Evolution, 58(7), 2004, pp. 1424–1433

QUANTIFYING PATTERNS IN THE EVOLUTION OF REPRODUCTIVE ISOLATION

TAMRA C. MENDELSON,[1,2] BRIAN D. INOUYE,[3,4] AND MARK D. RAUSHER[5,6]
[1]*Department of Biology, University of Maryland, College Park, Maryland 20742*
[2]*E-mail: tm178@umail.umd.edu*
[3]*Department of Biological Science, Florida State University, Tallahassee, Florida 32306-1100*
[4]*E-mail: binouye@bio.fsu.edu*
[5]*Department of Biology, Duke University, Durham, North Carolina 27708*
[6]*E-mail: mrausher@duke.edu*

Abstract.—We present a likelihood-based statistical method for examining the pattern or rate of evolution of reproductive isolation. The method uses large empirical datasets to estimate, for a given clade, the average duration of two phases in the divergence of populations. The first phase is a lag phase and refers to the period during which lineages diverge but no detectable reproductive isolation evolves. The second is an accumulation phase, referring to the period during which the magnitude of reproductive isolation between diverging lineages increases. The pattern of evolution is inferred from the relative durations of these two phases. Results of analyses of postzygotic isolation data indicate significant differences among taxa in the pattern of evolution of postzygotic isolation that are consistent with predictions based on genetic differences among these groups. We also examine whether the evolution of postzygotic isolation is best explained by either of two models for the rate of accumulation: a linear model or a quadratic function as may be suggested by recent studies. Our analysis indicates that the appropriateness of either model varies among taxa.

Key words.—Patterns of speciation, postzygotic isolation, reproductive isolation, sex chromosomes.

Received October 31, 2003. Accepted March 17, 2004.

A major goal of evolutionary biology is to elucidate mechanisms underlying the evolution of reproductive isolation, a process commonly equated with speciation (e.g., Dobzhansky 1937; Otte and Endler 1989; Coyne 1992; Howard and Berlocher 1998). The evolution of reproductive isolation is generally understood to begin when gene flow between interbreeding populations is restricted, primarily through geographic isolation, allowing traits in these populations to diverge (e.g., Dobzhansky 1937; Mayr 1963). What remains unclear, however, is how reproductive isolation accumulates over time once gene flow is restricted. The pattern of evolution of reproductive isolation is likely to be affected by underlying genetic mechanisms, including aspects of genetic architecture and the relative importance of adaptive versus neutral divergence in speciation (e.g., Orr 1995; Orr and Turelli 2001). Empirical investigations into the pattern of evolution of reproductive isolation are clearly integral to evaluating the relationship between pattern and process and may be particularly informative in a comparative context, if the pattern of evolution is compared among groups known to differ in one or more aspects of mechanism.

In this study, we used empirical data to quantify the pattern of evolution of postzygotic isolation (i.e., hybrid inviability and sterility) in three clades. Specifically, we have developed a likelihood-based statistical method that estimates for a given clade the average duration of two phases in the divergence of two populations (Fig. 1). The duration of the first phase is defined as the amount of time between initial genetic isolation and the onset of detectable reproductive isolation. This phase, referred to here as the ''lag phase,'' represents the early period of divergence during which genetic changes have no discernable effect on the magnitude of reproductive isolation. We include this phase because of the numerous examples of phenotypically and genetically distinct populations that are fully reproductively compatible (e.g., Hubbs 1955; Mayr 1963; Grant 1981; Coyne and Orr 1997; Mendelson 2003), suggesting that genetic divergence need not be associated with any significant accumulation of reproductive isolation. The second phase in our model begins when the magnitude of reproductive isolation is estimated to be greater than zero and represents the time during which reproductive isolation between two populations evolves from onset to completion. We refer to the second phase as the ''accumulation phase'' to signify an accumulation in the strength of reproductive isolation. The relative durations of these two phases represents the pattern of the evolution of reproductive isolation.

We are not the first to distinguish two separate phases in the divergence of populations. Dobzhansky argued that ''although the process of divergence is a gradual one, speciation in the strict sense, that is, the development of reproductive isolation, is a crisis that is passed relatively rapidly'' (1970, p. 368). A point on which Dobzhansky and others have been notably silent, however, is *when* in the divergence of two lineages the onset of reproductive isolation is expected to occur. The timing of the onset of reproductive isolation is as likely to be affected by underlying genetic mechanisms as the duration of speciation per se; our method allows us to detect differences among taxa in the relative durations of these two phases.

Estimating the pattern of evolution of reproductive isolation also allows us to explore recent predictions concerning the evolution of postzygotic isolation. Orr (1995) and Orr and Turelli (2001) demonstrated theoretically that, if postzygotic isolation evolves according to the Dobzhansky-Muller model of noncomplementary substitutions (Dobzhansky 1937; Muller 1940, 1942), the number of substitutions leading to postzygotic incompatibility between two lineages will accumulate at least as the square of the number of nucleotide substitutions. Orr and Turelli (2001), accounting for the stochasticity of the molecular clock, concluded that the number of postzygotic incompatibilities accumulates exactly as the

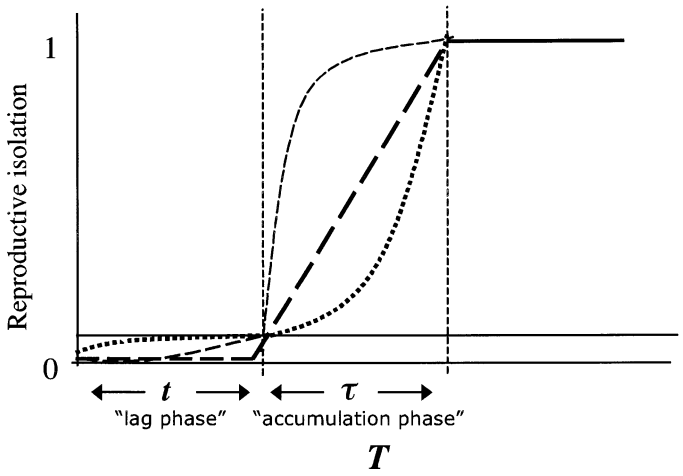

FIG. 1. Schematic representation of the evolution of reproductive isolation described by three different evolutionary trajectories. Vertical dashed lines denote the two phases of divergence for which the average duration is estimated by the model. Solid horizontal line represents the detectability threshold, that is, the magnitude of reproductive isolation determined by measurement technique to be demonstrably different than zero. Note the shape of the function describing either phase is not defined.

square of time since divergence (quadratic accumulation). These studies thus predict a snowballing of the number of postzygotic incompatibilities over time. It is unclear whether accumulation in the magnitude of postzygotic isolation as measured by hybrid fitness should mirror that of the number of Dobzhansky-Muller (D-M) incompatibilities, as the nature of the relationship between D-M incompatibilities and hybrid fitness is poorly understood. However, if the effects of incompatibilities on hybrid fitness are additive and the average magnitude of those effects do not change consistently over time, we would expect the pattern of accumulation of the magnitude of postzygotic isolation to reflect that of the number of incompatibilities.

Using our method with datasets from the literature, we address the following questions: (1) Do taxa differ (predictably) in the pattern of evolution of postzygotic isolation? (2) Does the magnitude of postzygotic isolation accumulate in a snowball fashion as predicted for the number of genetic incompatibilities (Orr and Turelli 2001)?

MATERIALS AND METHODS

Empirical data analyzed were derived from publications in which genetic distance is compared with the magnitude of reproductive isolation for multiple pairs of species in large taxa (Coyne and Orr 1989, 1997; Sasa et al. 1998; Presgraves 2002; similar studies include Tilley et al. 1990; Knowlton et al. 1993; Mendelson 2003). Genetic distance in these studies represents the amount of time two species have been evolving independently (i.e., divergence time; Avise 1994). Reproductive isolation is typically quantified on a scale from zero to one, where zero represents random mating or no isolation and one represents complete reproductive isolation. We applied our method to three such datasets on postzygotic isolation (see Fig. 2): (1) *Drosophila* (Coyne and Orr 1997); (2) anurans (Sasa et al. 1998); and (3) lepidopterans (Presgraves 2002).

In the first dataset, postzygotic isolation in *Drosophila*, genetic distance was estimated from published allozyme frequencies and was reported as Nei's genetic distance (*D*; Coyne and Orr 1989, 1997). The strength of postzygotic isolation was calculated as a composite isolation index, derived from estimates of hybrid inviability and hybrid sterility reported in the literature. Estimates of postzygotic isolation fell into five distinct categories, ranging from 0 (both sexes of both reciprocal crosses fully viable and fertile) to 4 (neither sex in either reciprocal cross viable or fertile). The authors divided these estimates by four to yield categories of total isolation ranging from zero to one (i.e., 0.00, 0.25, 0.50, 0.75, and 1.00). Data used in our analysis were species pairs for which both genetic distance and total postzygotic isolation indices were available ($n = 107$, see Appendix A available online at http://dx.doi.org/10.1554/03-632.1.s1).

The second dataset was derived from the literature survey by Sasa et al. (1998), which examined the evolution of postzygotic isolation in anurans. In that study, genetic distance was also derived from allozyme frequencies reported in the literature and calculated as Nei's *D*. Postzygotic isolation was calculated according to the same criteria as described for *Drosophila* and again was based on hybrid inviability and/or sterility of reciprocal crosses reported in the literature (Sasa et al. 1998). Thus, as in the *Drosophila* dataset, estimates of postzygotic isolation fell into five discrete categories ranging from zero to one. The number of species pairs for which both genetic distance and total postzygotic isolation indices were available was $n = 44$ (see Appendix B online).

In the Presgraves (2002) study of postzygotic isolation in lepidopterans, genetic distance was also calculated as Nei's *D*, derived primarily from allozyme data in the literature, but in some pairs distance was based on nucleotide sequence data (see Presgraves 2002, p. 1170, appendix 1). The strength of postzygotic isolation was quantified by the same method as in the *Drosophila* and anuran studies but also included estimates from unidirectional-cross data when reciprocal cross data were unavailable. The number of species pairs for which genetic distance and total isolation values were available was $n = 51$ (see Appendix C online).

Ideally, species pairs in the datasets should be statistically independent. A particular species (or subspecies) should be represented no more than once, and pairs should be phylogenetically independent (Felsenstein 1985); that is, the evolutionary branches connecting any particular pair of species should not overlap with those connecting any other species pair. If phylogenies are available, they can reveal which species pairs are phylogenetically independent (see Mendelson 2003); however, complying with a strict criterion of phylogenetic independence may be difficult if phylogenetic data are limited (Housworth and Martins 2001). Even if the phylogeny is known, using only independent species pairs may drastically limit the usable data. Because using more species pairs would probably improve estimates of the parameters of interest, we disregarded phylogenetic independence and used all available data for these analyses, understanding that results may be biased toward overrepresented groups. For heuristic purposes, we identified phylogenetically independent

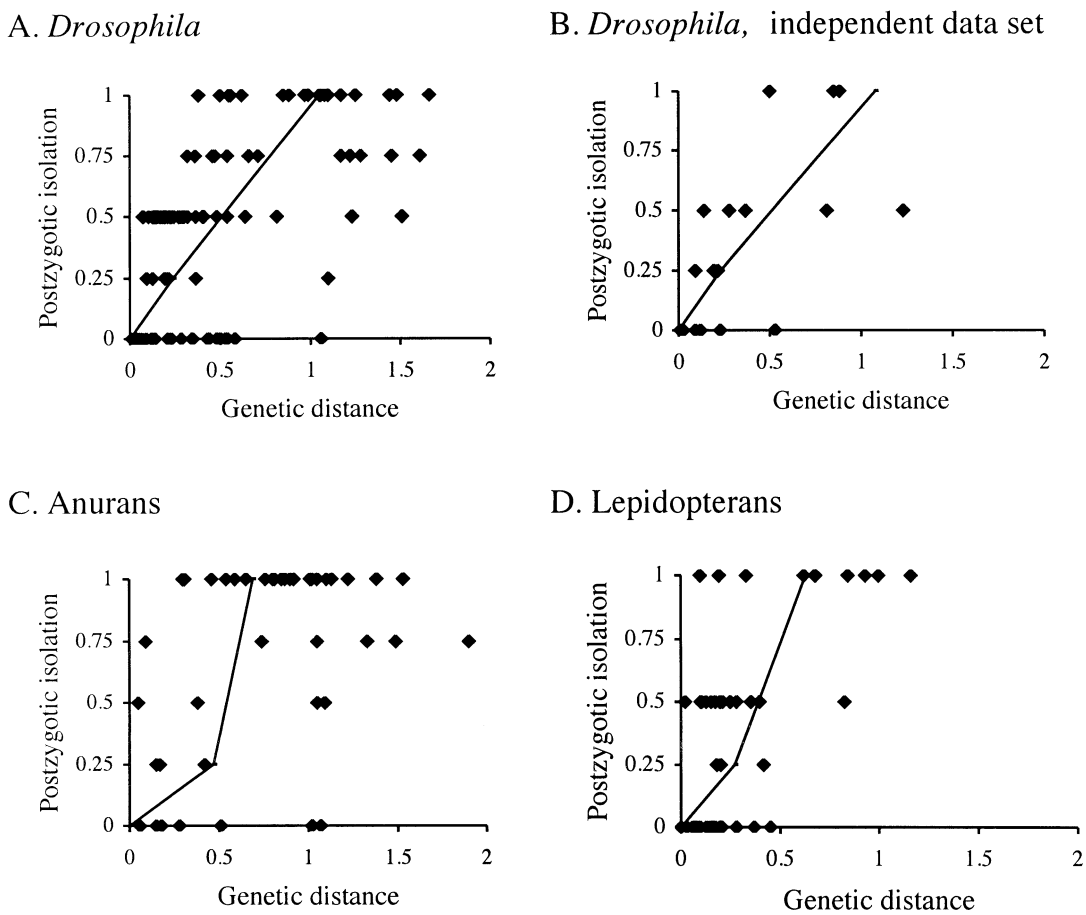

FIG. 2. Datasets examined. Each dataset consists of species pairs for which genetic distance (Nei's *D*, x-axis) is compared to the magnitude of postzygotic isolation. (A) *Drosophila* (n = 107), from Coyne and Orr (1989, 1997). (B) Phylogenetically independent species pairs of *Drosophila* (n = 17). (C) Anurans (n = 44), from Sasa et al. (1998). (D) Lepidopterans (n = 51), from Presgraves (2002). Lines depict estimated average trajectories for each phase of speciation and are depicted as linear for simplicity of comparison.

species pairs from the *Drosophila* dataset (Nurminsky et al. 1996; DeSalle and Brower 1997; Hilton and Hey 1997; Lemeunier et al. 1997; Gleason et al. 1998; O'Grady 1999; Harr et al. 2000; Rodriguez-Trelles et al. 2000; Schawaroch 2002; see Appendix D online). Results of analyses of this smaller dataset (n = 17) were compared with those of the full dataset.

The Model: Estimating the Pattern and Rate of Speciation

The statistical approach we adopted relies on the approximation of a relationship between the magnitude of reproductive isolation (*RI*) and genetic distance consisting of two joined segments or phases (Fig. 1). In this approximation, time zero represents the point in time at which gene flow between two populations becomes restricted. The first phase, extending from time zero to time t, represents the time during which reproductive isolation accumulates either very slowly or not at all. During this lag phase, the magnitude of *RI* is unlikely to be experimentally distinguishable from zero. The duration of the lag phase is represented by t in our model.

The second phase begins when the magnitude of reproductive isolation is detectable. This accumulation phase extends from time t to time $t + \tau$, where τ approximates a period of demonstrable accumulation of reproductive isolation. During this phase, *RI* is intermediate: it is measurably different from zero but not yet complete. We make no assumptions about the shape of the function describing the evolution of *RI* during this phase or during the lag phase. Our only assumption is that *RI* does not decrease once it has reached completion. In this approximation, the duration of the accumulation phase is represented by the value of τ. The ratio of the durations of the two phases, represented by $\theta = t/\tau$, describes the pattern of speciation.

Our model is based on a very simple concept, which is that the probability of observing a pair of species in the accumulation phase is proportional to the duration of that phase. If *RI* accumulates very rapidly once it begins, the probability that any randomly chosen pair of species will exhibit intermediate *RI* is very small. In a sample of species pairs, then, most will exhibit either zero or complete *RI*, and very few will exhibit intermediate *RI*. By contrast, if *RI* accumulates very slowly, the probability of catching a pair of species during this phase is much greater, and a greater fraction of sampled species pairs should exhibit intermediate values of *RI*.

The objective of the model is to estimate the average values of the parameters t, τ, and θ for pairs of species in a clade.

Because every species pair in a given dataset is unlikely to follow the same trajectory for the evolution of *RI*, we modeled the times t and τ as frequency distributions and used a maximum-likelihood analysis of the data to estimate the mean and variance of these distributions. Estimates of the average t and τ were calculated in units of genetic distance corresponding to the units used in the original dataset. These estimates therefore predict the average duration of each phase for a particular clade in terms of genetic divergence, rather than real time. Estimates of the pattern of speciation (θ), however, are dimensionless and thus more easily compared across taxa.

An exponential decay function was used to model the probability of the duration of the first phase:

$$\Pr[\text{duration of phase } 1 = t] = \gamma e^{-\gamma t} \tag{1}$$

where the parameter γ is estimated from the data. This function models the onset of the evolution of $RI > 0$ as a random process with equal probability of occurring in any generation after genetic isolation. With this distribution, the mean of t is estimated as $1/\gamma$. Therefore, the probability that a pair of species with genetic distance T is still in phase 1 (i.e., $RI = 0$) is given by

$$\Pr[RI = 0] = \Pr[t > T] = 1 - \int_0^T \gamma e^{-\gamma t}\, dt. \tag{2}$$

To describe the duration of the second phase, τ, we used a gamma distribution, given by

$$\Pr[\text{duration of phase } 2 = \tau] = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{\beta\tau}, \tag{3}$$

where α and β are estimated from the data. A gamma distribution was chosen because it can approximate a normal distribution while constraining τ to vary between zero and infinity for individual pairs of species. The mean of τ is estimated by α/β. The probability that a pair of species with genetic distance T is in phase 2 (i.e., *RI* is intermediate) is given by

$$\begin{aligned}\Pr[0 < RI < 1] &= \Pr[t < T \text{ and } t + \tau > T] \\ &= \int_0^T \gamma e^{-\gamma t}\left[1 - \int_0^{T-1} \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta\tau} d\tau\right] dt \\ &= T \int_0^T \gamma e^{-\gamma t}\left\{\frac{\Gamma[\alpha, \beta(T-t)]}{\Gamma(\alpha)}\right\} dt. \end{aligned} \tag{4}$$

Finally, the probability that complete reproductive isolation has evolved by time T is simply the probability that phases 1 and 2 have been completed, which, when equations (2) and (3) are combined, is

$$\begin{aligned}\Pr[RI = 1] &= \Pr[t + \tau < T] \\ &= \int_0^T \gamma e^{-\gamma t}\left[\int_0^{T-1} \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta\tau} d\tau\right] dt \\ &= \int_0^T \gamma e^{-\gamma t}\left\{1 - \frac{\Gamma[\alpha, \beta(T-t)]}{\Gamma(\alpha)}\right\} dt. \end{aligned} \tag{5}$$

With these probabilities, the likelihood of a given dataset is

$$\begin{aligned} L = {} & (\Pr[RI = 0 \mid T_1]\Pr[RI = 0 \mid T_2] \cdots \Pr[RI = 0 \mid T_i]) \\ & \times (\Pr[0 < RI < 1 \mid T_{i+1}]\Pr[0 < RI < 1 \mid T_{i+2}] \\ & \times \cdots \Pr[0 < RI < 1 \mid T_j])(\Pr[RI = 1 \mid T_{j+1}] \\ & \times \Pr[RI = 1 \mid T_{j+2}] \cdots \Pr[RI = 1 \mid T_N]), \end{aligned} \tag{6}$$

where $T_1 \ldots T_i$ are the genetic distances of pairs of species in the dataset that exhibit no reproductive isolation ($RI = 0$), $T_{i+1} \ldots T_j$ are the genetic distances of species pairs that exhibit intermediate reproductive isolation (here, $0.25 \leq RI \leq 0.75$), and $T_{j+1} \ldots T_N$ are the genetic distances of pairs of species that exhibit complete reproductive isolation ($RI = 1$). Estimates of the parameters α, β, and γ were then calculated by standard likelihood techniques that solve for the parameter values maximizing L. These values in turn define the average lag phase, t ($= 1/\gamma$), and the average duration of the accumulation phase, τ ($= \alpha/\beta$). The pattern of speciation is then $\theta = t/\tau = \beta/\alpha\gamma$.

Because postzygotic isolation was estimated in discrete categories in the datasets we examined, the lowest detectable degree of isolation is $RI = 0.25$, which corresponds to the inviability and/or sterility of all individuals of one sex of one reciprocal hybrid cross. Consequently, with these types of data, the estimated lag phase actually represents a combination of any true lag phase (a period during which there is no detectable *RI*) and the period during which *RI* accumulates but is less than 0.25. Using these data, it is therefore not possible to determine whether there is a true lag phase. However, it is possible, first, to determine whether and in what ways taxa differ in the pattern of speciation, and second, to determine whether the data are consistent with either a linear or exponential accumulation of *RI*.

If the accumulation of *RI* proceeds in a linear manner, then we expect $\theta = 1/3$; with linearity, 25% of the accumulation should occur during the lag phase, and 75% during the accumulation phase. The ratio of the two is $25/75 = 1/3$. By contrast, if accumulation is quadratic, that is, if *RI* is proportional to the square of genetic distance, then we expect $\theta = 1$. Accumulation faster than quadratic corresponds to $\theta > 1$.

Hypothesis Testing

For each dataset, we generated 95% confidence contours for t, τ, and θ by obtaining the combinations of α, β, and γ for which $ln(L)$ was within 3.9075 log-likelihood units of the maximum log-likelihood ($\alpha = 0.05$, df = 3, $2\chi^2 = 7.815$; $\chi^2 = 3.9075$). Points falling within these contours were taken as an approximate posterior probability distribution of t, τ, and θ for each clade. This process is equivalent to constructing a Bayesian posterior probability with a null (uninformative) prior for the distribution of these three variables. The approximated posterior distribution represents the probability that a given value of the variable of interest is the true value. The minimum and maximum values included in these confidence contours were used to represent confidence limits on estimates of t, τ, and θ.

The posterior probability distributions were used to test two types of hypotheses. First, we tested whether a particular parameter (t, τ, or θ) was equal for two given datasets by comparing the posterior distributions of t, τ, and θ across all

TABLE 1. Estimates of model parameter values for five datasets and significance of comparisons among them. (A) Estimates of model parameters (γ, α, and β) and parameters derived from them (t, τ, and θ). Numbers in parentheses indicate 95% confidence limits. (B) Upper diagonal: significance of difference in duration of lag phase (t). Lower diagonal: significance of difference in duration of accumulation phase (τ). (C) Significance of difference in θ. *$\phi < 0.05$; **$\phi < 0.01$; ***$\phi < 0.001$.

(A)

Dataset	γ	α	β	Lag phase (t)	Accumulation phase (τ)	Pattern ($\theta = t/\tau$)
Drosophila	4.18	1.95	2.41	0.24 (0.17–0.33)	0.81 (0.54–1.48)	0.296 (0.12–0.64)
Anurans	2.12	438.00	2000.07	0.47 (0.3–0.77)	0.22 (0.11–0.29)	2.14 (1.16–4.4)
Lepidopterans	3.66	1.42	3.97	0.27 (0.16–0.50)	0.36 (0.16–0.80)	0.76 (0.18–2.42)
Anurans excl. *Hyla versicolor*	2.06	425.50	1999.02	0.49	0.21	2.33
Drosophila independent dataset	4.33	1.38	1.62	0.23	0.85	0.27

(B)

	Drosophila	Anurans	Lepidopterans
Drosophila	—	***	ns
Anurans	***	—	*
Lepidopterans	***	*	—

(C)

	Drosophila	Anurans	Lepidopterans
Drosophila	—		
Anurans	***	—	
Lepidopterans	*	**	—

pairwise combinations of the three datasets. For these comparisons, we estimated the probability (ϕ) that the true value of θ (or t, or τ) for the dataset with the larger estimated value is greater than that with the lesser value, by calculating

$$\phi = \left[1 - \int P_a(\theta_a)P_b(\theta_b)d\theta_a\, d\theta_b, \qquad \theta_a > \theta_b\right], \quad (7)$$

where P_a and P_b are the posterior probability distributions for the datasets with the greater and lesser values, respectively, of the variable of interest. We used $\phi < 0.05$ to indicate a significant difference between datasets.

Second, we quantified the rate of accumulation of *RI* in each dataset by asking whether the data were consistent with either a null model of linear accumulation of *RI* or an accumulation proportional to the square of the genetic distance (quadratic accumulation). To test for linear accumulation, we evaluated the likelihood that $\theta = 1/3$. This null hypothesis is rejected if 1/3 lies outside the region of positive probability density for the posterior probability distribution of θ. Analogously, under the quadratic model, the durations of the lag and accumulation phases are expected to be equal, and therefore $\theta = 1$. This hypothesis is rejected if 1.0 lies outside the region of positive probability density for θ.

Alternative Statistical Approaches

The statistical analysis we have outlined assumes specific forms for the distributions of lag and accumulation times. Because these distributions are unknown, we chose an exponential decay and a gamma function because in our view they seemed reasonable and tractable. Nevertheless, one caveat to our analysis is that conclusions may differ if very different distributions are assumed.

Another alternative approach to the model would be the simple fitting of a curve to determine the shape of the relationship between *RI* and genetic distance. For example, one could fit both a linear and a quadratic regression to the data and determine which provided a better fit to the data. One limitation of such an approach is that, because *RI* can take a maximum value of 1.0, only species pairs for which $RI < 1$ could be used in the analysis. Yet species pairs for which $RI = 1$ provide key information about the trajectory of speciation, namely at what point it is expected to be complete. The straight regression approach is thus likely to be less powerful in discriminating between alternative hypotheses of accumulation. In addition, because the datasets used here take on only a few discrete values of *RI*, the regression residuals will not be normally distributed, making statistical comparisons of goodness of fit problematic. Moreover, if the accumulation phase is short, there will be few points with intermediate *RI*, making regression estimates unreliable. Finally, with appropriate data, our analysis allows us to determine whether there is a true lag phase preceding the onset of reproductive isolation and thus provides a more accurate estimation of the duration of speciation, that is, from onset to completion. A regression approach can accomplish neither of these tasks because it assumes there is no lag phase. We performed regression analyses on each of the three datasets examined here and found that they do not distinguish between linear and quadratic accumulation of *RI* and are therefore uninformative. Consequently, we present only the results of analysis with our model.

RESULTS

Comparing Parameter Values across Datasets

Maximum-likelihood estimates of the three parameters γ, α, and β were used to calculate estimates of the average duration of the lag phase (t), the average duration of the accumulation phase (τ), and the pattern of speciation (θ) for each of the three datasets (Table 1A). Estimated posterior probability distributions for average values of t, τ, and θ are given in Figure 3. The average duration of the lag phase for anurans was significantly greater than those for *Drosophila*

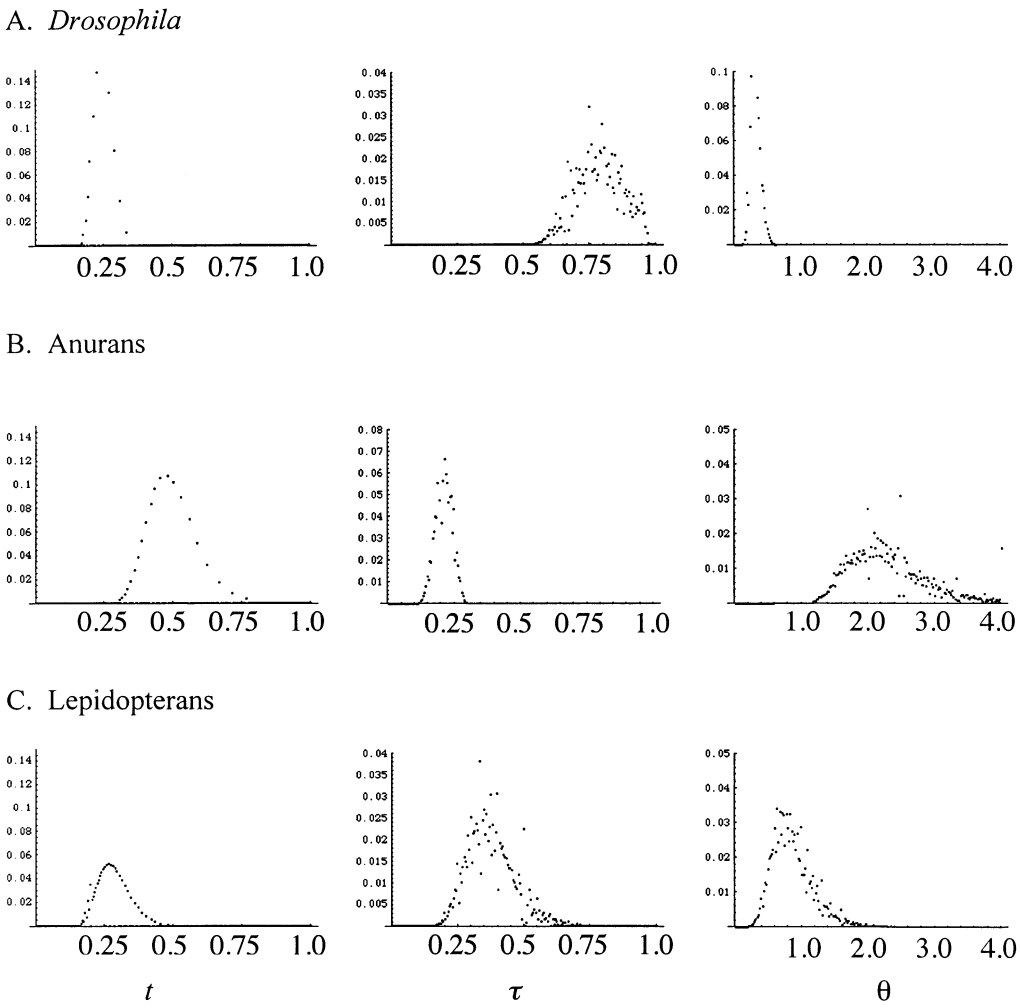

FIG. 3. Cumulative likelihood scores for values of t, τ, and θ, representing posterior probability distributions.

($\phi < 0.001$) and lepidopterans ($\phi = 0.012$), while the lag phases of *Drosophila* and lepidopterans did not differ significantly ($\phi = 0.46$). The duration of the accumulation phase in *Drosophila* was significantly greater than those of both lepidopterans ($\phi < 0.001$) and anurans ($\phi < 0.001$). The accumulation phase in lepidopterans was, in turn, greater than that in anurans ($\phi = 0.02$, Table 1B).

The taxa also differed in the pattern of speciation (θ), or the ratio of the two phases (Table 1C). The mean value of θ in anurans was significantly greater than those in *Drosophila* ($\phi < 0.001$) and lepidopterans ($\phi = 0.008$). For the insect taxa, despite similar lag phases, the pattern of evolution of postzygotic isolation was significantly different. Because of the greater duration of the second phase, *Drosophila* appeared to experience a shorter relative lag phase than did lepidopterans ($\phi = 0.01$).

Finally, the sum of the durations of the two phases is an estimate of the average magnitude of genetic divergence associated with complete postzygotic isolation. For *Drosophila*, postzygotic isolation appeared to evolve to completion by an average genetic distance of $D = 0.24 + 0.81 = 1.05$. In anurans, postzygotic isolation evolved to completion at an average genetic distance of $D = 0.47 + 0.22 = 0.69$. For lepidopterans, this value was $D = 0.27 + 0.36 = 0.63$. The value for lepidopterans is substantially shorter than the total duration of evolution of postzygotic isolation in *Drosophila* due to a significant difference in the duration of the accumulation phases for these two taxa.

Corrected Datasets

The analysis of postzygotic isolation in *Drosophila* using only phylogenetically independent species pairs yielded maximum-likelihood estimates of t and τ consistent with estimates from the full dataset ($t = 0.23$ versus 0.24, $\tau = 0.85$ versus 0.81, and $\theta = 0.27$ versus 0.35, for the independent and full datasets respectively; Table 1A, Fig. 2B). It seems, therefore, that estimates of parameters for the full dataset are not detectably biased by lack of phylogenetic independence.

One species in the anuran data set, *Hyla versicolor*, is a tetraploid species derived from the diploid *H. chrysocelis*. Because the tempo of polyploid speciation, predicted to be essentially instantaneous, may be very different than that of more typical speciation processes, a separate analysis excluding *H. versicolor* was also conducted.

The analysis of the anuran dataset excluding *H. versicolor*

yielded results consistent with those when *H. versicolor* was included (t = 0.49 versus 0.47, τ = 0.21 versus 0.22, θ = 2.33 versus 2.14, excluded versus included, respectively; Table 1A). Thus, although polyploid speciation may proceed along a very different evolutionary trajectory than normal speciation processes, this mode of speciation applies to only one species in the analysis and does not appear to affect the final result.

Evaluating Two Models of Accumulation

For postzygotic isolation in *Drosophila*, the region of positive probability encompassed $\theta = 1/3$, corresponding to the linear hypothesis, but not $\theta = 1.0$, corresponding to the quadratic hypothesis (Fig. 3A). This pattern indicates that the pattern of accumulation of *RI* is consistent with the linear model of accumulation, but not with the quadratic model.

The probability density for θ in anurans was zero for $\theta < 1.4$ (Fig. 3B), which means that both $\theta = 1/3$ and $\theta = 1.0$ have essentially zero probability of being the true values for this dataset. Our analysis therefore indicates that, for anurans, while neither the linear nor the quadratic model of accumulation of *RI* was appropriate, accumulation appears faster than quadratic and thus exhibits a snowball effect.

For lepidopterans, both $\theta = 1/3$ and $\theta = 1.0$ fell within the region of positive probability density for θ (Fig. 3C), indicating that this dataset is consistent with either the linear or quadratic accumulation model. We do note, however, that the posterior probability associated with θ less than or equal to 1/3 is very small ($\theta = 1/3$ lies in the extreme tail of the probability distribution), whereas $\theta = 1.0$ corresponds to the peak of the probability distribution. This difference suggests that in lepidopterans the quadratic accumulation model is more likely correct than the linear model. Accumulation of *RI* in lepidopterans therefore also appears to exhibit a snowball effect, though it is not as rapid as in anurans.

Discussion

Differences among Taxa in the Pattern of Speciation

A principal result of our analysis is that the three taxa differed substantially in the absolute duration of both the lag and accumulation phases and in the relative durations of the two phases, as measured by θ. Differences among taxa in absolute durations of each phase are perhaps not surprising; for example, they can be accounted for by differences among taxa in the proportion of substitutions that result in D-M incompatibilities or in the average effect of a D-M incompatibility on hybrid fitness. In addition, because Nei's genetic distance reflects an average per locus divergence, rather than the total number of substitutions differentiating two genomes, the rate differences detected here could also be caused by differences in the number of genes in the genomes of the different taxa.

By contrast, differences in the pattern of evolution (θ), which as a dimensionless parameter should control for the differences described above, require some explanation. One notable result is the significantly different pattern of evolution observed for the anurans. These results therefore do not support the conclusion of Sasa et al. (1998, p. 1816) that ''the rates of acquisition of postzygotic . . . isolation with respect to genetic distances are similar in frogs and in fruitflies.'' Our results suggest anurans experience a lag phase more than twice as long as the accumulation phase, whereas for the insect taxa the lag phase is a small fraction of the total duration of speciation.

One intriguing explanation for the difference detected between anurans and the insect taxa may lie in the nature of sex determination in these organisms. *Drosophila* and lepidopterans are both characterized by heterogametic sex determination (XY in *Drosophila*, ZW in lepidopterans), whereby sex-determining chromosomes are not homologous and therefore do not share the same genes. Differentiated sex chromosomes are thought to play a substantial role in the dynamics of postzygotic isolation (Muller 1940; Wu and Davis 1993; Laurie 1997). One consequence is that heterogametic F_1 hybrids (males in *Drosophila*, females in lepidopterans) experience an X-autosome imbalance, whereby half their autosomes lack a conspecific X (Z) chromosome. If epistatic interactions between autosomes and one of the sex chromosomes are important for viability or fertility, then heterogametic hybrids will suffer as substitutions appear on the sex chromosomes of diverging lineages, and the result will be early onset of postzygotic isolation.

In contrast, most species of anurans exhibit undifferentiated (homogametic) sex chromosomes (Schmid and Steinlein 2001). Hybrids of diverging lineages of anurans should therefore be balanced in terms of autosomes and sex chromosomes and should not suffer the same consequences as hybrids containing a degenerate sex chromosome. Relatively distantly related lineages of anurans may therefore be capable of producing viable and fertile hybrids despite a degree of differentiation on sex chromosomes that in *Drosophila* and lepidopterans would result in detectable reproductive isolation, here, complete sterility, and/or inviability of at least one sex of one reciprocal cross (RI = 0.25).

Differentiated sex chromosomes may also explain the relatively longer duration of the accumulation phase in the insect taxa. Coyne and Orr (1989, 1997) found that postzygotic isolation in *Drosophila* begins to evolve early, on average, reaching an intermediate magnitude (RI = 0.5) at relatively small genetic distances, but it then remains at that level for a significant amount of time before increasing again. The authors refer to this phenomenon as ''stalling'' and implicate the disproportionate effect of the X chromosome in male hybrids as the causal factor (Coyne and Orr 1989, p. 378), noting that hybrid female fitness is not detectably reduced until much greater genetic distances. This delayed onset of female inviability/sterility is a predictable result of Haldane's rule (i.e., the reduced fitness of the heterogametic sex in early stages of divergence; Haldane 1922), commonly observed in both *Drosophila* and lepidopterans. Although these causal explanations are speculative, the short lag phase and relatively long accumulation phase observed for the two insect taxa in the present study are consistent with results of Coyne and Orr's study and with predictions based on Haldane's rule. The difference observed between the insect taxa and anurans is consistent with differences in karyotype among these taxa.

The Snowball Effect

The second principal result of our analysis is that some, but not all, of the datasets exhibited a snowball effect for postzygotic isolation, whereby the magnitude of isolation accumulates at an increasing rate. Orr (1995) and Orr and Turelli (2001) demonstrated theoretically that the number of D-M incompatibilities between two lineages is expected to increase faster than linearly with the number of substitutions separating those lineages. If D-M incompatibilities involve primarily two loci, then accumulation is expected to be quadratic (Orr and Turelli 2001), whereas, if incompatibilities involve more than two loci, accumulation is expected to be even faster (Orr 1995). The likelihood of being able to test these predictions in the near future seems low: the low rate of discovery and genetic characterization of D-M incompatibilities and the difficulties associated with determining the exact number of substitutions that separate the genomes of two taxa render attempts to empirically estimate the form of the relationship between the number of substitutions and the number of incompatibilities exceedingly difficult.

It is possible, however, to determine whether the accumulation of the magnitude of reproductive isolation exhibits a snowball effect. Although it is not clear whether the pattern of evolution of hybrid fitness and that of D-M incompatibilities should be similar, we predict that if the effects of incompatibilities on hybrid fitness are additive and the average magnitude of those effects do not change consistently over time, the pattern of accumulation of *RI* should reflect the pattern of accumulation of incompatibilities. The absence of a snowball effect for *RI* would then suggest either that one of these assumptions is incorrect, or that some assumption of the Orr-Turelli model is possibly inappropriate.

Both the lepidopteran and the anuran data appeared to exhibit a snowball effect with respect to the accumulation of postzygotic isolation. For anurans, θ was substantially greater than 1.0, indicating *RI* accumulates faster than quadratically with genetic distance. For lepidopterans, the best estimate of θ was only slightly less than 1, suggesting the hypothesis that *RI* increases quadratically cannot be rejected. This result is consistent with expectations for the accumulation of D-M incompatibilities (Orr and Turelli 2001). The substantially faster rate observed in anurans may indicate a greater number of multilocus incompatibilities, or, as suggested above, a truly longer lag phase resulting from a lack of differentiated sex chromosomes.

Unlike in anurans and lepidopterans, postzygotic isolation in *Drosophila* did not appear to exhibit a snowball effect; indeed, the relationship between *RI* and genetic distance for *Drosophila* was estimated to be approximately linear. Again, we can only speculate about the causes of this relationship. One possibility, assuming the Orr-Turelli model is appropriate, is that the average effect of a D-M incompatibility on hybrid fitness decreases as incompatibilities accumulate, as would occur if fitness reductions caused by incompatibilities combine multiplicatively rather than additively (that is, $W = [1 - a]^n$, where W is hybrid fitness, $1 - a$ is the average proportional decrease in hybrid fitness caused by an incompatibility, and n is the number of incompatibilities). Such a possibility is suggested by the common assumption that survival components of fitness combine multiplicatively.

Another possible explanation for the lack of snowball effect observed in *Drosophila* is that substitutions resulting from selection may be more (or less) likely to result in incompatibility than those resulting from genetic drift (for example, Gavrilets 1999; Barbash et al. 2003; Presgraves et al. 2003). If so, then the pattern of accumulation of D-M incompatibilities (and by extension *RI*) may be influenced by temporal changes in the relative contributions of selection and drift to genetic divergence. Selective substitutions may be more frequent in the early period of divergence rather than in later periods. For example, a physical barrier separating two lineages may substantially alter the environment on either side of the barrier, leading to an initial period of rapid adaptive evolution followed by a reduced rate of evolutionary change as each lineage approached its adaptive peak. If substitutions due to selection are more likely to be involved in D-M incompatibilities than neutral substitutions, then an early period of intense selective divergence will raise the rate of accumulation of incompatibilities during that period, making the long-term accumulation look more linear than envisioned under the Orr-Turelli model.

Such possibilities are of course speculative, but the variation observed among datasets in the rate of accumulation requires explanation. Theoretical analyses exploring the expected relationship between the number of incompatibilities and the magnitude of postzygotic isolation, as well as those examining the consequences of altering some assumptions of the Orr-Turelli model, are likely to contribute to understanding this variation.

Suitability of Current Data

The strength of our conclusions must be tempered by limitations of our analysis. Two limitations pertain to the nature of the datasets used. First, the estimates of postzygotic isolation are discrete rather than continuous. Thus, species pairs with *RI* that is detectable but less than a value of 0.25 will be estimated to exhibit no isolation. The most obvious effect of this misclassification will be to overestimate the duration of the true lag phase, if it exists, and to underestimate the duration of the accumulation phase. This bias should not affect our two main conclusions, however, those being that taxa differ in the pattern of accumulation as reflected by θ and that only anurans and lepidopterans exhibit a snowball effect in the accumulation of *RI*. Neither of these conclusions depends on the assumption that species pairs with $RI = 0$ exhibit no detectable postzygotic isolation.

Second, the datasets analyzed do not consist of phylogenetically independent species pairs. A lack of independence may bias parameter estimates if some species or clades are represented disproportionately, and it violates the assumption of the likelihood analysis, potentially resulting in artificially low probability estimates. Analysis of the phylogenetically independent dataset for *Drosophila* suggests the first problem may not be a major one. Because estimates of the parameters t, τ, and θ were very similar for both datasets, there appears to be little bias associated with analyzing the full dataset. Although the second problem remains, we note that for most

comparisons differences are significant at a very high level (e.g., $\phi < 0.001$), suggesting that although the significance level may be somewhat inflated, our tests are probably detecting real differences. We recognize, however, that this issue will be settled definitively only when information on a substantial number of phylogenetically independent pairs is accumulated.

Coyne and Orr (1989) used an alternative approach to generate larger datasets of phylogenetically independent species pairs, but this approach cannot be used with our method. Coyne and Orr used the average of values of genetic distance and strength of isolation for two nonindependent species pairs as a new datapoint in the dataset (1989, pp. 371–372; see also Felsenstein 1985) and thus increase the sample size. By using average values, however, this approach may overestimate the number of species pairs exhibiting intermediate magnitudes of isolation. An overabundance of species pairs with intermediate reproductive isolation will in turn overestimate the duration of the accumulation phase, as the number of intermediate datapoints indicates the probability that any two species will be in that phase. Our method therefore requires knowledge of whether species pairs exhibit truly intermediate, complete, or no detectable reproductive isolation.

Conclusions

Using a method that estimates the durations of both early (lag) and subsequent (accumulation) phases of the evolution of reproductive isolation, we were able to quantify the pattern of speciation in three different organisms. Our analysis demonstrated that the relative durations of these two phases, the pattern of evolution, differed markedly among these groups in ways that are consistent with genetic differences. In addition, the rate of accumulation of reproductive isolation was found to differ among taxa, with lepidopterans and anurans exhibiting a quadratic and faster than quadratic rate of accumulation, respectively, and *Drosophila* exhibiting a more nearly linear rate of accumulation. The causes of this variation are unknown, but are likely to be revealed by detailed analysis of the genetic architecture of reproductive isolation.

Acknowledgments

We thank M. Gilchrist for assistance with programming. We are grateful to H. Allen Orr, M. Noor, N. Johnson, and an anonymous reviewer for providing helpful comments and insight. A. Thistle provided invaluable editorial advice. This work was partly funded by National Science Foundation grant MCB 0110596 to MDR.

Literature Cited

Avise, J. C. 1994. Molecular markers, natural history, and evolution. Chapman and Hall, New York.

Barbash, D. A., D. F. Siino, A. M. Tarone, and J. Roote. 2003. A rapidly evolving MYB-related protein causes species isolation in *Drosophila*. Proc. Natl. Acad. Sci. USA 100:5302–5307.

Coyne, J. A. 1992. Genetics and speciation. Nature 355:511–515.

Coyne, J. A., and H. A. Orr. 1989. Patterns of speciation in *Drosophila*. Evolution 43:362–381.

———. 1997. ''Patterns of speciation in *Drosophila*'' revisited. Evolution 51:295–303.

DeSalle, R., and A. V. Z. Brower. 1997. Process partitions, congruence, and the independence of characters: inferring relationships among closely related Hawaiian *Drosophila* from multiple gene regions. Syst. Biol. 46:751–764.

Dobzhansky, T. G. 1937. Genetics and the origin of species. Columbia Univ. Press, New York.

———. 1970. Genetics of the evolutionary process. Columbia Univ. Press, New York.

Felsenstein, J. 1985. Phylogenies and the comparative method. Am. Nat. 125:1–15.

Gavrilets, S. 1999. A dynamical theory of speciation on holey adaptive landscapes. Am. Nat. 154:1–22.

Gleason, J. M., E. C. Griffith, and J. R. Powell. 1998. A molecular phylogeny of the *Drosophila willistoni* group: conflicts between species concepts? Evolution 52:1093–1103.

Grant, V. 1981. Plant speciation. Columbia Univ. Press, New York.

Haldane, J. B. S. 1922. Sex-ratio and unisexual sterility in hybrid animals. J. Genet. 12:101–109.

Harr, B., B. Zangerl, and C. Schlötterer. 2000. Removal of microsatellite interruptions by DNA replication slippage: phylogenetic evidence from *Drosophila*. Mol. Biol. Evol. 17:1001–1009.

Hilton, H., and J. Hey. 1997. A multilocus view of speciation in the *Drosophila virilis* species group reveals complex histories and taxonomic conflicts. Genet. Res. 70:185–194.

Housworth, E. A., and E. P. Martins. 2001. Random sampling of constrained phylogenies: conducting phylogenetic analyses when the phylogeny is partially known. Syst. Biol. 50:628–639.

Howard, D., and S. H. Berlocher, eds. 1998. Endless forms: species and speciation. Oxford Univ. Press, New York.

Hubbs, C. L. 1955. Hybridization between fish species in nature. Syst. Zool. 4:1–20.

Knowlton, N., L. A. Weigt, L. A. Solorzano, D. K. Mills, and E. Bermingham. 1993. Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. Science 260:1629–1632.

Laurie, C. C. 1997. The weaker sex is heterogametic: 75 years of Haldane's rule. Genetics 147:937–951.

Lemeunier, F., S. Aulard, M. Arienti, J.-M. Jallon, M.-L. Cariou, and L. Tsacas. 1997. The *ercepae* complex: new cases of insular speciation within the *Drosophila ananassae* species subgroup (*melanogaster* group) and descriptions of two new species (Diptera: Drosophilidae). Ann. Entomol. Soc. Am. 90:28–42.

Mayr, E. 1963. Populations, species, and evolution. Harvard Univ. Press, Cambridge, MA.

Mendelson, T. C. 2003. Sexual isolation evolves faster than hybrid inviability in a diverse and sexually dimorphic genus of fish (Percidae: *Etheostoma*). Evolution 57:317–327.

Muller, H. J. 1940. Bearing of the *Drosophila* work on systematics. Pp. 185–286 *in* J. Huxley, ed. The new systematics. Clarenden Press, Oxford, U.K.

———. 1942. Isolating mechanisms, evolution and temperature. Biol. Symp. 6:71–125.

Nei, M. 1975. Molecular population genetics and evolution. Elsevier, New York.

———. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.

Nurminsky, D. I., E. N. Moriyama, E. R. Lozovskaya, and D. L. Hartl. 1996. Molecular phylogeny and genome evolution in the *Drosophila virilis* species group: duplications of the alcohol dehydrogenase gene. Mol. Biol. Evol. 13:132–149.

O'Grady, P. M. 1999. Reevaluation of phylogeny in the *Drosophila obscura* species group based on combined analysis of nucleotide sequences. Mol. Phylogenet. Evol. 12:124–139.

Orr, H. A. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics 139:1805–1813.

Orr, H. A., and M. Turelli. 2001. The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution 55:1085–1094.

Otte, D., and J. A. Endler, eds. 1989. Speciation and its consequences. Sinauer, Sunderland, MA.

Presgraves, D. 2002. Patterns of postzygotic isolation in Lepidoptera. Evolution 56:1168–1183.

Presgraves, D. C., L. Balagopalan, S. M. Abmayr, and H. A. Orr.

2003. Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. Nature 423: 715–719.

Rodriguez-Trelles, F., L. Alarcon, and A. Fontdevila. 2000. Molecular evolution and phylogeny of the *buzzatii* complex (*Drosophila repleta* group): a maximum-likelihood approach. Mol. Biol. Evol. 17:1112–1122.

Sasa, M. M., P. T. Chippindale, and N. A. Johnson. 1998. Patterns of postzygotic isolation in frogs. Evolution 52:1811–1820.

Schawaroch, V. 2002. Phylogeny of a paradigm lineage: the *Drosophila melanogaster* species group (Diptera: Drosophilidae). Biol. J. Linn. Soc. 76:21–37.

Schmid, M., and C. Steinlein. 2001. Sex chromosomes, sex-linked genes, and sex determination in the vertebrate class Amphibia. Pp. 143–176 *in* G. Scherer and M. Schmid, eds. Genes and mechanisms in vertebrate sex determination. Birkhäuser, Basel, Switzerland.

Tilley, S. G., P. A. Verrell, and S. J. Arnold. 1990. Correspondence between sexual isolation and allozyme differentiation: a test in the salamander *Desmognathus ochrophaeus*. Proc. Natl. Acad. Sci. USA 87:2715–2719.

Wu, C.-I., and A. W. Davis. 1993. Evolution of postmating reproductive isolation: the composite nature of Haldane's rule and its genetic bases. Am. Nat. 142:187–212.

Corresponding Editor: M. Noor