Phenotypic Covariation And Morphological Diversification In The Ruminant Skull

Annat Haber

University of Chicago, Committee on Evolutionary Biology, Chicago, IL

Dept. of Zoology, Tel Aviv University, Tel Aviv, Israel

Current address:

Michigan State University, BEACON Center for the Study of Evolution in Action

567 Wilson Rd, room 1450, East Lansing, MI 48824

annat22@gmail.com

Keywords: covariation, constraints, integration, macroevolution, diversification, Ruminantia

Online Appendix A: Additional details on methods and results

Online Appendix B: Simulations exploring the relationship between Average flexibility and rSDE

Dryad files: Data, phylogenetic hypotheses, and annotated R codes for all analyses

Submitted as an article

1

Abstract

Differences between clades in their diversification patterns result from a combination of extrinsic and intrinsic factors. In this study I examined the role of intrinsic factors on the morphological diversification of ruminants, and in particular the differences between bovids and cervids. Using skull morphology, which embodies many of the adaptations that distinguish bovids and cervids, I examined 132 of the 200 extant ruminant species. As a proxy for intrinsic constraints I quantified different aspects of the phenotypic covariation structure within species, and compared them with the among-species divergence patterns of their clades, using phylogenetic comparative methods. My results show that bovids have dispersed into a wider range of directions in morphospace than cervids, and that their overall disparity is higher. Within both bovids and cervids, most species divergence is well aligned with their phenotypic covariance matrices, and those that are better aligned have diverged further away from their ancestor. Bovids' greater disparity and wider range of dispersion is associated with a lower eccentricity of their within-population covariance matrices. These results are consistent with the role of intrinsic constraints in determining amount, range, and direction of dispersion, and demonstrate that intrinsic constraints can influence macroevolutionary patterns even as the covariance structure evolves.

Introduction

Diversity is distributed unevenly across the tree of life; a fact that is at the heart of many research programs in evolutionary biology. This unevenness is the result of both extrinsic and intrinsic factors. While extrinsic factors provide the opportunities for diversification, intrinsic factors determine the variation available for natural selection to work on, and therefore the potential of the population to respond to environmental changes and to take advantage of these opportunities. Studying this potential is therefore crucial for understanding macroevolutionary patterns. A common approach for studying the effect of intrinsic factors is to look at the structure of phenotypic and genetic (co)variation within the population, and compare it to patterns of divergence among populations (Lande 1979; Zeng 1988; Armbruster 1996; Schluter 1996; Arnold et al. 2001; Ackermann and Cheverud 2002; Hansen and Houle 2008; Hohenlohe and Arnold 2008; Marroig et al. 2009; Bolstad et al. 2014). In this study I examine the role of phenotypic covariation and intrinsic constraints in the morphological diversification of the ruminant skull, with particular emphasis on the contrast between bovids and cervids.

According to the quantitative genetics framework (Lande 1979), the rate and direction of short-term evolution depends on the relative alignment between selection and the within-population genetic (**G**) and phenotypic (**P**) covariation, as well as their strength. A closer alignment between the major axis of covariation and selection allows the population to evolve more rapidly in the direction of selection because more variation is available in this direction. A greater angle between them will divert the population further away from the optimum, slowing down its evolution. A direct extension of this framework to the macroevolutionary scale predicts that the divergence of populations will initially be biased by the major axis of their ancestral covariation structure – resulting in a close alignment between the within-population covariation and the among-population divergence – and that this effect will diminish with time (Zeng 1988; Armbruster 1996; Schluter 1996).

Hansen and Houle (2008) expanded this approach to quantify different aspects of the evolutionary potential and provide a null expectation for evaluating the extent of the bias (Hansen and Voje 2011; Bolstad et al. 2014). They defined the evolvability of a single trait as the

expected change in trait value in response to directional selection of unit strength per generation, measured as the mean-scaled trait variance. In a multivariate space, evolvability can vary in different directions and is therefore defined with respect to a specific direction; e.g., that in which selection is acting or that in which divergence has occurred ($e(\mathbf{d}_{SP})$ and $e(\mathbf{d}_{CL})$; see table 1 for details). The conditional evolvability is defined as the expected change in a given direction assuming stabilizing selection in the remaining directions, thus accounting for trait covariances as well as variances. When averaged across random directions these measures provide the null expectation of no bias ($\overline{e}$ and $\overline{c}$ in table 1).

At both the micro- and the macro-evolutionary scales, this framework has focused on the relative alignment between covariation and selection. It measures the amount of variation available in a specific direction and therefore the potential to evolve faster in that direction. However, as the phylogenetic scale and time span of the study increase it becomes more difficult to reconstruct selection and covariation at any given time, and more likely that either or both have changed repeatedly, so that their average estimation (or ancestral reconstruction) becomes less informative. The stability and predictability of **P** and **G** has been a particularly thorny issue when attempting to bridge the micro and macro scales, with the growing understanding that the relevant question is when and how they evolve, rather than if, and that the answer is complex and context-dependent (Arnold et al. 2008; Haber 2015; and references therein). Moreover, on the macro scale, selection is more likely to have shifted often enough and in directions different enough (Uyeda et al. 2011; Jones et al. 2012), presenting the populations with a more diverse array of challenges and opportunities. The probability of going extinct or being outcompeted before reaching the optimum also becomes increasingly relevant (Jablonski 2007; Gomulkiewicz and Houle 2009; Jones et al. 2012). Therefore, we might expect that the long-term evolutionary success of lineages and clades would be largely determined – and therefore better predicted – by their ability to respond effectively to a wide range of selective pressures (Vermeij 1973; Liem 1973; Draghi and Wagner 2008), rather than their ability to respond faster in any particular direction. In this context, 'responding effectively' means reaching close enough to the optimum to avoid extinction or competitive displacement, until selection changes again.

I consider here four measures that could reflect the potential to respond to a wide range of selective pressures: $\bar{e}$, $\bar{c}$, $\bar{f}$, and rSDE(**P**) (see table 1 for details). The average evolvabiliy ($\bar{e}$) and conditional evolvabiliy ($\bar{c}$) developed by Hansen and Houle (2008) measure the variance available in any random direction on average, and thus the average potential to respond to any selection vector. A higher value of $\bar{e}$ or $\bar{c}$ can be interpreted as a greater potential to respond to a wider range of selection vectors, where $\bar{c}$ accounts also for the possibility of stabilizing selection in the remaining directions. Both are proportional to total variance (i.e., matrix size), and $\bar{c}$ accounts also for trait covariances (i.e., matrix shape). Another measure developed based on Hansen and Houle (2008) is the average flexibility ($\bar{f}$; Marroig et al. 2009; Roland 2009), which measures how much the population response is deflected from the direction of selection, on average. A higher value means that the population is less biased by its major axis of covariation and is able to align more closely with a wider range of selection vectors (i.e., more flexible). $\bar{f}$ quantifies matrix shape only, not accounting for the orientation and size of the matrix.

Another metric that quantifies matrix shape only, albeit in a different way, is the relative standard deviation of the eigenvalues (rSDE(**P**) in table 1). This metric, and its variations, has been referred to as integration level in studies of morphological integration (Wagner 1984, Cheverud et al. 1989, Pavlicev et al. 2009, Haber 2011). Its use as a measure of integration has been criticized recently for being merely descriptive with no link to the theory of integration and evolution (Armbruster et al. 2014). However, here I use it to quantify matrix shape – as originally suggested by Van Valen (1974) – and reinterpret it in the context of Hansen and Houle's (2008) model. Since evolvability is defined as the (mean-scaled) variance in a given direction, the eigenvalues of a matrix measure the evolvabilities along its eigenvectors. Therefore, rSDE(**P**) measures the variation in the ability to respond in different directions. Because rSDE(**P**) is standardized by the total variance of the matrix, it captures matrix shape only. When the variation in the sample is distributed more evenly in different directions, matrix shape is less eccentric, and rSDE(**P**) is lower. By definition, matrices with a lower eccentricity are less biased by their major axis of covariation. Therefore, we would expect rSDE(**P**) to have a

tight negative correlation with $\overline{f}$. Indeed, preliminary simulations indicate that this relationship persists for a wide range of matrices that differ in their number of variables, mean correlation, and heterogeneity (see online appendix B). To the extent that the observed covariation matrix reflects underlying integrating factors, a more eccentric matrix would reflect a body-plan that is more strongly integrated overall. However, the link between rSDE(**P**) and integration is not the focus of this study. Instead, I focus here on the link between rSDE(**P**) and macroevolution.

Which variational property best predicts macroevolutionary patterns is still an open question (Hansen 2012; Bolstad et al. 2014). Because $\overline{e}$ and $\overline{c}$ are proportional to matrix size (i.e., total variance), a direct extrapolation to macroevolution assumes a linear relationship between the probability of responding effectively and the variance available to selection at any given time. We might, however, imagine a situation where the relationship is not linear, involving a threshold of variance that needs to be exceeded in order to avoid extinction and/or diverge successfully. When the variation is distributed more evenly across the different dimensions, there would be more directions in which the threshold is likely to be exceeded (for a given matrix size). Therefore, for a given matrix size, a population with a lower eccentricity would be able to respond effectively in more directions, and would be more likely to persist, evolve, and diversify in the long run, regardless of its orientation. This prediction relies on several additional assumptions (see *Discussion*). The extent to which they are warranted in practice and under what conditions is not known, and developing and exploring a comprehensive quantified model is beyond the scope of this study. Here I explore an empirical case in which these additional considerations are likely to be relevant.

Ruminants comprise the bulk of Artiodactyla (Vrba and Schaller 2000; Theodor et al. 2005), and consist of Tragulidae (chevrotains) and Pecora (giraffes, pronghorn, musk deer, bovids, and cervids). Bovids and cervids are currently the most species-rich families within Ruminantia (143 and 51 extant species, respectively). Yet, their taxonomic richness is not too high to be reasonably sampled within the scope of this study, allowing for both good taxonomic coverage and large within-species samples. In addition, their phylogenetic history is relatively well understood (Hernández Fernández and Vrba 2005; Price et al. 2005; Gilbert et al. 2006; Marcot 2007; Agnarsson and May-Collado 2008; Decker et al. 2009). Their fossil record is

exceptionally good for vertebrates (Geist 1985, 1987; Janis 1989; Vrba and Schaller 2000; Marcot 2004; Pitra et al. 2004; Ropiquet and Hassanin 2005; Theodore et al. 2005; Gillbert et al. 2006; Janis 2008; Heywood 2010), providing ample evidence that ruminants have faced a wide range of selective pressures due mostly to global changes such as Plio-Pleistocene glacial cycles and the opening of grasslands. At the same time, the fossil record reveals that cervids have had about the same opportunities as bovids to exploit open habitats and arid grasslands at least since the late Miocene (Janis 2008; Heywood 2010). Yet, no cervid has evolved into a true grazer. Cervids are mostly browsers, occupying habitats with dense vegetation, whereas bovids exhibit a high diversity of feeding adaptations and a wider range of body size and shape (Allard et al. 1992; Spencer 1997; Sinclair 2000; Marcot 2004; Janis 2008; Heywood 2010). These differences have left researchers puzzled as to why cervids have not taken advantage of the same opportunities as bovids (Janis 2008; Heywood 2010). My results support the role of intrinsic constraints in the morphological diversification of both bovids and cervids, and suggest that the greater diversification of bovids is best explained by their lower within-population eccentricity. I discuss these results in the context of ruminant history and the extrapolation from micro- to macroevolution.

Materials and Methods

Data collection and preparation

A total of 2857 skulls were included in this study, representing 5 out of 8 species of tragulids, 3 out of 7 extant species of moschids, 34 out of 51 extant species of cervids, and 87 out of 143 extant species of bovids. All 19 extant cervid genera are represented except *Przewalskium albirostris*, which has been endangered for the last few decades. Of the 50 extant bovid genera, 40 are represented. The remaining 10 consist of only 1 to 3 species each and are all nested within otherwise well-represented clades. All domesticated species were excluded. Taxonomic assignments of specimens were standardized following Grubb (2002). Only prime age adults were measured, as determined by their dentition and cranial sutures. Three-dimensional coordinates were recorded for 43 landmarks using MicroScribe MLX (online table A1). Landmark definition was based on the standard measurements recommended by von den Driesch (1976), as well as other studies of the artiodactyl skull (Janis 1990; Mendoza et al. 2002; Semprebon et al. 2004). Data processing involved unifying the dorsal and ventral aspects of the skull, averaging the left and right, and identifying outliers (see online Appendix A for details). Variances due to measurement error were estimated based on repeated measurements of selected specimens and were found to be between 1 to 3 orders of magnitude lower than the inter-specimen variance (see online Appendix A).

Two datasets of interlandmark distances were created. For the first set ('ILtes'), 107 interlandmark distances were defined based on Delaunay tessellation of the symmetric mean configuration of all specimens, function *delaunayn* in the R package *geometry* (Barber et al. 2012). This procedure maximizes coverage while minimizing redundancy and crossing over spatial modules. The second set ('IL32') included 32 ILMDs selected based on comparability with other studies (e.g., von den Driesch 1974; Janis 1990; Marroig and Cheverud 2004; Mendoza et al. 2002) and interpretability in the context of either function or putative modules (see online table A2). Both datasets were corrected for variation due to subspecies and sex by adding to each value within a subsample the difference between the grand mean and the subsample mean (Sokal and Rohlf 1995; Marroig and Cheverud 2004). No significant effect of sex and subspecies was found on either the orientation or eccentricity of the covariance matrix

for most species. Species for which I found a substantial effect, even if non-significant, were reduced to the largest single subspecies sample. A MANOVA test revealed no significant interaction between sex and subspecies for mean shape.

Quantifying within-population covariation

A **P** matrix was calculated for each species with more than 27 specimens, including 2 species of tragulids, 13 species of cervids, and 32 species of bovids (see fig. 2 for details). Three subspecies of white-tailed and mule deer were analyzed separately, bringing the total number of cervid samples to 17. Variables were scaled by the species mean. Scaling by the mean mitigates the isometric effect of size variation among individuals (within and across species), as well as scale differences among variables, and ensures that the evolvability measures are meaningful in the context of Lande's (1979) equation (Hansen et al. 2011). A non-weighted bending procedure was applied to each matrix to ensure that it is positive-definite (Jorjani et al. 2003; Pavlicev et al. 2009). Each sample was jackknifed specimen-by-specimen to identify aberrant specimens.

The average evolvability ($\bar{e}$), average conditional evolvability ($\bar{c}$), and average flexibility ($\bar{f}$) were calculated using simulations of random selection vectors (Marroig et al. 2009; Rolian 2009; Bolstad et al. 2014; see table 1). rSDE(**P**) was calculated based on the eigenvalues (see table 1) and adjusted for the effect of sample size by subtracting the mean of a 1000 permutations from the observed value. This is equivalent to the correction suggested by Wagner (1984), but applicable to covariances as well as correlations. Confidence intervals were estimated using a non-parametric bootstrap procedure with a bias correction (BCa; DiCiccio and Efron 1996; Carpenter and Bithell 2000) and 999 iterations. The BCa correction was necessary because the pseudovalues distribution is expected to be biased upward and to depend on its mean when bounded by zero and/or one.

The evolution of rSDE(**P**), ($\bar{e}$), and ($\bar{c}$) was explored by fitting several Ornstein-Uhlenbeck models, as well as a Brownian motion model, in order to identify nodes where a shift in the clade's typical values has occurred (interpreted here as selective regime; Hansen 1997; Butler and King 2004). Using R package *maticce* (Hipp and Escudero 2010), I tested five nodes: all ruminants excluding tragulids (node number 8 in Hernández Fernández and Vrba 2005);

Bovidae (node 15); Cervidae (node 25); Cervinae (node 27); and Caprinae (node 139). All possible combinations of the five putative transitions were tested, resulting in 32 alternatives of the OU model and one Brownian motion model (see table 2). Values were ln-transformed for this analysis, and branch length was scaled by tree height. Models were selected based on the relative weights of their AICc scores (Anderson et al. 2000). The maximum likelihood estimate for each regime was estimated as an average of the 32 models, weighted by their AICc weights (Anderson et al. 2000).

Quantifying morphological dispersion

The among-population divergence was measured by the overall amount, rate, and eccentricity of dispersion (table 1). These measures were calculated separately for each clade – Bovidae, Cervidae, Cervinae, and Caprinae, as above – based on its species means. Species means were scaled by the clade's phylogenetically-weighted mean (the inferred root state), assuming a Brownian motion model (function *fastAnc* in the R package *phytools*; Revel 2012). The overall dispersion was calculated as the mean squared morphological distance between each species and the clade's phylognetically-weighted mean (equivalent to Foote 1993's disparity). To facilitate comparisons, the expected distribution under Brownian motion model was generated by simulating a 1000 datasets (function *sim.char* in the R package *Geiger*; Harmon et al. 2008) and recalculating disparity for each simulated dataset. The rate parameters for these simulations were provided by the clade's evolutionary rate matrix ($\mathbf{D}_{IC}$ in table 1). $\mathbf{D}_{IC}$ is the covariance matrix of independent contrasts between species means, assuming a Brownian-motion model, and therefore represents the coevolution of traits while taking into account phylogeny (Revell 2007; Revell and Harmon 2008). The rate of dispersion is the trace of $\mathbf{D}_{IC}$. The eccentricity of dispersion is the relative standard deviation of the eigenvalues of $\mathbf{D}_{IC}$ (rSDE($\mathbf{D}_{IC}$) in table 1). Thus, this metric measures how evenly species have diverged in morphospace. Confidence intervals for the rates and eccentricity of dispersion were estimated using a parametric bootstrap procedure with 500 replications, using the original sample size and assuming a multivariate normal distribution.

Evolvability in the direction of divergence

Species divergence was compared with the amount of mean-scaled phenotypic variance (evolvability) available in the direction of divergence, following Hansen and Houle (2008) and Hansen and Voje (2011). I first calculated the amount of variation that each species has along the eigenvectors of its clade's evolutionary rate matrix ($e(\mathbf{d}_{CL})$; see table 1 for details), using its observed **P**. Species whose **P** matrices are more closely aligned with their clade's diversification would have higher evolvabilities along the first eigenvector of $\mathbf{D}_{IC}$ than along other eigenvectors. I then calculated the amount of variation available for species in the direction of their own divergence from their clade's inferred root state (($e(\mathbf{d}_{SP})$ and $c(\mathbf{d}_{SP})$ in table 1; Hansen and Voje 2011), based on the clade's average **P** ($\mathbf{P}_{AV}$). Ideally, **P** here would be averaged over the clade's history while accounting for phylogeny. However, previous analyses (Haber 2015) indicated that the phylogenetic structure of **P** within each clade has been obliterated substantially. Therefore, the simple average – i.e., assuming 'white noise' model – was preferred over other evolutionary models. The observed evolvabilities were compared to the expected values for random directions ($\overline{e}$ and $\overline{c}$; table 1), also calculated here based on $\mathbf{P}_{AV}$ (Hansen and Houle 2008, Hansen and Voje 2011). Higher evolvabilities indicate species that have diverged closer to the direction of $\mathbf{P}_{AV}$ than expected by chance (Hansen and Voje 2011).

All analyses were repeated for three different phylogenetic hypotheses derived from the literature. Because results were essentially the same regardless of phylogeny, I present below results based on the phylogeny of Hernández Fernández and Vrba (2005) only. Results based on the other two phylogenies are presented in online appendix A. All analyses were carried out using R v.3.0.2 (R development Core Team 2013). The phylogenetic trees were generated and manipulated using packages *ape* (Paradis et al. 2004) and *phytools* (Revell 2012). All R scripts, data, and phylogenetic hypotheses are available on Dryad (###).

Results

Within-population covariation

The two datasets (ILtes, IL32) yield largely the same results (online fig. A3), indicating that the smaller dataset captures most of the relevant information with its 32 variables. Values based on the IL32 datasets are consistently lower than those based on the ILtes datasets,

probably due to a lower overall level of redundancy in the data, and the signal is often weaker. Adjusting rSDE(**P**) for sample size has little effect (online fig. A4).

Average flexibility has a tight inverse correlation with the within-population eccentricity (rSDE(**P**)); average evolvability ($\bar{e}$) is positively but loosely correlated with rSDE(**P**); and average conditional evolvability ($\bar{c}$) is not correlated with rSDE(**P**), nor with $\bar{e}$ (fig. 1). $\bar{c}$ has yielded different results depending on shrinking tolerance, method of matrix inversion, and number of variables (online figure A3), and should therefore be regarded with caution. All rSDE(**P**) values deviate from zero after adjusting for sample size (fig.2), indicating that all species deviate significantly from the null expectation of a random covariation structure. Bovids are in general less variable than cervids (fig. 2); most bovid species do not differ from each other substantially, with the striking exception of Caprinae, whereas cervids show considerable variation throughout the clade and even among conspecific subspecies.

The model fitting results indicate that rSDE(**P**), $\bar{e}$, and $\bar{c}$ have not evolved following a Brownian motion process, but rather as a multi-regime OU process (table 2). The most strongly supported transition for rSDE(**P**) is at the base of Caprinae; all models that do not include a shift at this node (models 17-32) have the lowest AICc weights. The best-supported model for all datasets (model 14 in table 2) includes a shift at the base of Cervidae as well as Caprinae. In addition, there is strong evidence against a transition at the base of bovids plus cervids (node 8FV in table 2), as well as at the base of all bovids. Therefore, bovids and cervids probably do not share the same typical rSDE(**P**) value. Bovids and tragulids, on the other hand, have the same typical value, which most likely characterized their ruminant ancestor as well. Model 10, which includes a shift for cervines in addition to cervids and caprines, is also relatively well supported (see also online tables A3-A5). Therefore, there is some evidence that the typical rSDE(**P**) value of cervines is higher than that of other cervids. The best-supported model for $\bar{e}$ includes a transition for both bovids and caprines (model 8) and none for other clades (table 2), implying that bovids have deviated from the ancestral ruminant state rather than cervids. The best-supported model for $\bar{c}$ includes only a transition for caprine (model 16). Models 14 and 15 are well supported in IL32 (table 2) but not in ILtes (online table A3).

The parameter estimates for the best-supported models are given in table 3, along with their weighted averages. The shifts inferred for rSDE(**P**) involves an increase for caprines and cervids (and possibly cervines). The shifts inferred for $\overline{e}$ involves a decrease for bovids and an increase for caprines relative to cervids and the ancestral state. The shifts inferred for $\overline{c}$ involves a decrease for caprines based on IL32 and an increase based on ILtes. The alpha estimates for rSDE(**P**) are rather high, yielding a phylogenetic half-life of almost 8% of total tree height (calculated as log(2)/alpha; Hansen 1997). This implies that half of the phylogenetic structure has been obliterated within less than 4 My, out of the 50 My of ruminant history. The sigma-squared estimates for the best OU models are about 4 times higher than that of the Brownian motion model, implying a very strong stabilizing effect. The alpha sigma-squared estimates are even higher for $\overline{e}$ and $\overline{c}$.

Morphological dispersion

The position of both species of *Alces* in morphospace deviates substantially from all other cervids (online fig. A2). Therefore, all measures were repeated for cervids with and without *Alces*. Bovids' disparity is at the higher end of the expected range based on a Brownian motion model, especially when caprines are excluded, whereas cervids' disparity is at the lower end, especially when *Alces* is excluded (fig. 3). Bovids' rate of dispersion is somewhat lower than that of cervids, although this difference is not significant and disappears when *Alces* is excluded (fig. 3). Cervids' eccentricity of dispersion is significantly higher than that of bovids, especially when *Alces* is excluded (fig. 3). Caprines' eccentricity and total disparity are lower than that of all bovids, but their rate of dispersion is higher. These findings are somewhat more pronounced based on ILtes and the alternative phylogenies (online figures A6-A7)

Evolvability in the direction of divergence

Most species have substantially higher evolvabilities ($e(\mathbf{d}_{CL})$) along the first eigenvector of their clade's evolutionary rate matrix, $\mathbf{D}_{IC}$, than along other eigenvectors (fig.4). The first eigenvector is associated with 76%-85% of the variation in $\mathbf{D}_{IC}$. The minimal overlap between the range of evovlabilities along the first eigenvector and those along other eigenvectors indicate that the vast majority of species have their **P** matrix more closely aligned with their $\mathbf{D}_{IC}$ than expected by chance. At the same time, the wide range of values along the first eigenvector

indicates that species differ greatly in their alignment. Cervids' median value is only slightly higher than that of bovids, and their range is somewhat wider based on IL32 (though not based on ILtes; online fig. A8). Caprines have the same median and range as bovids. Excluding caprines from bovids (online fig. A10) and *Alces* from cervids (online fig. A11) made no difference.

Most of the observed evolvabilities in the direction of species divergence – $e(\mathbf{d}_{SP})$ and $c(\mathbf{d}_{sp})$ – are closer to the maximum value than to the expected value for random directions (fig. 5). The average evolvability of $\mathbf{P}_{AV}$ is 0.31% of trait mean for bovids, 0.19% for cervids, and 0.06% for caprines, whereas the observed $e(\mathbf{d}_{SP})$ are above 0.4%, 0.6%, and 0.29%, respectively. The average conditional evolvability of $\mathbf{P}_{AV}$ is 0.08% for bovids, 0.05% for cervids, and 0.01% for caprines, whereas all the observed $c(\mathbf{d}_{sp})$ are an order of magnitude higher. Many of the bovid $c(\mathbf{d}_{sp})$ values are over 1%, whereas most cervid $c(\mathbf{d}_{sp})$ values are below 0.5%. When caprines are excluded from bovids, all bovid $c(\mathbf{d}_{sp})$ values are lower than 1% (online fig A15). When *Alces* is excluded from cervids, more $c(\mathbf{d}_{sp})$ values exceed 0.5% (online fig A17), thus narrowing the gap between bovids and cervids.

There is a distinct non-linear association between the observed $e(\mathbf{d}_{sp})$ and $c(\mathbf{d}_{sp})$ values and the amount of morphological divergence, d^2 (fig. 5). Species whose divergence is more closely aligned with the direction of maximum evolvability managed to diverge further away from the clade's ancestral state. The curve flattens close to the maximum value of evolvability at about 150% distance, in both bovids and cervids. Only species whose divergence is well aligned with the direction of the major axis of their $\mathbf{P}_{AV}$ (i.e., their observed $e(\mathbf{d}_{SP})$ value is close to the maximum) managed to diverge beyond that threshold.

Discussion

The role of intrinsic constraints in determining macroevolutionary patterns is still an open question in evolutionary biology. The structure of variation and covariation within the population has been used for studying the role of intrinsic constraints on generational scale, but it is still not clear which properties of that structure best predict evolutionary outcomes on the macro scale, and under what conditions. The question of why some clades are more diverse than others is of particular interest for macroevolutionary studies. The contrast between bovids and cervids is such an example. Previous studies have provided ample evidence that bovids are substantially more diverse than cervids, taxonomically and ecologically (Allard et al. 1992; Spencer 1997; Sinclair 2000, Grubb 2002; Marcot 2004; Janis 2008; Heywood 2010). Based on discrete characters of teeth, Marcot (2004) has shown that bovids have been more diverse morphologically at least since the late Miocene (about 10 Mya). Here I provide additional evidence – based on their skull morphology – that bovids have a higher disparity than cervids and have dispersed into a wider range of directions in morphospace. So far, studies of the fossil record have not been able to explain these differences based on extrinsic factors (Janis 2008; Heywood 2010). My findings suggest that intrinsic constraints have played an influential role in the diversification of ruminants in general, and in the differences between bovids and cervids in particular.

The role of intrinsic constraints was assessed here by comparing various properties of the within-population covariation (**P** and $\mathbf{P}_{AV}$; see table 1) with those of the among-population divergence ($\mathbf{D}_{IC}$), where **P** is assumed to reflect the potential to evolve and diversify and $\mathbf{D}_{IC}$ reflects the actual divergence that has occurred. Comparing **P** and $\mathbf{D}_{IC}$ in terms of their relative alignment reveals that most species have their **P** matrices more closely aligned with the major axis of their clade's $\mathbf{D}_{IC}$ than with any other direction (fig. 4). Cervids have lower evolvabilities along directions in which they have diversified less than bovids, and higher evolvabilities in directions in which they have diversified more than bovids (online fig. A12), again implying that intrinsic constraints have influenced the directions in which bovids and cervids have diversified. In addition, species whose divergence is more closely aligned with the direction of maximum evolvability have diverged further away from their clade's ancestral state (fig. 5). At the same

time, there is a great variation in how well the **P** matrices of different species are aligned with divergence, implying also a great variation among species in the orientation of their **P** matrices (fig. 4). This is in accord with a previous analysis (Haber 2015), which found that closely-related taxa differ in the orientation of their **P** matrices more than expected from their phylogenetic distance (assuming a single-rate Brownian motion model) and that variation among bovid species in their matrix orientation is similar to that among cervid species.

The positive signal in figures 4 and 5 is consistent with a role of intrinsic constraints in ruminant diversification. However, without knowing the exact direction of selection throughout the ruminant history, it is impossible to say whether the close alignment between **P** and $\mathbf{D}_{IC}$ is because covariation has biased divergence or because selection has consistently pushed in the direction of covariation. That said, in the time span included here – 25 My for bovids and 20 My for cervids – it is not likely that selection would be pushing in the direction of **P** for that long unless **P** itself has aligned with selection, and such alignment would likely be due to its constraining effect (Jones et al. 2003; Revell 2007; Arnold et al. 2008). Therefore, although not conclusive, these results suggest that the orientation and size of **P** have likely played a substantial role in determining the direction and magnitude of ruminant divergence. Thus, this study joins others (e.g., Hansen and Voje 2010) in demonstrating that even when covariance structure evolves relatively rapidly, intrinsic constraints could still bias the evolution and divergence of populations. At the same time, these analyses reveal little to no difference between bovids and cervids, and therefore do not provide a good explanation for the finding that bovids' disparity is higher and more evenly distributed than that of cervids.

The analyses presented in figures 4 and 5 focus on the relative alignment between covariation and selection, accounting mostly for matrix orientation and size. As the time span and phylogenetic scale of the study increase, it becomes more difficult to reconstruct this relative alignment at any given time, and more likely that either or both have changed enough to obscure their relative contribution. Moreover, from theoretical standpoint, additional considerations might become more relevant on the macro scale, such as the heterogeneity of selection and the probabilities of extinction and speciation (Vermeij 1973; Liem 1973; Jablonski 2007; Gomulkiewicz and Houle 2009; Jones et al. 2012). Therefore, we might expect

macroevolutionary patterns to be predicted by properties that determine the potential of populations to respond effectively to a wide range of challenges. Here, I considered four such measures: $\overline{e}$, $\overline{c}$, $\overline{f}$, and rSDE(**P**) (see table 1 for details). Each of these measures can be thought of as capturing – in its own way – the potential of the population to respond relatively effectively to a wider range of selective pressures. However, $\overline{c}$ has yielded different results depending on shrinking tolerance, method of matrix inversion, and number of variables (online figure A3), due to its high sensitivity to sample size (relative to number of variables) and strength of covariance. It is therefore considered less reliable in this study, although this is not necessarily the case for conditional evolvabilities along specific directions. $\overline{f}$ has a tight inverse correlation with rSDE(**P**) (fig. 1), in accord with expectations based on the number of variables and heterogeneity of the matrices (see online Appendix B). Therefore, in this study at least, these two metrics capture effectively the same information regarding the potential to evolve into a wide range of directions. $\overline{e}$ is positively but loosely correlated with rSDE(**P**) (fig. 1). Since rSDE(**P**) is scaled by matrix size, this result reflects more than just a scaling association between the mean and variance of the eigenvalues. Instead, it suggests that as matrix size increases, variance tends to be added more in directions with larger eigenvalues (thus increasing the variance of the eigenvalues) rather than randomly or evenly.

The model-fitting results indicate that these properties have evolved within a relatively limited range, which has shifted during the phylogenetic history of ruminants (fig. 2 and table 2). The main shifts are found at the base of cervids and caprines for rSDE(**P**), and at the base of bovids and caprines for $\overline{e}$ (online fig. A5). The high sigma-squared and alpha values (table 3) imply a great lability and little phylogenetic signal among closely-related taxa. Yet, the fact that the observed patterns are best explained by an Ornstein-Uhlenbeck model with multiple optima implies that these properties have been relatively constrained at the family and subfamily scale in spite of a great variation at the lower scales (Hansen 1997; Butler and King 2004). A good fit to the OU model can also occur when a character evolves following Brownian motion with reflective barriers (Harmon et al. 2010). rSDE(**P**) is indeed bound between 0 and 1, and $\overline{e}$ is bound by zero (even if log-transformed for the model fitting), but there is no reason to

expect them to be bound within a different range for different subclades. Therefore, these findings suggest that rSDE(**P**) and $\bar{e}$ have been under selection, either directly or indirectly, likely due to their association with morphological diversification.

The maximum likelihood values inferred for rSDE(**P**) and $\bar{e}$ imply that bovids and cervids do not share the same typical value for either of them. However, the value of $\bar{e}$ inferred for bovids is very close to that of cervids, and even slightly lower. Based on the current theory, we would expect the greater diversification of bovids to be associated with a higher average evolvability, rather than a lower one, if constraints matter. Therefore, $\bar{e}$ cannot explain the greater diversificaton of bovids. The values inferred for rSDE(**P**), on the other hand, are easier to interpret here, especially in association with rSDE($\mathbf{D}_{IC}$). The value of rSDE(**P**) inferred for bovids – as well as the ruminant ancestor – is lower than that of cervids (and independently, caprines; fig. 2 and table 3), suggesting a greater potential to respond effectively to a wider range of selective pressures. Accordingly, bovids have diversified into a wider range of directions in morphospace (a lower rSDE($\mathbf{D}_{IC}$)), in addition to having a higher overall disparity (fig. 3). The finding that cervids have the same overall rate of evolution as bovids even though their disparity is substantially lower (fig. 3) further suggests that cervids' morphological evolution has been less diffusive than that of bovids, reverting back to previously-explored areas more often. Therefore, these findings support a scenario in which bovids have been able to take advantage of various ecological opportunities because they have retained the low within-population eccentricity (and higher flexibility) that they had inherited from their ancestor. Cervids, on the other hand, have taken the path of increasing eccentricity, which has impeded their ability to take advantage of the same opportunities and channeled their diversification.

Caprines dispersion is as spherical as that of other bovids (fig. 3) even though they have the highest within-population eccentricity as a clade (fig. 2 and table 3). This implies that caprines' evolutionary history does not follow the predicted association between rSDE(**P**) and rSDE($\mathbf{D}_{IC}$). However, this apparent conflict may be explained by their early history. Caprines have been identified as one of the most striking examples among ruminants of an explosive

radiation followed by high specialization (Geist 1985, 1987; Gentry 2000; Vrba and Schaller 2000; Ropiquet and Hassanin 2005). Ropiquet and Hassanin (2005) point out that caprine diversification into tribes concurs with the definition of adaptive radiation given by Schluter (2000): they have diversified rapidly soon after the origination of their common ancestor in the late Miocene; they show great morphological variation that follows a distinct environmental gradient; and their origination is associated with a key innovation - short and stocky metacarpals - that has enabled them to invade mountainous habitats not yet occupied by other herbivores. Thus, the exceptional diversification of caprines is largely due to an early burst of rapid radiation. It is possible, therefore, that the lower eccentricity – and higher flexibility – of their common ancestor had allowed it to take advantage of the opening of a new kind of environment, diversify rapidly, and then specialize into a wide range of climatic conditions. According to this scenario, the high eccentricity that characterizes some caprines today has evolved only after their early burst of diversification, along with their subsequent specialization.

The results presented here support the idea that the long-term evolutionary success of lineages and clades is affected by their ability to respond effectively to a wide range of selective pressures (Vermeij 1973; Liem 1973; Draghi and Wagner 2008). In addition, this study suggests that the best predictor of that ability, at least in this case, is their within-population matrix shape, measured as eccentricity (and/or average flexibility). However, the association found here between rSDE(**P**) and macroevolutionary patterns, could also be a mere coincidence. More empirical and theoretical work is required in order to establish the universality of these findings and the validity of the additional assumptions underlying this association. Some empirical evidence comes from research on morphological integration, to the extent that rSDE(**P**) reflects the magnitude of integration. For example, Lower integration was found to be associated with greater functional divergence in humans compared to apes (Rolian 2009; Rolian et al. 2010; Grabowski et al. 2011) and higher module disparity in Carnivora (Goswami and Polly 2011). Some relevant insights come from Draghi and Wagner (2008), who show that populations whose distribution of mutational effects is less eccentric can adapt faster to a wider range of circumstances.

Several assumptions underlie the suggested association between matrix shape and macroevolution. The main assumption is that selection changes frequently enough and in directions different enough from the major axis of **P** (or **G**), thus providing the opportunities for testing and realizing the potential reflected in matrix shape. This is ultimately an empirical question. Recent metanalyses provide some evidence that evolutionary patterns across broad taxonomic and time scales follow a model of erratic peak movements within fixed limits (Estes and Arnold 2007; Uyeda et al. 2011). In addition, simulations could reveal how heterogeneous selection needs to be in order for the potential to respond to a wide range of selective pressures to become more relevant than the potential to respond faster in any particular direction. Another assumption is that of a threshold that needs to be exceeded in order for the population to persist long enough for an effective response. Gomulkiewicz and Houle (2009) developed formulae for calculating this threshold, accounting for the structure of the fitness landscape and population dynamics. On the macro scale, other factors probably contribute as well (Jablonski 2005, 2007, 2008). A yet another assumption is that matrix shape (but not necessarily orientation or size) remains effectively stable throughout the clade's history. This, again, is an empirical question for which there is currently too little data. However, there is some evidence for it in the case of ruminants: matrix orientation has varied within 50% of its possible range (Haber 2015), whereas eccentricity has varied within only 33% (figs. 1 and 2). Simulations also indicate that eccentricity tends to vary less than orientation under most combinations of genetic architecture and selection regimes (Jones et al. 2003, 2012).

The association found here between matrix shape and ruminants' macroevolutionary patterns implies a trade-off between the ability to respond effectively in a wide range of directions and the ability to respond faster along a specific direction. This, in turn, implies a possible dissociation between macro- and microevolution, where matrix eccentricity is associated with the range and amount of clade dispersion on the macro scale, influenced by the heterogeneity of selection, while matrix orientation and size are associated with the direction and magnitude of population divergence relative to the direction and strength of selection on the generational scale. Again, comprehensive simulations could reveal the range of parameters and conditions necessary for this trade-off to occur. In addition, matrix shape could be linked to

516 macroevolution through taxonomic richness, if the probability of reproductive isolation
517 increases when the population moves in directions further away from the main axis of
518 covariation (regardless of magnitude of change) as opposed to moving along a narrow range of
519 directions. This would be in accord with other studies that point to a decoupling between
520 population-level processes and macroevolutionary patterns (Jablonski 2007 and references
521 therein; Rabosky and Matute 2013). Although the focus here is on macroevolutionary patterns,
522 the same rationale should hold for short-term evolution in highly dynamic environments, an
523 issue that has become increasingly relevant with the advance of global climate change (Kopp
524 and Matuszewski 2013). Therefore, studying variational properties that reflect the potential to
525 respond to a wide range of selective pressures could be useful also for understanding short-
526 term population response to environments with low predictability, as well as macroevoluiton.

Acknowledgements

I owe many thanks to my advisors and committee members for their invaluable support, guidance, and advice: Mark Webster, Leigh Van Valen, David Jablonski, Michael Foote, Peter Wagner, Ian Dworkin, Jeff Connor, and Tamar Dayan. This study has benefited greatly from discussions with Miriam Zelditch, Thomas Hansen, Gabriel Marroig, Mihaela Pavlicev, and David Bapst. I thank the following people and institutions for access to their collections and databases: D. Wilson and L. Gordon, The Smithsonian Institution; M. Schulenberg, Field Museum of Natural History; E. Westwig, American Museum of Natural History; R. Sabin, Natural History Museum, London UK; M. Herman, Quex Museum; T. Sharib, Tel Aviv University; F. Mayer, Museum für Naturkunde, Berlin DE; H. van Grouw, Naturalis; C. Conroy, Museum of Vertebrate Zoology at Berkeley; M. Flannery, California Academy of Sciences; J. Chupaska, Museum of Comparative Zoology at Harvard; J. Dines, Natural History Museum of Los Angeles County. Funding for this project was provided by the National Science Foundation (DDIG DEB-0709750), University of Chicago Hinds Fund, University of Chicago Women's Board Travel Award, AMNH Collection Study Grant, Dan David Fellowship, Center for Absorption in Science at the Ministry of Absorption, Israel, Council of Higher Education, Israel, and BEACON Center for the Study of Evolution in Action at Michigan State University. This material is based in part upon work supported by the National Science Foundation under Cooperative Agreement No. DBI-0939454. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Literature cited

Ackermann, R. R., and J. M. Cheverud. 2002. Discerning evolutionary processes in patterns of tamarin (genus *Saguinus*) craniofacial variation. American Journal of Physical Anthropology 117:260-271.

Agnarsson, I., and L. J. May-Collado. 2008. The phylogeny of Cetartiodactyla: The importance of dense taxon sampling, missing data, and the remarkable promise of cytochrome b to provide reliable species-level phylogenies. Molecular Phylogenetics and Evolution 48:964 - 985.

Allard, M. W., M. M. Miyamoto, L. Jarecki, F. Kraus, and M. R. Tennant. 1992. DNA systematics and evolution of the artiodactyl family Bovidae. Proceedings of the National Academy of Sciences of the United States of America 89:3972-3976.

Anderson, D. R., K. P. Burnham, and W. L. Thompson. 2000. Null hypothesis testing: problems, prevalence, and an alternative. The Journal of Wildlife Management 64:912-923.

Armbruster, W. S. 1996. Evolution of floral morphology and function: an integrative approach to adaptation, constraint, and compromise in *Dalechampia* (Euphorbiaceae). Pp. 241–272 *in* D. Lloyd and S. C. H. Barrett, eds. Floral Biology. Chapman & Hall, New York.

Armbruster, W. S., C. Pélabon, G. H. Bolstad, and T. F. Hansen. 2014. Integrated phenotypes: understanding trait covariation in plants and animals. Philosophical Transactions of the Royal Society B: Biological Sciences 369

Arnold, S. J., R. Bürger, P. A. Hohenlohe, B. C. Ajie, and A. G. Jones. 2008. Understanding the evolution and stability of the G-matrix. Evolution 62:2451-2461.

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

Barber, C. B., K. Habel, R. Grasman, R. B. Gramacy, A. Stahel, and D. C. Sterratt. 2014. geometry: Mesh generation and surface tesselation.

Bolstad, G. H., T. F. Hansen, C. Pélabon, M. Falahati-Anbaran, R. Pérez-Barrales, and W. S. Armbruster. 2014. Genetic constraints predict evolutionary divergence in Dalechampia blossoms. Philosophical Transactions of the Royal Society B: Biological Sciences 369

Butler, M. A., and A. A. King. 2004. Phylogenetic comparative analysis: A modeling approach

for adaptive evolution. American Naturalist 164:683-695.

Carpenter, J., and J. Bithell. 2000. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. Statistics in Medicine 19:1141-1164.

Cheverud, J. M., G. P. Wagner, and M. M. Dow. 1989. Methods for the comparative analysis of variation patterns. Systematic Zoology 38:201-213.

Decker, J. E., J. C. Pires, G. C. Conant, S. D. McKay, M. P. Heaton, K. Chen, A. Cooper, J. Vilkki, C. M. Seabury, A. R. Caetano, G. S. Johnson, R. A. Brenneman, O. Hanotte, L. S. Eggert, P. Wiener, J.-J. Kim, K. S. Kim, T. S. Sonstegard, T. Van, Curt P., H. L. Neibergs, J. C. McEwan, R. Brauning, L. L. Coutinho, M. E. Babar, G. A. Wilson, M. C. McClure, M. M. Rolf, J. Kim, R. D. Schnabel, and J. F. Taylor. 2009. Resolving the evolution of extant and extinct ruminants with high-throughput phylogenomics. Proceedings of the National Academy of Sciences 106:18644-18649.

DiCiccio, T. J., and B. Efron. 1996. Bootstrap Confidence Intervals. Statistical Science 11:189-212.

Draghi, J., and G. P. Wagner. 2008. Evolution of evolvability in a developmental model. Evolution 62:301-315.

Estes, S., and S. Arnold. 2007. Resolving the Paradox of Stasis: Models with Stabilizing Selection Explain Evolutionary Divergence on All Timescales. The American Naturalist 169:227-244.

Foote, M. 1993. Discordance and concordance between morphological and taxonomic diversity. Paleobiology 19:185-204.

Geist, V. 1985. On evolutionary patterns in the Caprinae with comments on the punctuated mode of evolution, gradualism and a general model of mammalian evolution. Pp. 15-30 *in* S. Lovari, ed. The Biology and Management of Mountain Ungulates. Croom Helm Ltd, London.

Geist, V. 1987. On speciation in Ice Age mammals, with special reference to cervids and caprids. Canadian Journal of Zoology 65:1067–1084.

Gentry, A. W. 2000. The ruminant radiation. Pp. 11-25 *in* E. S. Vrba and G. B. Schaller, eds. Antelopes, Deer, and Relatives: Fossil Record, Behavioral Ecology, Systematics and

Conservation. Yale University Press, New Haven.

Gilbert, C., A. Ropiquet, and A. Hassanin. 2006. Mitochondrial and nuclear phylogenies of Cervidae (Mammalia, Ruminantia): Systematics, morphology, and biogeography. Molecular Phylogenetics and Evolution 40:101-117.

Gomulkiewicz, R., and D. Houle. 2009. Demographic and Genetic Constraints on Evolution. The American Naturalist 174:E218-E229.

Grubb, P. 2002. Order Artiodactyla. Pp. 637-722 *in* D. E. Wilson and D. M. Reeder, eds. Mammal Species of the World. John Hopkins University Press, Baltimore.

Haber, A. 2011. A Comparative Analysis of Integration Indices. Evolutionary Biology 38:476-488.

Hansen, T. F. 1997. Stabilizing Selection and the Comparative Analysis of Adaptation. Evolution 51:1341-1351.

Hansen, T. F. 2012. Adaptive landscapes and macroevolutionary dynamics. Pp. 205-226 *in* R. Calsbeek and S. E., eds. The adaptive Landscape in Evolutionary Biology. Oxford University Press, Oxford.

Hansen, T. F., and D. Houle. 2008. Measuring and comparing evolvability and constraint in multivariate characters. Journal of Evolutionary Biology 21:1201-1219.

Hansen, T. F., C. Pélabon, and D. Houle. 2011. Heritability is not Evolvability. Evol Biol Evolutionary Biology 38:258-277.

Hansen, T. F., and K. L. Voje. 2011. Deviation from the line of least resistance does not exclude genetic constraints: a comment on Berner et al. (2010). Evolution 65:1821-1822.

Harmon, L. J., J. B. Losos, J. T. Davies, R. G. Gillespie, J. L. Gittleman, B. W. Jennings, K. H. Kozak, M. A. McPeek, F. Moreno-Roark, T. J. Near, A. Purvis, R. E. Ricklefs, D. Schluter, I. Schulte, J. A., O. Seehausen, B. L. Sidlauskas, O. Torres-Carvajal, J. T. Weir, and A. Ø. Mooers. 2010. Early bursts of body size and shape evolution are rare in comparative data. Evolution 64:2385-2396.

Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. GEIGER: investigating evolutionary radiations

. Bioinformatics 24:129-131.

Hernández Fernández, M., and E. S. Vrba. 2005. A complete estimate of the phylogenetic relationships in Ruminantia: a dated species-level supertree of the extant ruminants. Biological Review 80:269-302

Heywood, J. J. N. 2010. Explaining patterns in modern ruminant diversity: contingency or constraint? Biological Journal of the Linnean Society 99:657-672.

Hipp, A. L., and M. Escudero. 2010. MATICCE: mapping transitions in continuous character evolution. Bioinformatics 26:132-133.

Hohenlohe, P. A., and S. J. Arnold. 2008. MIPoD: A Hypothesis-Testing Framework for Microevolutionary Inference from Patterns of Divergence. The American Naturalist 171:366-385.

Jablonski, D. 2005. Mass extinctions and macroevolution. Paleobiology 31:192-210.

Jablonski, D. 2007. Scale and hierarchy in macroevolution. Palaeontology 50:87-109.

Jablonski, D. 2008. Biotic interactions and macroevolution: extensions and mismatches across scales and levels. Evolution 62:715-739.

Janis, C. M. 1989. A climatic explanation for patterns of evolutionary diversity in ungulate mammals. Palaeontology 32:463-481.

Janis, C. M. 1990. Correlation of cranial and dental variables with dietary preferences in mammals: a comparison of macropodoids and ungulates. Memoirs of the Queensland Museum 28:349-366.

Janis, C. M. 2008. An evolutionary history of browsing and grazing ungulates. Pp. 21-45 *in* I. J. Gordon and H. H. T. Prins, eds. The Ecology of Browsing and Grazing. Springer-Verlag, Berlin, Germany.

Jones, A. G., S. J. Arnold, and R. Bürger. 2003. Stability of the G-Matrix in a Population Experiencing Pleiotropic Mutation, Stabilizing Selection, and Genetic Drift. Evolution 57:1747-1760.

Jones, A. G., R. Bürger, S. J. Arnold, P. A. Hohenlohe, and J. C. Uyeda. 2012. The effects of stochastic and episodic movement of the optimum on the evolution of the G-matrix and the response of the trait mean to selection. Journal of Evolutionary Biology 25:2210-2231.

Jorjani, H., L. Klei, and U. Emanuelson. 2003. A Simple Method for Weighted Bending of Genetic (Co)variance Matrices. Journal of Dairy Science 86:677-679.

Kopp, M., and S. Matuszewski. 2014. Rapid evolution of quantitative traits: theoretical perspectives. Evolutionary Applications 7:169-191.

Lande, R. 1979. Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry. Evolution 33:402-416.

Liem, K. F. 1973. Evolutionary Strategies and Morphological Innovations: Cichlid Pharyngeal Jaws. Systematic Zoology 22:425-441.

Marcot, J. D. 2004. Evolutionary Radiations of Ruminantia. Pp. 278. Committee on Evolutionary Biology. University of Chicago, Chicago, Illinois.

Marcot, J. D. 2007. Molecular Phylogeny of Terrestrial Artiodactyls: Conflicts and Resolution. Pp. 4-18 *in* D. R. Prothero and S. E. Foss, eds. The Evolution of Artiodactyla. The Johns Hopkins University Press, Baltimore.

Marroig, G., and J. M. Cheverud. 2004. Cranial evolution in sakis (*Pithecia*, Platyrrhini) I: Interspecific differentiation and allometric patterns. American Journal of Physical Anthropology 125:266-278.

Marroig, G., L. Shirai, A. Porto, F. de Oliveira, and V. De Conto. 2009. The Evolution of Modularity in the Mammalian Skull II: Evolutionary Consequences. Evolutionary Biology 36:136-148.

Mendoza, M., C. M. Janis, and P. Palmqvist. 2002. Characterizing complex craniodental patterns related to feeding behaviour in ungulates: a multivariate approach. Journal of Zoology 258:223-246.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20:289-290.

Pavlicev, M., J. Cheverud, and G. Wagner. 2009. Measuring Morphological Integration Using Eigenvalue Variance. Evolutionary Biology 36:157-170.

Pitra, C., J. Fickela, E. Meijaard, and P. C. Groves. 2004. Evolution and phylogeny of old world deer. Molecular Phylogenetics and Evolution 33:880–895.

Price, S. A., O. R. P. Bininda-Emonds, and J. L. Gittleman. 2005. A complete phylogeny of the

whales, dolphins and even-toed hoofed mammals (Cetartiodactyla). Biological Reviews 80:445–473.

Rabosky, D. L., and D. R. Matute. 2013. Macroevolutionary speciation rates are decoupled from the evolution of intrinsic reproductive isolation in Drosophila and birds. Proceedings of the National Academy of Sciences 110:15354-15359.

Revell, L. J. 2007. Testing the genetic constraint hypothesis in a phylogenetic context: a simulation study. Evolution 61:2720-2727.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217-223.

Revell, L. J., and L. J. Harmon. 2008. Testing quantitative genetic hypotheses about the evolutionary rate matrix for continuous characters. Evolutionary Ecology Research 10:311-331.

Rolian, C. 2009. Integration and Evolvability in Primate Hands and Feet. Evolutionary Biology 36:100-117.

Ropiquet, A., and A. Hassanin. 2005. Molecular phylogeny of caprines (Bovidae, Antilopinae): the question of their origin and diversification during the Miocene. Journal of Zoological Systematics & Evolutionary Research 43:49-60.

Schluter, D. 1996. Adaptive Radiation Along Genetic Lines of Least Resistance. Evolution 50:1766-1774.

Schluter, D. 2000. The Ecology of Adaptive Radiation. Oxford University Press, New York, USA.

Semprebon, G., C. M. Janis, and N. Solounias. 2004. The diets of the Dromomerycidae (Mammalia: Artiodactyla) and their response to Miocene vegetational change. Journal Of Vertebrate Paleontology 24:427-444.

Sinclair, A. R. E. 2000. Adaptation, niche partitioning, and coexistance of African Bovidae: clues to the past *in* E. S. Vrba and G. B. Schaller, eds. Antelopes, Deer, and Relatives: Fossil Record, Behavioral Ecology, Systematics, and Conservation. Yale University Press.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry. W. H. Freeman and company, New York.

Spencer, L. M. 1997. Dietary adaptations of Plio-Pleistocene Bovidae: implications for hominid habitat use. Journal of Human Evolution 32:201-228.

Theodor, J. M., K. D. Rose, and J. Erfurt. 2005. Artiodactyla *in* K. D. Rose and D. J. Archibald, eds. The Rise of Placental Mammals. John Hopkins University Press, Baltimore.

Uyeda, J. C., T. F. Hansen, A. J. Arnold, and J. Pienaar. 2011. The million-year wait for macroevolutionary bursts. Proceedings of the National Academy of Sciences Proceedings of the National Academy of Sciences 108:15908-15913.

Van Valen, L. M. 1974. Multivariate structural statistics in natural history. Journal of Theoretical Biology 45:235-247.

Von den Driesch, A. 1976. A guide to the measurement of animal bones from archaeological sites. Peabody Museum Bulletin 1:1-136.

Vrba, E. S., and G. B. Schaller. 2000. Antelopes, Deer, and Relatives: Fossil Record, Behavioral Ecology, Systematics, and Conservation. Yale University Press

Wagner, G. P. 1984. On the eigenvalue distribution of genetic and phenotypic dispersion matrices: Evidence for a nonrandom organization of quantitative character variation. Journal of Mathematical Biology 21:77-95.

Zeng, Z.-B. 1988. Long-Term Correlated Response, Interpopulation Covariation, and Interspecific Allometry. Evolution 42:363-374.

Table 1. Parameters and metrics included in this study, their symbols, and their interpretations

Symbol	Measure		Meaning
$\mathbf{P}$	Pooled within-population phenotypic V/CV matrix; based on individuals and corrected for sex and subspecies mean		Variation and covariation of traits among individuals within a species
$\mathbf{P}_{AV}$	Average pooled within-population phenotypic V/CV matrix; based on individuals and corrected for species mean, as well as sex and subspecies		Variation and covariation of traits among individuals within a species, averaged across species within a clade.
$\mathbf{D}_{IC}$	Among-population phenotypic V/CV matrix; based on species means and corrected for phylogenetic correlation assuming Brownian-motion model (i.e., V/CV of independent contrasts; Revell et al. 2007)		Evolutionary rate matrix; rate of evolution and coevolution of traits among species within a clade
λ	An eigenvalue of $\mathbf{P}$ or $\mathbf{D}_{IC}$		
β	A vector of p elements, drawn randomly from a normal distribution with mean of zero and SD of one, normed to a unit length		Selection gradient; direction of selection operating on each trait independently of other traits
Average evolvability $\bar{e}$	The average mean-scaled phenotypic variance in $\mathbf{P}$; Proportional to matrix size	$E\left[\beta^t \mathbf{P} \beta\right]$ or $E\left[\lambda(\mathbf{P})\right]$	The expected ability of the population to respond to selection in a random direction independently of other directions
Average conditional evolvability $\bar{c}$	The average mean-scaled phenotypic variance in $\mathbf{P}$, accounting for trait covariances; Includes both matrix size and matrix shape	$E\left[\left(\beta^t \mathbf{P}^{-1} \beta\right)^{-1}\right]$	The expected ability of the population to respond to selection in a random direction, assuming stabilizing selection in other directions
Average flexibility $\bar{f}$	Cosine of the angle between β and the response vector, averaged over random vectors; Matrix shape only	$E\left[\frac{\beta^t \mathbf{P} \beta}{\sqrt{\Sigma(\mathbf{P}\beta)^2}}\right]$	The expected ability of the population to track in the direction of selection; reflecting the range of random directions into which the population can potentially evolve with relative precision
Within-pop. Eccentricity rSDE($\mathbf{P}$)	Relative standard deviation of the eigenvalues of $\mathbf{P}$*; Matrix shape only	$\sqrt{\frac{p\sum(\lambda_i - \bar{\lambda})^2}{(p-1)\sum\lambda\sum\lambda}}$	The relative ability to track in different directions; reflecting the range of orthogonal directions into which the population can potentially evolve with relative precision
Among-pop. eccentricity rSDE($\mathbf{D}_{IC}$)	Relative standard deviation of the eigenvalues of $\mathbf{D}_{IC}$		Eccentricity of dispersion; the range of directions in space into which the species have diverged
$\mathbf{d}_{CL}$	An eigenvector of $\mathbf{D}_{IC}$		Direction of clade's dispersion
$\mathbf{d}_{SP}$	Difference vector between a species mean and its clade's phylogenetically-weighted mean (inferred root state) assuming Brownian motion, standardized by that mean and normed to a unit length.		Direction of morphological divergence between a species and its clade's ancestral state
d^2	Squared distance between a species mean and its clade's inferred root	$\mathbf{d}_{sp}{}^t \mathbf{d}_{sp}$	Amount of morphological divergence between a species and its clade's ancestral state
Disparity	Average distance between species means and the clade's inferred root	$\frac{\sum d^2}{(N_{sp}-1)}$	Total amount of clade's dispersion
	Trace of $\mathbf{D}_{IC}$		Rate of dispersion
$e(\mathbf{d}_{SP})$	Mean-scaled phenotypic variance in the direction of species divergence	$\mathbf{d}_{sp}{}^t \mathbf{P}_{av} \mathbf{d}_{sp}$	The variation available for species, on average, in the direction of their divergence
$c(\mathbf{d}_{SP})$	Mean-scaled phenotypic variance in the direction of species divergence, accounting for trait covariances	$E\left[\frac{1}{\mathbf{d}_{sp}{}^t \mathbf{P}_{av}{}^{-1} \mathbf{d}_{sp}}\right]$	The variation available for species, on average, in the direction of their divergence assuming stabilizing selection in other directions
$e(\mathbf{d}_{CL})$	Mean-scaled phenotypic variance available along the eigenvectors of $\mathbf{D}_{IC}$	$\mathbf{d}_{CL}{}^t \mathbf{P} \mathbf{d}_{CL}$	The variation available for each species in the direction of their clade's diversification

* p is number of traits

** N_{SP} is number of species

Table 2. The different models describing the evolution of of rSDE(**P**), and $\overline{e}$. Based on the IL32 dataset. The best-supported models are in bold. K is the number of parameters in the model.

		Nodes included in the model (1)					rSDE(**P**)		$\overline{e}$	
Model No.	K	8FV*	Cervidae	Cervinae	Bovidae	Caprinae	AICc scores	AICc weights	AICc scores	AICc weights
1	8	1	1	1	1	1	93.24	0.0096	11.11	0.0045
2	7	0	1	1	1	1	90.42	0.0392	8.28	0.0186
3	7	1	0	1	1	1	90.42	0.0392	8.28	0.0186
4	6	0	0	1	1	1	90.07	0.0466	6.08	0.0561
5	7	1	1	0	1	1	91.79	0.0197	8.31	0.0184
6	6	0	1	0	1	1	89.10	0.0759	5.62	0.0707
7	6	1	0	0	1	1	89.10	0.0759	5.62	0.0707
8	5	0	0	0	1	1	89.87	0.0517	**3.50**	**0.2032**
9	7	1	1	1	0	1	90.42	0.0392	8.28	0.0186
10	6	0	1	1	0	1	**87.74**	**0.1497**	8.51	0.0166
11	6	1	0	1	0	1	94.17	0.0060	10.12	0.0074
12	5	0	0	1	0	1	91.65	0.0212	9.35	0.0110
13	6	1	1	0	0	1	89.10	0.0759	5.62	0.0707
14	5	0	1	0	0	1	**86.54**	**0.2728**	5.96	0.0594
15	5	1	0	0	0	1	96.02	0.0024	8.69	0.0152
16	4	0	0	0	0	1	93.73	0.0075	7.81	0.0236
17	7	1	1	1	1	0	98.92	0.0006	12.79	0.0020
18	6	0	1	1	1	0	96.22	0.0022	10.09	0.0075
19	6	1	0	1	1	0	96.22	0.0022	10.09	0.0075
20	5	0	0	1	1	0	95.01	0.0039	7.97	0.0218
21	6	1	1	0	1	0	97.43	0.0012	10.12	0.0074
22	5	0	1	0	1	0	94.85	0.0043	7.54	0.0270
23	5	1	0	0	1	0	94.85	0.0043	7.54	0.0270
24	4	0	0	0	1	0	94.65	0.0047	5.51	0.0747
25	6	1	1	1	0	0	96.22	0.0022	10.09	0.0075
26	5	0	1	1	0	0	93.69	0.0076	9.58	0.0097
27	5	1	0	1	0	0	96.03	0.0024	10.18	0.0072
28	4	0	0	1	0	0	93.73	0.0075	9.15	0.0121
29	5	1	1	0	0	0	94.85	0.0043	7.54	0.0270
30	4	0	1	0	0	0	92.43	0.0144	7.14	0.0330
31	4	1	0	0	0	0	96.83	0.0016	8.50	0.0167
32	3	0	0	0	0	0	94.76	0.0045	7.43	0.0285
BM	2	0	0	0	0	0	102.01	0.0001	36.35	0.0000

*8FV – Node No. 8 in Hernández Fernández and Vrba (2005), including Bovidae plus Cervidae without Tragulidae

Table 3. Maximum likelihood estimates for the best-supported models for the evolution of rSDE(**P**), and $\bar{e}$ (see table 2 and online table A3). Nodes for which a transition was inferred are in bold. Estimates for Brownian motion (BM) model are also given. Average values are weighted by AIC weights. All data are mean-scaled. Branch length is scaled by tree height. Theta values of rSDE(**P**) are adjusted for sample size as in fig. 1.

			Theta						
	Dataset	Model No.	8FV	Cervidae	Cervinae	Bovidae	Caprinae	Alpha	Sigma-squared
Within-pop. Eccentricity rSDE(**P**)	IL32 *	14	0.07	**0.15**	0.15	0.07	**0.16**	9.67	5.23
		10	0.07	**0.13**	**0.20**	0.07	**0.16**	9.61	5.06
		Average	0.09	**0.14**	0.16	0.07	**0.15**	8.96	4.91
		BM			0.08			--	1.38
	ILtes †	14	0.15	**0.23**	0.23	0.15	**0.31**	9.48	2.48
		10	0.15	**0.21**	**0.28**	0.15	**0.31**	9.21	2.36
		Average	0.17	**0.21**	0.24	0.15	**0.31**	8.64	2.29
		BM			0.16			--	0.66
Average evolvability $\bar{e}$	IL32 *	8	0.0057	0.0057	0.0057	**0.0048**	**0.0061**	70.45	7.13
		Average	0.0065	0.0056	0.0057	**0.0049**	**0.0057**	113.62	12.17
		BM			0.0055			--	0.38
	ILtes †	8	0.0050	0.0050	0.0050	**0.0040**	**0.0060**	34.28	3.69
		14	0.0040	**0.0050**	0.0050	0.0040	**0.0060**	33.59	3.75
		Average	0.0054	0.0049	0.0050	**0.0040**	**0.0059**	32.88	3.57
		BM			0.0046			--	0.37

* Reduced dataset of 32 interlandmark distances

† Tessellation-based dataset of 107 interlandmark distances

Figure legends

Figure 1: Comparison of **P**-matrix properties based on the IL32 dataset, mean-scaled. See table 1 for details on how these properties are measured and interpreted. Solid line indicates regression line. Grey bars indicate 95% confidence intervals based on non-parametric bootstrap with BCa correction.

Figure 2. The relative standard deviation of eigenvalues (rSDE(**P**)) calculated for the IL32 dataset, mean-scaled. Values were adjusted for sample size by subtracting the mean permutation value from the observed value. Therefore, zero represents the expected value for a random matrix. Sample sizes are given in parentheses. Phylogenetic relationships follow Hernandez-Fernandez and Vrba (2005). Vertical dashed lines represent the maximum likelihood estimates (see tables 2 and 3).

Figure 3. Measures of diversification for the four clades that involve a shift in their typical within-population eccentricity (see figure 1 and table 2), based on the IL32 dataset. Their inferred rSED(**P**) and $\overline{e}$ values are also shown (see table 3). All metrics are proportional to the clade's phylogenetically-weighted mean. Bars in the top panel (disparity) represent the expected distribution under Brownian motion model (95% interval). Bars in the lower two panels (rates and eccentricity) represent 95% confidence interval. Analyses involving Cervidae were repeated with (cross) and without (circle) *Alces*. Analyses involving Bovidae were repeated with (cross) and without (circle) Caprinae.

Figure 4: Evolvabilities in directions of clade's diversification ($e(\mathbf{d}_{CL})$), measured as the amount of variation that each species has in its **P**-matrix in the direction of its clade's eigenvectors. Eigenvectors are ordered by the relative size of their eigenvalues, indicated along the x-axis.

Figure 5: Evolvabilities ($e(\mathbf{d}_{sp})$) and conditional evolvabilities ($c(\mathbf{d}_{sp})$) in the direction of species divergence, calculated based on the average within-species covariance matrix ($\mathbf{P}_{AV}$), plotted against the amount of divergence between each species and its clade's ancestral state (d^2). Everything is proportional to trait mean. d^2 is ln-transformed for clarity. Horizontal bars mark the minimum and maximum evolvabilities (solid), the expected evolvability ($\overline{e}$, dashed) and conditional evolvability ($\overline{c}$, dotted). Caprines are highlighted in black in the top graph. Based on IL32

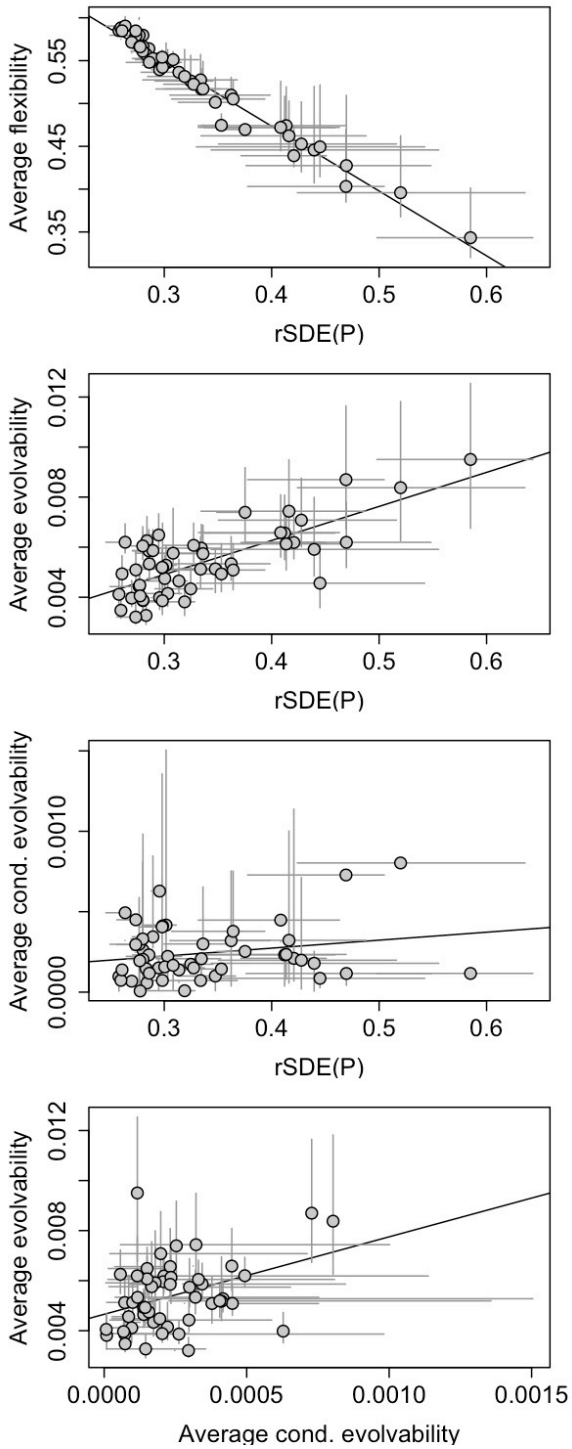

Figure 1: Comparison of **P**-matrix properties based on the IL32 dataset, mean-scaled. See table 1 for details on how these properties are measured and interpreted. Solid line indicates regression line. Grey bars indicate 95% confidence intervals based on non-parametric bootstrap with BCa correction.

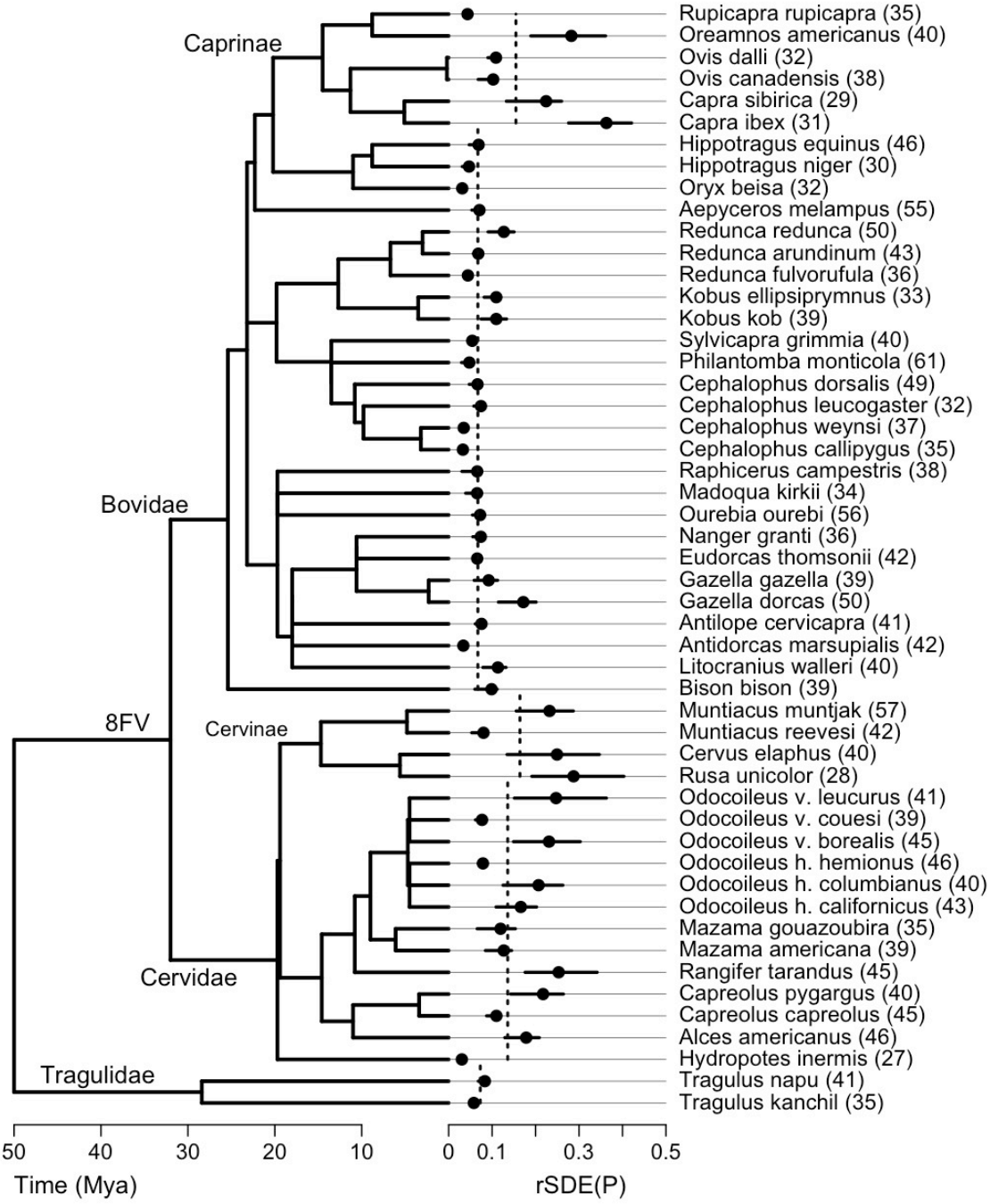

Figure 2. The relative standard deviation of eigenvalues (rSDE(**P**)) calculated for the IL32 dataset, mean-scaled. Values were adjusted for sample size by subtracting the mean permutation value from the observed value. Therefore, zero represents the expected value for a random matrix. Sample sizes are given in parentheses. Phylogenetic relationships follow Hernandez-Fernandez and Vrba (2005). Vertical dashed lines represent the maximum likelihood estimates (see tables 2 and 3).

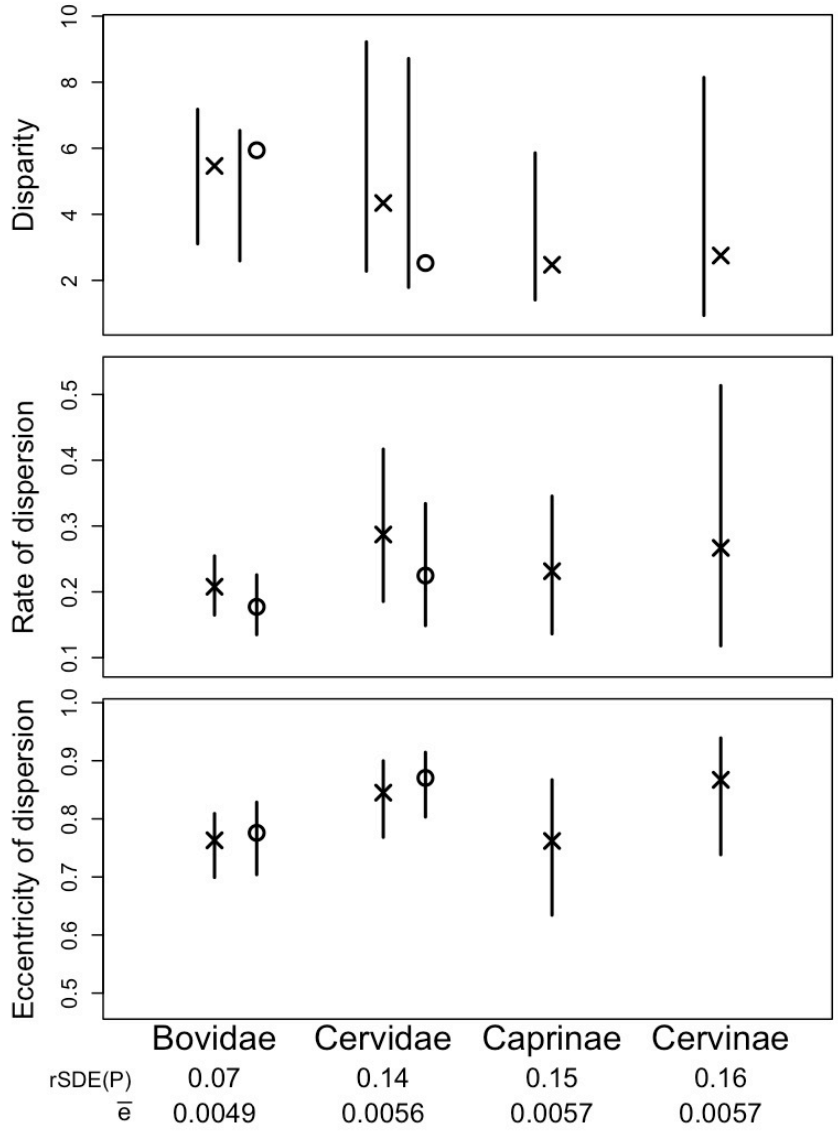

Figure 3. Measures of diversification for the four clades that involve a shift in their typical within-population eccentricity (see figure 1 and table 2), based on the IL32 dataset. Their inferred rSED(**P**) and $\bar{e}$ values are also shown (see table 3). All metrics are proportional to the clade's phylogenetically-weighted mean. Bars in the top panel (disparity) represent the expected distribution under Brownian motion model (95% interval). Bars in the lower two panels (rates and eccentricity) represent 95% confidence interval. Analyses involving Cervidae were repeated with (cross) and without (circle) *Alces*. Analyses involving Bovidae were repeated with (cross) and without (circle) Caprinae.

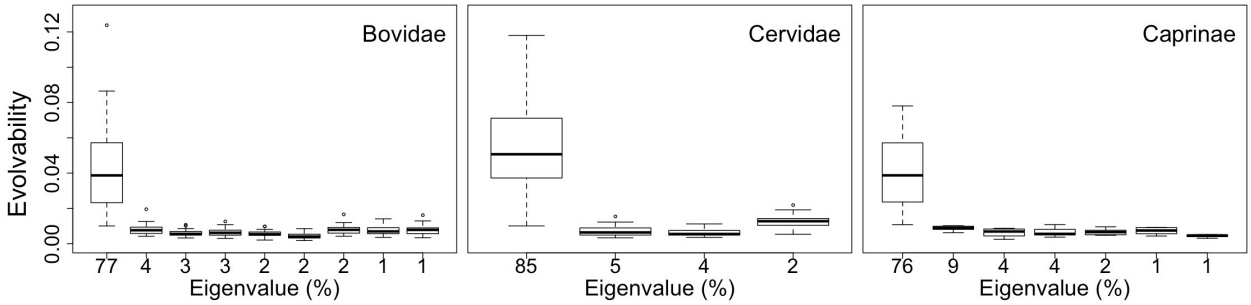

Figure 4: Evolvabilities in directions of clade's diversification ($e(\mathbf{d}_{CL})$), measured as the amount of variation that each species has in its **P**-matrix in the direction of its clade's eigenvectors. Eigenvectors are ordered by the relative size of their eigenvalues, indicated along the x-axis.

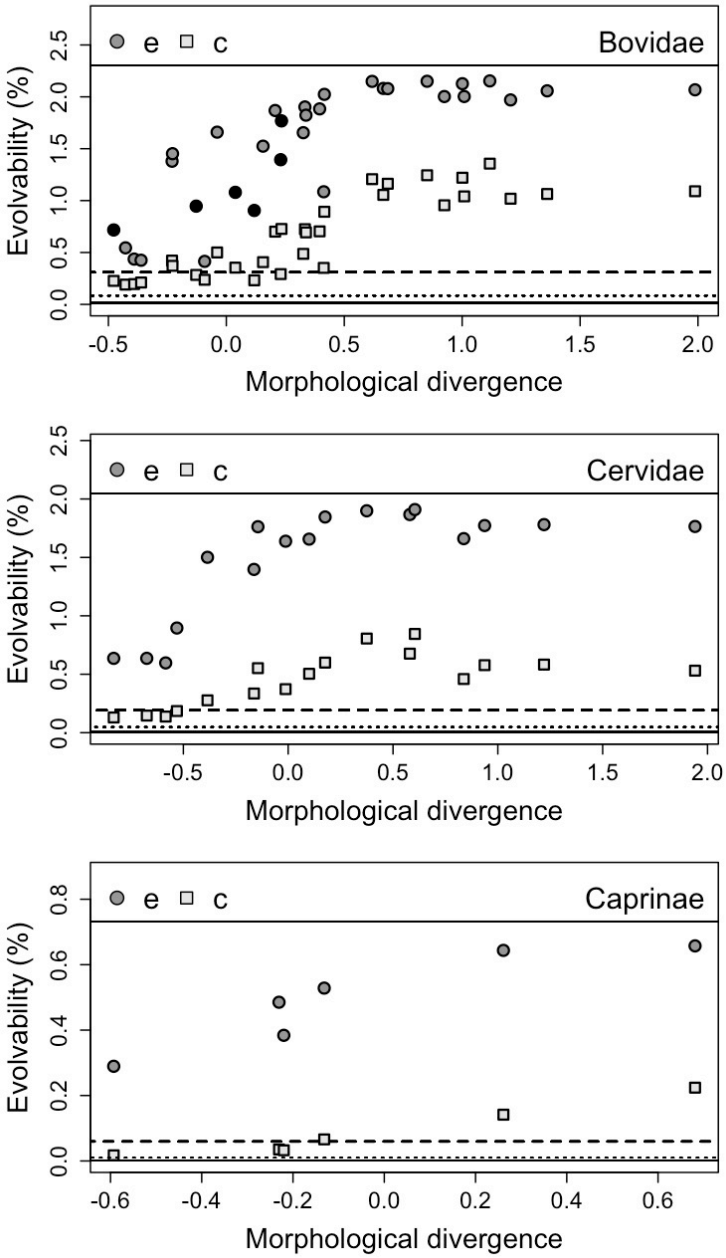

Figure 5: Evolvabilities ($e(\mathbf{d}_{sp})$) and conditional evolvabilities ($c(\mathbf{d}_{sp})$) in the direction of species divergence, calculated based on the average within-species covariance matrix ($\mathbf{P}_{AV}$), plotted against the amount of divergence between each species and its clade's ancestral state (d^2). Everything is proportional to trait mean. d^2 is ln-transformed for clarity. Horizontal bars mark the minimum and maximum evolvabilities (solid), the expected evolvability ($\overline{e}$, dashed) and conditional evolvability ($\overline{c}$, dotted). Caprines are highlighted in black in the top graph. Based on IL32