REVIEW ARTICLE

Bridging Performance and Adaptive Landscapes to Understand Long-Term Functional Evolution

Monique Nouailhetas Simon*
Daniel S. Moen
Department of Integrative Biology, Oklahoma State
University, Stillwater, Oklahoma 74078

ABSTRACT

Understanding functional adaptation demands an integrative framework that captures the complex interactions between form, function, ecology, and evolutionary processes. In this review, we discuss how to integrate the following two distinct approaches to better understand functional evolution: (1) the adaptive landscape approach (ALA), aimed at finding adaptive peaks for different ecologies, and (2) the performance landscape approach (PLA), aimed at finding performance peaks for different ecologies. We focus on the Ornstein-Uhlenbeck process as the evolutionary model for the ALA and on biomechanical modeling to estimate performance for the PLA. Whereas both the ALA and the PLA have each given insight into functional adaptation, separately they cannot address how much performance contributes to fitness or whether evolutionary constraints have played a role in form-function evolution. We show that merging these approaches leads to a deeper understanding of these issues. By comparing the locations of performance and adaptive peaks, we can infer how much performance contributes to fitness in species' current environments. By testing for the relevance of history on phenotypic variation, we can infer the influence of past selection and constraints on functional adaptation. We apply this merged framework in a case study of tur-tle shell evolution and explain how to interpret different possible outcomes. Even though such outcomes can be quite complex, they represent the multifaceted relations among function, fitness, and constraints.

Keywords: adaptive peaks, evolutionary constraints, functional modeling, Ornstein-Uhlenbeck model, performance peaks.

*Corresponding author; email: monique.nouailhetas@gmail.com.

Introduction

The striking diversity in organismal design found in nature must in part be driven by diversity in function. Evolution of form-function relationships reflects functional adaptation when changes in ecological conditions lead to changes in selection regimes that favor different functional demands associated with distinct phenotypes (Wainwright 2007; Arbour et al. 2019; Moen 2019; Grossnickle 2020; Morinaga and Bergmann 2020; Pigot et al. 2020; Friedman et al. 2021). Understanding the functional adaptation that has evolved over long timescales involves studies of morphometrics, biomechanics, physiology, performance evolution, and evolutionary ecology (Taylor and Thomas 2014). However, when applied in isolation, these related but diverse scientific fields may be too narrow to properly understand the complexity of functional adaptation and its effect on diversity in phenotypes. For example, most functional systems are composed of several elements that must properly interact with each other to execute their particular function (Lauder 1981, 1996; Wainwright et al. 2005; Wainwright 2007; Collar et al. 2014; Goswami et al. 2016; Rothier et al. 2022). Therefore, such elements must respond in concert to selection, highlighting the need for a multivariate approach to understand functional evolution. Moreover, the same elements may be part of different functional systems, and evolutionary changes in one system may compromise the functions of other systems (Ghalambor et al. 2003; Walker 2007; Holzman et al. 2011; Bergmann and McElroy 2014). These complex relations between form and function make studying functional adaptation a challenging enterprise that demands an integrative framework to fully understand its underlying processes.

Two such integrative frameworks have been the primary approaches employed to understand the ecological context of the long-term evolution of form-function relationships. First, the adaptive landscape approach (ALA) combines morphological, physiological, and/or behavioral data (thought to be functionally relevant) on multiple species; a phylogeny of those species; and a model of adaptive evolution to estimate a landscape based on fitness. Second, the performance landscape approach (PLA) combines biomechanical models and a theoretical or empirically derived morphospace to estimate a landscape based on performance. The PLA has been increasingly used in recent years, and its advocates often find results that contrast with those found when using the ALA for studying functional evolution (Stayton 2019*a*; Olsson et al. 2020; Polly 2020). Yet understanding how the model of adaptive evolution

(specifically, the model based on the Ornstein-Uhlenbeck [OU] process) used in the ALA works is not trivial. Thus, misconceptions are widespread in evolutionary biomechanics and in related fields (Hansen 2014; Moen et al. 2022; Grabowski et al. 2023). Moreover, different implementations of the OU process to understand adaptation, each of which conveys different assumptions, are not readily comparable. Thus, clarifying how the adaptation model used in the ALA works and how it relates to functional adaptation is timely.

In this review, we focus on these two distinct frameworks and show how they can be integrated to infer the complex interactions of morphology and function with selection and evolutionary constraints (fig. 1). Although the PLA directly incorporates biomechanical constraints, it overlooks other constraints that influence functional evolution. Conversely, the ALA accommodates genetic and phylogenetic constraints but only as unmeasured factors that cause phenotypes to deviate from the inferred adaptive peaks. Bridging the approaches will thus help improve how we understand the complex nature of the evolution of form-function relationships (fig. 1). To this end, we first discuss the theoretical bases of both the ALA and the PLA, highlighting their strengths and limitations. Focusing on recent papers, we then synthesize what each approach contributes to understanding functional adaptation. We do not intend to extensively review the literature but to rather exemplify the insights that each approach gives in isolation. Next, we present the rationale for combining the PLA and the ALA. We show how to interpret when performance peaks coincide with adaptive peaks and when they do not. We also discuss interpretation of when species' phenotypes match adaptive peaks and when they do not. We then use published data to show how the merged framework works in a case study. Finally, we outline the knowledge gaps that our new framework can fill when studying functional adaptation, as well as directions for further integration.

The Adaptive Landscape Approach Based on the Ornstein-Uhlenbeck Process

The adaptive landscape is a conceptual framework widely used to understand how different modes of selection (e.g., stabilizing, directional, and correlated selection) have influenced phenotypic evolution (Simpson 1944; Arnold et al. 2001; Calsbeek and Svensson 2012). When focusing on a single population and on short timescales, the major features of an adaptive (or selection) surface can be estimated with a quadratic function of the relationship between phenotypes of individuals (e.g., morphology, physiology, behavior, and kinematics) and individual relative fitness. Such a function can detect the slope and curvature of adaptive peaks (Lande 1979; Lande and Arnold 1983). Alternatively, more complicated functions, such as a cubic spline function (Schluter 1988) or a projection pursuit regression (Schluter and Nychka 1994), can be used to analyze multiple traits. The quadratic function can also be used to estimate an adaptive landscape for a single species, linking average phenotypes with average relative fitness (Lande and Arnold 1983; Arnold et al. 2001; Arnold 2003). When applied at microevolutionary scales, these methods for estimating adaptive surfaces and landscapes can characterize evolution over short timescales.

At longer timescales encompassing millions of years and the emergence of multiple species, adaptive evolution is best characterized by phylogenetic comparative methods (PCMs). With PCMs, we can understand the evolution of the macroevolutionary adaptive landscape, which describes shifts in adaptive peaks with changes in ecological conditions through evolutionary time. PCMs are often used to test how morphology, physiology, or functional properties vary as a function of ecology while explicitly accounting for the phenotypic similarity among species due to shared ancestry (Garland et al. 2005; O'Meara 2012; Rezende and Diniz-Filho 2012; Garamszegi 2014). Yet they can also be used to model change in adaptive landscapes over time (Hansen et al. 2008; Hohenlohe and Arnold 2008; Hansen 2012). The most appropriate model to study long-term adaptation is based on the OU process, which describes the attraction of phenotypes to a central value—the adaptive optimum (Hansen and Martins 1996; Hansen 1997). Similar to stabilizing selection, the attraction to the optimum can be viewed as a rubber band that pulls phenotypes to it, with the strength of the pull being proportional to the distance of phenotypes from the optimum (Felsenstein 1988; Butler and King 2004). The OU process also has a stochastic element analogous to genetic drift and modeled as a Brownian motion (BM) process (Hansen and Martins 1996).

Although the OU process can be used in many ways in evolutionary analyses (Hansen and Martins 1996; Hansen 2014; Cooper et al. 2016), we focus here on what has been called the Hansen model, named after Hansen's (1997) original development of a model of long-term adaptive evolution using the OU process (Butler and King 2004; Beaulieu et al. 2012). In this model, species' means tend to evolve toward a primary adaptive peak, defined as the optimal phenotype for their current (focal) environment. This peak may be shared by the entire clade, or species with different ecologies can each evolve toward an ecology-specific peak (Hansen 1997; Butler and King 2004). Hence, when using this model, researchers hypothesize selective regimes for all species analyzed based on biological factors thought to influence species' fitness. The idea is that these factors influence the position of primary adaptive peaks, causing a systematic effect on species' phenotypes that adapt to the factors (Hansen 1997).

The Hansen OU model has been the most used and most developed application of the OU process in macroevolutionary studies (see review in Moen et al. 2022). Such developments include the ability to estimate different rates of adaptation to the primary adaptive peak (Beaulieu et al. 2012), to characterize continuous environments (Hansen et al. 2008), to accommodate multivariate phenotypes (Bartoszek et al. 2012; Clavel et al. 2015), and to algorithmically detect primary adaptive peaks based on the phenotypic data (e.g., Ingram and Mahler 2013; Uyeda and Harmon 2014; Khabbazian et al. 2016; Bastide et al. 2018). All such extensions can be considered variants on the Hansen OU model and are distinct from other uses of the OU process in comparative methods (e.g., to model residual error around a regression line; Hansen 2014; Cooper et al. 2016). Thus, for brevity, we use the term "OU model" hereafter to mean using the OU process to study macroevolutionary adaptation. Moreover, for simplicity, we focus on discrete selective environments explicitly specified by researchers, even
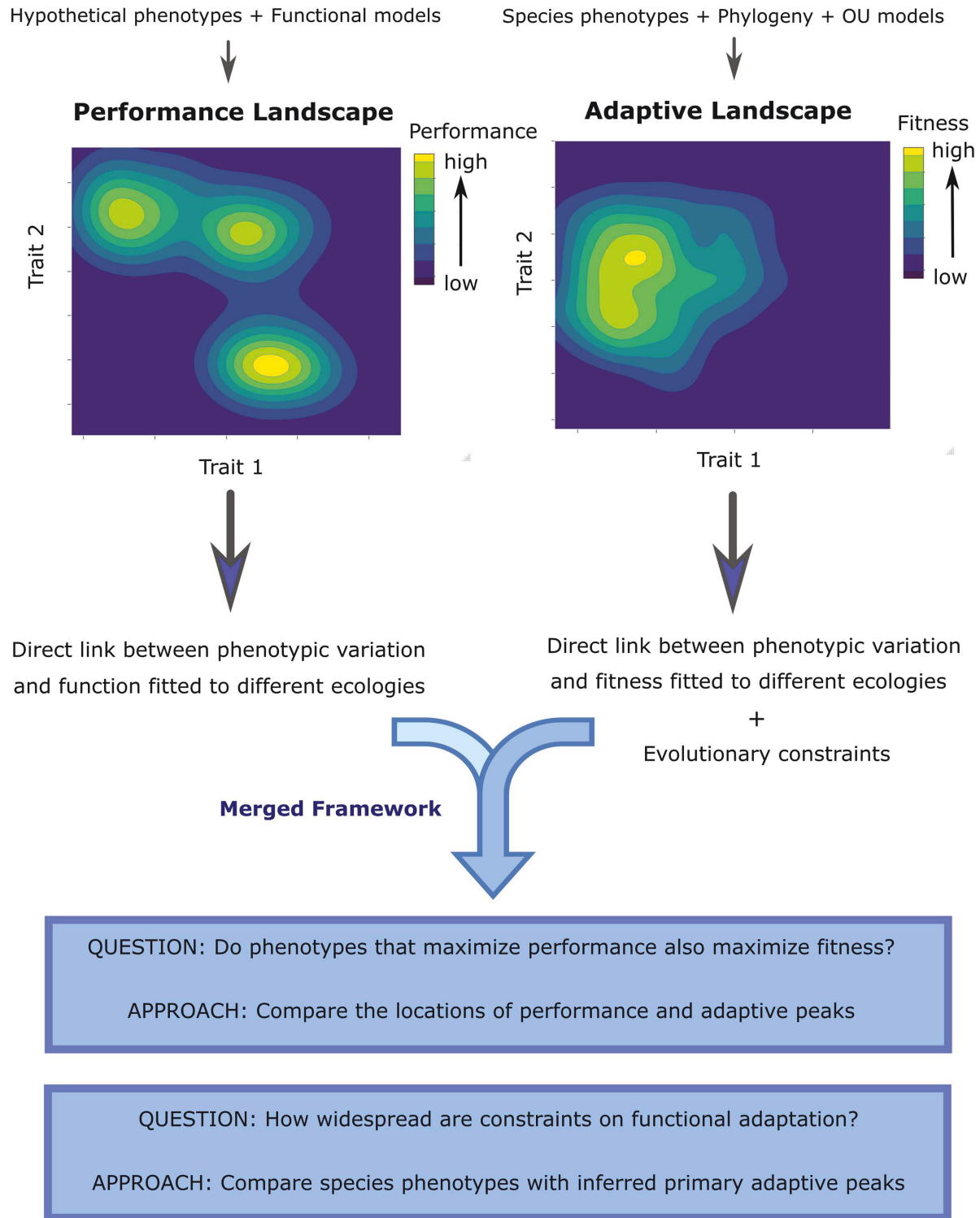
Figure 1. Merged performance-adaptive landscape framework. We integrate the performance and adaptive landscape approaches to address two key open questions in functional adaptation, which are highlighted in the bottom boxes. The logic is to use variation in phenotypic traits of species differing in ecology (e.g., different habitats) as inputs for biomechanical or physiological models to estimate performance peaks and then to use the same phenotypic variation in Ornstein-Uhlenbeck (OU) models to test for ecology-specific adaptive peaks. If performance and adaptive peaks match, we may infer that performance in the studied functions contributes to fitness. To test for a role of constraints, such as past selective regimes influencing current phenotypes, we suggest quantifying the contributions of history and species-specific factors to phenotypic variation across species.

though most extensions of Hansen's (1997) original development of this model can be incorporated into our framework.

Likewise, we do not focus here on OU models that algorithmically search for shifts in adaptive peaks (e.g., Ingram and Mahler 2013; Uyeda and Harmon 2014; Khabbazian et al. 2016; Bastide et al. 2018). These methods obviate the need for hypotheses of selective regimes. However, these implementations generally find several adaptive peaks that may not be easily related to the ecology (or functional properties) of species that share them (e.g., Mahler et al. 2013; Stayton 2019a). The interpretation of such results in the context of Hansen's (1997) original derivation is also unclear, given that such peaks may include phenotypic distinctiveness because of historical contingency (Moen et al. 2016). Moreover, having (sometimes many) hypotheses of selection regimes can be a stronger test of adaptation (Scales and Butler 2016a, 2016b; Swiderski and Zelditch 2022). Therefore, as stated above, we focus on Hansen's (1997) original OU model for the ALA framework, as it is firmly grounded on a deep understanding of how adaptation proceeds over long evolutionary timescales. That said, we could imagine future analyses leveraging methods that search for peaks or even combining algorithmic peak searching with ecological hypotheses (e.g., Swiderski and Zelditch 2022).

Adaptation toward the primary peak depends on a balance among selective pressures (e.g., ecological or functional trade-offs) and constraints (e.g., developmental, genetic, and mechanical constraints) simultaneously acting on the focal traits (Hansen 1997, 2014). These factors are often unmeasured but usually shared by closely related species, causing the observed phenotypes of such species to deviate similarly (i.e., in similar directions and magnitudes) from the primary peak. This shared deviation emphasizes the importance of PCMs to properly study adaptation. The balance of evolutionary forces may change over long periods of time, as ecological changes shift selection pressures related to specific functional demands on traits. Such changes can shift the location of the primary peak in phenotypic space. Changes in constraints on evolution may also occur because of shifts in the pattern of selection, influencing how fast phenotypes can track shifts in adaptive peaks (Hansen 2012). Therefore, the OU model is best described as representing the long-term dynamics of adaptive peaks on the evolutionary landscape rather than the movement of species among fixed peaks (Hansen 1997, 2012; Butler and King 2004; Hohenlohe and Arnold 2008).

Benefits and Limitations of Ornstein-Uhlenbeck Adaptation Models

The OU model provides many distinct advantages to studying phenotypic evolution. First, it simultaneously models adaptation to selective optima and accounts for deviations from such optima based on other (unmeasured) factors influencing phenotypes (Hansen 1997, 2014). Such joint modeling allows researchers to understand the relevance of past and current adaptations for promoting phenotypic and functional diversity in long-term evolution while also accounting for evolutionary constraints that may limit diversity. Such constraints may stem from limited genetic variation in traits, conserved allometric patterns across species, or other factors that limit evolutionary change (Hansen 1997).

Therefore, the interplay between selection and constraints on trait evolution is an essential part of the OU model.

Second, the OU model accounts for the time that a lineage has adapted toward primary optima of both past and current environments. Other PCMs implicitly assume that lineages have always been evolving in their present environments, thus failing to account for imperfect convergence due to history among lineages (currently) in the same environment (Hansen 2014). The explicit modeling of historical environmental changes in the OU model reflects the idea that although species' phenotypes can track these adaptive shifts, they may take time to fully adapt to a new environment. Thus, species may not be at the primary optimum of their current environment because of this lag in adaptation (Hansen et al. 2008; Moen et al. 2016; Toljagić et al. 2018). Consequently, even though the current phenotype of a species may not be the same as the primary optimum of a study's focus, that phenotype represents a species-specific optimal value given all factors affecting the focal trait's evolution (Hansen 1997, 2014). That is, species' phenotypes reflect the best possible phenotypes, in terms of average fitness, considering all influences on their evolutionary trajectory. For example, species-specific optima may be influenced by other functional demands that are relevant to fitness, and these demands are likely shared by closely related species (Hansen 2014; Moen et al. 2022).

Third, OU models are flexible. One can test the evidence for one or multiple optima (Butler and King 2004; Beaulieu et al. 2012) associated with different environments or selective regimes (e.g., different diets or habitats). Multivariate versions of the model have also been developed (Bartoszek et al. 2012, 2023; Clavel et al. 2015, 2019), so that several traits can be analyzed simultaneously, accounting for interactions among traits that may constrain adaptation. These extensions have rarely been leveraged by researchers studying functional adaptation, despite the potential advantages that they have over more common BM-based approaches (Hansen et al. 2008; Clavel and Morlon 2020).

OU models also have limitations. First, they demand a more complex interpretation than many other PCMs because the model has more parameters (e.g., the strength of pull toward the optimum [α], the phenotypic optimum [θ]). Although θ can be robustly estimated even with few taxa (Martins 2000; Ho and Ané 2014; Cressler et al. 2015), α can be hard to estimate at any phylogeny size (Beaulieu et al. 2012; Ho and Ané 2013, 2014). However, the consequences for model selection associated with α estimation are overemphasized in general (Moen et al. 2022; Grabowski et al. 2023). The stationary variance—a compound parameter of the BM variance and the pull toward optima ($\sigma^2/2\alpha$)—can be more reliably estimated than α alone (Ho and Ané 2013, 2014). More importantly, accurately estimating α primarily affects model selection when comparing constant-rate BM to a single-optimum OU model (Boettiger et al. 2012; Cooper et al. 2016; Grabowski et al. 2023), as α is the primary parameter that differs between these models in most software implementations (O'Meara and Beaulieu 2014). In contrast, α matters much less for distinguishing multioptimum OU models, given that statistical power in this context depends on species number and difference in locations of adaptive optima (Cressler et al. 2015). Our emphasis

here is on estimating peak location using multioptimum OU models, such that tens of species may be enough to achieve good statistical power (Moen et al. 2022). Still, a good practice is to use parametric bootstrapping to estimate confidence intervals (CIs) of the parameters (Boettiger et al. 2012; Cressler et al. 2015) or even to make only qualitative statements about differences in the rate of adaptation (Beaulieu et al. 2012).

As a second potential limitation, models that specify discrete selective regimes (e.g., herbivorous vs. carnivorous diet) require an estimation of ancestral regimes at internal nodes, and these estimates can be inaccurate (Cunningham 1999; Losos 1999, 2011). However, such inaccuracy affects analyses less if values of α are high (Moen et al. 2022), and uncertainty in ancestral states may be incorporated into the analysis using Bayesian approaches (e.g., Price et al. 2015; Corn et al. 2021). Finally, multivariate OU methods have shown poor statistical performances when analyzing considerably fewer species than traits (Adams and Collyer 2018). Recent methods may remove such limitations in single-peak OU models (Clavel and Morlon 2020), but no such solution yet exists for the multioptimum approach that we describe here. Overall, even though the OU model of adaptation has clear strengths, use of these models demands adequate implementation and careful interpretation of their parameters (see review in Moen et al. 2022).

The Performance Landscape Approach Based on Biomechanical Models

Almost 40 years ago, Arnold (1983) introduced a framework in which the effects of phenotypes on fitness are mediated by functional performance, the level of executing an ecologically relevant task in nature (e.g., capturing prey or escaping from predators). This paradigm placed performance in the center of evolutionary biology. It proposed partitioning the action of selection into the following two parts: the effects of phenotypes on performance (the performance gradients) and the effects of performance on fitness (the selection gradients). Then 20 years later, Arnold (2003) formalized the idea of a performance landscape—the relation between average trait values and performance of species—whose peaks can be estimated similarly to those of the phenotypic adaptive landscape (Lande 1979; Arnold 2003). However, quantifying Arnold's (2003) performance landscape at a macroevolutionary scale poses great logistical challenges, demanding the measurement of phenotypes and performance in many species.

The PLA proposes reducing the challenge of large-scale data collection with two key practices. First, hypothetical phenotypic spaces are created by systematically varying trait values so as to mimic or even go beyond observed variation across species. Second, biomechanical or functional modeling is used to estimate performance across this hypothetical trait space. Hence, the PLA estimates a performance landscape, based on hypothetical phenotypes and modeled performance, that is different from Arnold's performance landscape, which is based on empirical phenotypes and performance values (e.g., Simon et al. 2019). Moreover, the PLA landscape differs from Arnold's performance adaptive landscape, which maps empirical performance to fitness. Consequently, even though maxi-

mal performance values are estimated in the PLA performance landscapes, performance gradients are more realistic in Arnold's performance landscape, which reflects the empirical landscape. Nonetheless, if a biomechanical model has been validated empirically, the estimated performance landscape may be close to Arnold's performance landscape.

With the PLA, one systematically varies phenotypes (e.g., link lengths in lever systems) to create a phenotypic space that reflects mechanical properties of organismal design that are known to be relevant to functional performance (e.g., Hulsey and Wainwright 2002; Tseng 2013; Martinez and Wainwright 2019; Olsson et al. 2020). One can also use geometric morphometrics to produce hypothetical body shapes (Stayton 2011, 2019a; Tseng 2013; Dumont et al. 2014) and even to measure traits directly from the hypothetical shapes (e.g., linear measurements and bone curvature measured on hypothetical shape warps; Dickson and Pierce 2019) to construct morphospaces. Functional traits, such as mechanical advantage (transmission efficiency of input to output force) and frontal area (reflecting hydrodynamics; Dumont et al. 2014; Dickson and Pierce 2019), can then be estimated from these trait spaces. Moreover, performance of hypothetical shapes can be estimated by using biomechanical models, such as using the suction-induced force-field model to estimate prey capture success in fishes (e.g., Olsson et al. 2020). In the specific case of mechanical load and stress, a reference finite-element (FE) model of one shape (e.g., a focal species) can be warped to correspond to the shape of other specimens, avoiding the time-intensive development of a separate FE model for each specimen (Stayton 2009; Polly et al. 2016).

Because performance models in the PLA can be computationally intensive, trait dimensionality is often reduced by projecting traits and shapes onto principal components (PCs), which are then used as axes in the performance landscape. Hence, the development of geometric morphometrics and the increased power of computers have made the PLA recently flourish as an approach to study biomechanical and performance evolution at a macroevolutionary scale. That said, the PLA could also be extended to accommodate physiological models to estimate performance, including complex relations between physiology, performance, and behavior (Careau and Garland 2012).

Benefits and Limitations of the Performance Landscape Approach

Some features unique to the PLA may enhance understanding of functional adaptation over long temporal scales. First, the approach can find multiple performance peaks, each associated with a distinct phenotype (e.g., Stayton 2019a; Olsson et al. 2020). These performance peaks may relate to differing ecologies, such as distinct prey types (Olsson et al. 2020) or habitats (Stayton 2011; Polly et al. 2016), as in OU models. Unique to the PLA, however, is that multiple performance peaks may also be estimated for the same ecological niche (e.g., cluster of peaks for the same prey type; Olsson et al. 2020). Such findings suggest a many-to-one mapping of phenotype to similar performance for the same ecological state. Finding that species occupy these multiple peaks may indicate that phenotypic diversification has been driven by other aspects of

ecology that also affect the focal phenotypic traits (Wainwright 2007). Alternatively, evolutionary history—such as different evolutionary trajectories—may partly determine which strategy a species adopts (Alfaro et al. 2004, 2005). In these cases of multiple performance peaks for the same ecological state, a single highest performance optimum may exist, but local performance peaks also occur, each with (somewhat) lower levels of performance for the same functional task (Olsson et al. 2020).

A second key advantage of the PLA is that it does not require extensive phenotypic data to estimate the landscape. Instead, biomechanical principles may be used to estimate performance of phenotypes in all regions of morphospace, even those unoccupied by extant or extinct organisms (Stayton 2009, 2011; Olsson et al. 2020). The PLA can also characterize performance valleys that suggest biomechanical constraints limiting morphological and functional diversity (Martin et al. 2019; Stayton 2019a). These valleys may be harder to find when using phenotypes of only extant species, as such poor performance regions potentially result in low fitness and thus represent phenotypes unexpected in nature.

A final strength of the PLA is that performance landscapes for different functional demands can be combined using a set of weights for functions that reflects their expected relative contribution to fitness. These weights are not estimated using fitness. Instead, they mostly have been estimated in one of three ways. First, weights for a given ecology have been estimated as the average weights for species with that ecology, with each species' weights corresponding to those that produce peak estimated performance for that species (Stayton 2019a). Second, maximum likelihood has been used to calculate the weights that maximize the height of the peak that species occupy on the combined performance landscape (Polly et al. 2016; Dickson and Pierce 2019; Polly 2020). Third, one can calculate all possible combined performance peaks by simulating all combinations of relative weights and then finding performance peaks that are closest to observed phenotypes (Stayton 2019a). Overall, these combined performance landscapes can be used to explore evolutionary changes in the relative importance of different functions as a change in the relative weights associated with each function.

The PLA also has clear limitations. Most obviously, the approach can be applied only to systems for which a strong biomechanical understanding of specific functional traits exists (e.g., Olsson et al. 2020). This limitation may be alleviated by using FE analysis, given that the technique is well established in engineering, relies on easily tested biological assumptions (e.g., bone material properties), and has been validated in different biological systems (Dumont et al. 2009; Bright and Rayfield 2011; Stayton 2018). More generally, if a modeled function depends on complex sets of quantitative traits, extracting performance data for all theoretical shapes across the morphospace may demand intensive computational effort (Stayton 2019a). The biomechanical model itself can also be quite complicated and influenced by multiple factors that vary across species (e.g., suction-induced force-field model; Holzman et al. 2011). Therefore, applying the PLA to a broad range of organisms may be feasible for only simple functions determined by a few easily measured morphological traits. Yet studying simple functions may exclude other functions that are relevant to fitness, which also may

limit the utility of the PLA because of the widespread existence of functional trade-offs (Wainwright 2007; Garland et al. 2022).

Another complication of the PLA is its current reliance on PCs to represent the phenotypic space, which is a technique commonly used to reduce computational burden (e.g., Dickson and Pierce 2019; Stayton 2019a; Olsson et al. 2020). PC analysis is a technique for reducing dimensionality in a dataset. It reorders variation to produce a new set of axes of decreasing variation; these axes are composed of combinations of the original traits weighted by their correlation with the specific PC axis (Jolliffe and Cadima 2016). By eliminating many of these PC axes, researchers can reduce dimensionality (e.g., from hundreds of dimensions to a few dimensions in geometric morphometric studies). However, these combinations of traits may actually explain little performance variation across species (e.g., Stayton 2019a), and higher performance peaks may exist along unanalyzed morphological dimensions (i.e., those dropped in the PC analysis). This means that the link could be quite weak between performance as estimated by the PLA and actual performance for a given (original) phenotype. In turn, such a mismatch could hinder interpretation of macroevolutionary patterns. Still, it is reasonable to assume that axes of higher phenotypic variance will relate to axes of higher functional diversity, such that we expect phenotypic PCs to often show high functional relevance.

Comparing the Contributions of the ALA and the PLA to Understand Functional Adaptation

In the context of functional adaptation, the ALA based on the OU model has mainly been used to test for the existence of one or more optima across species that share similar selective regimes (e.g., Anderson et al. 2014; Dumont et al. 2014; Hagey et al. 2017; Kilbourne and Hutchinson 2019; Moen 2019; Rader et al. 2020; Swiderski and Zelditch 2022). Most studies have found that a multiple-peak OU model is the best fit for the evolution of morphological, kinematic, and performance traits in a diverse array of ecological contexts and organisms (e.g., feeding mechanics in fishes [Hulsey et al. 2019], swimming-associated morphology in fishes [Friedman et al. 2021], foraging-associated morphology in piscivorous birds [Eliason et al. 2020]). This general pattern suggests that long-term functional evolution is driven by shifts in selective regimes associated with specific functional demands that become relevant when species differ in diet, behavior, or habitat, among other factors.

The OU model has also been applied to directly test whether functional trade-offs have an adaptive basis, in which conflicting selection pressures have optimized different functional properties in species that use distinct kinematic strategies (e.g., Arbour and López-Fernández 2013; Anderson et al. 2014; Corn et al. 2021; Burress and Muñoz 2023). For instance, bottom-feeding fishes show different optima for morphology and kinematic traits compared to species that feed in the water column (Arbour and López-Fernández 2013). The same pattern was also found when contrasting the feeding kinematics of fishes that use suction to capture prey with those of fishes that bite to capture prey (Corn et al. 2021). However, some expected functional trade-offs may be alleviated

by many-to-one mapping of divergent morphologies occupying the same kinematic (e.g., Friedman et al. 2021), functional (e.g., Dumont et al. 2014; Pigot et al. 2020), or performance peak (e.g., Moen 2019). These results indicate that conflicting selective pressures do not always result in different optima related to functional trade-offs, especially if the functions are underlain by many traits whose influence on various functions differs.

Similar to the ALA, the PLA has been used to understand the complexity of form-function relationships in terms of number and height of performance peaks over a wide range of theoretical phenotypic variation (Tseng 2013; Stayton 2019a; Olsson et al. 2020). However, the PLA also focuses on the occurrence of performance valleys, which may bound functional diversity by limiting the full occupation of trait space (Hulsey and Wainwright 2002; Tseng 2013). Beyond describing the major features of the performance landscape, practitioners use the PLA to test whether the distribution of species is centered around the theoretical performance peaks. They do this by projecting phenotypes of species onto the performance landscape (Stayton 2011, 2019a; Tseng 2013; Polly et al. 2016; Dickson and Pierce 2019; Dickson et al. 2021; Holzman et al. 2022). A match between distributions of observed phenotypes and performance peaks favors a role of natural selection on the functions studied in shaping phenotypes in long-term evolution (Stayton 2019a). Yet most studies have found an incomplete match (Tseng 2013; Polly et al. 2016; Dickson and Pierce 2019; Stayton 2019a; Holzman et al. 2022). This lack of correspondence suggests that functional performance is not always maximized by selection in all species. Other unstudied functional demands (Dickson and Pierce 2019; Stayton 2019a) or other factors beyond functional optimization likely also influence current phenotypes. These latter factors include genetic and phylogenetic constraints (Polly et al. 2016; Holzman et al. 2022).

The PLA has also been used to elucidate the link between conflicting functional demands and fitness by the estimation of the combined performance landscape (Polly et al. 2016; Dickson and Pierce 2019; Stayton 2019a; Polly 2020). These studies have found that the relative weights for performance of different functions vary among species with different ecologies, which suggests shifts in the strength of selection for each function with changes in the balance of functional demands. For example, Polly et al. (2016) combined performance landscapes for turtle shell shape based on shell strength and hydrodynamic efficiency. They found that performance peaks always occurred along a single line, called the Pareto front, indicating optimization of a trade-off (Farnsworth and Niklas 1995; Shoval et al. 2012). In this case, the same shape cannot be highly resistant to crushing stress and also hydrodynamically efficient. Moreover, Polly et al. (2016) showed that only about half of the variation in turtle shell shape was explained by the line defining the Pareto front, indicating that other factors cause shell shape to deviate from the optimal trade-off values.

From this synthesis of the general inferences on functional adaptation drawn from the ALA and the PLA, we can infer that different selective regimes shape some diversity in form, as evidenced by the empirical support for multiple adaptive peaks (in the ALA). Yet optimization of specific functions is not enough to fully explain this diversity, as shown by the partial mismatches

between extant phenotypes and performance peaks (in the PLA). A key unresolved issue is understanding how much selection to optimize functional performance is able to influence diversity in form when multiple constraints are at play. As we delineate in the next section, merging performance and adaptive landscapes can give insights into this issue by indicating whether past selection from ancestral regimes and constraints has caused species' phenotypes to deviate from the predicted adaptive peaks based on ecological factors thought to influence performance.

A Merged Performance-Adaptive Landscape Framework

A conceptual framework combining performance and adaptive landscapes is not a new idea. Arnold (2003) discussed the intriguing scenario in which the long-term dynamics of the adaptive landscape are constrained by the shape of the performance landscape. In this scenario, the adaptive peaks would move along performance lines of least resistance, multivariate directions in which changes in phenotypes would not result in reduced performance and therefore reduced fitness (Arnold 2003). This scenario is consistent with many-to-one mapping of phenotype to performance, with different phenotypes occupying different performance peaks that match adaptive peaks. However, this scenario seems more likely when the relative importance of functions for fitness does not change over time. Alternatively, the balance of functional demands may itself be dynamic, changing when selective regimes change (Hansen 1997; Butler and King 2004). For example, the balance could change if a specific function that had low relevance to fitness in an environment (e.g., swimming in a terrestrial species) became more relevant with the invasion of a new environment (e.g., aquatic environment). Hence, an alternative hypothesis is that the (combined) performance landscape also changes over evolutionary time, potentially causing shifts in adaptive peaks if the new functions are relevant to fitness.

The performance landscape is generally interpreted as a functional manifestation of the adaptive landscape (Tseng 2013), bridging variation in biomechanical systems and fitness (Martin and Wainwright 2013; Martin et al. 2019; Muñoz 2019). However, attempting to compare the peaks inferred from the combined performance landscape with adaptive peaks estimated using OU models has been challenging (Stayton 2019a; Olsson et al. 2020; Polly 2020). Any divergence between the performance peaks and the adaptive peaks has been interpreted as evidence that the latter is not well suited to study the evolution of biomechanics and performance (Stayton 2019a, 2019b; Polly 2020; Olsson et al. 2020). Yet this interpretation overlooks differences between primary and species-specific adaptive peaks in the OU framework. For example, advocates of the PLA state that whereas the PLA can detect multiple performance peaks for a given ecological niche (Olsson et al. 2020), OU models can detect only a single adaptive peak for that niche. Although this is true, this property of the OU method is not an intrinsic flaw. Rather, it stems from using coarse ecological states (e.g., arboreal, aquatic, and terrestrial states) that may poorly capture the ecological complexity found among multiple species assigned to a single state (Leroi et al. 1994). In some cases, more finely defining ecological states can improve the fit of primary adaptive peaks to species (e.g., finding that semiaquatic and aquatic

species each have their own primary adaptive peaks; Moen et al. 2016). Moreover, combining multiple ecological axes (e.g., diet, habitat; Swiderski and Zelditch 2022) when comparing models can be fruitful.

More generally, other factors beyond those influencing performance in focal regime states and functions will always affect the phenotypes of species to some extent (Hansen 1997). The OU model accommodates this complexity by design through adopting the concept of species-specific optima that are determined by myriad demands (Hansen 1997), not just those examined in a given study. Thus, these models were not developed to find multiple peaks for the same ecological niche. Rather, they find a single primary peak that species' phenotypes would match if (1) the focal trait was under long-term selection for only the specific functional demands associated with this primary peak and (2) no constraints or trade-offs affected the phenotypes in question (Hansen 1997, 2014). Therefore, in contrast to the PLA, the ultimate goal of OU models is not to find all possible peaks associated with a specific ecological state. Instead, OU models test whether that state influences the position of a primary optimum in the adaptive landscape. In other words, the key goal of the OU model is to uncover ecological factors that have influenced phenotypes through fitness by using species as replicates of evolution in that environment while taking their similarity due to history into account.

In contrast to previous work attempting to combine performance and adaptive landscapes, the merged performance-adaptive landscape framework that we propose takes advantage of this key distinction between primary adaptive peaks and species-specific peaks to infer the role of constraints on functional adaptation to current environments (fig. 1). We also capitalize on the OU parameter α to infer whether adaptation to ancestral environments influences phenotypes of species adapting to current environments (this includes the related concept of phylogenetic half-life, the time over which half of the ancestral influence is lost during adaptive evolution; Hansen et al. 2008). With this new conceptual framework, we can address the following two key open questions: (1) Do phenotypes that maximize performance always maximize fitness? (2) How widespread are constraints on functional adaptation, particularly ancestral constraints related to past selective environments? The first question can be answered by comparing the locations of performance and adaptive peaks. This comparison can reveal whether adaptive peaks are mostly linked to the specific functional demands directly studied in the PLA (if peaks match) or whether other functions are relevant to fitness (if peaks do not match). The second question can be answered by comparing species' phenotypes with inferred primary adaptive optima. Below, we describe the rationale behind the merged framework. We also provide a case study to exemplify how our approach works in practice.

Comparing the Locations of Performance and Adaptive Peaks

To infer whether performance in the studied function is relevant to fitness, we propose comparing the hypothetical phenotypes that maximize performance in the PLA with the observed species' phenotypes that maximize fitness in the ALA. This peak comparison can be done in three steps. First, one selects phenotypic

traits that can be used as inputs in biomechanical (or physiological) models for estimating the performance landscape. Such landscapes can be different landscapes for different functions or a combined performance landscape that integrates performance across functions (Stayton 2019a, 2019b; Polly 2020). The modeling then translates variation in phenotypes to variation in performance. To find the performance peaks for the different ecological factors thought to have shaped phenotypes, more than one method can be used. A performance peak can be estimated in a combined performance landscape, as we described above when introducing the PLA. Another method is to use an algorithm, such as a gradient ascent analysis, to find local performance optima (Olsson et al. 2020). Both methods can find more than one performance peak for a given ecological state.

When comparing peaks, one next estimates statistical support for various OU models with the same phenotypic traits but uses mean values of real species. At least one of these models must use the same ecological factors (e.g., omnivore vs. herbivore vs. carnivore) as the PLA to define selective regimes. The OU model showing the best fit to the phenotypic data (i.e., highest Akaike information criterion weight) should include estimates of the primary adaptive optima specific to at least some of the ecological states in the PLA to compare performance and adaptive peaks. This optimal model may have separate primary peaks for each state of the ecological factor analyzed in the PLA, or it may lump some levels into the same peak (e.g., omnivore with carnivore). The latter scenario could happen if performance in one of the studied functions is most relevant to fitness for multiple states of the ecological factor, such as cutting force for omnivores and carnivores or crushing force for omnivores and herbivores. In such a scenario, because other functions that contribute less to fitness can be used to estimate a combined performance landscape, different performance peaks may still occur for different ecologies, even though these other functions contribute less information for inferring distinct primary adaptive peaks.

The final step is to compare the locations of the performance and primary adaptive peaks for a specific ecological state. If the number of peaks differs between types, we suggest comparing all sets of comparable peaks, even if more than one performance peak would be compared with a single adaptive peak (e.g., one performance peak each for two ecological states compared with a single adaptive peak for those states combined). Because the PLA assumes that performance is directly proportional to fitness, the simplest expectation is that performance peaks will match primary adaptive peaks. Alternatively, if performance peaks are far from the primary adaptive peak, it would suggest that phenotypic diversity may be associated with diversity in performance but that performance in the studied functions contributes little to fitness, at least in species' current environments. For example, this scenario could have occurred when performance in the studied functions was relevant in past selective environments, such that past selection actually drove the current phenotypic diversity. Performance peaks estimated when using just one or a few functions may also not match adaptive peaks, because the latter would reflect trade-offs that were not considered in the performance landscape. When estimating a combined performance landscape, in which weights determine the relative importance of multiple functions, a mismatch between

performance and adaptive peaks may indicate error in the estimation of the weights.

In practice, even when maximum performance maximizes fitness, the locations of both peaks will likely differ because of estimation error. Thus, we can use 95% CIs to statistically compare the peak locations in phenotypic space. For the primary adaptive peaks, one can use standard errors or a bootstrap approach (for more details, see the case study below). In the current PLA framework, no explicit measure of uncertainty is produced for performance peak location, so we see no way to calculate 95% CIs for these peaks. However, we suggest also testing whether the region that encompasses the peak performance and 5% lower performance (i.e., a region of highest performance instead of just a single point; Stayton 2019b) overlaps the 95% CI of the primary adaptive peak. Overlap in the 95% CI for these two peaks would mean that the phenotype that maximizes performance and fitness is likely the same (or at least very similar).

This comparison of peaks can then be used to test the following key assumption of the PLA: the fitness driving evolution of the focal phenotypes depends on the modeled performance in the studied functions, such that changes in performance in those functions will impact fitness. If the performance peak does not overlap the 95% CI of the relevant adaptive peak, we would conclude that this key assumption may not hold. Such a result could be driven by three reasons. First, the biomechanical modeling could produce incorrect estimates of performance (e.g., the models include unrealistic assumptions or parameter values). Second, performance in the focal function(s) could have a limited impact on fitness, and this impact may vary across species. Instead, other functions unaccounted for in a given analysis may be more relevant to mean fitness for that specific ecology. Third, suboptimal performance, at least for some functions, may suffice to guarantee survival and reproduction. These last two possibilities, moreover, may stem from trade-offs with performance in functions unexamined in a study. Thus, several observed phenotypes, which could still be close to primary adaptive peaks, may not be optimal for maximizing performance of the studied functions. This latter possibility brings us to the comparison of observed species' phenotypes to primary adaptive peaks.

Comparing Species' Phenotypes with Primary Adaptive Optima

To test the role of historical constraints on functional adaptation, we suggest comparing observed phenotypes of extant species to the inferred primary adaptive peaks. Here, it is worth reemphasizing that the OU framework distinguishes between primary peaks (i.e., the phenotype maximizing fitness only if that specific ecological state influenced the peak location and if no constraints existed) and species-specific peaks (i.e., as indicated by the species' current phenotypes, which reflect other selective environments and constraints; Hansen 1997). Hence, observed phenotypes that deviate from the primary peak are inferred to be under the influence of constraints that hinder functional adaptation to their current (focal) environment. This logic applies to only the ALA because deviations of observed phenotypes from performance

peaks would not directly inform us about constraints, given that performance peaks are not inferred using phylogenetic information and that nonbiomechanical constraints are not modeled in the PLA framework.

To study constraints, we recommend using two related analyses developed by Moen et al. (2016). The first analysis compares the proportion of variance in observed phenotypes around primary adaptive optima that is due to history of adaptation to ancestral environments versus due to species-specific factors (e.g., drift, measurement error, and myriad selective factors affecting the phenotype). The variance due to history is based on deviation of the mean phenotype (i.e., across species) of a specific ecology from the inferred primary adaptive optimum for that ecology. The variance explained by species-specific factors is based on the deviations of individual species' phenotypes from their ecology-specific mean (i.e., the average phenotype of species with a given ecology; for calculations, see Moen et al. 2016). Note that Moen et al. (2016) called these random factors because their collective effect on phenotypes across species was expected to be random with respect to the focal factors affecting phenotypes (Hansen 1997). We expect a higher proportion of phenotypic variance due to history if most species recently changed selective regime. However, an effect of history on current phenotypes could still exist even if species have spent more time in their current environments, particularly if the rate of adaptation is low (e.g., as a high half-life would suggest) because of many constraints (Moen et al. 2016). Such constraints may be genetic (e.g., low genetic variation in the direction of selection) or associated with functional trade-offs. For instance, the adaptive peak may not be easily reached because it was not aligned with directions of high genetic variance, in which the response to selection is faster and of greater magnitude (genetic lines of least resistance; Schluter 1996; Arnold et al. 2001; Melo et al. 2016).

However, we expect a higher proportion of phenotypic variance due to species-specific factors if species have been evolving in their current regimes for a long time and if the rate of adaptation is high (Moen et al. 2016). In such a scenario, the mean phenotype for a specific ecology would be close to the primary adaptive optimum, but individual species' adaptive optima (i.e., their current phenotypes) would differ from that primary optimum. Several factors may contribute to differences in species that evolved in a similar selective regime. Examples include other (unstudied) ecological factors that may influence some species more than others (Hansen and Bartoszek 2012), lumping of diverse ecological states into a single selective regime (Leroi et al. 1994), and different responses to selection due to different strengths of developmental and genetic constraints (Melo et al. 2016).

The second analysis of historical constraints tests whether species that have spent less time in their current environment show a higher deviation from their primary adaptive peak (Moen et al. 2016). A negative relationship between time and deviation from the peak would suggest a lag of adaptation. The first step in this analysis is to use the phylogeny to quantify how long species have been evolving in their current selective regimes (e.g., ecological states). The second step is to calculate distances of species' phenotypes to the inferred primary adaptive optimum of their current regime.

The final step is to use a phylogenetic least squares regression (Martins and Hansen 1997; Revell 2010) to test whether the time for adaptation explains variation in distance from the optimum. Using a phylogenetic method for this regression is important to account for nonindependence between species that share common ancestors and a common regime origin. We recommend log transforming time because the relevance of past selective regimes is modeled as decreasing exponentially over time in the OU model (Hansen 1997; Moen et al. 2016).

Merging Frameworks to Investigate Turtle Shell Evolution

We applied the merged framework to examine functional adaptation of the turtle shell, building on a recent study by Stayton (2019a). We tested whether aquatic and terrestrial lifestyles have influenced the position of both performance and adaptive peaks. Moreover, we examined how much evolution toward these peaks has influenced the phenotypic distribution of extant species. Stayton (2019a) compared several performance and adaptive peaks that were not based on these states. Therefore, our approach of using only two performance and primary adaptive peaks (one of each for aquatic and terrestrial lifestyles) was thus fundamentally different from the analysis of these data by Stayton (2019a). We provided the code and analysis tutorial (using R markdown) for all analyses described in this section in the supplemental files. All analyses were conducted in the R programming environment (R Core Team 2022).

We analyzed the 237 turtle species sampled on the ultrametric time-calibrated consensus phylogeny of McLaughlin and Stayton (2016) that was previously used by Stayton (2019a). Following that study and Stayton et al. (2018), we used 53 three-dimensional shell landmarks, aligned those coordinates using a generalized Procrustes analysis (Zelditch et al. 2012), projected them into a Euclidean tangent space, and then projected each specimen onto the first two PCs (PC1 and PC2) of those coordinates. To calculate the same PC scores as Stayton et al. (2018) and Stayton (2019a), we used all 274 species in the landmark dataset for these initial operations. PC1 separated turtles with more flattened shells from turtles with taller shells, whereas PC2 separated turtles by bridge length and plastron size (Stayton 2019a). For the PLA analysis, we used estimated performance in three functions (shell strength, hydrodynamic drag, and a metric of self-righting ability) to construct performance surfaces that were then combined into a single performance landscape by using a set of relative weights for each function (for further details, see Stayton 2019a). We estimated performance peaks on the combined performance landscape for aquatic and terrestrial species separately with the average relative weights provided by Stayton (2019a) for each type.

For the ALA analysis, we first estimated ancestral states of the ecological states (aquatic vs. terrestrial states) using maximum likelihood (Schluter et al. 1997) via the ace function of the R package ape (ver. 5.0; Paradis and Schliep 2019). The aquatic state was the most likely state for nearly all the deep ancestral nodes, except the node giving rise to the family Testudinidae, a family of all terrestrial species (fig. A1). We next compared a series of multivariate OU models for PC1 and PC2 using the R package ouch (Butler and King

Table 1: Model comparison of turtle shell evolution

Model	k	logLik	AICc	Weight
BM	5	1,120	−2,229	.00
OU1	8	1,139	−2,262	.00
OU2	10	1,154	−2,288	1.00

Note. We compared the following three models: constant-rate Brownian motion (BM), single-peak Ornstein-Uhlenbeck (OU1; one peak for all species), and two-peak Ornstein-Uhlenbeck (OU2; one peak for aquatic species and another peak for terrestrial species) models. All models examined multivariate evolution (with both principal components being used as a combined phenotype) across 237 turtle species. OU2 was the most highly supported model. AICc = corrected Akaike information criterion; k = number of parameters; logLik = log likelihood; weight = AICc model weight.

2004). We found that the shell data best supported a two-peak model (separate primary optima for aquatic and terrestrial species; table 1). However, the peaks were clearly separated for only PC1 (table A1). We obtained the 95% CIs of the two primary adaptive peak positions by parametric bootstrapping with 1,000 iterations (Scales et al. 2009).

We next examined whether these 95% CIs included their respective performance peaks. Whereas the aquatic performance and adaptive peaks matched (fig. 2A), the terrestrial performance peak was displaced to PC1 and PC2 values that were greater than the 95% CI of the terrestrial adaptive peak (fig. 2B). However, if we consider the region of highest performance as being within 95% of maximum performance (Stayton 2019b), the terrestrial performance and adaptive peaks matched (table A1), given that a large portion of the performance landscape showed high performance (fig. 2B). These results indicated that all three functions (shell strength, hydrodynamic drag, and self-righting ability) were relevant to fitness in aquatic and terrestrial species, although the contribution of the first function has been more relevant to fitness in terrestrial species, whereas hydrodynamics have been more relevant in aquatic species.

Although performance and adaptive peaks matched, the phenotypes of several species were far from inferred primary adaptive peaks (fig. 3A). This discord suggests lags in adaptation and/or constraints. The phylogenetic half-lives inferred by the two-peak OU model were 0.136 and 0.104 for PC1 and PC2, respectively, suggesting fast adaptation given the unit length of the phylogeny. We acknowledge, however, that adaptation could be slower, especially for PC1, because of high uncertainty in α values, leading to uncertainty in half-lives (see table A1). Yet because several terrestrial lineages recently transitioned from aquatic ancestors (fig. A2), past selective regimes may still be reflected in current phenotypes.

We tested this latter hypothesis by quantifying the importance of history versus species-specific factors for phenotypes (Moen et al. 2016). We tested such effects separately for aquatic and terrestrial lineages. We first performed this test with all species included and then performed it with species from the family Testudinidae excluded, given that the common ancestor of this family changed habitats much earlier than other terrestrial lineages (around 57 mya; see fig. A2). We found that the proportion of phenotypic
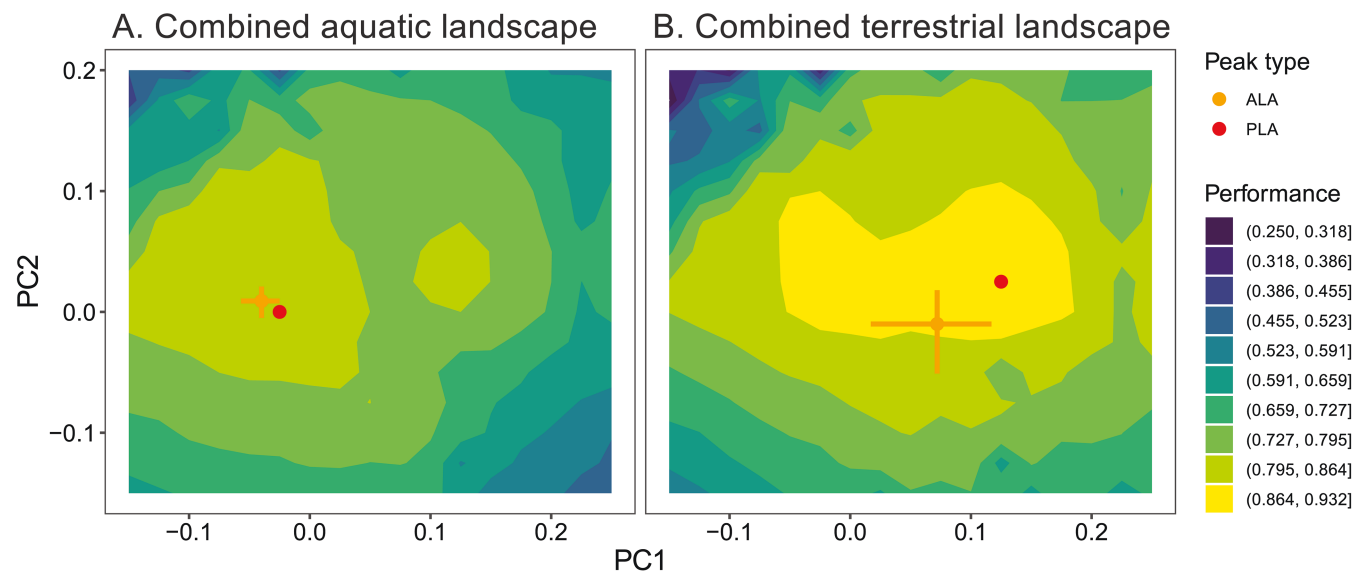
Figure 2. Combined performance landscapes of turtle shell shape for aquatic and terrestrial habitats. The aquatic landscape (*A*) and the terrestrial landscape (*B*) were estimated using a specific set of average weights estimated for the functions measured in hypothetical shell shapes (Stayton 2019*a*; aquatic landscape: shell strength = 0.47, hydrodynamic drag = 0.38, self-righting ability = 0.15; terrestrial landscape: shell strength = 0.68, hydrodynamic drag = 0.18, self-righting ability = 0.16). Red dots mark the position of the performance peak. Orange dots and bars represent the position of primary adaptive peaks and their 95% confidence intervals, respectively. ALA = adaptive landscape approach; PC1 = principal component 1; PC2 = principal component 2; PLA = performance landscape approach.

variation explained by history was much lower in aquatic lineages than in terrestrial lineages, but only when excluding Testudinidae (table 2). This result is consistent with extant turtles being ancestrally aquatic (fig. A2), as lineages that have been aquatic for their entire history have had nearly 10 half-lives to approach the aquatic primary adaptive peak.

In contrast, most terrestrial lineages have been terrestrial for little of their total evolutionary history (fig. A2). Even Testudinidae, the oldest terrestrial lineage, had only around 2.5 half-lives to approach the terrestrial primary peak. This suggests that the observed phenotypes of terrestrial lineages are still influenced by past adaptation to an aquatic environment. Hence, we next tested whether lineages farther from the primary optimum of their current habitat have spent less time in that habitat. Surprisingly, time did not explain variation in distance from the optimum (fig. 3*B*; table A2). Thus, the strong effect of history that we found for terrestrial lineages (table 2) is not generally related to too little time for adaptation after transitioning from an aquatic ancestor (consistent with Stayton 2019*b*). Instead, other constraints (e.g., genetic or developmental constraints) may have limited movement to the primary adaptive peak, or other niche aspects could be influencing terrestrial species' phenotypes. One possibility is that terrestrial species may have higher genetic integration of phenotypic traits that are relevant to performance in the functions measured, consequently showing high genetic variance accumulated in only one or a few directions. Because the response to selection depends on both the location of the optimum and the distribution of genetic variance (Lande 1979), if selection toward the optimum is not aligned with these directions

of high genetic variance, the response to selection may have deviated from the direction of selection, causing a lag in adaptation (Arnold et al. 2001; Melo et al. 2016).

Our worked example demonstrates the new knowledge that is possible as a result of using the merged framework. First, we showed that the different positions of adaptive peaks between aquatic and terrestrial species are related to terrestrial species having a higher relative contribution of shell strength and lower relative contribution of hydrodynamic drag than aquatic species. This inference of which specific functions are essential to each selective regime can be done only by merging the PLA and the ALA. Second, we found evidence that constraints unrelated to biomechanical limitations or adaptations to past environments influenced terrestrial lineages that have recently transitioned to the terrestrial environment. This inference of which types of constraints may limit functional evolution could only be achieved with the merged framework.

Using Theoretical or Real Performance as a Trait in Ornstein-Uhlenbeck Models

Part of the logic behind the merged framework can also be used by directly analyzing performance in an OU framework. Some studies carry out such analyses with modeled performance (as is typically done in the PLA) but then test for multiple adaptive optima in the ALA framework (e.g., Dumont et al. 2014; Rader et al. 2020). For example, Dumont et al. (2014) studied functional adaptation of the skulls of bat species with distinct diets. The authors used a reference skull of a single species and then used computer-aided design
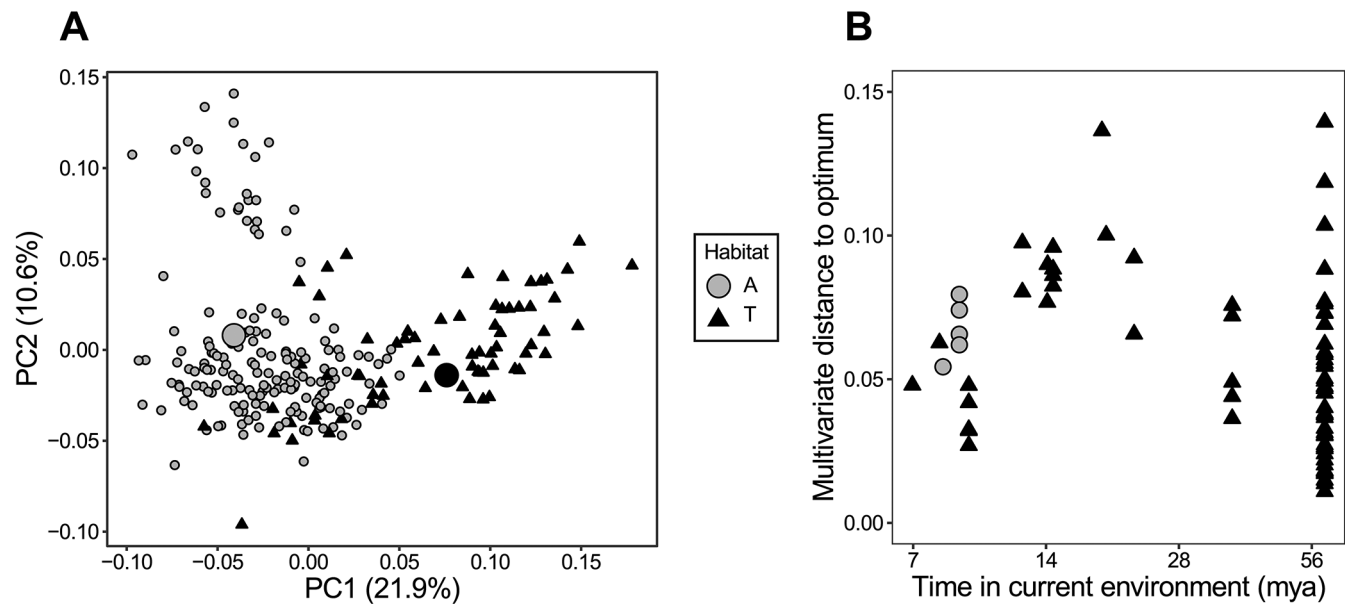
Figure 3. Species phenotypes and inferred primary adaptive optima for aquatic and terrestrial habitats. *A*, Each point is the phenotype of an individual species in shape space, and large circles correspond to the positions of estimated primary adaptive optima. Several terrestrial species (triangles) are closer to the aquatic optimum (large gray circle) than to the terrestrial optimum (large black circle). Aquatic species are represented by the small circles. *B*, Relationship between time spent in current habitat and multivariate distance from the primary adaptive optimum for that habitat. Time is plotted in original units (mya) but on a logged scale. Multivariate distance is in principal component (PC) space, so it has no units. This analysis was done only for lineages that switched habitats since the root of the tree; such transitions were mostly from aquatic to terrestrial habitats (triangles), but some secondary transitions back to aquatic habitats (circles) are included. The accumulated points on the far right of the plot correspond to species in the family Testudinidae. Note that analyses with phylogenetic generalized least squares account for this (and all other) nonindependence.

software to morph the reference to have a similar palate width and length as the skulls of other extant species. They then estimated mechanical advantage and von Mises stress using FE analysis on these hypothetical skulls. Diet hardness (hard or soft), which should determine the importance of force transmission and/or

breaking resistance of the skull, dictated their expectation of where each species should be located in the performance landscape. Dumont et al.'s (2014) results indicated that even though some of the dietary zones (e.g., nectivore, frugivore, and short-faced bats) imposed distinct selection regimes on biomechanical

Table 2: Variance partitioning of turtle shell shape into history and species-specific factors

Lineages	History	Species-specific factors	Total	Prop. history	Prop. species-specific factors
Aquatic lineages:					
PC1	.0002	.0010	.0012	.18	.82
PC2	.0000	.0019	.0019	.02	.98
Multivariate	.0003	.0029	.0032	.08	.92
Terrestrial lineages:					
PC1	.0000	.0028	.0028	.00	1.00
PC2	.0001	.0009	.0010	.14	.85
Multivariate	.0001	.0037	.0038	.04	.96
Terrestrial lineages (excluding Testudinidae):					
PC1	.0026	.0007	.0034	.79	.21
PC2	.0000	.0009	.0009	.00	1.00
Multivariate	.0026	.0016	.0043	.61	.39

Note. Estimated variance for history is based on deviations of aquatic or terrestrial mean phenotypes from their respective inferred adaptive optimum. Variance for species-specific factors is based on deviations of individual species from ecology-specific mean phenotypes. Total variance is the sum of the first two variances. Proportions of phenotypic variance due to history (prop. history) and to species-specific factors (prop. species-specific factors) are each based on component variance divided by the total variance for each row. Prop. history is much higher for terrestrial lineages for principal component 1 (PC1) when excluding Testudinidae, which can be seen in figure 3*A*, as many terrestrial species fall between the aquatic and terrestrial primary adaptive peaks. PC2 = principal component 2.

skull traits, some species with divergent diets were under similar selection regimes.

These results advanced knowledge on selection for diet-associated performance. Yet an explicit interpretation of OU parameters, such as the phylogenetic half-life, can give further insight into the functional adaptation of bat skulls. For example, phylogenetic half-life values in the optimal OU models were all below 1 myr, which corresponds to only 2.85% of the tree length, when using α values in Dumont et al.'s (2014) table 2. Such low half-lives indicate that functional adaptation has been very fast and under very low or no constraints (Hansen 1997), despite some diet types (namely, frugivores) not showing distinct peaks for functional traits such as mechanical advantage. Thus, by combining the PLA and ALA modeling, we can infer that functional adaptation of bat skulls was driven by divergence in diet but that it was also fast and under low ancestral constraints. Such an interpretation is only possible when using the ALA at its full potential (e.g., examining half-lives to understand the rate of adaptation).

We also recognize that good biomechanical or physiological models are often unavailable for estimating the performance landscape. Indeed, this was a key motivation for our analysis of the well-developed turtle shell system of Stayton (2009, 2011, 2018, 2019a, 2019b; see also Polly et al. 2016; Stayton et al. 2018; Polly 2020). However, when such models are lacking, one can combine morphology, physiology, observed (rather than hypothetical) performance, and OU models to better understand functional adaptation. In contrast to the approach that we outlined above, primary adaptive peaks here are estimated for both morphology/physiology and performance, and the concordance of these peaks (or lack thereof) can be compared. In the case of swimming and jumping performances in frogs, distinct morphological phenotypes associated with each of six microhabitats can produce a similar (maximum) jumping performance (Moen 2019). In contrast, swimming performance varies according to the importance of swimming in different microhabitats (i.e., there is one OU peak for species that rarely swim and another peak for species that regularly swim). Hence, Moen (2019) showed that morphological peaks do not always match performance peaks, emphasizing the relevance of including performance in analyses of functional adaptation, be it hypothetical or real performance.

Conclusions

We have shown that integrating the PLA with the ALA leads to a deeper understanding of functional adaptation by filling important gaps. First, the merged framework clarifies whether maximizing performance of specific functions has resulted in increased fitness. Given that the same phenotypic traits may be involved in different functions, it is essential to know which functions affect fitness the most. Yet this knowledge is possible only because uncertainty in the position of adaptive peaks can be measured, allowing comparisons of the locations of performance and adaptive peaks. Incorporating uncertainty in the position of performance peaks would be an interesting avenue for future development of the PLA. For example, instead of a single set of weights, one can use a range of relative weights to combine different performance surfaces (Stayton 2019a).

Moreover, a wider range of the performance landscape (e.g., 3%–5% lower performance than the highest value; Stayton 2019b) can be used for estimating performance peaks, which we found to be important in terrestrial turtle shells. One can also consider a range of parameter values for biomechanical and physiological models to produce uncertainty in the output performance.

Second, by considering the historical perspective of evolution in OU modeling, we can infer the frequency of limits to functional adaptation that are due to constraints and past selection in ancestral environments. Species' phenotypes may deviate from the primary optimum for a variety of reasons, and the merged framework can explicitly test for some of those factors (e.g., history vs. species-specific factors). However, we acknowledge that performance and adaptive landscapes may also change because of functional innovations (Wainwright 2007). How much functional innovations can alleviate constraints on functional adaptation is an important question unincorporated in the merged framework. Nonetheless, the use of the merged framework in a wider range of taxa may better elucidate how much selection on performance actually translates to shifts in adaptive peaks and to phenotypic diversity when multiple constraints are at play. Most functional systems are complex and composed of many interrelated traits, which are subjected to diverse evolutionary processes themselves. Thus, our integrative framework can contribute to elucidating this multifaceted nature of adaptation in form and function.

Acknowledgments

Literature Cited

Adams D.C. and M.L. Collyer. 2018. Multivariate phylogenetic comparative methods: evaluations, comparisons, and recommendations. Syst Biol 67:14–31.

Alfaro M.E., D.I. Bolnick, and P.C. Wainwright. 2004. Evolutionary dynamics of complex biomechanical systems: an example using the four-bar mechanism. Evolution 58:495–503.

———. 2005. Evolutionary consequences of many-to-one mapping of jaw morphology to mechanics in labrid fishes. Am Nat 165:E140–E154.

Anderson P.S.L., T. Claverie, and S.N. Patek. 2014. Levers and linkages: mechanical trade-offs in a power-amplified system. Evolution 68:1919–1933.

Arbour J.H., A.A. Curtis, and S.E. Santana. 2019. Signatures of echolocation and dietary ecology in the adaptive evolution of skull shape in bats. Nat Commun 10:2036.

Arbour J.H. and H. López-Fernández. 2013. Ecological variation in South American geophagine cichlids arose during an early burst of adaptive morphological and functional evolution. Proc R Soc B 280:20130849.

Arnold S.J. 1983. Morphology, performance and fitness. Am Zool 23:347–361.

———. 2003. Performance surfaces and adaptive landscapes. Integr Comp Biol 43:367–375.

Arnold S.J., M.E. Pfrender, and A.G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macro-evolution. Pp. 9–32 in A.P. Hendry and M.T. Kinnison, eds. Microevolution: rate, pattern, process. Springer, Dordrecht, Netherlands.

Bartoszek K., J. Fuentes-González, V. Mitov, J. Pienaar, M. Piwczynski, R. Puchalka, K. Spalik, and K.L. Voje. 2023. Model selection performance in phylogenetic comparative methods under multivariate Ornstein-Uhlenbeck models of trait evolution. Syst Biol 72:syac079. https://doi.org/10.1093/sysbio/syac079.

Bartoszek K., J. Pienaar, P. Mostad, S. Andersson, and T.F. Hansen. 2012. A phylogenetic comparative method for studying multivariate adaptation. J Theor Biol 314:204–215.

Bastide P., C. Ané, S. Robin, and M. Mariadassou. 2018. Inference of adaptive shifts for multivariate correlated traits. Syst Biol 67:662–680.

Beaulieu J.M., D.-C. Jhwueng, C. Boettiger, and B.C. O'Meara. 2012. Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. Evolution 66:2369–2383.

Bergmann P.J. and E.J. McElroy. 2014. Many-to-many mapping of phenotype to performance: an extension of the F-matrix for studying functional complexity. Evol Biol 41:546–560.

Boettiger C., G. Coop, and P. Ralph. 2012. Is your phylogeny informative? measuring the power of comparative methods. Evolution 66:2240–2251.

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

Burress E.D. and M.M. Muñoz. 2023. Functional trade-offs asymmetrically promote phenotypic evolution. Syst Biol 72:150–160.

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

Calsbeek R. and E. Svensson. 2012. The adaptive landscape in evolutionary biology. Oxford University Press, Oxford.

Careau V. and T. Garland Jr. 2012. Performance, personality, and energetics: correlation, causation, and mechanism. Physiol Biochem Zool 85:543–571.

Clavel J., L. Aristide, and H. Morlon. 2019. A penalized likelihood framework for high-dimensional phylogenetic comparative methods and an application to new-world monkeys brain evolution. Syst Biol 68:93–116.

Clavel J., G. Escarguel, and G. Merceron. 2015. mvMORPH: an R package for fitting multivariate evolutionary models to morphometric data. Methods Ecol Evol 6:1311–1319.

Clavel J. and H. Morlon. 2020. Reliable phylogenetic regressions for multivariate comparative data: illustration with the MANOVA and application to the effect of diet on mandible morphology in phyllostomid bats. Syst Biol 69:927–943.

Collar D.C., P.C. Wainwright, M.E. Alfaro, L.J. Revell, and R.S. Mehta. 2014. Biting disrupts integration to spur skull evolution in eels. Nat Commun 5:5505.

Cooper N., G.H. Thomas, C. Venditti, A. Meade, and R.P. Freckleton. 2016. A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies. Biol J Linn Soc 118:64–77.

Corn K.A., C.M. Martinez, E.D. Burress, and P.C. Wainwright. 2021. A multifunction trade-off has contrasting effects on the evolution of form and function. Syst Biol 70:681–693.

Cressler C.E., M.A. Butler, and A.A. King. 2015. Detecting adaptive evolution in phylogenetic comparative analysis using the Ornstein-Uhlenbeck model. Syst Biol 64:953–968.

Cunningham C.W. 1999. Some limitations of ancestral character-state reconstruction when testing evolutionary hypotheses. Syst Biol 48:665–674.

Dickson B.V., J.A. Clack, T.R. Smithson, and S.E. Pierce. 2021. Functional adaptive landscapes predict terrestrial capacity at the origin of limbs. Nature 589:242–245.

Dickson B.V. and S.E. Pierce. 2019. Functional performance of turtle humerus shape across an ecological adaptive landscape. Evolution 73:1265–1277.

Dumont E.R., I.R. Grosse, and G.J. Slater. 2009. Requirements for comparing the performance of finite element models of biological structures. J Theor Biol 256:96–103.

Dumont E.R., K. Samadevam, 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.

Eliason C.M., L. Straker, S. Jung, and S.J. Hackett. 2020. Morphological innovation and biomechanical diversity in plunge-diving birds. Evolution 74:1514–1524.

Farnsworth K.D. and K.J. Niklas. 1995. Theories of optimization, form and function in branching architecture in plants. Funct Ecol 9:355–363.

Felsenstein J. 1988. Phylogenies and quantitative characters. Annu Rev Ecol Syst 19:445–471.

Friedman S.T., S.A. Price, and P.C. Wainwright. 2021. The effect of locomotion mode on body shape evolution in teleost fishes. Integr Org Biol 3:obab016.

Garamszegi L.Z. 2014. Modern phylogenetic comparative methods and their application in evolutionary biology. Springer, Heidelberg.

Garland T., Jr., A.F. Bennett, and E.L. Rezende. 2005. Phylogenetic approaches in comparative physiology. J Exp Biol 208:3015–3035.

Garland T., Jr., C.J. Downs, and A.R. Ives. 2022. Trade-offs (and constraints) in organismal biology. Physiol Biochem Zool 95:82–112.

Ghalambor C.K., J.A. Walker, and D.N. Reznick. 2003. Multi-trait selection, adaptation, and constraints on the evolution of burst swimming performance. Integr Comp Biol 43:431–438.

Goswami A., M. Randau, P.D. Polly, V. Weisbecker, C.V. Bennett, L. Hautier, and M.R. Sánchez-Villagra. 2016. Do developmental constraints and high integration limit the

evolution of the marsupial oral apparatus? Integr Comp Biol 56:404–415.

Grabowski M., J. Pienaar, K.L. Voje, S. Andersson, J. Fuentes-González, B.T. Kopperud, D.S. Moen, M. Tsuboi, J. Uyeda, and T.F. Hansen. 2023. A cautionary note on "A cautionary note on the use of Ornstein Uhlenbeck models in macroevolutionary studies." Syst Biol 72:syad012. https://doi.org/10.1093/sysbio/syad012.

Grossnickle D.M. 2020. Feeding ecology has a stronger evolutionary influence on functional morphology than on body mass in mammals. Evolution 74:610–628.

Hagey T.J., J.C. Uyeda, K.E. Crandell, J.A. Cheney, K. Autumn, and L.J. Harmon. 2017. Tempo and mode of performance evolution across multiple independent origins of adhesive toe pads in lizards. Evolution 71:2344–2358.

Hansen T.F. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351.

———. 2012. Adaptive landscapes and macroevolutionary dynamics. Pp. 205–226 in E. Svensson and R. Calsbeek, eds. The adaptive landscape in evolutionary biology. Oxford University Press, Oxford.

———. 2014. Use and misuse of comparative methods in the study of adaptation. Pp. 351–379 in L.Z. Garamszegi, ed. Modern phylogenetic comparative methods and their application in evolutionary biology. Springer, Heidelberg.

Hansen T.F. and K. Bartoszek. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. Syst Biol 61:413–425.

Hansen T.F. and E.P. Martins. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. Evolution 50:1404–1417.

Hansen T.F., J. Pienaar, and S.H. Orzack. 2008. A comparative method for studying adaptation to a randomly evolving environment. Evolution 62:1965–1977.

Ho L.S.T. and C. Ané. 2013. Asymptotic theory with hierarchical autocorrelation: Ornstein-Uhlenbeck tree models. Ann Stat 41:957–981.

———. 2014. Intrinsic inference difficulties for trait evolution with Ornstein-Uhlenbeck models. Methods Ecol Evol 5:1133–1146.

Hohenlohe P.A. and S.J. Arnold. 2008. MIPoD: a hypothesis-testing framework for microevolutionary inference from patterns of divergence. Am Nat 171:366–385.

Holzman R., D.C. Collar, R.S. Mehta, and P.C. Wainwright. 2011. Functional complexity can mitigate performance trade-offs. Am Nat 177:E69–E83.

Holzman R., T. Keren, M. Kiflawi, C.H. Martin, V. China, O. Mann, and K.H. Olsson. 2022. A new theoretical performance landscape for suction feeding reveals adaptive kinematics in a natural population of reef damselfish. J Exp Biol 225:jeb243273.

Hulsey C.D., M.E. Alfaro, J. Zheng, A. Meyer, and R. Holzman. 2019. Pleiotropic jaw morphology links the evolution of mechanical modularity and functional feeding convergence in Lake Malawi cichlids. Proc R Soc B 286: 20182358.

Hulsey C.D. and P.C. Wainwright. 2002. Projecting mechanics into morphospace: disparity in the feeding system of labrid fishes. Proc R Soc B 269:317–326.

Ingram T. and D.L. Mahler. 2013. SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike information criterion. Methods Ecol Evol 4:416–425.

Jolliffe I.T. and J. Cadima. 2016. Principal component analysis: a review and recent developments. Philos Trans R Soc B 374:20150202.

Khabbazian M., R. Kriebel, K. Rohe, and C. Ané. 2016. Fast and accurate detection of evolutionary shifts in Ornstein-Uhlenbeck models. Methods Ecol Evol 7:811–824.

Kilbourne B.M. and J.R. Hutchinson. 2019. Morphological diversification of biomechanical traits: mustelid locomotor specializations and the macroevolution of long bone cross-sectional morphology. BMC Evol Biol 19:37.

Lande R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution 33:402–416.

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

Lauder G.V. 1981. Form and function: structural analysis in evolutionary morphology. Paleobiology 7:430–442.

———. 1996. The argument from design. Pp. 55–92 in M.R. Rose and G.V. Lauder, eds. Adaptation. Academic Press, San Diego.

Leroi A.M., M.R. Rose, and G.V. Lauder. 1994. What does the comparative method reveal about adaptation? Am Nat 143:381–402.

Losos J.B. 1999. Uncertainty in the reconstruction of ancestral character states and limitations on the use of phylogenetic comparative methods. Anim Behav 58:1319–1324.

———. 2011. Seeing the forest for the trees: the limitations of phylogenies in comparative biology. Am Nat 177:709–727.

Mahler D.L., T. Ingram, L.J. Revell, and J.B. Losos. 2013. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. Science 341:292–295.

Martin C.H., J.A. McGirr, E.J. Richards, and M.E. St. John. 2019. How to investigate the origins of novelty: insights gained from genetic, behavioral, and fitness perspectives. Integr Org Biol 1:obz018.

Martin C.H. and P.C. Wainwright. 2013. Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. Science 339:208–211.

Martinez C.M. and P.C. Wainwright. 2019. Extending the geometric approach for studying biomechanical motions. Integr Comp Biol 59:684–695.

Martins E.P. 2000. Adaptation and the comparative method. Trends Ecol Evol 15:296–299.

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. Am Nat 149:646–667.

McLaughlin C.J. and C.T. Stayton. 2016. Convergent evolution provides evidence of similar radiations in shell shape in the turtle

families Emydidae and Geoemydidae. Herpetologica 72:120–129.

Melo D., A. Porto, J.M. Cheverud, and G. Marroig. 2016. Modularity: genes, development, and evolution. Annu Rev Ecol Evol Syst 47:463–486.

Moen D.S. 2019. What determines the distinct morphology of species with a particular ecology? the roles of many-to-one mapping and trade-offs in the evolution of frog ecomorphology and performance. Am Nat 194:E81–E95.

Moen D.S., E. Cabrera-Guzmán, I.W. Caviedes-Solis, E. González-Bernal, and A.R. Hanna. 2022. Phylogenetic analysis of adaptation in comparative physiology and biomechanics: overview and a case study of thermal physiology in treefrogs. J Exp Biol 225:jeb243292.

Moen D.S., H. Morlon, and J.J. Wiens. 2016. Testing convergence versus history: convergence dominates phenotypic evolution for over 150 million years in frogs. Syst Biol 65:146–160.

Morinaga G. and P.J. Bergmann. 2020. Evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards. Proc R Soc B 287:20200192.

Muñoz M.M. 2019. The evolutionary dynamics of mechanically complex systems. Integr Comp Biol 59:705–715.

Olsson K.H., C.H. Martin, and R. Holzman. 2020. Hydrodynamic simulations of the performance landscape for suction-feeding fishes reveal multiple peaks for different prey types. Integr Comp Biol 60:1251–1267.

O'Meara B.C. 2012. Evolutionary inferences from phylogenies: a review of methods. Annu Rev Ecol Evol Syst 43:267–285.

O'Meara B.C. and J.M. Beaulieu. 2014. Modelling stabilizing selection: the attraction of Ornstein-Uhlenbeck models. Pp. 381–393 in L.Z. Garamszegi, ed. Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice. Springer, Heidelberg.

Paradis E. and K. Schliep. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528.

Pigot A.L., C. Sheard, E.T. Miller, T.P. Bregman, B.G. Freeman, U. Roll, N. Seddon, C.H. Trisos, B.C. Weeks, and J.A. Tobias. 2020. Macroevolutionary convergence connects morphological form to ecological function in birds. Nat Ecol Evol 4:230–239.

Polly P.D. 2020. Functional tradeoffs carry phenotypes across the valley of the shadow of death. Integr Comp Biol 60:1268–1282.

Polly P.D., C.T. Stayton, E.R. Dumont, S.E. Pierce, E.J. Rayfield, and K.D. Angielczyk. 2016. Combining geometric morphometrics and finite element analysis with evolutionary modeling: towards a synthesis. J Vertebr Paleontol 36:e1111225.

Price S.A., S.T. Friedman, and P.C. Wainwright. 2015. How predation shaped fish: the impact of fin spines on body form evolution across teleosts. Proc R Soc B 282:20151428.

R Core Team. 2022. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/.

Rader J.A., T.L. Hedrick, Y. He, and L.D. Waldrop. 2020. Functional morphology of gliding flight. II. Morphology

follows predictions of gliding performance. Integr Comp Biol 60:1297–1308.

Revell L.J. 2010. Phylogenetic signal and linear regression on species data. Methods Ecol Evol 1:319–329.

Rezende E.L. and J.A.F. Diniz-Filho. 2012. Phylogenetic analyses: comparing species to infer adaptations and physiological mechanisms. Compr Physiol 2:639–674.

Rothier P.S., M.N. Simon, G. Marroig, A. Herrel, and T. Kohlsdorf. 2022. Development and function explain the modular evolution of phalanges in gecko lizards. Proc R Soc B 289:20212300.

Scales J.A. and M.A. Butler. 2016a. Adaptive evolution in locomotor performance: how selective pressures and functional relationships produce diversity. Evolution 70:48–61.

———. 2016b. The relationship between microhabitat use, allometry and functional variation in the eyes of Hawaiian Megalagrion damselflies. Funct Ecol 30:356–368.

Scales J.A., A.A. King, and M.A. Butler. 2009. Running for your life or running for your dinner: what drives fiber-type evolution in lizard locomotor muscles? Am Nat 173:543–553.

Schluter D. 1988. Estimating the form of natural selection in a quantitative trait. Evolution 42:849–861.

———. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50:1766–1774.

Schluter D. and D. Nychka. 1994. Exploring fitness surfaces. Am Nat 143:597–616.

Schluter D., T. Price, A.Ø. Mooers, and D. Ludwig. 1997. Likelihood of ancestor states in adaptive radiation. Evolution 51:1699–1711.

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 phenotype space. Science 336:1157–1160.

Simon M.N., R. Brandt, T. Kohlsdorf, and S.J. Arnold. 2019. Bite performance surfaces of three ecologically divergent Iguanidae lizards: relationships with lower jaw bones. Biol J Linn Soc 127:810–825.

Simpson G.G. 1944. Tempo and mode in evolution. Columbia University Press, New York.

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 to analyze both. Evolution 63:1348–1355.

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

———. 2018. Warped finite element models predict whole shell failure in turtle shells. J Anat 233:666–678.

———. 2019a. Performance in three shell functions predicts the phenotypic distribution of hard-shelled turtles. Evolution 73:720–734.

———. 2019b. Performance surface analysis identifies consistent functional patterns across 10 morphologically divergent terrestrial turtle lineages. Integr Comp Biol 59:346–357.

Stayton C.T., L.F. O'Connor, and N.M. Nisivoccia. 2018. The influence of multiple functional demands on morphological

diversification: a test on turtle shells. Evolution 72:1933–1949.

Swiderski D.L. and M.L. Zelditch. 2022. Complex adaptive landscape for a "simple" structure: the role of trade-offs in the evolutionary dynamics of mandibular shape in ground squirrels. Evolution 76:946–965.

Taylor G.K. and A.L.R. Thomas. 2014. Evolutionary biomechanics: selection, phylogeny and constraint. Oxford University Press, Oxford.

Toljagić O., K.L. Voje, M. Matschiner, L.H. Liow, and T.F. Hansen. 2018. Millions of years behind: slow adaptation of ruminants to grasslands. Syst Biol 67:145–157.

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

Uyeda J.C. and L.J. Harmon. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. Syst Biol 63:902–918.

Wainwright P.C. 2007. Functional versus morphological diversity in macroevolution. Annu Rev Ecol Evol Syst 38:381–401.

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? Integr Comp Biol 45:256–262.

Walker J.A. 2007. A general model of functional constraints on phenotypic evolution. Am Nat 170:681–689.

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