each entire species (not local population size) over the time span that the group has been evolving. Local population sizes of *Chrysemys picta* often exceed 1,000 (Wilbur, 1975), and the

minimum viable population size of the species has been estimated at about 7,500 individuals (Reed et al., 2003). The clade has been evolving for about 40 million years (Warnock et al., 2015) and generation length in *C. picta* is on average about 10 years (Wilbur, 1975). If we start with these values and assume h^2 to be about 0.5 (Cheverud, 1988), then a phenotypic variance of about 2×10^{-7}, or a total range of within-species variation of about 0.001, is needed on each PC axis in order to account for the variation among species (Fig. 6A). Real within-species variation in shell shape is probably larger than this. Angielczyk et al.'s (2011) study of variation in emydid plastron shape found that within-species variation in several species marginally overlapped with other species on the first two PCs of morphospace (e.g., *Emys blandingii* and *Emys orbicularis*). If carapace shape is proportionally as variable, then we can estimate a within-species range of variation of about 0.01 for each species, which is equivalent to a phenotypic variance of about 6×10^{-6}. Our value of 7,500 individuals is likely a gross underestimate of the breeding population for *C. picta* across its range, however. If heritability is higher at $h^2 = 0.75$ and if population size is a more reasonable 100,000 breeding individuals per species, then drift could have produced a wider range of phenotypes than is observed (ca. 0.05 units on each PC axis; Fig. 6).

The estimated rate of drift therefore suggests that selection is acting to keep emydid shell shape closer to the Pareto front than expected from drift alone, similar to the scenario depicted in Fig. 3C. In that scenario, selection gradients are relatively weak compared to drift, allowing for substantial departure from the optimum value while still keeping the distribution near the peak. Selection plus drift in nearly equal proportions would therefore be enough to explain the observed shell shapes if it were not for the observation that they are not distributed randomly with respect to the Pareto front.

Phylogenetic inertia is unlikely to explain the departure of shell shape from the Pareto front. Phylogenetic inertia refers to the inheritance of traits from an ancestor, implying that selection has not had the opportunity to transform the ancestral form to its current functional optimum. Given favorable conditions of high heritability and small population size, drift alone can account for the occupation of all of the morphospace in our PCA plots without selection. This suggests that selection (which by definition is faster than drift) has had ample time to optimize species along the Pareto front if it was going to do so. Furthermore, the Pareto front has not only been reached but crossed independently at least 11 times by different branches (Fig. 6A). Finally, there is little phylogenetic structure to trade-off weighting (w_2; Fig. 6B), contrary to what would be expected if selection were acting slowly enough that species had not had time to reach the optimum. Common ancestry alone, therefore, does not explain why some taxa lie away from the Pareto front.

Within-species phenotypic and genetic covariances influence the path of phenotypic evolution (Arnold et al., 2001) and can therefore affect the distribution of species around the optimum. Such covariances can arise from developmental integration or modularity, pleiotropy, or correlated selection. Within-species covariance structures are unknown for our taxa (they form part of the error term in Eq. (4)), but they are insufficient to explain the failure of species to reach the optimum. Although these covariances would affect the path taken by the phenotype as it approaches the selective peak (cf. Fig. 3A, Supplementary Data Movie 1), they do not prevent the phenotype from reaching the

←FIGURE 5. Estimating trade-offs for real taxa using maximum likelihood. **A**, positions of aquatic (A) and terrestrial (T) species in morphospace. *Chrysemys picta* is highlighted in red (see Supplementary Data Fig. 1 for key to taxa). Two hypothetical selective factors are shown in blue at right angles to the Pareto front; **B**, likelihood curve of the proportional weight of the cross-sectional area performance surface (w^2 from Eq. (5)) for *C. picta*. At the optimal value of $w^2 = 0.84$, *C. picta* (star) is higher on the peak, **D**, than for other values of w^2 (**C**, **E**). Black dots mark the peaks. Insets show shell morphology at each corner of the morphospace.
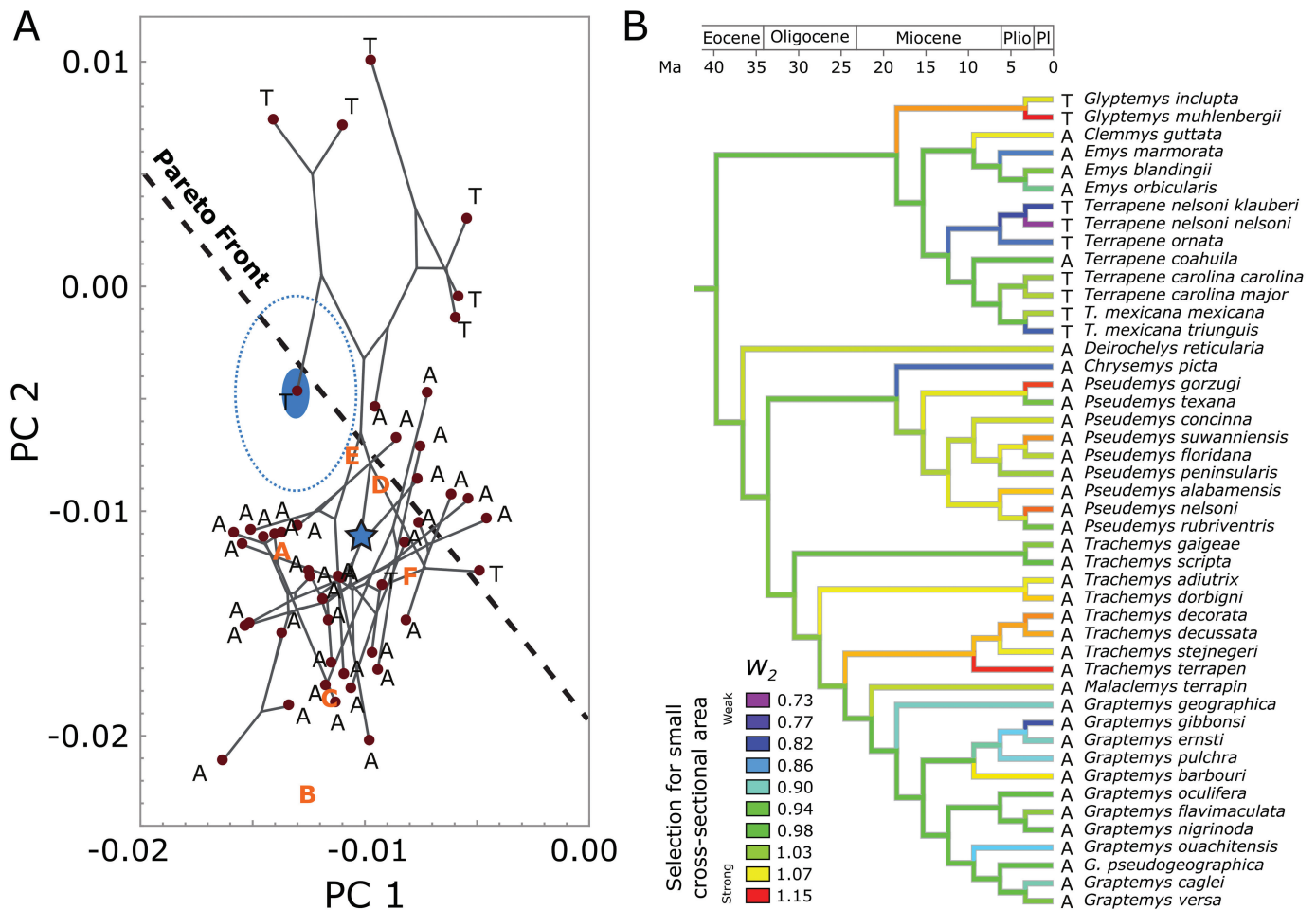
FIGURE 6. Evolution of shell shape and phylogenetic change in selection for hydrodynamic performance. **A**, phylogenetic tree projected into principal component space provides an estimate of the paths shell shape has taken during emydid evolution. Star marks the root of the tree; solid blue ellipse shows the within-species variation needed to explain divergence by drift; dotted blue ellipse shows estimated within-species variation (see text for details). Black symbols represent extant taxa (A = aquatic; T = terrestrial; see Supplementary Data Fig. 1 for key to taxa). Orange symbols represent fossil taxa (A = *Echmatemys septaria*; B = *Echmatemys* sp.; C = *Echmatemys wyomingensis*; D = *Emys blandingii*; E = *Pseudograptemys inornata*; F = *Glyptemys valentinensis*); **B**, mapping of w_2, the proportional weight of selection on cross-sectional area, on a phylogenetic tree. A value of 1.0 indicates that selection for cross-sectional area is proportionally as strong as selection for shell strength. Taxonomy of the genus *Emys* is based on Feldman and Parham (2002); taxonomy of the genus *Terrapene* is based on Martin et al. (2013).

peak per se as long as there is heritable variance in all directions of the morphospace. The fact that taxa lie in both directions from the Pareto front indicates that this appropriate variance does exist in our study system.

Given the inability of drift, phylogenetic inertia, and genetic covariances to explain the departure of emydid shell shape from the Pareto front, the best interpretation is that additional selective factors are acting on shell shape. We did not address other potential functions, but our results provide interesting patterns for future investigation. Note that the aquatic taxa appear to be distributed at right angles to the Pareto front toward the bottom left corner of morphospace and the terrestrial taxa are distributed at right angles toward the top right corner (Fig. 5A). This distribution shows that on average aquatic turtles have carapaces that are more shallow posteriorly with a smaller posterior opening than expected from optimization of cross-sectional area and strength alone, whereas terrestrial turtles have more anteroposteriorly symmetrical doming of the carapace. The carapace shape of aquatic turtles might permit greater forelimb and head maneuverability while swimming, easier access to the surface for breathing in aquatic turtles, or better leverage for hauling out of

the water, all of which would be irrelevant to terrestrial turtles (Pace et al., 2001). Conversely, a domed morphology might provide greater ability for an upside down terrestrial turtle to right itself (Domokos and Várkonyi, 2008; Stayton, 2011), which would be less of a concern to a dominantly aquatic turtle. The unexpected departures in shell shape also change the surface area–to-volume ratios of the shells, which would have important implications for heat transfer (Bramble, 1974). Aquatic emydids bask more frequently than terrestrial emydids, and a relatively flat shell could facilitate rapid heat exchange while they are out of the water. Testing these hypotheses is beyond the scope of this analysis, but they could be investigated using the same quantitative framework that we used here to test the roles of shell strength and cross-sectional area. Results from other functional or physiological analyses could be used to describe the expected performance of shell shape relative to these functional roles, which could then be used to estimate the selection gradients needed to assess their contribution relative to the factors we assessed here.

Estimating Function from Form in Fossils—Our study is based on extant taxa, but the same methods can be applied to clades

that are entirely extinct. Indeed, GM and FEA are capable of quantifying shape variation and functional performance in extinct taxa, particularly in comparison to closely related modern forms, and in broad investigations of the influence of size/shape on simulated function. Our framework does not require many individuals of the same taxon because it does not rely on estimating within-species variation, but it should be noted that FEA requires at least one complete, undistorted (or retrodeformed) specimen that preserves the morphology of interest for most or all taxa in a clade.

In addition to direct application of the method to completely extinct taxa, it can be used to make inferences about fossil taxa using performance surfaces derived from their living relatives. For example, we used GM to project the carapace shapes of six fossil emydids into the shape space defined by our living species (Fig. 6; also see Supplementary Data Fig. 1). Once inside the morphospace, the same likelihood procedure can be applied to estimate the w_2 parameter for each fossil that describes the strength of selection on hydrodynamic efficiency relative to selection on shell strength (Appendix 1; also see Supplementary Data Fig. 3).

The first specimen we considered is a fossil occurrence of the extant species *Emys blandingii* (FMNH PR 428; late Pleistocene or Recent of Pulaski County, Indiana). Its w_2 value (1.008) is nearly identical to its living conspecific (1.012), indicating that both are equally optimized for an aquatic lifestyle. An extinct emydid species, *Glyptemys valentinensis*, from the middle Miocene (Barstovian) of Nebraska (Holman and Fritz, 2001) is represented by UNSM 76233. The w_2 value of this specimen, 1.087, suggests strong selection for a hydrodynamically efficient cross-sectional shape. Interestingly, the w_2 value of USNM 76233 is intermediate between the living species *G. insculpta* ($w_2 = 1.060$) and *G. muhlenbergii* ($w_2 = 1.151$), although it is closer to the former. This is consistent with Holman and Fritz's (2001) conclusion that the overall shell morphology of *G. valentinensis* was more similar to *G. insculpta* than *G. muhlenbergii*. The phylogenetic position of *G. valentinensis* has not been formally examined, but Holman and Fritz (2001) suggested that it might be the common ancestor of both extant *Glyptemys* species or a direct ancestor of *G. insculpta* after its lineage diverged from that of *G. muhlenbergii*. If the former scenario is correct, then it would imply that the selective regimes of the extant *Glyptemys* species diverged in 'opposite' directions with that of *G. muhlenbergii* favoring greater streamlining and that of *G. insculpta* favoring less streamlining (Angielczyk and Feldman [2013] noted the possibility of a similar divergence in the species' plastron shape ontogenies). If the latter scenario is correct, *G. valentinensis* would have less direct importance for our understanding of evolution of shell shape in *G. muhlenbergii*, but it would still imply relaxed selection for streamlining in *G. insculpta*. A reduction in the importance of streamlining in *G. insculpta* relative to its ancestors is consistent with it having a more terrestrial lifestyle than many emydids (see above), although selection for streamlining still influences it much more strongly than the terrestrial box turtles in the genus *Terrapene*.

An older fossil (FMNH PR 589) is a nearly complete shell of *Pseudograptemys inornata* from the late Eocene (Chadronian) of South Dakota. Material currently assigned to *P. inornata* has been referred to both *Graptemys* and *Chrysemys* in the past, but Hutchinson (1996) erected the genus *Pseudograptemys* to reflect the fact that these specimens cannot be assigned with certainty to either extant genus. Indeed, Hutchinson (1996) questioned whether *Pseudograptemys* was an emydid at all because it possesses some characters of the shell that also occur within Geoemydidae, and Joyce et al. (2013) chose to avoid it as a calibration point for Emydidae because of this problem. Despite this taxonomic uncertainty, we consider *Pseudograptemys inorata* to be a useful addition to our analysis because it provides insight into the likely shell shape of turtles near the base of the emydid radiation. Specimen FMNH PR 589 has a w_2 of 0.96, which falls toward the more streamlined end of the range of variation displayed by modern members of *Graptemys*, and is somewhat more streamlined than *Chrysemys picta*. The w_2 value at the base of the tree (Fig. 6) is 1.009, reconstructed from the living taxa using a Brownian motion model of evolution. Although the reconstructed weight is higher than *P. inorata*, it is fairly close given the range of w_2 values in other taxa (Appendix 1) and well within the broad confidence intervals of Brownian motion reconstructions of deep nodes (cf. Polly, 2001; Gómez-Robles et al., 2013).

Perhaps more interesting are the results from three *Echmatemys* specimens, which have notably different w_2 weights despite presumably being very closely related. In North America, *Echmatemys* first appears in the early Eocene (Wasatchian; Hutchinson, 1998), and following Hirayama (1984) it is usually considered to be a basal geoemydid. However, Hirayama (1984) noted that the genus may well be a polyphyletic plesion, so the species we consider here could be basal geoemydids, basal testugurians, or even basal testudinoids. *Echmatemys septaria* is the most common 'geoemydid' in the middle Eocene Green River Basin, which was dominated by large inland lakes, and is found in specifically lacustrine environments (Bartels, 1993; Zonneveld et al., 2000). The specimen of *Ech. septaria* we examined (FMNH P 27242) has a w_2 of 0.913, in the bottom quarter of our modern sample, indicating that selection for hydrodynamic efficiency was comparatively low but similar to *Emys orbicularis*, the European pond turtle (Appendix 1). In contrast, *Ech. wyomingensis* (FMNH UC 1429), another Green River species that is associated with fluvial environments well upstream from lake margins (Bartels, 1993; Zonneveld et al., 2000), has a w_2 value of 1.076. This value falls in the top quarter of our extant sample and is similar to *Graptemys barbouri*, which is associated with clear, fluvial environments (Ewert et al., 2006; Appendix 1). The third specimen we considered (FMNH PR 98) is only identified as *Echmatemys* sp. It was collected in the Green River Formation in Garfield County, Colorado, but the available provenance data do not include detailed information about the specimen's depositional environment and potential habitat. This specimen has the highest w_2 value of our *Echmatemys* specimens, 1.096, which is similar to the value for the extant *Pseudemys suwaniensis*, a species that typically inhabits rivers with slow to moderate currents but spends little time out of water aside from basking and nesting (Ernst et al., 1994). The performance surface parameters for *Echamtemys* are consistent with habitat interpretations based on available sedimentological and independent functional morphological data, and they hint that there may have been divergence in shell shape among species living in different flow regimes similar to that observed for extant emydids (e.g., Lubcke and Wilson, 2007; Rivera, 2008; Rivera and Stayton, 2011; Selman, 2012; Ennen et al., 2014; Rivera et al., 2014).

A final interesting observation about the fossil turtles is that their w_2 values all fall within the range observed for the extant species we examined. This implies that the selective pressures related to shell strength and hydrodynamic efficiency experienced by these taxa were not dramatically different than those experienced by extant species and that they were faced with similar trade-offs. Moreover, the fact that the fossils fall within the morphospace defined by the extant species, with the exception of the *Echmatemys* sp. specimen (FMNH PR 98), which falls just outside at the bottom of the plot (Fig. 6), indicates that the fossil species responded to these selective pressures in ways that are broadly similar to their extant relatives. It has long been recognized that there is a general correspondence between habitat preference and shell shape in turtles (e.g., Romer, 1956; Zangerl, 1969; Claude et al., 2003; Renous et al., 2008), and our analysis raises the question of whether this consistency stems from intrinsic constraints associated with the unique chelonian Bauplan or

relative stasis in the selective pressures operating on that Bauplan. Study of turtle species with morphologies that run counter to the trend, such as the highly flattened, terrestrial pancake tortoise *Malacochersus tornieri*, as well as fossil species that fall outside of the range of extant turtle morphospace, may help to resolve which of these alternatives are more likely.

CONCLUSIONS

Geometric morphometrics and finite element analysis are both important tools for studying the function and evolution of morphology, and can provide unique insights when applied to fossil taxa. Each technique is powerful by itself, but when combined in the context of quantitative evolutionary theory they can provide an innovative framework for testing hypotheses about the relative importance of specific functions in the evolution of a clade. Geometric morphometrics provides a sophisticated system for quantifying complex morphological shapes, for analyzing morphology in a phylogenetic context, and for estimating the response of morphology to selection and drift. Finite element analysis allows the performance of morphology in specific functional contexts to be quantified, thus providing a way to predict performance even in extinct species. In combination, these tools allow functional hypotheses to be translated into equations that describe the expected direction of evolution due to selection, which in turn facilitates analysis of the role of particular functional factors in the evolutionary history of a clade. Importantly, this framework permits the trade-offs between different functions to be quantified, as well as estimation of the relative strength of selection for each.

We used this framework to show that selection for structural strength and hydrodynamic performance has played a role in the evolution of shell shape in emydid turtles and that the trade-off between these factors has constrained their morphological diversity. However, this framework also allowed us to demonstrate that these two functional factors are relatively weak compared to genetic drift and that functions that have not yet been assessed are also likely to play an important role in the evolution of this clade. Although these results could have been approximated with either GM or FEA alone, it is the quantitative evolutionary framework linking the two that provides the predictive power necessary for quantitative testing of the importance of alternative functional factors in the evolutionary history of the clade. Shell strength and hydrodynamic efficiency have had a quantifiable impact on the evolution of emydids, but the drift-prone pattern observed here does not achieve the mathematical precision between function and form envisioned by D'Arcy Thompson. Instead, the combination of drift and multiple factors influencing shell shape indicate that a many-to-one mapping of form to function (e.g., Alfaro et al., 2005; Wainwright, et al., 2005; Wainwright, 2007; see Stayton [2011] for a specific application to turtles) is a more reasonable expectation.

Our quantitative framework can be applied to any study in which the performance of a structure can be measured or estimated, especially those in which measured values bear on performance in different or changing environments. Here, those functions were carapace strength (measured by FEA), which is arguably important for most turtles in all environments, and the hydrodynamic efficiency reflected by shell cross-sectional area, which is more important in aquatic environments. In principle, the FEA and cross-sectional area data could be replaced with any other functional or physiological variables of interest, giving the approach broad applicability beyond the specific synthesis of GM and FEA that we attempted here.

There has been recent interest in how the results of FEA analyses correlate with shape differences, and the evolutionary implications of those correlations. For example, Pierce et al. (2008) examined mechanical performance of the crocodilian skull in relation to bite force and hydrodynamic efficiency and concluded that selective factors related to foraging and feeding had a strong influence on the evolution of skull shape. Tseng (2008; also see Tseng, 2013) looked at skull and mandible strength in carnivorans in relation to bone cracking behaviors, concluding that stress dissipation from loads on the third and fourth premolars was an important selective factor for durophagous feeding. Dumont et al. (2014) studied selection for low stress and mechanical efficiency in New World leaf-nosed bats and found that selection for mechanical advantage was much more important in their evolutionary history. The quantitative evolutionary framework we described here allows such studies to go one step further and mathematically describe the expected evolutionary trajectories and outcomes of selection for functional performance and to quantify the contribution of these factors, phylogenetic history, and genetic drift to the evolutionary history of a group.

LITERATURE CITED

Abdala, F., T. Jashashvili, B. S. Rubidge, and J. van den Heever. 2014. New material of *Microgomphodon oligocynus* (Eutherapsida, Therocephalia) and the taxonomy of southern African Bauriidae; pp. 209–231 in C. F. Kammerer, K. D. Angielczyk, and J. Fröbisch (eds.), Early Evolutionary History of the Synapsida. Springer, Dordrecht.

Adams, D. C., and M. L. Collyer. 2009. A general framework for the analysis of phenotypic trajectories in evolutionary studies. Evolution 63:1143–1154.

Adams, D. C., F. J. Rohlf, and D. E. Slice. 2004. Geometric morphometrics: ten years of progress after the revolution. Italian Journal of Zoology 71:5–16.

Adams, D. C., F. J. Rohlf, and D. E. Slice. 2013. A field comes of age: geometric morphometrics in the 21st century. Hystrix 24:7–14.

Alfaro, M. E., D. I. Bolnick, and P. C. Wainwright. 2005. Evolutionary consequences of many-to-one mapping of jaw morphology to mechanics in labrid fishes. American Naturalist 165:E140–E164.

Anderson, P. S. L., and E. J. Rayfield. 2012. Virtual experiments, physical validation: dental morphology at the intersection of experiment and theory. Journal of the Royal Society Interface 9:1846–1855.

Anderson, P. S. L., P. G. Gill, and E. J. Rayfield. 2011. Modeling the effects of cingula structure on strain patterns and potential fracture in tooth enamel. Journal of Morphology 272:50–65.

Angielczyk, K. D., and C. R. Feldman. 2013. Are diminutive turtles miniaturized? The ontogeny of plastron shape in emydine turtles. Biological Journal of the Linnean Society 108:727–755.

Angielczyk, K. D., and H. D. Sheets. 2007. Investigation of simulated tectonic deformation in fossils using geometric morphometrics. Paleobiology 33:125–148.

Angielczyk, K. D., C. R. Feldman, and G. R. Miller. 2011. Adaptive evolution of plastron shape in emydine turtles. Evolution 65:377–394.

Arbour, J. H., and C. M. Brown. 2014. Incomplete specimens in geometric morphometric analyses. Methods in Ecology and Evolution 5:16–26.

Arnold, S. J. 1983. Performance surfaces and adaptive landscapes. American Zoologist 23:347–361.

Arnold, S. J. 2003. Performance surfaces and adaptive landscapes. Integrative and Comparative Biology 43:367–375.

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

Bartels, W. S. 1993. Niche separation of fluvial and lacustrine reptiles from the Eocene Green River and Bridger Formations of Wyoming. Journal of Vertebrate Paleontology 13(3, Supplement):25A.

Benoit, J., T. Lehmann, M. Vatter, R. Lebrun, S. Merigeaud, L. Costeur, and R. Tabuce. 2015. Comparative anatomy and three-dimensional geometric morphometric study of the bony labyrinth of Bibymalagasia (Mammalia, Afrotheria). Journal of Vertebrate Paleontology 35:e930043, DOI:10.1080/02724634.2014.930043.

Berg, H. C. 1993. Random Walks in Biology. Princeton University Press, Princeton, New Jersey, 142 pp.

Bhullar, B.-A., Z. S. Morris, E. M. Sefton, A. Tok, M. Tokita, B. Namkoong, J. Camacho, D. A. Burnham, and A. Abzhanov. 2015. A molecular mechanism for the origin of a key evolutionary innovation, the bird beak and palate, revealed by an integrative approach to major transitions in vertebrate history. Evolution 69:1665–1677.

Böhmer, C., O. W. M. Rauhut, and G. Wörheide. 2015. Correlation between *Hox* code and vertebral morphology in archosaurs. Proceedings of the Royal Society B Biological Sciences 282:20150077.

Bookstein, F. L. 1987. Random walk and the existence of evolutionary rates. Paleobiology 13:446–464.

Bookstein, F. L. 1989. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence 11:567–585.

Bookstein, F. L. 1991. Morphometric Tools for Landmark Data. Cambridge University Press, Cambridge, U.K., 435 pp.

Bookstein, F. L. 1996a. Biometrics, biomathematics and the morphometric synthesis. Bulletin of Mathematical Biology 58:313–365.

Bookstein, F. L. 1996b. Standard formulae for the uniform shape component in landmark data; pp. 153–168 in L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice (eds.), Advances in Morphometrics. Plenum Press, New York.

Bookstein, F. L. 1997. Landmark methods for forms without landmarks: localizing group differences in outline shape. Medical Image Analysis 1:225–243.

Bookstein, F. L. 1998. A hundred years of morphometrics. Acata Zoologica Academiae Scientiarum Hungaricae 44:7–59.

Bookstein, F. L. 2013a. Allometry for the twenty-first century. Biological Theory 7:10–25.

Bookstein, F. L. 2013b. Random walks as a null model for high-dimensional morphometrics of fossil series: geometrical considerations. Paleobiology 39:52–74.

Bookstein, F. L. 2014. Measuring and Reasoning: Numerical Inference in the Sciences. Cambridge University Press, Cambridge, U.K., 535 pp.

Bramble, D. M. 1974. Emydid shell kinesis: biomechanics and evolution. Copeia 1974:707–727.

Brassey, C. A., R. N. Holdaway, A. G. Packham, J. Anné, P. L. Manning, and W. I. Sellers. 2013. More than one way of being a Moa: differences in leg bone robustness map divergent evolutionary trajectories in Dinornithidae and Emeidae (Dinornithiformes). PLoS ONE 8(12):e82668.

Bright, J. A. 2014. A review of paleontological finite element models and their validity. Journal of Paleontology 88:760–769.

Bright, J. A., and E. J. Rayfield. 2011. Sensitivity and ex vivo validation of finite element models of the domestic pig cranium. Journal of Anatomy 219:456–471.

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

Button, D. J., E. J. Rayfield, and P. M. Barrett. 2014. Cranial biomechanics underpins high sauropod diversity in resource-poor environments. Proceedings of the Royal Society B Biological Sciences 281:1795. DOI: 10.1098/rspb.2014.2114.

Castanera, D., J. Colmenar, V. Sauqué, and J. I. Canudo. 2015. Geometric morphometric analysis applied to theropod tracks from the Lower Cretaceous (Berriasian) or Spain. Palaeontology 58:183–200.

Cheverud, J. M. 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36:499–516.

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

Claude, J. 2006. Convergence induced by plastral kinesis and phylogenetic constraints in Testudinoidea: a geometric morphometric assessment. Fossil Turtle Research, Volume 1. Russian Journal of Herpetology 13(Supplement):34–45.

Claude, J. 2008. Morphometrics in R. Springer, New York, 317 pp.

Claude, J., E. Paradis, H. Tong, and J. C. Auffray. 2003. A geometric morphometric assessment of the effects of environment and cladogenesis on the evolution of the turtle shell. Biological Journal of the Linnean Society 79:485–501.

Clough, R. W. 1990. Original formulation of the finite element method. Finite Elements in Analysis and Design 7:89–101.

Collyer, M. L., and D. C. Adams. 2013. Phenotypic trajectory analysis: comparison of shape change patters in evolution and ecology. Hystrix 24:75–83.

Cox, P. G., M. J. Fagan, E. J. Rayfield, and N. Jeffery. 2011. Finite element modelling of squirrel, guinea pig and rat skulls: using geometric morphometrics to assess sensitivity. Journal of Anatomy 219:696–709.

Cullen, T. M., D. Fraser, N. Rybczynski, and C. Schröder–Adams. 2014. Early evolution of sexual dimorphism and polygyny in Pinnipedia. Evolution 68:1469–1484.

Curtis, N., M. E. H. Jones, S. E. Evans, P. O'Higgins, and M. J. Fagan. 2013. Cranial sutures work collectively to distribute strain throughout the reptile skull. Journal of the Royal Society Interface 10:86. DOI:10.1098/rsif.2013.0442.

Curtis, N., M. E H. Jones, A. K. Lappin, S. E. Evans, P. O'Higgins, and M. J. Fagan. 2010. Comparison between in vivo and theoretical bite performance: using multi-body modelling to predict muscle and bite forces in a reptile skull. Journal of Biomechanics 43:2804–2809.

Dieleman, J., B. Van Bocxlear, C. Manntschke, D. W. Nyingi, D. Adriaens, and D. Verschuren. 2015. Tracing functional adaptation in African cichlid fishes through morphometric analysis of fossil teeth: exploring methods. Hydrobiologia 755:73–88.

Domokos, G., and P. Várkonyi. 2008. Geometry and self-righting of turtle shells. Proceedings of the Royal Society B Biological Sciences 275:11–17.

Drake, A. G., M. Coquerelle, and G. Colombeau. 2015. 3D Morphometric analysis of fossil canid skulls contradicts the suggested domestication of dogs during the late Paleolithic. Scientific Reports 5:8299. DOI:10.1038/srep08299.

Dryden, I. L., and K. V. Mardia. 1998. Statistics Shape Analysis. John Wiley & Sons, Chichester, U.K., 376 pp.

Dumont, E. R., I. R. Grosse, and G. L. Slater. 2009. Requirements for comparing the performance of finite element models of biological structures. Journal of Theoretical Biology 256:96–103.

Dumont, E. R., J. Piccirillo, and I. R. Grosse. 2005. Finite-element analysis of biting behavior and bone stress in the facial skeletons of bats. Anatomical Record A 283A:319–330.

Dumont, E. R., J. L. Davis, I. R. Grosse, and A. M. Burrows. 2011. Finite element analysis of performance in the skulls of marmosets and tamarins. Journal of Anatomy 218:151–162.

Dumont, E. R., K. Samedevam, I. Grosse, O. M. Warsi, B. Baird, and L. M. Davalos. 2014. Selection for mechanical advantage underlies multiple cranial optima in New World leaf-nosed bats. Evolution 68:1436–1449.

Endler, J. A. 1986. Natural Selection in the Wild. Princeton University Press, Princeton, New Jersey, 336 pp.

Ennen, J. R., M. E. Kalis, A. L. Patterson, B. R. Kreiser, J. E. Lovich, J. Godwin, and C. P. Qualls. 2014. Clinal variation or validation of a subspecies? A case study of the *Graptemys nigrinoda* complex (Testudines: Emydidae). Biological Journal of the Linnean Society 111:810–822.

Ernst, C. H., J. E. Lovich, and R. W. Barbour. 1994. Turtles of the United States and Canada. Smithsonian Institution Press, Washington, D.C., 578 pp.

Ewert, M. A., P. C. H. Pritchard, and G. E. Wallace. 2006. *Graptemys barbouri*, Barbour's Map Turtle; pp. 260–272 in P. A. Meylan (ed.),

Biology and Conservation of Florida Turtles. Chelonian Research Monographs 3.

Fabre, A. C., R. Cornette, A. Perrard, D. M. Boyer, G. V. Prasad, J. J. Hooker, and A. Goswami. 2014. A three-dimensional morphometric analysis of the locomotory ecology of *Deccanolestes*, a eutherian mammal from the Late Cretaceous of India. Journal of Vertebrate Paleontology 34:146–156.

Falkingham, P. L. 2012. Acquisition of high-resolution 3D models using free, open-source, photogrammetric software. Palaeontologia Electronica 15.1.1T:1–15.

Fearon, J. L., and D. J. Varricchio. 2015. Morphometric analysis of the forelimb and pectoral girdle of the Cretaceous ornithopod dinosaur *Oryctodromeus cubicularis* and implications for digging. Journal of Vertebrate Paleontology 35:4, e936555, DOI: 10.1080/02724634.2014.936555

Feldman, C. R., and J. F. Parham. 2002. Molecular phylogenetics of emydine turtles: taxonomic revision and evolution of shell kinesis. Molecular Phylogenetics and Evolution 22:388–398.

Felsenstein, J. 1988. Phylogenies and quantitative characters. Annual Review of Ecology and Systematics 19:445–471.

Fish, J. F., and C. T. Stayton. 2014. Morphological and mechanical changes in the juvenile red-eared slider (*Trachemys scripta elegans*) shells during ontogeny. Journal of Morphology 275:391–397.

Fisher, R. A. 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford, U.K., 272 pp.

Fitton, L. C., J. F. Shi, M. J. Fagan, and P. O'Higgins. 2012. Masticatory loadings and cranial deformation in *Macaca fascicularis*: a finite element analysis sensitivity study. Journal of Anatomy 221:55–68.

Fitton, L. C., M. Prôa, C. Rowland, V. Toro-Ibacache, and P. O'Higgins. 2015. The impact of simplifications on the performance of a finite element model of a *Macaca fascicularis* cranium. Anatomical Record 298:107–121.

Fortuny, J., J. Marcé-Nogué, E. Heiss, M. Sanchez, M., L. Gil, and À. Galobart. 2015. 3D Bite modeling and feeding mechanics of the largest living amphibian, the Chinese Giant Salamander *Andrias davidianus* (Amphibia: Urodela). PLoS ONE 10(4):e0121885.

Gaffney, E. S., and P. A. Meylan. 1988. A phylogeny of turtles; pp. 157–219 in M. J. Benton (ed.), The Phylogeny and Classification of Tetrapods. Clarendon Press, Oxford, U.K.

Gavrilets, S. 2004. Fitness Landscapes and the Origin of Species. Princeton University Press, Princeton, New Jersey, 476 pp.

Geraads, D. 2014. How old is the cheetah skull shape? The case of *Acinonyx pardinensis* (Mammalia, Felidae). Geobios. DOI:10.1016/j.geobios.2013.12.003

Gill, P. G., M. A. Purnell, N. Crumpton, K. Robson Brown, N. J. Gostling, M. Stampanoni, and E. J. Rayfield. 2014. Dietary specializations and diversity in feeding ecology of the earliest stem mammals. Nature 512:303–305.

Gingerich, P. D. 1993. Quantification and comparison of evolutionary rates. American Journal of Science 293A:453–478.

Gómez-Robles, A., J. M. Bermúdez de Castro, J.-L. Arsuaga, E. Carbonell, and P. D. Polly. 2013. No known hominin species matches the expected dental morphology of the last common ancestor of Neanderthals and modern humans. Proceedings of the National Academy of Sciences 110:18196–18201.

Goswami, A., and P. D. Polly. 2010. Methods for studying morphological integration, modularity and covariance evolution; pp. 213–243 in J. Alroy and G. Hunt (eds.), Quantitative Methods in Paleobiology. Paleontological Society Papers, Volume 16.

Goswami, A., W. J. Binder, J. Meachen, and F. R. O'Keefe. 2015. The fossil record of phenotypic integration and modularity: a deep-time perspective on developmental and evolutionary dynamics. Proceedings of the National Academy of Sciences of the United States of America 112:4891–4896.

Gower, J. C. 1975. Generalized Procrustes analysis. Psychometrika 40:33–51.

Grassi, L., N. Hraiech, E. Schileo, M. Ansaloni, M. Rochette, and M. Viceconti. 2011. Evaluation of the generality and accuracy of a new mesh morphing procedure for the human femur. Medical Engineering & Physics 33:112–120.

Green, W. D. K. 1996. The thin-plate spline and images with curving features; pp. 79–87 in K. V. Mardia, C. A. Gill, and I. L. Dryden (eds.), Image Fusion and Shape Variability. University of Leeds Press, Leeds, U.K.

Gröning, F., M. J. Fagan, and P. O'Higgins. 2011. The effects of the periodontal ligament on mandibular stiffness: a study combining finite

element analysis and geometric morphometrics. Journal of Biomechanics 44:1304–1312.

Gröning, F., M. Fagan, and P. O'Higgins. 2012. Modeling the human mandible under masticatory loads: which input variables are important? The Anatomical Record 295:853–863.

Gröning, F., J. A. Bright, M. J. Fagan, and P. O'Higgins. 2012. Improving the validation of finite element models with quantitative full-field strain comparisons. Journal of Biomechanics 45:1498–1506.

Gunz, P., P. Mitteroecker, and F. L. Bookstein. 2005. Semilandmarks in three dimensions; pp. 73–98 in D. E. Slice (ed.), Modern Morphometrics in Physical Anthropology. Kluwer Academic, New York.

Gunz, P., P. Mitteroecker, S. Neubauer, G. Weber, and F. Bookstein. 2009. Principles for the virtual reconstruction of hominin crania. Journal of Human Evolution 57:48–62.

Haldane, J. B. S. 1924. A mathematical theory of natural and artificial selection. Part I. Transactions of the Cambridge Philosophical Society 23:363–372.

Head, J. J., and P. D. Polly. 2015. Evolution of the snake body form reveals homoplasy in amniote *Hox* gene function. Nature 520:86–89.

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

Hirayama, R. 1984. Cladistic analysis of batagurine turtles (Batagurinae: Emydidae: Testudinoidea); a preliminary result. Studia Geologica Salamanticensia (Studia Palaeocheloniologica) Volumen Especial 1:141–157.

Holman, J. A., and U. Fritz. 2001. A new emydine species from the medial Miocene (Barstovian) of Nebraska, USA with a new generic arrangement for the species of *Clemmys* sensu McDowell (1964) (Reptilia: Testudines: Emydidae). Zoologische Abhandlungen, Staatliches Museum für Tierkunde Dresden 51:321–343.

Hunt, G. 2006. Fitting and comparing models of phyletic evolution: random walks and beyond. Paleobiology 32:578–601.

Hutchison, J. H. 1996. Testudines; pp. 337–353 in D. R. Prothero and R. J. Emry (eds.), The Terrestrial Eocene–Oligocene Transition in North America. Cambridge University Press, Cambridge, U.K.

Hutchison, J. H. 1998. Turtles across the Paleocene/Eocene Epoch boundary in west-central North America; pp. 401–408 in M.-P. Aubry, S. G. Lucas, and W. A. Berggren (eds.), Late Paleocene–Early Eocene Climatic and Biotic Events in the Marine and Terrestrial Records. Columbia University Press, New York.

Jones, D., A. R. Evans, E. J. Rayfield, K. K. W. Siu, and P. C. J. Donoghue. 2012a. Testing microstructural adaptation in the earliest dental tools. Biology Letters 8:952–955.

Jones, D., A. R. Evans, K. K. W. Siu, E. J. Rayfield, and P. C. J. Donoghue. 2012b. The sharpest tools in the box? Quantitative analysis of conodont element functional morphology. Proceedings of the Royal Society B 279:2849–2854.

Jones, K. E., K. D. Rose, and J. M. G. Perry. 2014. Body size and premolar evolution in the early–middle Eocene euprimates of Wyoming. American Journal of Physical Anthropology 153:15–28.

Joyce, W. G., J. F. Parham, T. R. Lyson, R. C. M. Warnock, and P. C. J. Donoghue. 2013. A divergence dating analysis of turtles using fossil calibrations: an example of best practices. Journal of Paleontology 87:612–634.

Kendall, D. G. 1977. The diffusion of shape. Advances in Applied Probability 9:428–430.

Kendall, D. G. 1984. Shape-manifolds, Procrustean metrics and complex projective spaces. Bulletin of the London Mathematical Society 16:81–121.

Kendall, D. G. 1985. Exact distributions for shapes of random triangles in convex sets. Advances in Applied Probability 17:308–329.

Kimura, M. 1954. Process leading to quasi-fixation of genes in natural populations due to random fluctuations of selection intensities. Genetics 39:280–295.

Klingenberg, C. P. 2010. Evolution and development of shape: integrating quantitative approaches. Nature Reviews Genetics 11:623–635.

Klingenberg, C. P. 2013. Visualizations in geometric morphometrics: how to read and how to make graphs showing shape changes. Hystrix 24:15–24.

Klingenberg, C. P. 2015. Analyzing fluctuating asymmetry with geometric morphometrics: concepts, methods, and applications. Symmetry 7:843–934.

Klingenberg, C. P., and G. S. McIntyre. 1998. Geometric morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with Procrustes methods. Evolution 52:1363–1375.

Kowalewski, M., and P. Novack-Gottshall. 2010. Resampling methods in paleontology; pp. 19–54 in J. Alroy and G. Hunt (eds.), Quantitative Methods in Paleobiology. Paleontological Society Special Papers 16.

Krauss, S., E. Monsongeo-Ornan, E. Zelzer, P. Fratzl, and R. Shahar. 2009. Mechanical function of complex three-dimensional suture joining the bony elements in the shell of the red-eared slider turtle. Advanced Materials 21:407–412.

Kupczik, K., C. A. Dobson, M. J. Fagan, R. H. Crompton, C. E. Oxnard, and P. O'Higgins. 2007. Assessing mechanical function of the zygomatic region in macaques: validation and sensitivity testing of finite element models. Journal of Anatomy 210:41–53.

Lande, R. 1976. Natural selection and random drift in phenotypic evolution. Evolution 30:314–334.

Lande, R. 1986. The dynamics of peak shifts and the pattern of morphological evolution. Paleobiology 12:343–354.

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

Lawing, A. M., and P. D. Polly. 2010. Geometric morphometrics: recent applications to the study of evolution and development. Journal of Zoology 280:1–7.

Ledoux, L., and M. Boudadi-Maligne. 2015. The contribution of geometric morphometric analysis to prehistoric ichnology: the example of large canid tracks and their implication for the debate concerning wolf domestication. Journal of Archaeological Science 61:25–35.

Levens, R. 1962. Theory of fitness in a heterogeneous environment. I. The fitness set and adaptive function. American Naturalist 891:361–373.

Lourenço, J. M., J. Claude, N. Galtier, Y. Chiari. 2012. Dating cryptodiran nodes: origin and diversification of the turtle superfamily Testudinoidea. Molecular Phylogenetics and Evolution 62:496–507.

Lubcke, G. M., and D. S. Wilson. 2007. Variation in shell morphology of the Western Pond Turtle (Actinemys marmorata Baird and Girard) from three aquatic habitats in northern California. Journal of Herpetology 41:107–114.

Macho, G. A., and I. R. Spears. 1999. Effects of loading on the biochemical behavior of molars of Homo, Pan, and Pongo. American Journal of Physical Anthropology 109:211–227.

Maddison, W. P. 1991. Squared-change parsimony reconstruction of ancestral states for continuous-valued characters on a phylogenetic tree. Systematic Biology 40:304–314.

Magwene, P. M., and J. Socha. 2013. Biomechanics of turtle shell: how whole shells fail in compression. Journal of Experimental Biology 319:86–98.

Mallon, J. C., and J. S. Anderson. 2014. Implications of beak morphology for the evolutionary paleoecology of the megaherbivorous dinosaurs from the Dinosaur Park Formation (upper Campanian) of Alberta, Canada. Palaeogeography, Palaeoclimatology, Palaeoecology 394:29–41.

Manly, B. F. J. 1991. Randomization and Monte Carlo Methods in Biology. Chapman and Hall, New York, 281 pp.

Martin, B. T., N. P. Bernstein, R. D. Birkhead, J. F. Koukl, S. M. Mussmann, and J. S. Placyk, Jr. 2013. Sequence-based molecular phylogenetics and phylogeography of the American box turtles (Terrapene spp.) with support from DNA barcoding. Molecular Phylogenetics and Evolution 68:119–134.

Martínez-Pérez, C., E. J. Rayfield, M. A. Purnell, and P. C. J. Donoghue. 2014. Finite element, occlusal, microwear and microstructural analyses indicate that conodont microstructure is adapted to dental function. Palaeontology 57:1059–1066.

Martins, E. P., and T. F. Hansen. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. American Naturalist 149:646–667.

Martín-Serra, A., B. Figueirido, and P. Palmqvist. 2014. A three-dimensional analysis of morphological evolution and locomotor performance of the carnivoran forelimb. PloS ONE 9:e85574.

McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community ecology from functional traits. Trends in Ecology and Evolution 21:178–185.

McKinney, M. L. 1990. Classifying and analysing evolutionary trends; pp. 28–58 in K. J. McNamara (ed.), Evolutionary Trends. University of Arizona Press, Tucson.

Meachen, J. A., F. R. O'Keefe, and R. W. Sadleir. 2014. Evolution in the sabre-tooth cat, Smilodon fatalis, in response to Pleistocene climate change. Journal of Evolutionary Biology 27:714–723.

Meloro, C., A. Hudson, and L. Rook. 2015. Feeding habits of extant and fossil canids as determined by their skull geometry. Journal of Zoology 295:178–188.

Mitteroecker, P., and P. Gunz. 2009. Advances in geometric morphometrics. Evolutionary Biology 36:235–247.

Mitteroecker, P., and S. M. Huttegger. 2009. The concept of morphospaces in evolutionary and developmental biology: mathematics and metaphors. Biological Theory 4:54–67.

Mitteroecker, P., P. Gunz, M. Bernhard, K. Schaefer, and F. L. Bookstein. 2004. Comparison of cranial ontogenetic trajectories among great apes and humans. Journal of Human Evolution 46:679–698.

Mooers, A. Ø., and D. Schulter. 1999. Reconstructing ancestor states with maximum likelihood support for one- and two-rate models. Systematic Biology 48:623–633.

Murdock, D. J. E., E. J. Rayfield, and P. C. J. Donoghue. 2014. Functional adaptation underpinned the evolutionary assembly of the earliest vertebrate skeleton. Evolution and Development 16:354–361.

Neenan, J. M., M. Ruta, J. A. Clack, and E. J. Rayfield. 2014. Feeding biomechanics in Acanthostega and across the fish-tetrapod transition. Proceedings of the Royal Society B Biological Sciences 281:20132689. DOI: 10.1098/rspb.2013.2689.

O'Higgins, P. 2000. The study of morphological variation in the hominid fossil record: biology, landmarks and geometry. Journal of Anatomy 197:103–120.

O'Higgins, P., and N. Milne. 2013. Applying geometric morphometrics to compare changes in size and shape arising from finite elements analyses. Hystrix 24:126–132.

O'Higgins, P., S. N. Cobb, L. C. Fitton, F. Gröning, R. Phillips, J. Liu, and M. J. Fagan. 2011. Combining geometric morphometrics and functional simulation: an emerging toolkit for virtual functional analyses. Journal of Anatomy 218:3–15.

O'Higgins, P., L. C. Fitton, R. Phillips, J. Shi, J. Liu, F. Gröning, S. N. Cobb, and M. J. Fagan. 2012. Virtual functional morphology: novel approaches to the study of craniofacial form and function. Evolutionary Biology 39:521–535.

O'Keefe, F. R., W. J. Binder, S. R. Frost, R. W. Sadlier, and B. Van Valkenburgh. 2014. Cranial morphometrics of the dire wolf, Canis dirus, at Rancho La Brea: temporal variability and its links to nutrient stress and climate. Palaeontologia Electronica 17.1.17A:1–23.

Olsen, A. M., and M. W. Westneat. 2015. StereoMorph: an R package for the collection of 3D landmarks and curves using a stereo camera set-up. Methods in Ecology and Evolution 6:351–356.

Pace, C. M., R. W. Blob, and M. W. Westneat. 2001. Comparative kinematics of the forelimb during swimming in the red-eared slider (Trachemys scripta) and spiny softshell (Apalone spinifera) turtles. Journal of Experimental Biology 204:3261–3271.

Pagel, M. D. 1992. A method for analysis of comparative data. Journal of Theoretical Biology 156:431–442.

Panagiotopoulou, O. 2009. Finite element analysis (FEA): applying an engineering method to functional morphology in anthropology and human biology. Annals of Human Biology 36:609–623.

Parr, W. C. H., S. Wroe, U. Chamoli, H. S. Richards, M. R. McCurry, P. D. Clausen, and C. McHenry. 2012. Toward integration of geometric morphometrics and computational biomechanics: new methods for 3D virtual reconstruction and quantitative analysis of finite element models. Journal of Theoretical Biology 310:1–14.

Pierce, S. E., K. D. Angielczyk, and E. 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 269:840–864.

Pierce, S. E., K. D. Angielczyk, and E. J. Rayfield. 2009. Shape and mechanics in thalattosuchian (Crocodylomorpha) skulls: implications for feeding behaviour and niche partitioning. Journal of Anatomy 215:555–576.

Piras, P., G. Sansalone, L. Teresi, T. Kotsakis, P. Colangelo, and A. Loy. 2012. Testing convergent and parallel adaptations in talpids humeral mechanical performance by means of geometric morphometrics and finite element analysis. Journal of Morphology 273:696–711.

Piras, P., G. Sansalone, L. Teresi, M. Moscato, A. Profico, R. Eng, T. C. Cox, A. Loy, P. Colangelo, and T. Kotsakis. 2015. Digging adaptation in insectivorous subterranean eutherians. The enigma of

Mesoscalops montanensis unveiled by geometric morphometrics and finite element analysis. Journal of Morphology 276:1157–1171.

Polly, P. D. 2001. Paleontology and the comparative method: ancestral node reconstructions versus observed node values. American Naturalist 157:596–609.

Polly, P. D. 2004. On the simulation of the evolution of morphological shape: multivariate shape under selection and drift. Palaeontologia Electronica 7.2.7A:1–28.

Porro, L. B., K. A. Metzger, J. Iriarte-Diaz, and C. F. Ross. 2013. In vivo bone strain and finite element modeling of the mandible of *Alligator mississippiensis*. Journal of Anatomy 223:195–227.

Raff, R. A. 1996. The Shape of Life: Genes, Development, and the Evolution of Animal Form. University of Chicago Press, Chicago, 520 pp.

Raup, D. M. 1977. Stochastic models in evolutionary paleobiology; pp. 59–78 in A. Hallam (ed.), Patterns of Evolution as Illustrated in the Fossil Record. Elsevier Science, Amsterdam.

Rayfield, E. J. 2005. Aspects of comparative cranial mechanics in the theropod dinosaurs *Coelophysis, Allosaurus* and *Tyrannosaurus*. Zoological Journal of the Linnean Society 144:309–316.

Rayfield, E. J. 2007. Finite element analysis and understanding the biomechanics and evolution of living and fossil organisms. Annual Review of Earth and Planetary Sciences 35:541–576.

Rayfield, E. J. 2011. Strain in the ostrich mandible during simulated pecking and validation of specimen-specific finite element models. Journal of Anatomy 218:47–58.

Rayfield, E. J., D. B. Norman, C. C. Horner, J. R. Horner, P. M. Smith, J. J. Thomason, and P. Upchurch. 2001. Cranial design and function in a large theropod dinosaur. Nature 409:1033–1037.

Reed, D. H., J. J. O'Grady, B. W. Brook, J. D. Ballou, and R. Frankham. 2003. Estimates of minimum viable population sizes for vertebrates and factors influencing those estimates. Biological Conservation 113:23–34.

Rega, E. A., K. Noriega, S. S. Sumida, A. Huttenlocker, A. Lee, and B. Kennedy. 2012. Healed fractures in the neural spines of an associated skeleton of *Dimetrodon*: implications for dorsal sail morphology and function. Fieldiana Life and Earth Sciences 5:104–111.

Renous, S., F. de Lapparent de Broin, M. Depecker, J. Davenport, and V. Bels. 2008. Evolution of locomotion in aquatic turtles; pp. 97–138 in J. Wyneken, M. H. Godfrey, and V. Bels (eds.), Biology of Turtles. CRC Press, Boca Raton, Florida.

Rhee, H., M. F. Horstemeyer, Y. Hwang, H. Lim, H. El Kadiri, and W. Trim. 2009. A study on the structure and mechanical behavior of the *Terrapene carolina* carapace: a pathway to design bio-inspired synthetic composites. Materials Science and Engineering C 29:2333–2339.

Richmond, B. G., B. W. Wright, I. Grosse, P. C. Dechow, C. F. Ross, M. A. Spencer, and D. S. Strait. 2005. Finite element analysis in functional morphology. Anatomical Record A 283:259–274.

Rivera, G. 2008. Ecomorphological variation in shell shape of the freshwater turtle *Pseudemys concinna* inhabiting different aquatic flow regimes. Integrative and Comparative Biology 48:769–787.

Rivera, G., and C. T. Stayton. 2011. Finite element modeling of shell shape in the freshwater turtle *Pseudemys concinna* reveals a trade-off between mechanical strength and hydrodynamic efficiency. Journal of Morphology 272:1192–1203.

Rivera, G., and C. T. Stayton. 2013. Effects of asymmetry on the strength of the chelonian shell: a comparison of three species. Journal of Morphology 274:901–908.

Rivera, G., J. N. Davis, J. C. Godwin, and D. C. Adams. 2014. Repeatability of habitat-associated divergence in shell shape of turtles. Evolutionary Biology 41:29–37.

Rohlf, F. J. 1993. Relative warp analysis and an example of its application to mosquito wings; pp. 131–159 in L. F. Marcus, E. Bello, and A García-Valdecasas (eds.), Contributions to Morphometrics. Monografías del Museo Nacional de Ciencias Naturales, Madrid.

Rohlf, F. J. 1995. Multivariate analysis of shape using partial-warp scores; pp. 154–158 in K. V. Mardia and C. A. Gill (eds.), Proceedings in Current Issues in Statistical Shape Analysis. Leeds University Press, Leeds, U.K.

Rohlf, F. J. 1999. Shape statistics: Procrustes superimpositions and tangent spaces. Journal of Classification 16:197–223.

Rohlf, F. J., and F. L. Bookstein. 2003. Computing the uniform component of shape variation. Systematic Biology 52:66–69.

Rohlf, F. J., and L. F. Marcus. 1993. A revolution in morphometrics. Trends in Ecology and Evolution 8:129–132.

Rohlf, F. J., and D. E. Slice. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. Systematic Zoology 39:40–59.

Romer, A. S. 1956. Osteology of the Reptiles. University of Chicago Press, Chicago, 772 pp.

Ross, C. F., M. A. Berthaume, P. C. Dechow, J. Iriarte-Diaz, L. B. Porro, B. G. Richmond, M. Spencer, and D. S. Strait. 2011. In vivo bone strain and finite-element modeling of the craniofacial haft in catarrhine primates. Journal of Anatomy 218:112–141.

Sampson, P. D., F. L. Bookstein, H. Sheean, and E. L. Bolson. 1996. Eigenshape analysis of left ventricular outlines from contrast ventriculograms; pp. 131–152 in L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice (eds.), Advances in Morphometrics. NATO ASI Series, Series A: Life Science, New York.

Sansalone, G., T. Kotsakis, and P. Piras. 2015. *Talpa fossilis* or *Talpa europaea*? Using geometric morphometrics and allometric trajectories of humeral moles remains from Hungary to answer a taxonomic debate. Palaeontologia Electronica 18.2.42A:1–17.

Schwarz-Wings, D., A. C. Meyer, E. Frey, H.-R. Manz-Steiner, and R. Schumacher. 2009. Mechanical implications of pneumatic neck vertebrae in sauropod dinosaurs. Proceedings of the Royal Society B 277:11–17.

Selman, W. 2012. Intradrainage variation in population structure, shape morphology, and sexual dimorphism in the Yellow-blotched Sawback, *Graptemys flavimaculata*. Herpetological Conservation Biology 7:427–436.

Serrano-Fochs, S., S. De Esteban-Trivigno, J. Marcé-Nogué, J. Fortuny, and R. A. Fariña. 2015. Finite element analysis of the Cingulata jaw: an ecomorphological approach to armadillo's diets. PLoS ONE 10(4):e0120653.

Shi, J., N. Curtis, L. C. Fitton, P. O'Higgins, and M. J. Fagan. 2012. Developing a musculoskeletal model of the primate skull: predicting muscle activations, bite force, and joint reaction forces using multibody dynamics analysis and advanced optimisation methods. Journal of Theoretical Biology 310:21–30.

Shoval, O., H. Sheftel, G. Shinar, Y. Hart, O. Ramote, A. Mayo, E. Dekel, K. Kavanagh, and U. Alon. 2012. Evolutionary trade-offs, Pareto optimality, and the geometry of phenotypic space. Science 336:1157–1160.

Sigal, I. A., M. R. Hardisty, and C. M. Whyne. 2008. Mesh-morphing algorithms from specimen-specific finite element modeling. Journal of Biomechanics 41:1381–1389.

Simpson, G. G. 1944. Tempo and Mode in Evolution. Columbia University Press, New York, 237 pp.

Simpson, G. G. 1953. The Major Features of Evolution. Columbia University Press, New York, 434 pp.

Slater, G. J., L. J. Harmon, D. Wegmann, P. Joyce, L. J. Revell, and M. E. Alfaro. 2011. Fitting models of continuous trait evolution to incompletely sampled comparative data using approximate Bayesian computation. Evolution 66:752–762.

Slice, D. E. 2001. Landmark coordinates aligned by Procrustes analysis do no lie in Kendall's shape space. Systematic Biology 50:141–149.

Slice, D. E. 2005. Modern morphometrics; pp. 1–45 in D. E. Slice (ed.), Modern Morphometrics in Physical Anthropology. Kluwer Academic, New York.

Slice, D. E. 2007. Geometric morphometrics. Annual Review of Anthropology 36:261–281.

Spears, I. R., and G. A. Macho. 1998. Biomechanical behavior of modern human molars: implications for interpreting the fossil record. American Journal of Physical Anthropology 106:467–482.

Stayton, C. T. 2009. Application of thin-plate spline transformations to finite element models or, how to turn a bog turtle into a spotted turtle and analyze both. Evolution 63:1348–1355.

Stayton, C. T. 2011. Biomechanics on the half shell: functional performance influences patterns of morphological variation in the emydid turtle carapace. Zoology 114:213–223.

Stephens, P. R., and J. J. Wiens. 2003. Ecological diversification and phylogeny of emydid turtles. Biological Journal of the Linnean Society 79:577–610.

Stephens, P. R., and J. J. Wiens. 2008. Testing for evolutionary trade-offs in a phylogenetic context: ecological diversification and evolution of locomotor performance in emydid turtles. Journal of Evolutionary Biology 21:77–87.

Strait, D. S., B. G. Richmond, M. A. Spencer, C. F. Ross, P. C. Dechow, and B. A. Wood. 2007. Masticatory biomechanics and its relevance

to early hominid phylogeny: an examination of palatal thickness using finite-element analysis. Journal of Human Evolution 52:585–599.

Strait, D. S., Q. Wang, P. C. Dechow, C. F. Ross, B. G. Richmond, M. A. Spencer, and B. A. Patel. 2005. Modeling elastic properties in finite-element analysis: how much precision is needed to produce an accurate model? The Anatomical Record 283A:275–287.

Straight, D. S., I. R. Grosse, P. C. Dechow, A. L. Smith, Q. Wang, G. W. Weber, S. Neubauer, D. E. Slice, J. Chalk, B. G. Richmond, P. W. Lucas, M. A. Spencer, C. Schrein, B. W. Wright, C. Byron, and C. F. Ross. 2010. The structural rigidity of the cranium of *Australopithecus africanus*: implications for diet, dietary adaptations, and the allometry of feeding biomechanics. Anatomical Record 293:583–593.

Svensson, E., and R. Calsbeek. 2012. The Adaptive Landscape in Evolutionary Biology. Oxford University Press, Oxford, U.K., 376 pp.

Tallman, M., N. Amenta, E. Delson, S. R. Frost, D. Ghosh, Z. S. Klukkert, A. Morrow, and G. J. Sawyer. 2014. Evaluation of a new method of fossil retrodeformation by algorithmic symmetrization: crania of papionins (Primates, Cercopithecidae) as a test case. PloS One 9(7):e0100833.

Tanner, J. B., E. R. Dumont, S. T. Sakai, B. L. Lundrigan, and K. E. Holekamp. 2008. Of arcs and vaults: the biomechanics of bone-cracking in spotted hyaenas (*Crocuta crocuta*). Biological Journal of the Linnean Society 95:246–255.

Thomason, J. J. 1991. Cranial strength in relation to estimated biting forces in some mammals. Canadian Journal of Zoology 69:2326–2333.

Thompson, D'A. W. 1917. On Growth and Form. Cambridge University Press, Cambridge, U.K., 793 pp.

Tseng, Z. J. 2008. Bone-cracking capability in the skull of *Dinocrocuta gigantea* (Carnivora, Mammalia) reveals biomechanical convergence between hyaenids and percrocutids. Journal of Vertebrate Paleontology 28(3, Supplement):153A.

Tseng, Z. J. 2009. Cranial function in a late Miocene *Dinocrocuta gigantean* (Mammalia: Carnivora) revealed by comparative finite element analysis. Biological Journal of the Linnean Society 96:51–67.

Tseng, Z. J. 2013. Testing adaptive hypotheses of convergence with functional landscapes: a case study of bone-cracking hypercarnivores. PloS ONE 8:e65305.

Tseng, Z. J., and J. J. Flynn. 2015. Are cranial biomechanical simulation data linked to known diets in extant taxa? A method for applying diet–biomechanics linkage models to infer feeding capability of extinct species. PLoS ONE 10(4):e0124020.

Tseng, Z. J., and X. Wang. 2010. Cranial morphology of fossil dogs and adaptation for durophagy in *Borophagus* and *Epicyon* (Carnivora, Mammalia). Journal of Morphology 271:1386–1398.

Turtle Taxonomy Working Group [Van Dijk, P. P., J. B. Iverson, A. G. J. Rhodin, H. B. Shaffer, and R. Bour]. 2014. Turtles of the world, 7th edition: annotated checklist of taxonomy, synonymy, distribution, with maps, and conservation status. Chelonian Research Monographs 5:329–479.

Vega, C., and C. T. Stayton. 2011. Dimorphism in shell shape and strength in two species of emydid turtle. Herpetologica 67:397–405.

Wade, M. J., and S. Kalisz. 1990. The causes of natural selection. Evolution 44:1947–1955.

Wainwright, P. C. 2007. Functional versus morphological diversity in macroevolution. Annual Review of Ecology, Evolution, and Systematics 39:381–401.

Wainwright, P. C., M. E. Alfaro, D. I. Bolnick, and C. D. Hulsey. 2005. Many-to-one mapping of form to function: a general principle in organismal design? Integrative and Comparative Biology 45:256–262.

Wang, Q., S. A. Wood., I. R. Grosse, C. F. Ross, U. Zapata, C. D. Byron, B. W. Wright, and D. S. Straight. 2012. The role of sutures in biomechanical dynamic simulation of a macaque cranial finite element model: implications for the evolution of craniofacial form. The Anatomical Record 295:278–288.

Warnock, R. C. M., J. F. Parham, W. G. Joyce, T. R. Lyson, and P. C. J. Donoghue. 2015. Calibration uncertainty in molecular dating analyses: there is no substitute for the prior evaluation of time priors. Proceedings of the Royal Society B Biological Sciences 282:20141013. DOI: 10.1098/rspb.2014.1013.

Weber, G. W., and F. L. Bookstein. 2011. Virtual Anthropology: A Guide to a New Interdisciplinary Field. Springer, Vienna, Austria, 423 pp.

Weber, G. W., F. L. Bookstein, and D. S. Strait. 2011. Virtual anthropology meets biomechanics. Journal of Biomechanics 44:1429–1432.

Wiens, J. J., C. A. Kuczynski, and P. R. Stephens. 2010. Discordant mitochondrial and nuclear gene phylogenies in emydid turtles: implications for speciation and conservation. Biological Journal of the Linnean Society 99:445–461.

Wilbur, H. M. 1975. The evolutionary and mathematical demography of the turtle *Chrysemys picta*. Ecology 56:64–77.

Wiley, D. F., N. Amenta, D. A. Alcantara, D. Ghosh, Y. J. Kil, E. Delson, W. Harcourt-Smith, F. J. Rohlf, K. St. John, and B. Hamann. 2005. Evolutionary morphing; pp. 431–438 in Visualization, 2005. Papers presented at IEEE VIS 05, Minneapolis, MN. Institute of Electrical and Electronics Engineers, New York, New York, U.S.A.

Witmer, L. M. 1995. The Extant Phylogenetic Bracket and the importance of reconstructing soft tissues in fossils; pp. 19–33 in J. Thomason (ed.), Functional Morphology in Vertebrate Paleontology. Cambridge University Press, Cambridge, U.K.

Wolff, J. 1892. Das Gesetz der Transformation der Knochen. Hirschwald Verlag, Berlin, 152 pp.

Wright, S. 1932. The roles of mutation, interbreeding, crossbreeding, and selection in evolution. Proceedings of the 6th International Congress on Genetics 1:356–366.

Wright, S. 1968. Evolution and the Genetics of Populations, Volume 1: Genetic and Biometric Foundations. University of Chicago Press, Chicago, 469 pp.

Wroe, S. 2008. Cranial mechanics compared in extinct marsupial and extant African lions using a finite-element approach. Journal of Zoology 274:332–339.

Wroe, S., P. Clausen, C. McHenry, K. Moreno, and E. Cunningham. 2007. Computer simulation of feeding behavior in the thylacine and dingo as a novel test for convergence and niche overlap. Proceedings of the Royal Society B Biological Sciences 274:2819–2828.

Young, M. T., M. A. Bell, and S. L. Brusatte. 2011. Craniofacial form and function in Metriorhynchidae (Crocodylomorpha: Thalattosuchia): modelling phenotypic evolution with maximum-likelihood methods. Biology Letters 7:913–916.

Zangerl, R. F. 1969. The turtle shell; pp. 311–339 in C. Gans, A. D. Bellairs, and T. S. Parsons (eds.), The Biology of the Reptilia, Volume 1. Academic Press, London.

Zapata, U., K. Metzger, Q. Wang, R. M. Elsey, C. F. Ross, and P. C. Dechow. 2010. Material properties of mandibular cortical bone in the American alligator, *Alligator mississippiensis*. Bone 46:860–867.

Zelditch, M. L., D. L. Swiderski, and H. D. Sheets. 2012. Geometric Morphometrics for Biologists: A Primer, second edition. Academic Press, Amsterdam, 478 pp.

Zelditch, M. L., D. L. Swiderski, H. D. Sheets, and W. L. Fink. 2004. Geometric Morphometrics for Biologists: A Primer. Elsevier, Amsterdam, 443 pp.

Zienkiewicz, O. C., J. P. De S. R. Gago, and D. W. Kelly. 1983. The hierarchical concept in finite element analysis. Computers and Structures 16:53–65.

Zonneveld, J.-P., G. F. Gunnell, and W. S. Bartels. 2000. Early Eocene fossil vertebrates from the southwestern Green River Basin, Lincoln and Uinta Counties, Wyoming. Journal of Vertebrate Paleontology 20:369–386.

APPENDIX 1. Emydid species, their habitat preference or fossil status, scores on PC axes 1 and 2, maximum likelihood estimate of w_2 (the weight of the cross-sectional area performance surface) based on maximizing Eq. (5), and the log-likelihood of the estimate. Fossil specimens are shown in with their specimen numbers in parentheses. Species are listed from lowest w_2 value to highest. **Abbreviations**: **A**, aquatic; **F**, fossil; **T**, terrestrial.

Species	Habitat	PC 1	PC 2	w_2	Log-likelihood
Terrapene nelsoni nelsoni	T	−0.014109	0.007414	0.731	−0.0296
Terrapene ornata	T	−0.013032	−0.004658	0.807	−0.0586
Graptemys gibbonsi	A	−0.015849	−0.010941	0.822	−0.0381
Terrapene carolina major	T	−0.005961	−0.001346	0.836	−0.1169
Emys marmorata	A	−0.015067	−0.010822	0.841	−0.0388
Terrapene mexicana mexicana	T	−0.005828	−0.000443	0.846	−0.0007
Emys blandingii	A	−0.008594	−0.006751	0.849	−0.0315
Graptemys ouachitensis	A	−0.014556	−0.011114	0.871	−0.0298
Graptemys pulchra	A	−0.014046	−0.010977	0.888	−0.0256
Graptemys geographica	A	−0.015357	−0.015058	0.898	−0.0657
Graptemys ernsti	A	−0.013717	−0.010915	0.898	−0.0232
Graptemys caglei	A	−0.015153	−0.014933	0.904	−0.0623
Echmatemys septaria (FMNH P 27242)	F	−0.013825	−0.012348	0.913	−0.0318
Emys orbicularis	A	−0.013019	−0.010628	0.917	−0.0178
Pseudograptemys inornata (FMNH PR 589)	F	−0.010738	−0.007658	0.960	−0.0014
Graptemys nigrinoda	A	−0.016347	−0.021066	0.964	−0.1449
Trachemys gaigeae	A	−0.011642	−0.014808	0.965	−0.0241
Graptemys oculifera	A	−0.013717	−0.015376	0.966	−0.0498
Terrapene coahuila	A	−0.009561	−0.005314	0.971	−0.0009
Graptemys pseudogeographica	A	−0.012469	−0.012909	0.971	−0.0249
Pseudemys rubriventris	A	−0.011866	−0.013918	1.003	−0.0251
Pseudemys texana	A	−0.011208	−0.012875	1.008	−0.0168
Emys blandingii (FMNH PR 428)	F	−0.009625	−0.009113	1.008	−0.0015
Deirochelys reticularia	A	−0.007653	−0.008546	1.012	−0.0005
Graptemys versa	A	−0.011064	−0.012979	1.013	−0.0164
Trachemys scripta	A	−0.012535	−0.012629	1.021	−0.0274
Graptemys flavimaculata	A	−0.013417	−0.018638	1.024	−0.0703
Terrapene carolina carolina	T	−0.005480	0.003045	1.031	−0.0802
Pseudemys peninsularis	A	−0.007211	−0.004692	1.032	−0.0080
Terrapene nelsoni klauberi	T	−0.010988	0.007164	1.039	−0.0403
Terrapene mexicana triunguis	T	−0.009719	0.010060	1.041	−0.0323
Pseudemys floridana	A	−0.007532	−0.007116	1.044	−0.0015
Malaclemys terrapin	A	−0.011486	−0.016746	1.047	−0.0357
Clemmys guttata	A	−0.010960	−0.017208	1.051	−0.0001
Pseudemys concinna	A	−0.011755	−0.017739	1.054	−0.0438
Glyptemys insculpta	T	−0.009238	−0.013277	1.060	−0.0087
Trachemys adiutrix	A	−0.008234	−0.011381	1.064	−0.0022
Chrysemys picta	A	−0.015501	−0.011434	1.066	−0.0337
Trachemys stejnegeri	A	−0.007601	−0.010470	1.070	−0.0004
Graptemys barbouri	A	−0.011324	−0.018469	1.073	−0.0440
Echmatemys wyomingensis (FMNH UC 1429)	F	−0.011356	−0.018764	1.076	−0.0462
Trachemys dorbigni	A	−0.010636	−0.017864	1.081	−0.0345
Pseudemys alabamensis	A	−0.009708	−0.016282	1.083	−0.0206
Glyptemys valentinensis (UNSM 76233)	F	−0.007976	−0.012792	1.087	−0.0038
Trachemys decussata	A	−0.006155	−0.009224	1.091	−0.0011
Echmatemys sp. (FMNH PR 98)	F	−0.012922	−0.023002	1.096	−0.1026
Pseudemys suwanniensis	A	−0.009441	−0.017052	1.097	−0.0222
Trachemys decorata	A	−0.008155	−0.014824	1.100	−0.0086
Pseudemys nelsoni	A	−0.005387	−0.009432	1.110	−0.0021
Pseudemys gorzugi	A	−0.009826	−0.020173	1.123	−0.0406
Trachemys terrapen	A	−0.004553	−0.010328	1.134	−0.0025
Glyptemys muhlenbergii	T	−0.004883	−0.012648	1.151	−0.0001