

Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa

Daniel J. Funk^{*†}, Patrik Nosil[‡], and William J. Etges[§]

^{*}Department of Biological Sciences, Vanderbilt University, Station B, Box 35-1634, Nashville, TN 37235-1634; [‡]Department of Biosciences, Behavioral Ecology Research Group, Simon Fraser University, Burnaby, BC, Canada V5A 1S6; and [§]Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701

Edited by Francisco J. Ayala, University of California, Irvine, CA, and approved January 10, 2006 (received for review October 4, 2005)

To what degree is the divergent adaptation responsible for life's phenotypic variety also responsible for generating the millions of species that manifest this variation? Theory predicts that ecological divergence among populations should promote reproductive isolation, and recent empirical studies provide support for this hypothesis in a limited number of specific taxa. However, the essential question of whether ecology plays a truly general role in speciation has yet to be systematically evaluated. Here we address this integral issue using an approach that adds an ecological dimension to comparative studies investigating the relationship between reproductive isolation and divergence time. Specifically, we quantify ecological divergence for >500 species pairs from eight plant, invertebrate, and vertebrate taxa and statistically isolate its association with reproductive isolation. This approach demonstrates a highly consistent and significant positive association between ecological divergence and reproductive isolation across taxa. This relationship was also observed across different aspects of ecological divergence and components of reproductive isolation. These findings are highly consistent with the hypothesis that ecological adaptation plays a fundamental and taxonomically general role in promoting reproductive isolation and speciation.

comparative methods | Coyne and Orr | divergent adaptation | ecological speciation | genetic distance

The hypothesis that ecological adaptation promotes biological diversification was an important element of early 20th century evolutionary thought. This notion was inherent in the idea that adaptive radiations resulted from access to new ecological resources (1, 2). Verbal models further explained how the adaptive fixation of alternative alleles in ecologically diverging populations might incidentally cause the reproductive isolation (RI) that promotes speciation (3–5). Because models predict that such ecological divergence (ED) can drive speciation in allopatry (3) as well as in sympatry (6), this hypothesis pertains across geographic scenarios. Nonetheless, it has been primarily over the last 15 years that explicit empirical studies of ecology's role in speciation have been conducted on natural populations, providing new insights into the mechanisms by which ED causes RI (6–16). However, although examples of ecologically driven RI in a few individual taxa have begun to accumulate, it remains unclear whether these cases represent the exception, reflecting the nonrandom selection of study taxa, or the rule. That is, the fundamental question of whether ED represents a taxonomically general contributor to speciation remains untested.

Addressing this question requires a comparative approach. One powerful comparative approach for the study of speciation was introduced by Coyne and Orr (17, 18), who plotted indices of RI against genetic distance using published data from each of dozens of pairs of *Drosophila* species. Treating molecular genetic distance (GD) as a surrogate for time, these authors evaluated the relationship between time and RI across species pairs to make inferences about the mechanisms and time course of

speciation. The results of Coyne and Orr have proven very influential and have spawned an increasing number of parallel studies on additional taxa (19–25) (referred to hereafter as C&O studies/data sets). In their recent book on speciation (26), Coyne and Orr have emphasized the great value of adopting such comparative approaches for identifying the factors that promote speciation while also arguing for the potential importance of ecological adaptation as one of these factors. However, no rigorous and broadly comparative analysis of ecology's role in speciation has previously been conducted.

The present study offers such an analysis. This analysis adopts a previously unused comparative approach that extends that of Coyne and Orr by adding an ecological dimension (9). We did so by quantifying the ED of individual pairs of populations/species, that is, the degree to which members of each such pair differ from each other with respect to traits that mediate their interactions with the environment. To calculate three such indices of ED, we collated published ecological data on three basic ecological traits (habitat, diet, and body size) for >500 species pairs from the eight previously published C&O data sets (17–25). We then incorporated these indices into analyses that evaluated the relationship between ED and RI (hereafter referred to as the “ecology–isolation association”) while statistically controlling for time. Because both ED and RI are expected to increase with time, this approach was necessary to isolate the contributions of ED *per se*. (See Fig. 1 for a diagrammatic explanation of this approach and *Materials and Methods* for the details of this approach.)

Our approach first quantifies the ecology–isolation association through “individual analyses” that treated all combinations of the particular components of RI and indices of ED that were available for each of the eight study taxa (Table 1). Each of these 24 analyses removed the effects of time by calculating residuals from a regression of RI on GD. The analyses then quantified the association between residual RI and ED through regression analysis. The ultimate reason for calculating these values of the ecology–isolation association was to determine whether they tend to be consistently positive across taxa. To do so, we first calculated the mean of the values for the individual analyses of each of the eight study taxa. Our “cross-taxon analyses” then tested whether these eight means, as a group, were significantly greater than zero.

Our study finds that, indeed, as the degree of ED between pairs of populations/species increases, so too does their degree

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Freely available online through the PNAS open access option.

Abbreviations: RI, reproductive isolation; ED, ecological divergence; GD, genetic distance; C&O, Coyne and Orr.

[†]To whom correspondence should be addressed. E-mail: daniel.j.funk@vanderbilt.edu.

© 2006 by The National Academy of Sciences of the USA

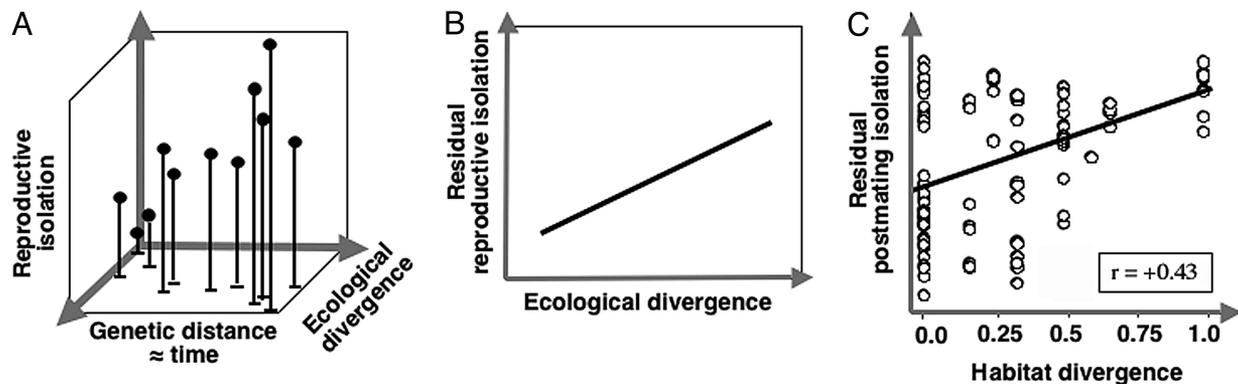


Fig. 1. Method illustration. (A) A hypothetical data set, illustrating the distribution of species comparisons (data points) with respect to the three aspects (axes) of evolutionary differentiation analyzed here. GD is used as a surrogate for time. (B) A hypothetical best-fit line illustrating the predicted positive association for the relationship between ED and residual RI upon statistical removal of the contributions of time by regression of RI on GD. (C) An individual analysis of actual data that illustrates the predicted pattern, in this case, of a positive association between habitat divergence and postmating isolation for the angiosperm data set (see Table 1).

of RI. To evaluate the robustness of this finding, we calculated association values using each of four different approaches to regression analysis that we applied to both unadjusted and phylogenetically adjusted data sets. We observed a consistently significant positive association between these two aspects of evolutionary divergence across disparate taxa and diverse approaches to analysis. This observation is highly consistent with the hypothesis that ecology plays a general role in speciation.

Results and Discussion

Our cross-taxon analyses of the unadjusted data provided the primary tests of our hypothesis, are sufficient and appropriate for its evaluation (27–30), and proved highly informative (Table 2). The most important result of our study was the finding that the “overall” mean association between ED and RI across the eight study taxa was indeed significantly greater than zero.

Table 1. Summary of individual analyses of the strength of the association between indices of ecological divergence and components of reproductive isolation

Taxon	Ecological trait	RI	<i>n</i>	Traditional analyses		Analyses of <i>a priori</i> residuals	
				Parametric analysis, <i>r</i>	Nonparametric analysis, ρ	Parametric analysis, <i>r</i>	Nonparametric analysis, ρ
Angiosperms	Habitat	Pre-	82	0.06	0.17	0.06	0.18
Angiosperms	Habitat	Post-	102	0.43	0.35	0.42	0.35
Angiosperms	Habitat	Total	54	0.15	0.24	0.16	0.24
Lepidoptera	Diet/hab	Post-	66	0.17	0.02	0.18	0.15
Lepidoptera	Diet/hab	Total	48	-0.14	-0.10	-0.14	-0.12
<i>Drosophila</i>	Diet/hab	Pre-	30	0.18	0.05	0.14	0.01
<i>Drosophila</i>	Diet/hab	Post-	35	-0.27	-0.22	-0.31	-0.26
<i>Drosophila</i>	Diet/hab	Total	50	0.31	0.03	0.27	0.26
Fishes	Habitat	Post-	37	0.12	0.03	0.12	0.01
Fishes	Diet	Post-	28	-0.16	-0.18	-0.15	-0.12
Fishes	Size	Post-	36	-0.01	-0.03	0.00	0.00
Darters	Habitat	Pre-	13	-0.03	-0.18	-0.03	-0.17
Darters	Habitat	Post-	9	-0.34	-0.28	-0.35	-0.27
Darters	Habitat	Total	7	0.27	0.21	0.27	0.14
Darters	Size	Pre-	13	0.53	0.45	0.56	0.31
Darters	Size	Post-	9	-0.13	-0.08	-0.14	-0.23
Darters	Size	Total	7	0.35	0.61	0.37	0.64
Frogs	Habitat	Post-	89	0.21	0.25	0.21	0.26
Birds	Habitat	Post-	121	0.10	0.06	0.12	0.08
Birds	Diet	Post-	114	0.13	0.11	0.16	0.15
Birds	Size	Post-	117	0.14	0.06	0.17	0.05
Doves	Habitat	Post-	12	0.16	0.09	0.35	0.23
Doves	Diet	Post-	12	0.39	0.26	0.40	0.42
Doves	Size	Post-	12	0.09	-0.11	0.09	-0.12

Presented are results from regression analyses for each data set. *n*, number of species comparisons analyzed. *r* and ρ values indicate the strength of the association between ED and RI after time (GD) has been statistically removed. Values in bold illustrate the predicted positive ED–RI association. These association values provide the data for the cross-taxon analyses (Table 2). Diet/hab, diet/habitat; Pre-, prezygotic RI; Post-, postzygotic RI; Total, total RI. Each column of association values refers to the particular method of regression analysis used to derive these values (see text for details). For Lepidoptera, “postzygotic isolation” data were in the form of hybrid inviability and “total isolation” represented a combined index of hybrid inviability and hybrid sterility (data for which were not provided in the original paper).

Table 2. One-sample, cross-taxon *t* tests of the association between ecological divergence and reproductive isolation across data sets

Comparison	<i>n</i>	Traditional analyses				Analyses of <i>a priori</i> residuals				
		Parametric analysis		Nonparametric analysis		Parametric analysis		Nonparametric analysis		
		Mean <i>r</i>	<i>P</i>	Mean ρ	<i>P</i>	Mean <i>r</i>	<i>P</i>	Mean ρ	<i>P</i>	
Overall	8	0.12	0.004 [†]	0.08	0.059*	0.127	0.006 [†]	0.11	0.018 [†]	
By ecological trait										
Habitat	8	0.11	0.006 [†]	0.06	0.102	0.129	0.012 [†]	0.09	0.046 [†]	
Diet	5	0.09	0.185	0.02	0.399	0.093	0.182	0.09	0.182	
Size	4	0.12	0.062*	0.06	0.281	0.131	0.052*	0.04	0.302	
By component of reproductive isolation										
Prezygotic	3	0.16	0.050 [†]	0.12	0.040 [†]	0.182	0.018 [†]	0.08	0.049 [†]	
Postzygotic	8	0.08	0.191	0.04	0.293	0.085	0.187	0.06	0.236	
Total	3	0.26	0.020 [†]	0.23	0.088*	0.251	0.018 [†]	0.30	0.013 [†]	

n, number of datasets available for use in an analysis. Tests compare the mean observed associations (*r* and ρ values) between ED and residual RI from the individual analyses (Table 1) with an absence of association (i.e., *r* and ρ = 0). Values in bold illustrate the predicted positive ED–RI association, independent of time. Each column of association values refers to the particular method of regression analysis used to derive these values (see text for details). *, *P* < 0.10; †, *P* < 0.05.

Statistical support for the predicted relationship was observed across methods of regression analysis (*P* = 0.004, 0.059, 0.006, and 0.018, respectively; Table 2). Indeed, mean associations were positive for seven of eight study taxa (*P* = 0.035, binomial test) for three of the methods of regression analysis and only slightly negative for the single dissenting data set in each case (most negative *r* = −0.03). Importantly, mean slopes were positive for all 28 cross-taxon *t* tests in Table 2, with *P* < 0.10 in 17 of 28 analyses. Significant or marginally nonsignificant statistical support for positive slopes was separately observed for the ecological traits of habitat and size and for prezygotic and total components of RI (Table 2).

Although phylogenetically adjusted data were not required to test the hypothesis evaluated here (see *Materials and Methods*), analyses of adjusted data sets nonetheless provided strong support for a positive ecology–isolation association, further illustrating the robustness of this result. Indeed, cross-taxon tests yielded significant results for all analyses of the overall comparison (*P* = 0.005, 0.021, 0.007, and 0.037 for the four regression analyses), and 27 of 28 mean slopes proved positive. (For complete results, see Tables 3 and 4, which are published as supporting information on the PNAS web site.) Notably, these results were obtained despite the small number of species comparisons of these reduced data sets (*n* ≤ 46 for all 24 individual analyses of the adjusted data; *n* ≤ 10 in 12 of 24 analyses). The parameter estimates from such small data sets might be expected to be “noisy” with respect to the underlying statistical signal, yet the signal was sufficiently strong to be detected.

In sum, our results demonstrate a highly consistent pattern of positive associations between ED and RI across taxa, ecological traits, components of RI, and approaches to handling data and conducting regression analyses. Our results are thus strongly in accordance with the hypothesis that ecological factors are quite consistent, rather than anecdotal, contributors to RI and, thus, to speciation.

It is important to note that the correlation between variables demonstrated by this study does not necessarily imply a causal relationship between them or indicate the direction of the possibly causal relationship, i.e., whether ecology drives RI or vice versa. However, it is also important to recognize that these limitations are general features of all comparative analyses. This is true, for example, of the seminal Coyne and Orr paper (17)

that inspired this study, in which the authors entertain alternative causes of the patterns they found. In this context, however, it is also relevant that theory and experimental work (cited above) have provided *a priori* demonstrations of the potential and actual causal contributions of ED to RI. Thus, although the power of comparative approaches to provide taxonomically general tests of hypotheses is necessarily offset by their more limited capacity to directly address causality, the present study is framed in the context of a mechanism that has already received considerable support of other kinds.

We do not argue that this study demonstrates that ecological adaptation is necessarily the most important driver of RI and promoter of speciation. Much recent evidence suggests that sexual selection, as one example, also plays an important role in speciation (26, 31). Indeed, one of the traits we evaluated, size, may evolve under sexual selection (31). However, size is also under selection as a function of climate, trophic habit, competitive interactions, etc., and is undoubtedly an ecologically important trait in many taxa. The observation that ecological selection’s role in body size is associated with RI in various freshwater fish systems (11) illustrates the reasons for our selection of this trait as a focus of study. Moreover, as demonstrated by the results from habitat divergence alone (Table 2), our findings hold even if size is excluded from the analyses.

We do argue that the present study demonstrates the varied and general contributions of ED to speciation in a heterogeneous set of eight taxa. Some previous approaches have been informatively applied to individual taxa to make comparative inferences about the role of natural selection in speciation (8, 13, 32, 33). However, a strength of the present study is that the taxonomic generality of our results cannot be attributed to a biased selection of study taxa because these taxa were determined by a previously published list of C&O studies (26) that were not conducted with ecological factors in mind. The data sets evaluated here thus represent a random taxonomic sample with respect to the hypothesis under testing. Assuming that these taxa are indeed representative of other animal and plant groups, our findings suggest that speciation is, in part, an inherently ecological process.

Further testing will require the continued accumulation of C&O studies from a variety of taxa and the compilation of additional ecological data from these taxa. Further testing should also include the identification and evaluation of hypoth-

eses presenting alternative explanations for the patterns we have identified here. Additionally, further inspection of our present findings highlights specific areas of potentially informative future research. For example, our observation of significant associations of ED with prezygotic but not postzygotic isolation is particularly intriguing in view of current debates on the relative importance of these two fundamental aspects of RI for ecological speciation (26). These sorts of investigations will continue to refine our understanding of what may be a rare example of a general evolutionary rule, in this case, of a fundamental and ecological cause of evolutionary diversification.

Materials and Methods

Data Sets. To avoid bias in our selection of study systems, we evaluated each data set cited by Coyne and Orr (17–25) in their review of comparative analyses of isolating barriers (ref. 26, pp. 72–81). These eight data sets treated, respectively, three angiosperm genera, Lepidoptera, *Drosophila*, teleost fishes, *Etheostoma* darters (percid fishes), frogs, birds, and doves/pigeons. Each of these data sets included information on GDs and components of RI (prezygotic, postzygotic, and/or total isolation) for many pairs of populations or species. To these data we added our quantitative indices of ED. We calculated these indices for each of three basic ecological traits: habitat, diet, and body size. More than 200 sources (see *Supporting Text*, which is published as supporting information on the PNAS web site) were consulted to provide the ecological data required to calculate these indices. These data were obtained for 80–100% of species pairs for each data set except *Drosophila* (45%), for which ecological information is limited. Each data set was represented by each ecological index for which sufficient data were available. Habitat divergence was quantified for all data sets. For analyses of the pigeon/dove data set (19), we deleted species pairs redundant with the bird data set (23).

Ecological Divergence Indices. To quantify habitat divergence for the species pairs of a particular data set, ecological variables (e.g., altitude, moisture, vegetation type, etc.) that varied among species' habitat descriptions were selected for evaluation. We attempted to identify as many such variables as possible for each data set. Values were assigned to each species pair for each variable for which information was available. Specifically, a value of 0, 1, or 2 was assigned depending on whether a pair of species had the same, overlapping, or nonoverlapping states for that variable, respectively. The habitat divergence index was then calculated for each species pair as the mean value of the variables scored for that pair. Although we chose variables so as to restrict overlap among them, any remaining redundancy did not compromise statistical analyses because we averaged their values, ensuring that the degrees of freedom were not affected. Diet divergence was based on the proportion of total diet items shared by the two species {diet divergence = $1 - (\text{number of items shared by species 1 and 2} / [\text{number of items used by species 1 only} + \text{number of items used by species 2 only} + \text{number of items shared by species 1 and 2}])$. For Lepidoptera and *Drosophila*, distinct diet items corresponded to different host plant families and genera, respectively. For the remaining data sets, diet items reflected predetermined categories (e.g., invertebrates, seeds, fruits, etc.). (See *Supporting Text* for a list of the ecological variables and diet categories scored for each data set.) Size divergence was calculated as the absolute value of the percentage size difference between compared species $\{[\text{size of species 1} - \text{size of species 2}] / [(\text{size of species 1} + \text{size of species 2}) / 2]\}$. All habitat variables, habitat divergence values, and diet categories were determined before any analyses and in the absence of species names or other data.

Statistical Evaluation: Obtaining Association Values from Individual Analyses. The interpretations of this study were based on quantifying the independent association of ED with RI. Values for this association were obtained from regression analyses of residuals. We adopted standard approaches for obtaining residuals and conducting our regression analyses (29). These analyses treated GD (a surrogate for time, as discussed above) and ED as independent variables and RI as the dependent variable. Specifically, residuals of RI on GD were calculated before determining the association of these residuals with ED by using regression analysis. These analyses thereby statistically controlled for time while evaluating the ecology–isolation association. We conducted such analyses for each of the 24 combinations of ecological trait and RI component for which appropriate data were available across the eight taxon-specific data sets (Table 1). We hereafter refer to these as the “individual analyses.” The parameter values (r or ρ) of the ecology–isolation association that were estimated by these analyses supplied the data used in the cross-taxon analyses (described below) that provide the critical findings of our study.

Before conducting these analyses, we first evaluated whether the ED–GD and RI–GD relationships of the data were linear, as assumed by linear regression. Deviations from linearity were detected in only a few cases, and for these cases we found that a quadratic regression provided a proper fit to the data. We thus applied linear and/or quadratic regression as appropriate to derive the residuals used in our individual analyses. Detailed discussion of the approaches and analyses we used to evaluate and incorporate nonlinearity is provided in *Supporting Text*.

Statistical Evaluation: Examining the Robustness of Results. We conducted these individual analyses using four different approaches to regression analysis (see Table 1) and two different ways of treating data from the original species comparisons (discussed below) to evaluate the robustness of our results. Analysis 1 represented standard parametric multiple regression analyses. Analysis 2 calculated residuals of RI versus GD by using parametric regression but then adopted a nonparametric approach in evaluating the Spearman rank correlation of these residuals with ED (29). Analyses 3 and 4 also represented parametric and nonparametric approaches, respectively, but adopted a different method for deriving and analyzing residuals. These latter analyses incorporated estimates of ED and RI that had each been separately corrected for genetic divergence *a priori*. This correction was accomplished by generating residuals of ED versus GD as well as residuals of RI versus GD for use in the regression analyses. These analyses thus used an alternative method of regression that evaluated the association between two sets of residuals rather than between residual RI and raw ED values (32).

The purpose of our individual analyses was parameter (i.e., r or ρ) estimation rather than significance testing. Because parameter estimation does not require statistical independence (29), all analyses described above appropriately used phylogenetically unadjusted data sets, which offered the greatest amount of data while still presenting unbiased parameter estimates (27, 28, 30). However, as a further evaluation of the robustness of our findings, we nonetheless conducted a second suite of analyses. These analyses applied the four approaches to regression analysis just described to data that had been phylogenetically adjusted to alleviate possible nonindependence among species pairs due to patterns of shared ancestry (34). For these analyses, data for each study taxon were phylogenetically adjusted by using the methods applied in the original C&O studies.

Statistical Evaluation: Analysis of Cross-Taxon Trends. Our principal goal in this study was to evaluate whether the association between ED and RI is consistently positive when these associ-

