Speciation along a shared evolutionary trajectory

Ned A. Dochtermann and Marjorie D. Matocq

A major goal in evolutionary biology is to identify the patterns and processes associated with divergence among groups of organisms. Describing and quantifying patterns of divergence is necessarily multifaceted because groups of organisms can diverge from one another in several ways. For example, groups may exhibit phenotypic divergence, wherein they differ from one another in mean trait values of phenotypic characters ranging from morphology to behavior. Groups may also exhibit genotypic divergence wherein genetic differentiation has occurred. Additionally, groups may diverge in their potential evolutionary responses to selection, placing them on distinct evolutionary trajectories. This third type of divergence, which we will refer to as response divergence, has been suggested but has rarely been explicitly tested and compared with phenotypic and genotypic divergence of natural populations (Cheverud and Marroig 2007).

Response divergence, as defined here, is largely an outcome of differences among groups in the patterns of how traits vary and covary. Trait variances and covariances are of central importance in evolutionary biology (Steppan et al. 2002) because these relationships among traits, both phenotypically and genetically, determine how populations may respond to selection (Lande and Arnold 1983). Currently many methods are available for comparing how populations differ in trait variances and covariances (reviewed by Roff et al. 2012, Aguirre et al. 2014). One criticism of most methods, however, is that they are often detached from the general body of evolutionary theory and so inferences can be difficult to tie to evolutionary outcomes regarding population divergence and selection responses (Aguirre et al. 2014).

Group differences in (co)variances connect to differences in selection responses because how traits covary within groups and relative differences in trait variability among groups change the evolutionary trajectories available, that is, response divergence. For example, if 2 groups are evolving toward the same optimum and from the same starting point, how traits are correlated and how much variation is present within traits will affect how the populations move along a fitness landscape and how many generations are required to reach an optimal phenotype. While these populations will eventually reach the same optima (if stationary), the amount of time required to do so can differ greatly as will mean genotypes and
Response divergence can best be understood in the context of the multivariate response to selection. The evolutionary change in trait means as a response to linear selection (denoted as the vector $\Delta z$) can be estimated as (Lande 1979; Lande and Arnold 1983):

$$\Delta z = G\beta$$  \hspace{1cm} (1)

where $G$ is a matrix with additive genetic variances of traits along its diagonal and additive genetic covariances off the diagonal and $\beta$ is the selection gradient (i.e., selection differentials standardized by trait (co)variances). In contrast to phenotypic divergence and genotypic divergence, which are current attributes of populations, response divergence corresponds to differences in $\Delta z$, the expected change of trait means in response to selection. The set of values in $\Delta z$ corresponds to a location in multivariate space and, for the same $\beta$, groups may end up with different genotypic and phenotypic trait means even after just one generation of selection due to differences in $G$. Such a difference represents a divergence in evolutionary trajectories, placing groups on different evolutionary paths.

Whether groups exhibit response divergence in this sense will largely be based on how the (co)variation summarized by $G$ is oriented in multivariate space. For example, for any $G$ there will be a direction in multivariate space in which the greatest additive genetic variance is oriented (i.e., “$g_{\text{max}}$”, Schluter 1996). $g_{\text{max}}$ influences selection responses because selection will not initially be simply in the direction of greatest fitness increase but instead will also be influenced by the direction of greatest variation (Figure 1B; this curving effect continues until $g_{\text{max}}$ is oriented to the optima). The influence of $g_{\text{max}}$ on evolutionary processes can be quite profound, with differences in $g_{\text{max}}$ explaining a substantial amount of variation in patterns of morphological divergence (e.g., Schluter 1996). Differences in $g_{\text{max}}$ can, therefore, be indicative of response divergence. However, differences in $g_{\text{max}}$ may not sufficiently demonstrate response divergence as other directions in multivariate space may contain substantive amounts of variation and likewise impose effects on evolutionary responses. Thus, an overall demonstration of response divergence at the level of $G$ becomes necessary.

Cheverud and Marroig (2007) and Calsbeek and Goodnight (2009) have, however, argued that the statistical demonstration of statistical differences in phenotypes and in $G$ do not necessarily correspond to divergence in available evolutionary trajectories. As a result, determining how phenotypic and genotypic divergence relate to response divergence can be difficult to determine (Cheverud and Marroig 2007). Indeed, the correspondence between phenotypic, genotypic, and response divergence has rarely been assessed (but see de Oliveira et al. 2009, Porto et al. 2009).

While not typically explicitly articulated, the concept of response divergence is consistent with many existing approaches to evaluating differences among population $G$‘s. For example, one approach of comparing matrices is to challenge $G$ matrices from two or more groups with random $\beta$‘s (i.e., multiplying the $G$ matrices by $\beta$ vectors, equation 1) and compare the response vectors ($\Delta z$‘s). This approach—known as a “random skews” analysis (Cheverud and Marroig 2007)—then estimates the correlation between responses allowing a pairwise estimation of response divergence. An analogous approach, “selection skews”, incorporates phenotypic information (Calsbeek and Goodnight 2009). Similarly, the angle between $g_{\text{max}}$ vectors for each pair of groups will estimate the degree of response divergence imposed by one direction of available variation. Comparisons of $G$ matrix eigenvalues achieves a similar function (Kranzowski 1979; Roff et al. 2012; Aguirre et al. 2014), while more complicated analyses can reveal additional details of how matrices might differ (e.g., “tensor analysis”: (Hine et al. 2009, Aguirre et al. 2014); “hierarchical analysis” (Arnold and Phillips 1999; Roff et al. 2012)). Of this variety of approaches, random and selection skews approaches along with a calculation of the angle between $g_{\text{max}}$ vectors best test for the presence of response divergence while alternative analyses can identify specific differences among populations in how traits vary and covary.

To address correspondence among the phenotypic, genotypic, and response divergence and to understand the potential evolutionary dynamics between recently diverged species in secondary contact, we sought to determine whether previously identified phenotypic and genotypic differences between 2 species of woodrats (Neotoma fuscipes and Neotoma macrotis) were associated with response divergence. We tested for response divergence in skull morphology between the 2 species using 1) random skews analysis, which tests for congruence between groups in response to random selection gradients; and 2) a test of whether most morphological
variation was oriented differently in multivariate space for the 2 species. We also 3) used hierarchical (Flury) analyses, and 4) modified tensor analyses to determine relevant matrix differences that might contribute to response divergence.

Materials and Methods

To test for interspecific phenotypic, genotypic, and response divergence we used data previously collected from *N. fuscipes* and *N. macrotis* (Matocq and Murphy 2007). These two species are sister lineages that diverged from a common ancestor approximately 2 million years ago (Matocq 2002b) and share 2 major contact zones: one in the Salinas Valley of California (Matocq 2002a) and the other in the foothills of the central Sierra Nevada of California (Matocq and Murphy 2007).

Matocq and Murphy (2007) measured 11 craniodental traits for 82 individuals captured across the contact zone in the Sierra Nevada of California (Figure 2). We adjusted measurements for differences due to allometry (Matocq and Murphy 2007) and estimated covariances among these craniodental traits. Phenotypic covariances often correspond to genetic covariances (Cheverud 1988; Roff 1996; Dochtermann 2011), a common and often appropriate assumption for morphological traits like skull morphology (Cheverud 1996; de Oliveira et al. 2009; Porto et al. 2009) but see (Willis et al. 1991), and thus we used phenotypic variances and covariances in lieu of genetic variances and covariances to test for evolutionary divergence. Unfortunately, because these were wild caught individuals and we used phenotypic measure, these data could confound differences between the populations in gene by environment interactions with genetic differences in morphology. Because of our final results we do not consider such confounding to significantly affect our inferences.

To determine whether the 2 species exhibited response divergence, we conducted a “random skewers” analysis (Cheverud and Marroig 2007). Using phenotypic covariance matrices we determined whether the 2 *Neotoma* species have diverged to the point of now being on different evolutionary trajectories. The random skewers approach tests for differences among groups in response to selection (i.e., $\Delta z$; equation 1) when exposed to the same selection gradient (i.e., $\mathbf{b}$; equation 1). This difference is tested by applying a set of randomly generated selection gradients (here $N=1,000$) to the covariance matrices estimated for each group. We then evaluated the vector correlation ($r^*$, i.e., $r^* = \mathbf{B} \mathbf{b}$) between groups among the set of responses to determine whether the response trajectories for the 2 species were shared. Following Calsbeek and Goodnight (2009) we were primarily interested in whether this correlation differed from 1, which was tested via randomization using the RAND.SKEWER function of Roff et al. (2012) in the R programming language.

To determine whether variation in skull morphology is similarly oriented for the 2 species, we conducted a test of whether “the multivariate direction of greatest (variation)” differed between the 2 species (Schluter 1996). This “direction of greatest variation” is typically estimated from estimates of genetic parameters as $\hat{p}_{max}$ but here was estimated from phenotypic data, i.e., as $p_{max}$. To test whether $p_{max}$ differed between the species we first calculated $r^*$ between the estimates for each species. Next, using randomization ($N=1,000$), we estimated the null distribution of $r^*$ values to calculate the $P$ value for the observed $r^*$ versus a null expectation of 1.

To further explore potential differences between the 2 species in trait variances and covariances, we also conducted a hierarchical analysis following Arnold and Phillips (1999) and a partial tensor analysis following Hine et al. (2009). Hierarchical analyses assess a range of possible ways in which matrices might differ (see Figure 2 in Roff 2000) and which can be indicative of whether selection and drift have influenced population divergence. In a hierarchical analysis whether or not 2 populations exhibit common principal components versus unrelated structures (e.g., orthogonal greatest eigenvectors) is first tested. Next, a hypothesis of proportional (co)variance is tested versus unrelated structure, followed by a test between matrix equality versus a null expectation of 1, which is a jump-up approach; Arnold and Phillips 1999). Hierarchical testing was conducted via randomization testing ($N=500$) using the MATRIX.TESTS function of Roff et al. (2012). Tensor analyses (Hine et al. 2009) likewise compares matrix structure but reveals which particular traits are primarily contributing to matrix differences. Here we estimated a fourth order covariance tensor which we then summarized in matrix form and extracted the resulting 66
Table 1. Hierarchical analysis results

<table>
<thead>
<tr>
<th>Hierarchical level</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equality</td>
<td>0.312</td>
</tr>
<tr>
<td>Proportionality</td>
<td>0.634</td>
</tr>
<tr>
<td>CPC</td>
<td>0.716</td>
</tr>
</tbody>
</table>

\( P \) values are from randomization and indicate whether the hierarchical level being evaluated cannot be rejected on the basis of comparing its fit versus that of unrelated structure. Here \( N. \ fuscipes \) and \( N. \ mactrots \) exhibited equivalent covariance matrices.

element diagonal. We used randomization testing \((N = 1,000)\) to test which traits showed differences between the two populations in variances or covariances. Because this approach required testing for differences in 66 parameters so to control for family-wise error we a \textit{a priori} set our critical \( P \) value to 0.001.

Finally, because random skewers analysis and the high vector correlation between \( p_{max} \) ’s for the 2 species demonstrated a lack of response divergence (see below), we compared the placement of the 2 species along a shared \( p_{max} \). To do so we first conducted a principal components reduction of the phenotypic scores of the entire dataset, pooling \( N. \ mactrots \) and \( N. \ fuscipes \). The first component, corresponding to their shared \( p_{max} \) explained 57% of the observed variation (the second \( p_2 \), only explained 9.7%). We then used an unequal variance \( t \)-test to compare \( p_2 \) scores between the 2 species. This tested for whether the 2 species differed in their location along \( p_{max} \) and \( p_2 \), and thus whether they have diverged along this shared evolutionary trajectory as is generally predicted (Schluter 1996).

Results

Despite previous demonstration of phenotypic and genetic divergence (Matocq and Murphy 2007), \( N. \ fuscipes \) and \( N. \ mactrots \) did not differ according to a random skewers analysis. The average response vector correlation \((r)\) was 0.895, which did not differ from 1 \((P = 0.114)\). Likewise, the majority of variation \((p_{max})\) was oriented in the same direction for the 2 species \((r^2 = 0.977, P = 0.509)\). Consistent with these results, hierarchical analysis indicated that the covariance matrices for the 2 species were equal \((P = 0.312, \text{Table 1})\). Tensor analysis likewise did not detect differences between \( N. \ mactrots \) and \( N. \ fuscipes \) (Table 31).

Finally, despite the lack of response divergence found via random skewers analysis and the comparison of \( p_{max} \) vectors and, consistent with the findings of Matocq and Murphy (2007), the 2 species have phenotypically diverged along this shared evolutionary trajectory \((r_{max} = 10.66, \ P < 0.001; \text{Figure 3})\). The 2 species did not differ in their second principal component \((p_2, r_{4.35} = 1.60, P = 0.11; \text{Figure 3})\).

Discussion

Combined, these results suggest that \( N. \ fuscipes \) and \( N. \ mactrots \) do not exhibit response divergence. While the hierarchical analyses and examination of the fourth-order covariance tensor are consistent with the general conclusion, that the 2 species do not demonstrate response divergence is primarily based on results from the random skewers analysis and comparison of the phenotypic equivalent of \( p_{max} \) \((i.e., \ p_{max})\). The first of these analyses demonstrated that the correlation in response vectors for the 2 species was 0.895 and did not significantly differ from 1. The \( p_{max} \) for each species were similarly highly correlated \((r^2 = 0.977)\). However, as our sample sizes were quite small for estimating covariance matrices for 11 traits, it is possible that we were simply unable to detect a biologically significant effect with statistical significance.

Three main pieces of information suggest otherwise. First, the actual observed values from the random skewers test and the comparison of \( p_{max} \) vectors suggest strong similarity between the 2 populations. Second, we used bootstrapping to estimate a 95% confidence interval \((CI)\) around the \( r^2 \) for the 2 species’ \( p_{max} \). The lower bound for \( r^2 \) was 0.72 which, while representing a much greater angle between the vectors still indicates that variation was largely oriented in the same direction \((\text{Supplementary Text 2})\). Finally, as for \( p_{max} \) we again used bootstrapping to estimate the 95% CI around both the average response vector correlation \((r^2 \text{ 95% CI: } 0.70: \ 0.86)\), again indicating relatively high correspondence. Combined, this suggests that our failure to detect statistically significant differences is due to an absence of biologically significant differences. That is, despite significant divergence in other genetically and for mean morphology, the two species are morphologically differentiating along a shared evolutionary axis.

As proposed by Cheverud and Marroig (2007) and Caliebeek and Goodnight (2009), our findings suggest that statistically significant phenotypic and genotypic differences are not necessarily representative of response divergence. This conclusion is important to keep in mind when comparing population, species, or other levels of grouping. In particular, researchers often—and reasonably—infer that mean differences indicate divergence. However, as was the case here with wood rats, such a demonstration is not necessarily indicative of how populations may evolve in the future.

Matocq and Murphy (2007) suggested that the morphological differences between \( N. \ fuscipes \) and \( N. \ mactrots \) in the areas surrounding
the region of secondary contact of the two species may be due to character displacement arising from competition for resources and shifts in resource use. In particular, variation in the traits discussed here correspond to potential dietary differences between the species (Matocq and Murphy 2007), mirroring fine-scale diet divergences observed at contact zones between other closely related species of woodrats (Shurtliff et al. 2014). Our analyses are consistent with this explanation but suggest that divergence between the 2 species has been constrained or facilitated along the same evolutionary trajectory. This finding is generally consistent with earlier analyses which have likewise found divergence to primarily occur in the phenotypic directions with the greatest variation (Schluter 1996).

The highly concordant selection responses and evolutionary trajectories maintained across this species boundary have potentially important implications for the potential pace and direction of phenotypic evolution in response to interspecific interactions between these lineages when they come into contact. Possible implications in second-typic evolution in response to interspecific interactions between these important implications for the potential pace and direction of phenotypes maintained across this species boundary have potentially with the greatest variation (Schluter 1996).

whether found divergence to primarily occur in the phenotypic directions with the greatest variation (Schluter 1996).

Besides just differences via selection, the finding that the 2 species do not exhibit response divergence and the maintenance of evolutionary trajectories could also lead to an erosion of species differences and species boundaries due to drift. Just as selection responses are more rapid in directions with greatest variation, so too is drift greater in those directions. Under relaxed selection, our 2 species would primarily exhibit drift along the same phenotypic axes. Thus, possible divergence due to drift will be limited in phenotypic space relative to if there were no correlations among craniodental traits. Indeed, the 2 populations exhibit phenotypic divergence along \( p_{max} \) but not in the direction with the second most variation (\( p_2 \), Figure 3). Indeed, following the rationale of Roff (2000), our hierarchical analysis results are consistent with a hypothesis of drift (though selection cannot be rejected without additional data).

Whether and which of any of these inferences are appropriate is currently unclear. How an absence of divergence in evolutionary trajectories facilitates or constrains ecological and reproductive character displacement resulting from competition resulting in selection along this line could be accelerated under secondary contact. Such an outcome would thereby facilitate coexistence of the 2 species. Simultaneously, however, this character displacement would be constrained by the shared evolutionary trajectories. How this might affect the efficacy of reinforcing selection in secondary contact is currently unclear.

Acknowledgments

We thank Kevin Burls, Alan de Queiroz, Chris Feldman, Angela Hornsby, Peter Murphy, Derek Roff, and an anonymous reviewer for insightful criticisms and comments on earlier versions of this manuscript.

Funding

Funding was provided in part by a National Science Foundation grant to MDM (DEB-0952946) and by an ND EPSCoR grant to NAD.

Author contributions

N.A.D. and M.D.M. jointly framed the questions asked and wrote the paper. Statistical analyses were conducted by N.A.D.

Supplementary material

Supplementary material can be found at http://www.cz.oxfordjournals.org/.

References


