The Dissertation Committee for Alan Richard Lemmon

certifies that this is the approved version of the following dissertation:

Analytical, Computational, and Statistical Approaches to Studying Speciation

Committee:

_____
Mark Kirkpatrick, Supervisor

_____
David M. Hillis, Supervisor

_____
Thomas Juenger

_____
Lauren Ancel Meyers

_____
Tandy Warnow

Analytical, Computational, and Statistical Approaches to Studying Speciation

by

Alan Richard Lemmon, B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2007

This dissertation is dedicated to Emily Moriarty Lemmon, my inspiration.

Acknowledgments

I would like to thank all of the colleagues who helped me to grow as student and scientist. At the University of Colorado at Boulder, I thank Andrew Martin, Ruth Bernstein, Jorge Moreno, and especially Michael Grant for seeing promise in a naïve undergraduate. At the Free University of Brussels in Belgium, I would like to thank Michael Milinkovitch for a stimulating collaboration. At the University of Texas at Austin I thank Mark Kirkpatrick, David Hillis, Thomas Juenger, and David Hall for setting high standards and for helping me to achieve those standards. I would also like to thank participants in the IGERT program, as well as the graduate students in the Ecology, Evolution, and Behavior graduate program (especially Jeremy Brown), who provided support, feedback, and collaboration.

I would like to thank my friends and family, for their support and patience during the long process of becoming an independent biologist. I especially thank Emily Moriarty Lemmon, who's support was absolutely essential to my success as a graduate student.

Work presented in this dissertation was supported by National Science Foundation grants DEB-09973221, EF-0328594, & DGE-0114387, National Environmental Research Council grant NE/A/S.2002/00857, and an NSF Graduate Research Fellowship.

Analytical, Computational, and Statistical Approaches to Studying Speciation

Publication No. _____

Alan Richard Lemmon, Ph.D.
The University of Texas at Austin, 2007

Supervisors: Mark Kirkpatrick and David M. Hillis

Two of the most challenging goals of evolutionary biology are to reconstruct the evolutionary relationships among all extant species and to understand the process by which new species form. Accomplishing these goals will require accurate computational methods for reconstructing phylogenetic trees, general analytic models of speciation, and powerful statistical tools for studying the process of speciation in natural systems.

In the first chapter, I study the effects of improper model assumption on estimates of phylogeny. Using DNA sequence data simulated under a variety of models of sequence evolution, I demonstrate that use of oversimplified models can result in erroneous phylogeny estimates. This result suggests that if the models currently utilized are oversimplified then current estimates of phylogeny may be inaccurate and more complex models need to be developed and employed.

In the second and third chapters, I study one process thought to be important in completing the final stages of speciation: reinforcement. Using simulations of a hybrid zone, I show that the process of reinforcement can result in patterns other than reproductive character displacement. I also show that speciation by reinforcement is more likely when the genes involved in reproductive isolation are sex-linked.

In the fourth chapter, I develop a statistical method of quantifying the degree of isolation between species undergoing divergence. Using genotype data obtained from natural hybrid zones, this novel method can be used to estimate the fitness of hybrids during different stages of their life cycle. This approach offers a new approach to empirical biologists studying extrinsic postzygotic isolation in natural systems.

Table of Contents

Chapter 1: The Importance of Proper Model Assumption in Bayesian Phylogenetics 1

1.1 Introduction 2

1.2 Methods 4

1.3 Results 12

1.4 Discussion 19

Figures 25

References 35

Chapter 2: Reproductive Character Displacement is Not the Only Possible Outcome of Reinforcement 43

2.1 Introduction 44

2.2 Model 47

2.3 Quantifying Cline Shape 49

2.4 Results…..... 50

2.5 Discussion..... 52

Tables..... 56

Figures..... 58

References..... 62

Chapter 3: Reinforcement and the genetics of hybrid incompatibilities..... 64

3.1 Introduction..... 65

3.2 Analytic Model..... 67

3.3 Simulations..... 84

3.4 Discussion..... 90

Tables..... 94

Figures..... 100

References..... 104

Chapter 4: The Utility of Diagnostic Markers for Estimating Hybrid Fitness Components in Natural Hybrid Zones..... 109

4.1 Introduction..... 110

4.2 Methods..... 112

4.3 Results..... 120

4.3 Discussion..... 125

Tables..... 129

Figures..... 130

References..... 141

Appendix A: Model of Reinforcement..... 143

Appendix B: Asymptotic Change in Allele Frequencies Among Positions at a Locus..... 171

Appendix C: Associations Between Preference and Incompabitility Positions at QLE..... 174

Appendix D: Selection Coefficients for One- and Two-Locus Incompatibilities..... 185

Appendix E: The Effect of Sex Linkage of Two-Locus Incompatibilities on Reinforcement..... 191

Appendix F: Equilibrium Hybrid Class Frequencies..... 195

Consolidated References..... 203

Vita..... 219

Chapter 1

The importance of proper model assumption in Bayesian phylogenetics*

Abstract. We studied the importance of proper model assumption in the context of Bayesian phylogenetics, by examining more than 5000 Bayesian analyses and six nested models of nucleotide substitution. We found that model misspecification can strongly bias bipartition posterior probability estimates. These biases were most pronounced when rate heterogeneity was ignored. Moreover, the type of bias seen at a particular bipartition appeared to be strongly influenced by the lengths of the branches surrounding that bipartition. In the case of the Felsenstein zone, posterior probability estimates of bipartitions were biased when the assumed model was under-parameterized, but were unbiased when the assumed model was over-parameterized. In the case of the inverse-Felsenstein zone, however, both under-parameterization and over-parameterization lead to biased bipartition posterior probabilities, though the bias caused by over-parameterization was less pronounced and disappeared with increased sequence length. Model parameter estimates were also affected by model misspecification. Under-parameterization caused a bias in some parameter estimates, such as branch lengths and the gamma shape parameter, whereas over-parameterization caused a decrease in the precision of some parameter estimates. We caution researchers to assure that the most

appropriate model is assumed by employing both *a priori* model-choice methods and *a posteriori* model-adequacy tests.

*****Significant portions of this chapter have been previously published as Lemmon & Moriarty, 2004. Systematic Biology 53:265–277.

1.1 INTRODUCTION

Model choice is becoming a critical issue as the number of available models of nucleotide evolution increases rapidly. Moreover, recent studies have shown that adequate model choice is important, demonstrating that violations of model assumptions can produce biased results (Felsenstein, 1978; Huelsenbeck and Hillis, 1993; Yang et al., 1994; Swofford et al., 2001). If the model assumed is over-parameterized (too complex relative to the true underlying model), unnecessary sampling variance is introduced from estimation of extra parameters. This added variance may compromise phylogenetic accuracy (Cunningham et al., 1998). Cases in which the model assumed is under-parameterized (too simple relative to the true underlying model) are especially problematic for phylogeny estimation because of the phenomenon of "long branch attraction" where the confidence in estimation of an incorrect bipartition increases as more data are included (Swofford et al., 2001). Most studies of the importance of model choice have concentrated on the four-taxon case, often comparing maximum parsimony

and/or distance-based methods with maximum-likelihood methods under simple model assumptions (Felsenstein, 1978; Huelsenbeck and Hillis, 1993; Gaut and Lewis, 1995; Swofford et al., 2001).

Traditional likelihood, parsimony, and distance-based methods of phylogeny reconstruction are giving way as Bayesian approaches to phylogeny inference gain rapidly in popularity (Huelsenbeck et al., 2001). Traditional methods yield a single (best) tree, and the uncertainty of each clade is assessed through repeatability tests, such as the bootstrap. The end product of a Bayesian analysis is fundamentally different, consisting of a distribution of "best" trees with associated model parameters sampled in proportion to their posterior probabilities. Uncertainty in the phylogeny and parameter estimates is expressed in the posterior probability distribution. For a more detailed introduction to the use of Bayesian methods in phylogenetics, see Huelsenbeck et al., 2001.

Because Bayesian methods have only recently emerged at the forefront of phylogenetics, research concerning the proper application of these methods and the interpretation of their results is still inadequate (Huelsenbeck et al., 2002). Progress has been made with regard to the relationship between bipartition posterior probabilities and nonparametric bootstrap values, though the relative accuracy of the two measures is still being debated (Suzuki et al., 2002; Wilcox et al., 2002; Alfaro et al., 2003; Cummings et al., 2003; Douady et al., 2003; Erixon et al., 2003). Further exploration of at least three other questions is especially critical: 1) how sensitive are these analyses to prior probability assumptions? 2) what is the most appropriate way to check for convergence

and stationarity of Markov chains in the context of phylogenetics? and 3) how important is proper model assumption within the Bayesian framework?

We present here an analysis that addresses the third question. In particular, we investigate the effect of model misspecification on bipartition posterior probabilities, branch-length estimates, and other substitution-model parameter estimates. We do this by analyzing more than 5000 Bayesian runs, under a variety of nucleotide substitution models. To explore further how bipartition posterior probabilities are affected by model misspecification, we examine two special cases in which adequate model assumption is known to be important: the Felsenstein zone and the inverse-Felsenstein zone (Swofford et al., 2001). We conclude with a discussion of the importance of proper model assumption and what can be done to assure that the most appropriate model available is assumed.

1.2 METHODS

Data Set Simulation. We selected six nested models of nucleotide substitution for our analyses: JC (Jukes and Cantor, 1969), K2P (Kimura, 1980), HKY (Hasegawa et al., 1985), GTR (Lanave et al., 1984; Tavaré, 1986; Rodríguez et al., 1990), GTR+Γ (Steel et al., 1993; Yang, 1993), and GTR+Γ+I (Gu et al., 1995; Waddell and Penny, 1996). We simulated 100 replicate data sets of 1000 bp sequence length (Seq-Gen 1.2.5; Rambaut and Grassly, 1997) assuming each of these six substitution models and the

following parameter values (as appropriate for each model): transition/transversion ratio (κ) = 2.0, π_A = 0.35, π_C = 0.22, π_G = 0.18, π_T = 0.25, r_{CT} = 30.7, r_{CG} = 0.225, r_{AG} = 7.35, r_{AT} = 6.125, r_{AC} = 2.675, gamma shape parameter (α) = 0.67256, and proportion of invariable sites = 0.25. With the exception of the transition/transversion ratio and the proportion of invariable sites, all of these parameter values were obtained from a mitochondrial DNA analysis used to construct a phylogeny of North American chorus frogs (Genus: *Pseudacris*) (Moriarty and Cannatella, in press). We used data sets of 1000 bp because this length was typical of empirical data sets at the onset of this study.

The 30-taxon tree used to simulate the data sets (referred to hereafter as Tree 1) was generated using DNA-Sim (speciation rate = 10^{-4}; extinction rate = 10^{-5}), a program written by A.R.L. that assumes a birth-death process. The branch lengths were assigned in the following fashion: first, we numbered the internal branches from 0 to 26, choosing the order of the branches randomly. Second, each branch was assigned a branch length using the following equation: $f(x) = 10^{3x/26-4}$, where x is the number assigned to that branch. This method of assigning branch lengths assured that a wide range of bipartition posterior probabilities would result from our Bayesian analyses. The procedure was repeated for the 30 external branches, using the equation: $f(x) = 10^{3x/29-4}$. Tree 1 is given in Figure 1.1.

Bayesian Analyses. To investigate the effect of model misspecification on bipartition posterior probabilities, we performed 3600 Bayesian analyses using the program MrBayes 3.0b3 (Huelsenbeck, 2001; Huelsenbeck and Ronquist, 2001). For each of the

600 simulated data sets, we conducted six MrBayes searches, each assuming a different one of the nested models mentioned above. Thus, we examined 36 model combinations, 15 in which the assumed model was under-parameterized, 15 in which the assumed model was over-parameterized, and six in which the assumed model was appropriate relative to the model used to simulate the data sets. This design is depicted in Figure 1.2. We compared results from runs that used the same data set but were analyzed under different substitution models. Using nested substitution models allowed us to systematically test the effect of the presence or absence of each type of parameter on the estimation of bipartition posterior probabilities, branch lengths, and other model parameters. We limited our sampling to 100 replicates because of computational constraints. To assure that results obtained from 100 replicates were reliable, we analyzed an additional 400 replicates for one model combination (GTR+Γ+I -JC). We also assessed the sensitivity of our sample design using power analyses (see below).

We conducted extensive preliminary analyses to determine the sample size, sample interval, and burn-in period appropriate for our data sets. The goal of these preliminary analyses was to determine the conditions that minimized the amount of variation among independent Bayesian analyses run under identical conditions. We assumed default priors for all parameters except for the GTR rate matrix. For the GTR rate matrix, we found that assuming a flat prior yielded incorrect substitution rate estimates and poor convergence to the true posterior distribution. This phenomenon has been studied more extensively by Zwickl and Holder (unpublished data). We found that assuming an exponential prior (revmatpr = exponential(0.2)) allowed for reasonable

convergence to the true posterior distribution when the correct model was assumed (by reasonable convergence we mean that the maximum likelihood estimate of each parameter was at or very near the true value). Four MCMC chains with a temperature of 0.15 assured proper mixing. Each MrBayes run spanned 500,000 generations. We sampled every 25 generations, yielding 20,000 total samples per run. Based on our preliminary tests, we chose an appropriate burn-in time of 25,000 generations (1,000 samples). Thus, each run was analyzed using 19,000 post-burn-in samples.

Convergence Testing. We employed several methods to assure that our runs had converged on the posterior distribution and that we had collected enough samples to obtain reliable results. First, we examined the stationarity of likelihood scores for all 3600 runs performed. As the number of runs conducted in this study is large, however, we could not examine each one independently. Instead, we visualized the likelihood curves (generation plotted on the x-axis, log likelihood on the y-axis) for all 100 replicates of each model combination on the same graph, plotting only those samples in which the likelihood score was greater than in any previous sample (data not shown). By plotting the likelihood scores in this manner, we could quickly identify any runs that failed reach stationarity within the chosen burn-in period.

Second, we examined the convergence of bipartition posterior probabilities, maximum likelihood scores, and model parameter estimates. We expect two converged runs performed on the same data set and under the same model assumptions to produce very similar bipartition posterior probability distributions, maximum likelihood scores,

and model parameter estimates. Thus, a comparison of results from duplicate runs can be used to test for convergence. Due to computational constraints, however, we could not duplicate all 3600 runs. Instead, we chose to concentrate on the 8 model combinations in which the simulated and assumed models were either equivalent (e.g. JC-JC) or showed the greatest disparity (i.e. JC-GTR+Γ+I and GTR+Γ+I-JC). The duplicate runs were compared by observing the degree of correlation (across all 100 replicates) of bipartition posterior probabilities, maximum likelihood scores, and model parameter estimates. Checking for convergence in this way required an additional 800 Bayesian analyses.

Third, we checked the nature of tree space to assure that 500,000 generations allowed convergence upon and proper sampling of the true posterior distribution. To this end, we repeated the analysis of five randomly chosen replicates from each of the four extreme model combinations (JC-JC, JC-GTR+Γ+I, GTR+Γ+I-JC, and GTR+Γ+I-GTR+Γ+I), but allowed these chains to run for 5 million generations. We then compared the posterior distribution sampled in the shorter (500,000 generations) runs with that sampled in the longer (5 million generations) runs to determine whether shorter chains were prone to entrapment in local optima. Each of the 20 pairs of posterior distributions was compared using the Mesquite (Maddison and Maddison, 2003) module Tree Set Viz (Amenta and Klinger, 2002). Tree Set Viz uses multi-dimensional scaling to represent the relationships between topologies (in this case, the topologies in the posterior distribution) as a scatter of points in two-dimensional space. The software arranges the points such that they group according to the distance between the trees (distance between trees was calculated using Robinson-Foulds differences; Robinson and Foulds, 1981). In

addition to the visual comparisons, we compared the posterior distributions of topologies by examining the correlation of posterior probabilities of the topologies found in the shorter runs and the posterior probabilities of the topologies found in the longer runs.

Determining the Effects of Model Misspecification. We studied the effects of model misspecification on estimates of bipartition posterior probabilities, branch lengths, and other model parameters. Since we simulated the data sets, we know the true model parameter values and can compare those values with the estimates obtained through our Bayesian analysis. We do not, however, have true values for the bipartition posterior probabilities (though we do know which bipartitions are correct, we do not know their posterior probabilities a priori, as these will depend on the data set assumed). Consequently, we compared the bipartition posterior probability estimates obtained when an incorrect model was assumed with the estimates that were obtained when the correct model was assumed. The results of this procedure tell us the effect of model misspecification relative to the results that would have been obtained had the correct model been assumed. This comparison gives us a way to measure the bias in bipartition posterior probability estimates induced by model misspecification.

How might biased bipartition posterior probabilities affect conclusions regarding the relationships among taxa? To answer this question we specified a rule for deciding whether a particular observed bipartition was true based on comparison of the posterior probability of that bipartition to a predetermined threshold (the threshold was varied from 0.5 to 1.0). A threshold of 0.5, for example, implies that all bipartitions with a posterior

probability greater than or equal to 0.5 are accepted as true (i.e. all bipartitions in the majority-rule consensus tree are accepted). Since we know the true bipartitions, we can use the decision rule to estimate the probability of Type I and Type II error for each posterior distribution observed. The probability of Type I error was estimated as the proportion of true bipartitions observed that were rejected based on their posterior probability. Conversely, the probability of Type II error was estimated as the proportion of false bipartitions observed that were accepted as true. We compared the probabilities of Type I and Type II error across the 36 model combinations to see how model misspecification might affect conclusions about the relationships among taxa.

Model misspecification may negatively affect parameter estimation by either decreasing accuracy or decreasing precision (Cunningham et al., 1998). To assess how the accuracy of parameter estimates may be affected by model misspecification, we compared (for each parameter) the maximum likelihood estimate of the parameter obtained when the correct model was assumed, with the maximum likelihood estimate of the parameter obtained when the model was misspecified. In order to quantify the degree of bias for each parameter, we employed the two-tailed, paired-sample t test (Zar, 1999). In this case, we are testing whether the distribution of differences (value assuming correct model – value assuming incorrect model) is significantly different from zero. We calculated the P-value associated with the amount of bias observed for all applicable model combinations in which the model was misspecified. To assess how the precision of parameter estimates may be affected by model misspecification, we repeated these tests using the width of the 95% credible set from the posterior distribution of the

parameter as our measure of precision. For each t test performed, we estimated the minimum difference in accuracy and in precision that we are able to detect 99% of the time (β=0.01), given a level of significance of 0.01, and the variance estimated from the distribution of differences (Zar, 1999).

Robustness Tests. In order to test for robustness of our results, we performed a second set of analyses, assuming a different topology and set of model parameters. Because of time constraints, we focused on the four extreme model combinations: JC-JC, JC-GTR+Γ+I, GTR+Γ+I-JC, and GTR+Γ+I-GTR+Γ+I. The parameter values chosen are as follows: π_A = 0.313735, π_C = 0.285552, π_G = 0.18302, π_T = 0.217693, r_{CT} = 33.79102, r_{CG} = 0.55726, r_{AG} = 11.10442, r_{AT} = 3.44797, r_{AC} = 4.16568, gamma shape parameter (α) = 0.583564, and proportion of invariable sites = 0.454113. These values were obtained from a phylogenetic analysis of Siluriformes (D. Hillis, unpublished data). We created a 16-taxon tree (referred to hereafter as Tree 1.2, see Figure 1.3) containing structures that have been shown to be difficult to recover when the assumed substitution model is inappropriate: the structures found in the Felsenstein and inverse-Felsenstein zones (Swofford et al., 2001). We refer to these structures as Felsenstein structures and inverse-Felsenstein structures, respectively. More specifically, Tree 2 contained two Felsenstein structures (containing two long branches separated by a third, much shorter branch) and two inverse-Felsenstein structures (containing a pair of long branches that are adjacent to a pair of much shorter branches). We included two structures of each type in order to yield some variation in the difficulty of the problem, though we could not

perform an extensive analysis here. The branches separating the four structures were fairly long (0.5 substitutions per site), allowing us to avoid confounding effects of interactions among two or more structures. We assumed the same prior distributions, sample size, sample interval, and burn-in times that were described above. The 800 Bayesian analyses (400 unique runs, each duplicated) conducted under these conditions were analyzed in a fashion similar to that described above.

We constructed Tree 2 using two Felsenstein structures and two inverse-Felsenstein structures with the hopes of determining how the properties of a particular bipartition affect the type of bias produced by model misspecification. We can determine the effect of misspecification for each bipartition by comparing the posterior probability obtained for that bipartition when the model is misspecified, with the posterior probability obtained when the model is correctly specified. The proper test for this type of comparison, given that the distributions of bipartition posterior probabilities across the 100 replicates are neither normal, nor homoschedastic, is the non-parametric sign test (Zar, 1999). We employed this test for each of the bipartitions in Tree 2 that are part of either a Felsenstein structure, or an inverse-Felsenstein structure.

1.3 RESULTS

Bipartition Posterior Probabilities. Our main result is that model misspecification can strongly bias bipartition posterior probability estimates (Figure 1.4). Although over-

parameterization had no noticeable effect on bipartition posterior probability estimates, under-parameterization produced a strong bias. The bias observed for a particular bipartition depends on how well supported that bipartition is when the correct model is assumed: well-supported bipartitions tend to be overestimated whereas poorly supported bipartitions tend to be underestimated. Although this is the general trend, the effect of model misspecification on a particular bipartition is likely to be affected by the length of the branch at that bipartition, the length of the branches surrounding that bipartition (see below), and the data set used to infer the phylogeny. Results from an additional 400 replicates for the model combination GTR+Γ+I-JC (see Figure 1.5) suggest that increasing sampling efforts would not have affected our qualitative results regarding the effect of model misspecification on bipartition posterior probability estimates.

The largest bias in bipartition posterior probabilities was seen when the assumed model failed to incorporate rate variation across sites. This bias was especially pronounced when gamma distributed rate heterogeneity was neglected. We also observed that failing to account for unequal rates of base substitution (i.e. the transition bias or the GTR rate matrix) led to slightly biased bipartition posterior probability estimates. Inappropriately assuming equal base frequencies, however, had very little effect on bipartition posterior probability estimates under the conditions tested.

Bias caused by under-parameterization resulted in an increased incidence of Type II error (Figure 1.6). In other words, assuming an under-parameterized model led to the acceptance of a greater number of false bipartitions. This pattern was consistent across

all threshold values tested. Interestingly, the opposite trend occurred for Type I error: assuming an under-parameterized model resulted in the rejection of fewer true bipartitions. As one might expect, increasing the threshold resulted in an increase in Type I error and a decrease in Type II error. Over-parameterization had very little effect on the probability of either type of error.

Branch Lengths and Other Model Parameters. Branch lengths were also affected by model misspecification. Model under-parameterization led to underestimated branch lengths, especially for long branches (Figure 1.7). Failing to account for rate heterogeneity had the largest effect on branch-length estimates, though neglecting to include other parameters also produced slightly underestimated branch lengths. Over-parameterization produced little, if any, bias in branch-length estimates.

In many cases, parameter estimates were biased when the model assumed was under-parameterized (Figure 1.8). Estimation of rate matrix parameters, however, seems to be fairly robust to model misspecification, at least under the conditions tested. Nucleotide base frequencies were only biased when the GTR rate matrix was inappropriately neglected. Estimates of the gamma shape parameter appeared to be biased when the proportion of invariable sites was inappropriately included or ignored. The gamma shape parameter was also biased when the simulated model assumed homogeneity of rates across sites (Figure 1.9). This result is expected because the true value of α in these cases is infinity.

We observed decreased precision of some parameter estimates when the assumed model was over-parameterized (Figure 1.8). Tree length, base frequencies, and the gamma shape parameter showed a strong decrease in precision under some conditions. Estimates of rate matrix parameters and transition bias appeared to be more robust to changes in precision with model over-parameterization. Not surprisingly, under-parameterization tended to result in an increase in the precision of most parameter estimates.

Our model design allowed us to detect small changes in accuracy and precision for estimates of TL, κ, and base frequencies, and moderate changes for estimates of rate matrix parameters and α. The ranges in the minimum detectable difference estimates for our accuracy tests are as follows: TL: 0.015--0.030; κ: 0.096--0.104; π_A: 0.005--0.006; π_C: 0.004--0.007; π_G: 0.004--0.005; π_T: 0.004--0.007; r_{CT}: 2.969--3.466; r_{CG}: 0.072--0.099; r_{AG}: 0.799--0.886; r_{AT}: 0.605--0.760; r_{AC}: 0.304--0.351; α: 0.126--0.169. The range in the minimum detectable difference estimates for our precision tests are as follows: TL: 0.002--0.011; κ: 0.007--0.009; π_A: 4×10^{-4}--0.001; π_C: 4×10^{-4}--0.001; π_G: 4×10^{-4}--0.001; π_T: 4×10^{-4}--0.001; r_{CT}: 2.229--2.666; r_{CG}: 0.024--0.052; r_{AG}: 0.499--0.559; r_{AT}: 0.406--0.480; r_{AC}: 0.177--0.229; α: 0.152--0.174. These results suggest, for example, that we would be able to detect a difference less than or equal to 0.006 between the maximum likelihood estimate of π_A obtained when the model was correctly assumed and the estimate obtained when the model was misspecified. Given that the estimate for this parameter was always greater than 0.268 (for the applicable model combinations), we

would have a 99% chance of detecting a change on the order of 2.2% in the maximum likelihood estimate of π_A. Likewise, we would be able to detect a difference of 0.001 between the width of the 95% credible set of π_A obtained when the correct model was assumed and the width obtained when the model was misspecified. Given that the estimate for this parameter was always greater than 0.036 (for the applicable model combinations), we would have a 99% chance of detecting a change on the order of 2.8% in the width of the 95% credible set.

Convergence. We observed adequate convergence of bipartition posterior probabilities (Figure 1.4), and model parameter estimates (data not shown). When observing likelihood burn-in plots, we found that all runs reached stationarity before 25,000 generations, the chosen burn-in time. Good convergence of bipartition posterior probabilities can be observed in correlation plots of the duplicate runs (see the shaded diagonal of Figure 1.4). These plots demonstrate the congruence of bipartition posterior probability distributions between pairs of independent Bayesian analyses (model combinations not shown, JC-GTR+Γ+I and GTR+Γ+I-JC, demonstrated a pattern very similar to that seen in the diagonal of Figure 1.4). Posterior distributions of the substitution model parameters were congruent with the true parameter values (denoted by arrows in Figure 1.9), when the correct model was assumed. Parameter estimates were also very similar between independent runs under the same conditions (data not shown). Moreover, running the Markov chains for 5,000,000 generations instead of 500,000 did not substantially change the resulting sample taken from the posterior distribution of topologies; the posterior distributions of the shorter and longer runs were congruent in all

20 visual comparisons made using Tree Set Viz (data not shown). Lastly, we observed a strong correlation of topology posterior probabilities between long and short runs for all four model combinations tested (data not shown).

Robustness Tests. The results of the second set of analyses, which investigated Felsenstein and inverse-Felsenstein structures, agreed with those of the first set, though the bias was even more pronounced in this case (Figure 1.10). The one exception is that, in the second set of analyses, we were able to detect a bias in bipartition posterior probabilities caused by model over-parameterization under some conditions (see results below). This bias, however, was much less pronounced than the bias caused by under-parameterization. The branch lengths were also more strongly biased by under-parameterization under the second set of conditions, though this could be due to the fact that branches in Tree 2 (Figure 1.3) were longer than those in Tree 1 (Figure 1.1).

We were surprised to observe a slight bias in bipartition posterior probabilities when the assumed model was over-parameterized (Figure 1.10a). Other authors have found similar patterns when using simulated data sets of 1000 nucleotides, but found that the bias disappeared with increased sequence length (Swofford et al., 2001; Sullivan and Swofford, 2001). To determine if the bias we observed was also attributable to sequence length, we constructed data sets of length 5000, 10000, and 50000 nucleotides by successively concatenating randomly chosen data sets (without replacement) from the pool of 100 replicate data sets used in the robustness tests (see Methods). Upon analyzing these data sets in the same fashion as those containing 1000 nucleotides, we

found that increasing sequence length corrected the slight bias seen in the case of over-parameterization, but amplified the already large bias seen in the case of under-parameterization (data not shown).

After examining more carefully how model misspecification affected the posterior probabilities of each of the bipartitions, we found that posterior probabilities of bipartitions found in Felsenstein structures were biased by under-parameterization but not over-parameterization. For three of the four bipartitions found in Felsenstein structures, a significant number of replicates showed a decrease in the bipartition posterior probability when the assumed model was under-parameterized (one-tailed Paired Sign Test, α = 0.05; node 0: $P = 1.84 \times 10^{-1}$, node 2: $P = 3.32 \times 10^{-3}$, node 7: $P = 9.05 \times 10^{-8}$, and node 9: $P = 1.53 \times 10^{-17}$). We attribute the one exception to Type II error. When the assumed model was over-parameterized, however, no significant bias was observed for any of the four bipartitions (node 0: $P = 3.09 \times 10^{-1}$, node 2: $P = 7.95 \times 10^{-2}$, node 7: $P = 5.00 \times 10^{-1}$, and node 9: $P = 4.19 \times 10^{-1}$).

For bipartitions found in inverse-Felsenstein structures, we found that both over-parameterization and under-parameterization produced a bias in bipartition posterior probability estimates. The direction of the bias in each case depended upon whether the particular bipartition was adjacent to two long tips, or two short tips. When the assumed model was under-parameterized, a significant number of replicates showed a decrease in the posterior probabilities of bipartitions closest to two long tips (node 3: $P = 1.32 \times 10^{-25}$, node 10: $P = 9.05 \times 10^{-8}$) and a significant number of replicates showed an increase in

the posterior probabilities of bipartitions closest to two short tips (node 5: $P = 7.97 \times 10^{-29}$, node 12: $P = 7.89 \times 10^{-31}$). Over-parameterization produced the opposite bias for all nodes in the inverse-Felsenstein structures: a significant number of replicates showed an increase in the posterior probabilities of bipartitions closest to two long tips (node 3: $P = 4.43 \times 10^{-2}$, node 10: $P = 9.16 \times 10^{-5}$), whereas a significant number of replicates showed a decrease in the posterior probabilities of bipartitions closest to two short tips (node 5: $P = 1.20 \times 10^{-3}$, node 12: $P = 7.81 \times 10^{-27}$).

Note that the sign test does not tell us the magnitude of the bias, only the direction of the bias. Figure 1.10a makes it clear, however, that the bias in bipartition posterior probabilities due to under-parameterization is much more pronounced than that due to over-parameterization. We should also point out that the bipartition posterior probabilities for nodes 1, 4, 6, 8, and 11 were always at or very near 1.0 for all 100 replicates and all model combinations. This demonstrates that our results were not influenced by interactions among two or more structures.

1.4 DISCUSSION

Model under-parameterization can strongly bias estimates of bipartition posterior probabilities, branch lengths and other model parameters. The bias is especially severe when rate heterogeneity is neglected. This result is not surprising, as previous studies have demonstrated that ignoring rate heterogeneity among sites can bias topology

estimation (Kuhner and Felsenstein, 1994; Yang et al., 1994; Sullivan et al., 1995; Lockhart et al., 1996) and can lead to underestimation of branch lengths (Golding, 1983; Yang et al., 1994). Our results also agree with previous studies of model misspecification in that the bias seen in branch-length estimates increases disproportionately as branch length increases (Golding, 1983).

Results from our analyses of Type I and Type II error suggest that the best approach to assuring accurate and informative phylogenies is to employ a sufficiently complex model and to accept bipartitions as true only if their posterior probability is moderately high ($0.7 \leq$ decision threshold ≥ 0.9). Based on the results of this study, there appears to be little advantage to requiring posterior probabilities to be very near one when an adequate model is available. When an adequate model is not available, however, the conservative approach would be to accept bipartitions as true only if they have very high posterior probabilities (e.g. greater than 0.9). We should point out that these conclusions are based on results obtained under the particular set of conditions we examined here. Clearly, more research investigating the factors affecting error in phylogeny estimates is needed.

Model over-parameterization carries a cost as well: including unnecessary parameters can lead to decreased precision in estimates of branch lengths and other model parameters, as suggested by Cunningham et al. (1998). We have also seen, in the case of the inverse-Felsenstein zone, that model over-parameterization can sometimes lead to slightly biased estimates of bipartition posterior probabilities, though this bias is

expected to decrease with increased sequence length. There are two additional negative consequences of model over-parameterization that we have not investigated here. First, computation time is likely to increase rapidly with the complexity of the model assumed, especially when data sets are large (Lemmon and Milinkovitch, 2002). Second, over-parameterization may affect the convergence of Markov chains (see Huelsenbeck et al., 2002 for a discussion of convergence).

If appropriate model assumption is so important, how are we to identify an appropriate model? Two measures should be taken to assure that a proper model is assumed: First, before a Bayesian analysis is performed, one should identify the available model that best fits one's data set. Several methods facilitate this choice, including the likelihood-ratio test (Goldman, 1993), the Akaike information criterion (Akaike, 1974), and the Bayesian information criterion (Schwarz, 1974). Posada and Crandall (2001) compare the performance of these methods in detail, finding the likelihood-ratio test to be more accurate than the Akaike and Bayesian information methods under some conditions. Note that none of these methods tell us whether or not a particular model adequately describes a particular data set, but are only useful in choosing the best model among a set that may or may not contain an adequate model. In fact, there is some evidence that when all of the available models are inadequate, the hierarchical likelihood-ratio test performs poorly relative to other model selection methods (Minin et al., 2003). Another disadvantage of the commonly used likelihood-ratio test is that it is only appropriate for choosing among nested models. When unnested models are being compared, an alternative method should be used, such as parametric bootstrapping (Goldman, 1993) or

recently developed Decision Theory methods (Minin et al., in 2003). Both of these methods can also be used to test for absolute goodness-of-fit of the model to the data set.

Second, following a Bayesian analysis, one should perform model-adequacy tests to assure that the model assumed in the analysis adequately explains the observed data set. Model adequacy can be tested, for example, using posterior predictive distributions as described by Bollback (2002). This method uses the posterior probability distribution from a Bayesian analysis to simulate numerous data sets (the posterior predictive distribution). The simulated data sets are then compared to the observed data set (the one used in the Bayesian analysis). When the assumed model is adequate, the properties of the simulated data sets (e.g. the distribution of site patterns) will be congruent with the properties of the observed data set. When the assumed model is inadequate, however, the properties of the simulated data sets will not be congruent with those of the observed data set. In this case, the researcher should be very cautious when interpreting the results of a Bayesian analysis, and perhaps should continue to search for a more appropriate model with which to reanalyze the data set.

Numerous models have been developed that attempt to account for one or another complexity of sequence evolution, including: temporal variation in base frequencies (Lockhart et al., 1994; Galtier and Gouy, 1998), temporal variation in rates of evolution (Sanderson, 1997; Thorne et al., 1998; Huelsenbeck et al., 2000; Kishino et al., 2001), base-pairing interactions of RNA (Muse, 1995; Tillier and Collins, 1998), correlation between rates of adjacent sites (Felsenstein and Churchill, 1996), different rates of

synonymous and nonsynonymous substitutions (Goldman and Yang, 1994; Muse and Gaut, 1994), selection on protein coding regions (Halpern and Bruno, 1998; Nielsen and Yang, 1998), and insertions or deletions (McGuire et al., 2001). Consideration of these models in phylogenetic studies, however, has yet to become common practice. Three reasons likely contribute to the lack of use of these more complex models: 1) most of these models are not implemented in commonly used phylogenetic analysis packages (although recent versions of MrBayes have improved in this respect), 2) these models typically require much greater computational efforts, 3) commonly used model selection methods (such as hierarchical likelihood-ratio tests) are restricted to nested models, and 4) justification for the use of more complex models is still insufficient.

Given the well-demonstrated cost to model misspecification and the general reluctance of systematists to consider the use of more complex models, how should we proceed? First of all, the adequacy of available models should to be assessed on a broad-scale using real data sets. Although the intuition of many systematists suggests that our models are inadequate, no large-scale test of model adequacy has been performed. If we discover that the current set of available models is inadequate with respect to most real data sets, then more research should be conducted to identify the properties of sequence evolution that are inappropriately being ignored, and models should be developed to account for those complexities. Second, we should develop computationally feasible and broadly applicable (i.e. not restricted to nested models) methods of model choice and subsequently test the absolute and relative performance of these methods. Determining the absolute and relative performance of model choice methods is especially important

since different methods are likely to disagree on which model is most appropriate. For example, the hierarchical likelihood-ratio test and Akaike information criterion for model choice only chose the same model approximately 25% of the time when several hundred empirical data sets were analyzed (A. R. Lemmon, unpublished data). Third, enough models should be incorporated into phylogenetic analysis packages to assure that adequate models are available for most real data sets.

The goals of this study were to determine the affects of model misspecification on the estimation of phylogeny and substitution model parameters in the context of Bayesian phylogenetics. The results of this study are congruent with those from previous studies of model choice outside of the Bayesian context (Golding, 1983; Kuhner and Felsenstein, 1994; Yang et al., 1994; Sullivan et al., 1995; Lockhart et al., 1996), and therefore underscore the importance of proper model assumption. Given the bias that may result from under-parameterization and the imprecision that may result from over-parameterization, we strongly caution researchers to refrain from choosing models haphazardly (e.g. by assuming the most complex model that is computationally feasible or by assuming the model that happens to be the default in their favorite phylogenetic inference package). Careful consideration of the caveats resulting from studies of the importance of proper model choice will enable systematists to have greater confidence in their choice of models and estimates of phylogeny.

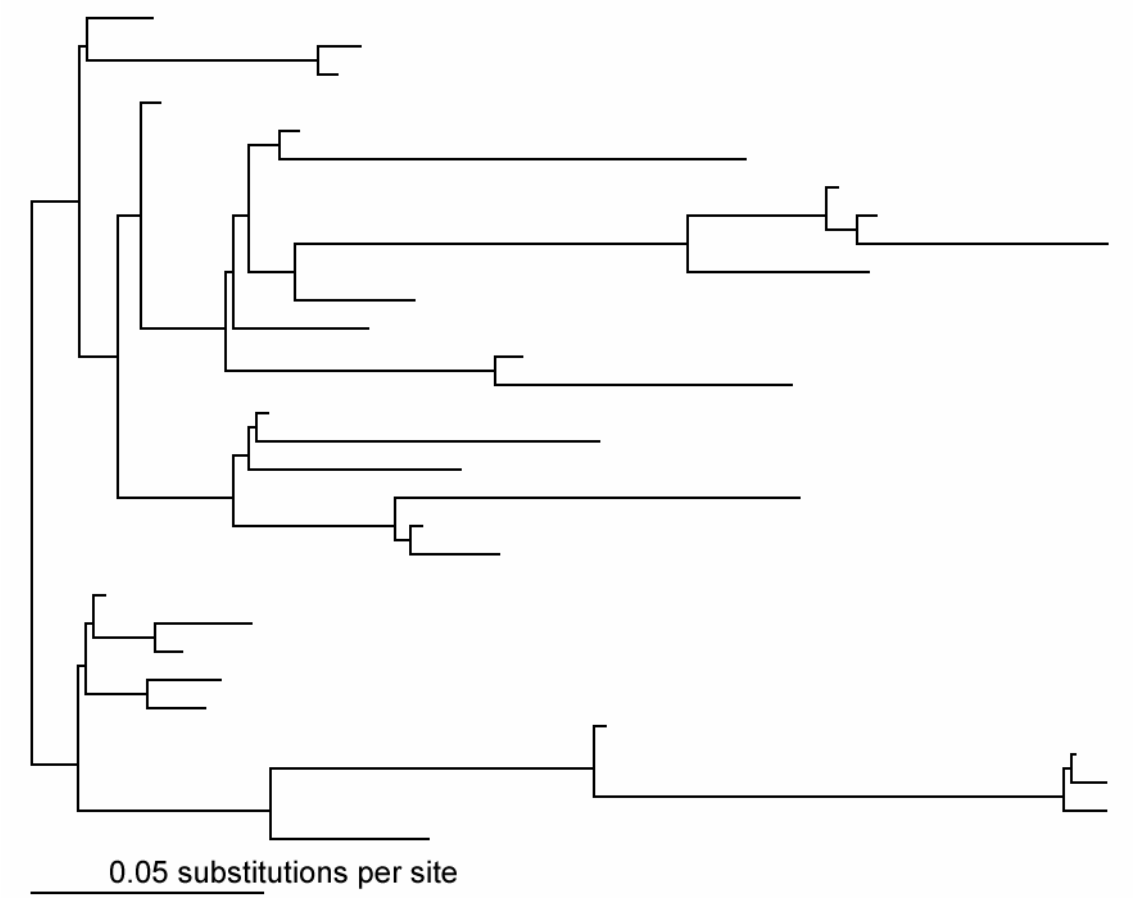

Figure 1.1. Tree 1, the tree used to simulate data sets for the first set of analyses. The topology for this tree was generated using a birth-death process. The branch lengths were assigned randomly (see text for details).

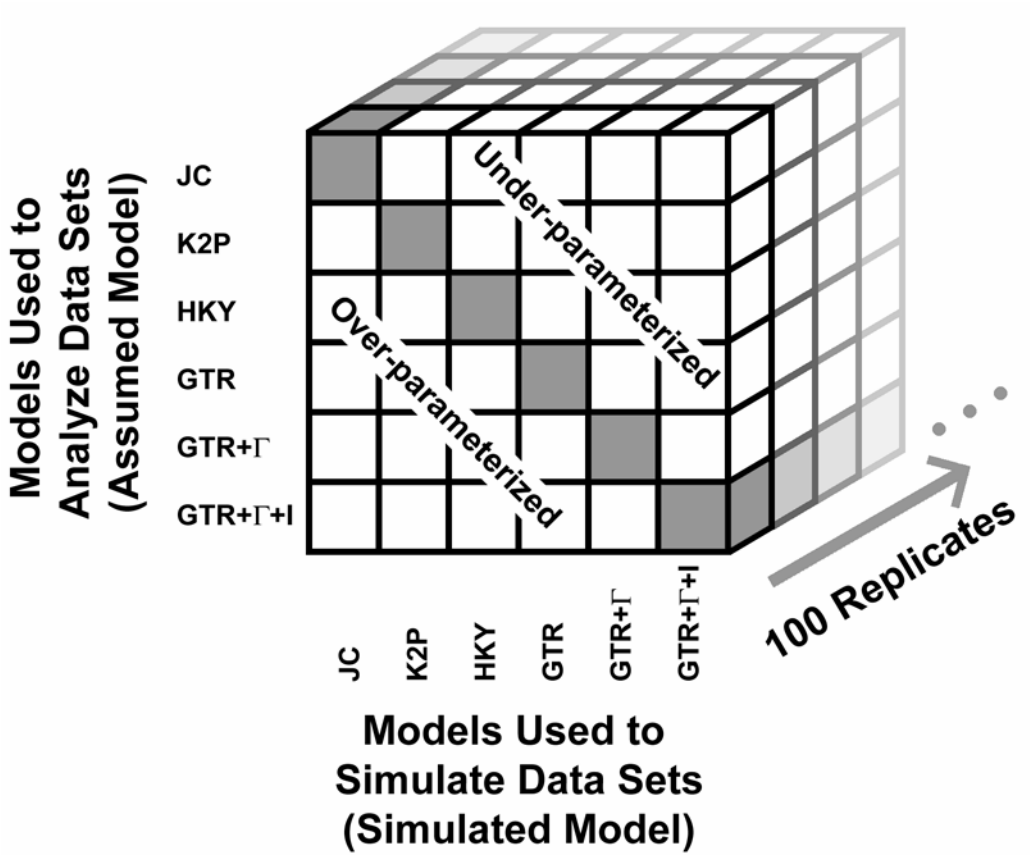

Figure 1.2. Study design of 36 model combinations. Shaded squares represent the model combinations in which the assumed model matches the simulated model. Model combinations above the diagonal contain an assumed model missing one or more parameters of the simulated model. Model combinations below the diagonal contain an assumed model including one or more parameters not present in the simulated model.

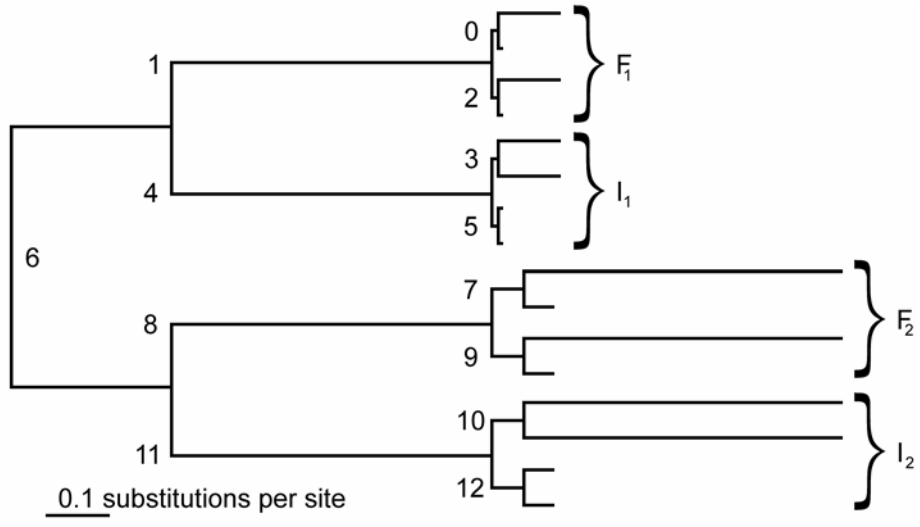

Figure 1.3. Tree 2, the tree used to simulate data sets for the second set of analyses. This tree contains two Felsenstein structures (F_1 and F_2) and two inverse-Felsenstein structures (I_1 and I_2). Internal nodes are labeled for reference in the Results section.

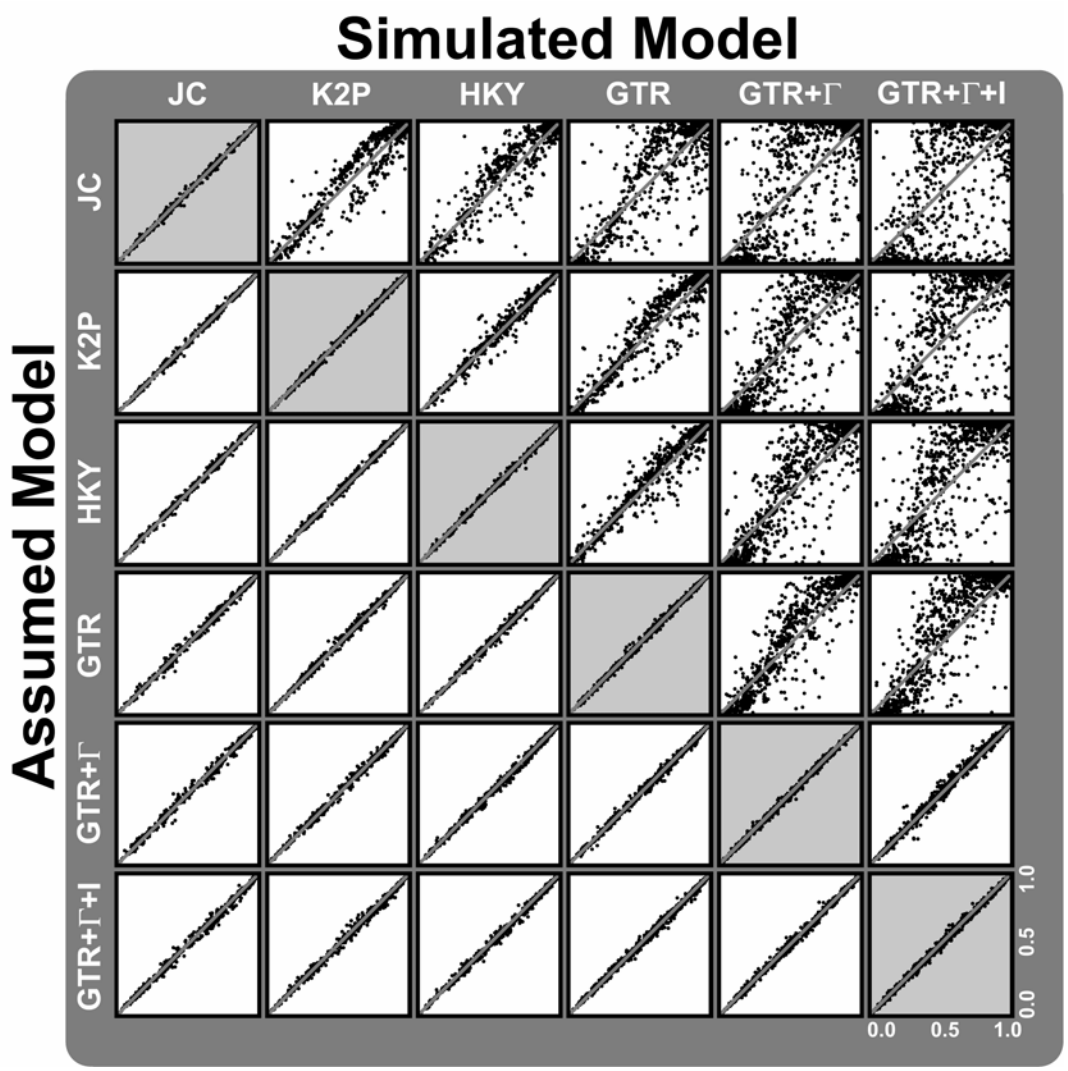

Figure 1.4. The effect of model misspecification on bipartition posterior probability estimates. The six graphs on the shaded diagonal demonstrate the convergence of the bipartition posterior probabilities for 100 pairs of independent runs when the correct model is assumed. Each of the 30 unshaded graphs compare the bipartition posterior probabilities obtained when the correct model was assumed (plotted on the x axis) with those obtained when an incorrect model was assumed (plotted on the y axis). The posterior probabilities of all 27 true bipartitions are plotted on the same graph for all 100 replicates, yielding 2700 points per graph. To determine the effect of model misspecification involving a single type of parameter, compare a graph in the shaded diagonal with a graph either directly above (under-parameterized) or directly below (over-parameterized). Note that only the bipartitions found in the true tree (Tree 1) are represented.

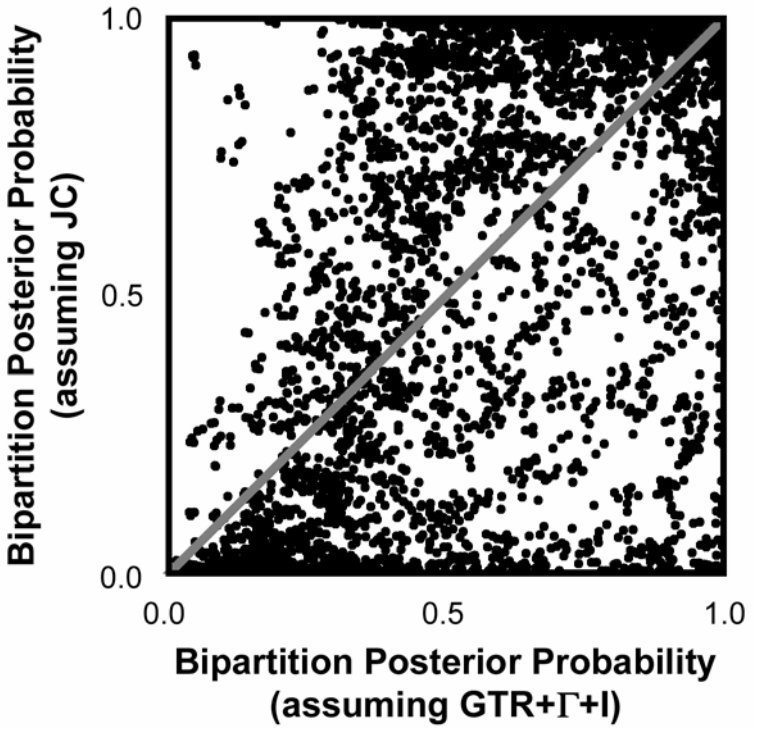

Figure 1.5. Graph including 400 additional replicates for the model combination GTR+Γ+I-JC. Compare this graph to the graph in the upper right-hand corner of Figure 1.4.

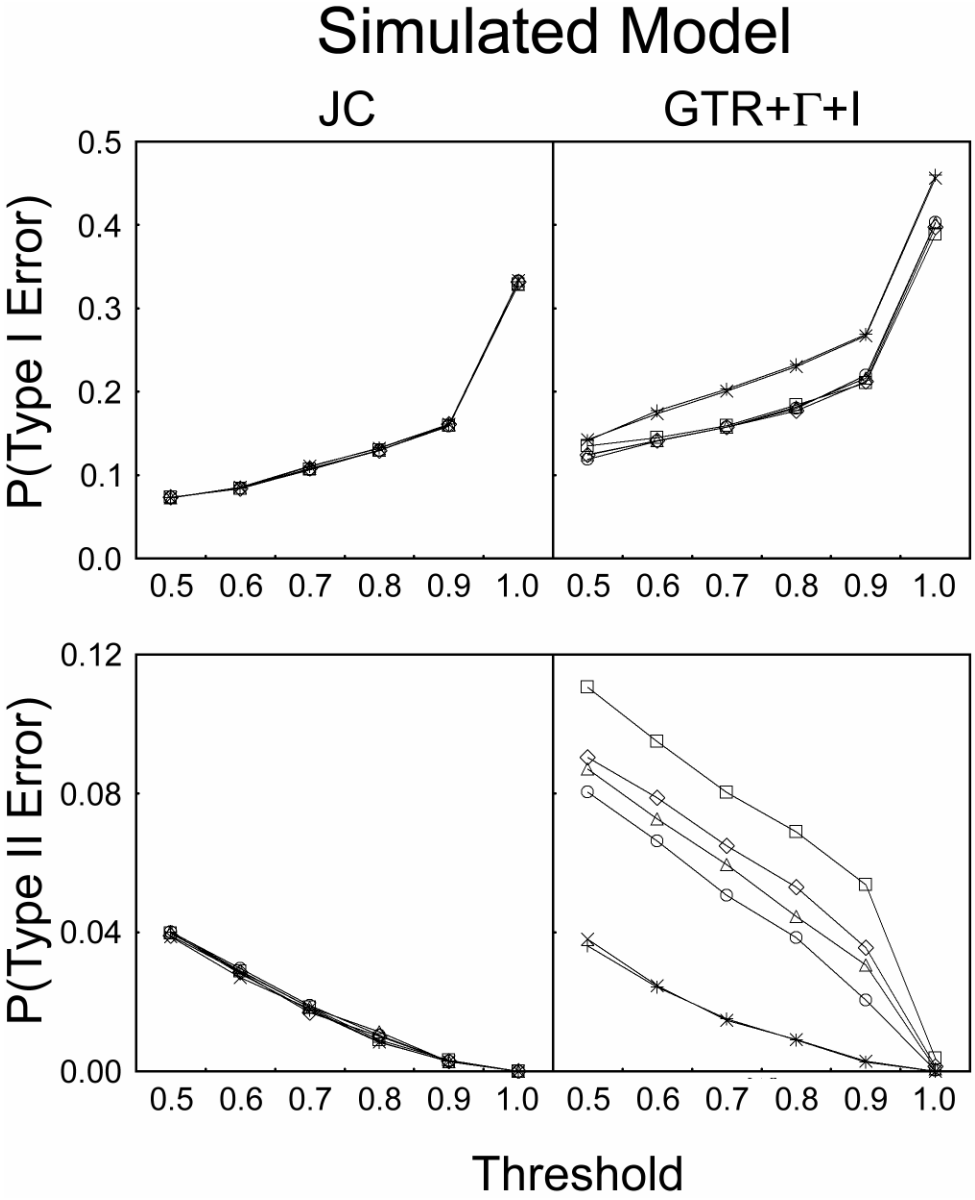

Figure 1.6. The effect of model misspecification on Type I and Type II error rates for hypothesis tests using bipartition posterior probabilities. The threshold is the posterior probability below which a particular bipartition is rejected. Type I error was calculated as the proportion of true bipartitions observed that were rejected based on their posterior probability. Conversely, Type II error was calculated as the proportion of false bipartitions observed that were accepted as true. The assumed model is represented by squares (JC), diamonds (K2P), triangles (HKY), circles (GTR), exes (GTR+Γ), and pluses (GTR+Γ+I). Each point represents the average across 100 replicates for the model combination depicted.

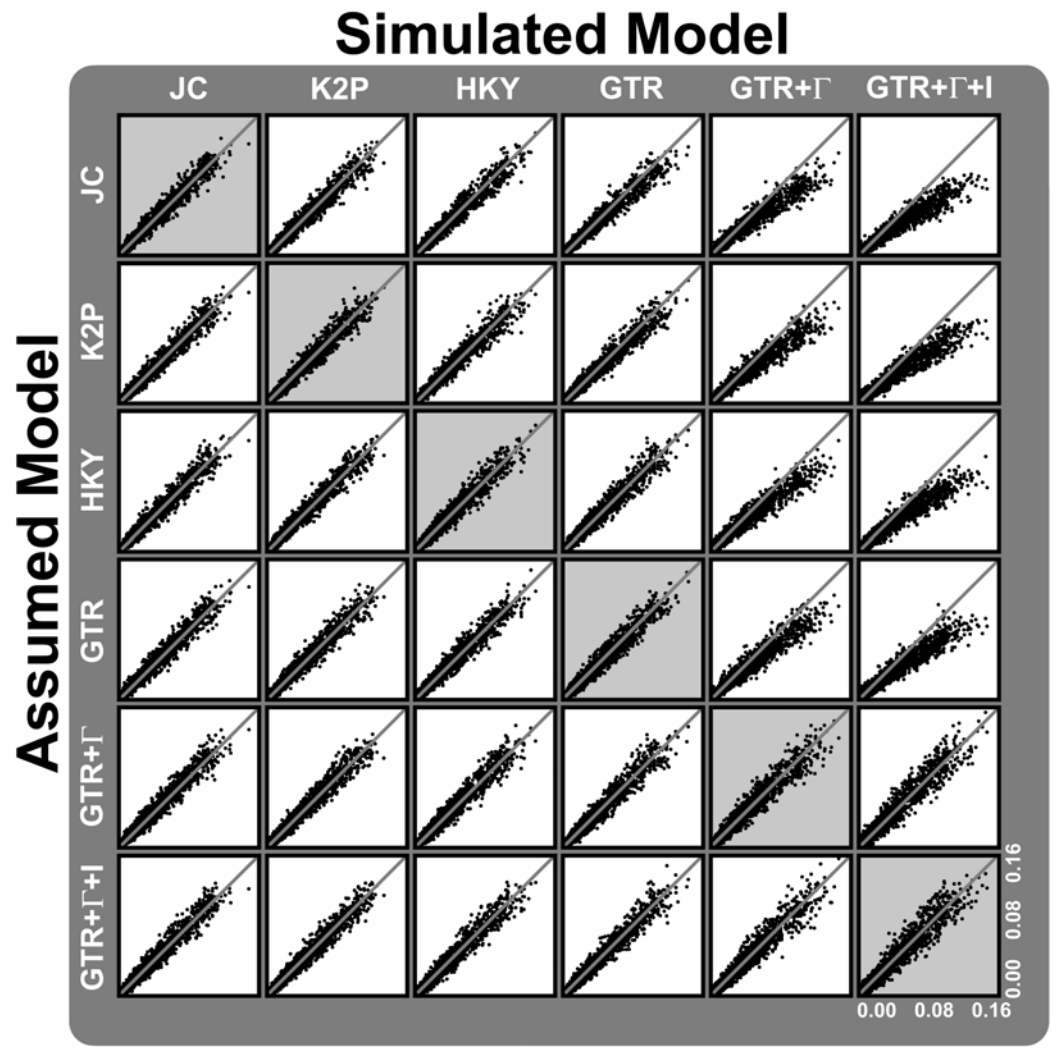

Figure 1.7. The effect of model misspecification on branch-length estimates. The format for this figure is the same as in Figure 1.4, except that the values plotted are the maximum likelihood estimates of the branch lengths (when the maximum likelihood tree did not contain a particular true internal branch, no value is plotted). Note that only interior branches found in the true tree are represented here. Plots of external branches demonstrate very similar results.

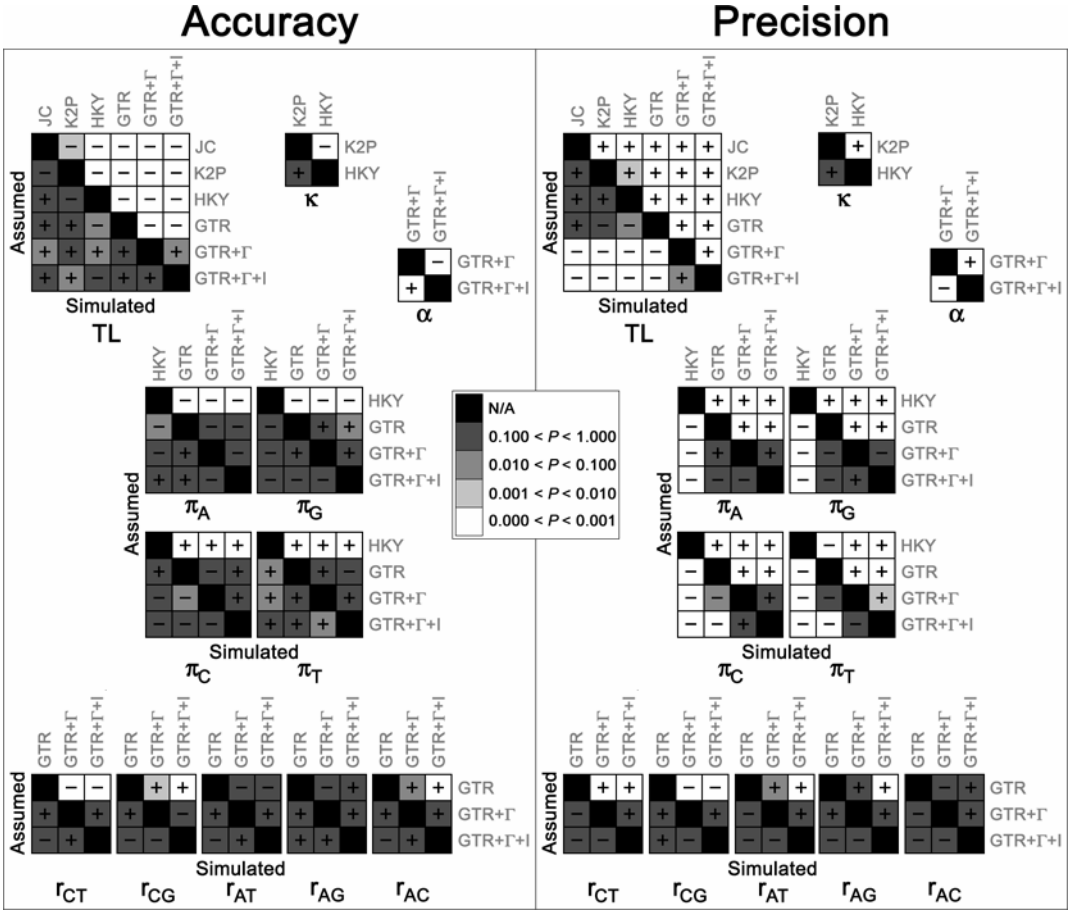

Figure 1.8. The effect of model misspecification on the accuracy and precision of parameter estimates. In the left panel, we compare the maximum likelihood estimate of each parameter obtained assuming the correct model with the estimate obtained assuming an incorrect model (see Methods for details). A plus (+) indicates that model misspecification produced a positive bias, whereas a minus (-) indicates a negative bias. In the right panel, we compare the width of the 95% credible set of each parameter obtained assuming the correct model with the width obtained assuming an incorrect model. A plus (+) indicates that model misspecification increased precision of the parameter estimate, whereas a minus (-) indicates that model misspecification decreased precision. The boxes are shaded according to the *P*-value obtained from a two-tailed paired-sample *t* test across 100 replicates.

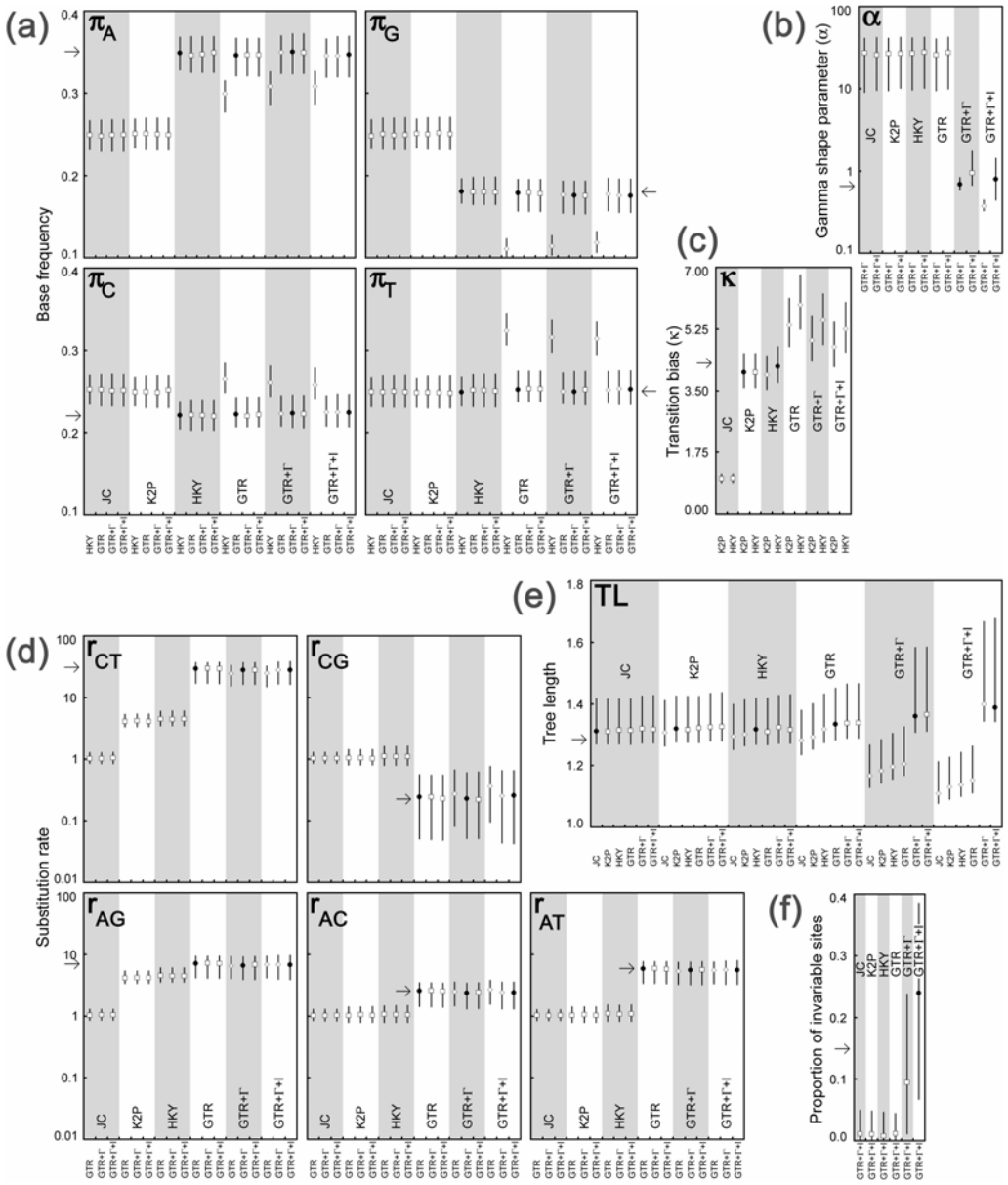

Figure 1.9. The effect of model misspecification on estimates of (a) base frequencies, (b) the gamma shape parameter (α), (c) the transition bias (κ), (d) substitution rates, (e) tree length, and (f) the proportion of invariable sites. The assumed model is labeled on the x-axis and the simulated model is labeled in the alternating panels; each panel groups the model combinations that share the same simulated model. The maximum likelihood estimate (averaged across 100 replicates) is plotted using either solid circles (●), open squares (□), or open diamonds (◇), representing model combinations in which the assumed model is either correct, over-parameterized, or under-parameterized, respectively. Vertical bars represent the range of the 95% credible set, also averaged across the 100 replicates. The effect of model misspecification can be observed by comparing points represented by solid circles with other points within the same panel. Arrows indicate true parameter values (but are only pertinent to those model combinations where the simulated model includes the parameter of interest).

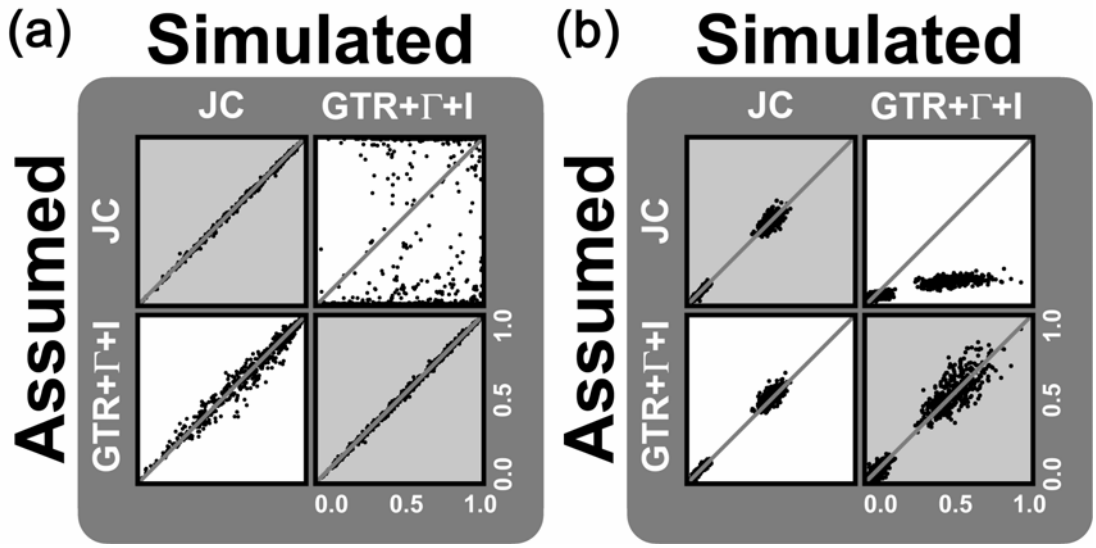

Figure 1.10. The effect of model misspecification on (a) bipartition posterior probability and (b) branch-length estimates for the second set analyses, which focus on Felsenstein and inverse-Felsenstein structures. The formats for (a) and (b) are the same as in Figures 1.4 and 1.7, respectively. Note the extreme effect of model under-parameterization on bipartition posterior probabilities and branch-length estimates.

References

Akaike, H. 1974. A new look at statistical model identification. IEEE Trans. Autom. Contr. 19:716-723.

Alfaro, M. E., S. Zoller, and F. Lutzoni. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Molecular Biology and Evolution 20:255-266.

Amenta, N., and J. Klinger. 2002. Case Study: Visualizing sets of evolutionary trees. IEEE Information Visualization. Pages 71-74.

Bollback, J. P. 2002. Bayesian model adequacy and choice in phylogenetics. Molecular Biology and Evolution 19:1171-1180.

Cummings, M. P., S. A. Handley, D. S. Myers, D. L. Reed, A. Rokas, and K. Winka. 2003. Comparing bootstrap and posterior probability values in the four-taxon case. Systematic Biology 52:477-487.

Cunningham, C. W., H. Zhu, and D. M. Hillis. 1998. Best-fit maximum-likelihood models for phylogenetic inference: Empirical tests with known phylogenies. Evolution 52:978-987.

Douady, C. J., F. Delsuc, Y. Boucher, W. F. Doolittle, and E. J. P. Douzery. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. Molecular Biology and Evolution 20:248-254.

Erixon, P., B. Svennblad, T. Britton, and B. Oxelman. 2003. The reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Systematic Biology 52:665-673.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Systematic Zoology 27:401-410.

Felsenstein, J., and G. A. Churchill. 1996. A hidden Markov model approach to variation among sites in rate of evolution. Molecular Biology and Evolution 13:93–104.

Galtier, N., and M. Gouy. 1998. Inferring pattern and process: Maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. Molecular Biology and Evolution 15:871–879.

Gaut, B. S., and P. O. Lewis. 1995. Success of maximum likelihood phylogeny inference in the four-taxon case. Molecular Biology and Evolution 12:152–162.

Golding, G. B. 1983. Estimates of DNA and protein sequence divergence: An examination of some assumptions. Molecular Biology and Evolution 1: 125-142.

Goldman, N. 1993. Statistical tests of models of DNA substitution. Journal of Molecular Evolution 36:182-198.

Goldman, N, and Z. H. Yang. 1994. Codon-based model of nucleotide substitution for protein-coding DNA sequences. Molecular Biology and Evolution 11:725-736.

Gu, X., Y-X. Fu, and W-H. Li. 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. Molecular Biology and Evolution 12: 546-557.

Halpern, A. L., and W. J. Bruno. 1998. Evolutionary distances for protein-coding sequences: Modeling site-specific residue frequencies. Molecular Biology and Evolution 15: 910-917.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22:160–174.

Huelsenbeck, J. P. 2001. MrBayes 3.0b3: Bayesian inference of phylogeny. Distributed by the author. Department of Biology, Univ. Rochester, Rochester, New York.

Huelsenbeck, J. P., and D. M. Hillis. 1993. Success of phylogenetic methods in the four-taxon case. Systematic Biology 42:247–264.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Huelsenbeck, J. P., B. Larget, and D. Swofford. 2000. A compound Poisson process for relaxing the molecular clock. Genetics 154:1879-1892.

Huelsenbeck, J. P., B. Larget, R. E. Miller, and F. Ronquist. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Systematic Biology 51:673-688.

Huelsenbeck, J. P., F. Ronquist, R. Nielsen, and J. P. Bollback. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294:2310-2314.

Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. In: *Mammalian Protein Metabolism* (H. N. Munro, ed.), pp. 21–132. Academic Press, New York.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16:111-120.

Kishino, H., J. L. Thorne, and W. J. Bruno. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Molecular Biology and Evolution 18:352-361.

Kuhner, M. K., and J. Felsenstein. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Molecular Biology and Evolution 11:459–468.

Lanave, C., G. Preparata, C. Saccone, and G. Serio. 1984. A new method for calculating evolutionary substitution rates. Journal of Molecular Evolution 20:86–93.

Lemmon, A. R., and M. C. Milinkovitch. 2002. The metapopulation genetic algorithm: an efficient solution for the problem of large phylogeny estimation. Proceedings of the National Academy of Sciences U.S.A. 99:10516-10521.

Lockhart, P. J., A. W. D. Larkum, M. A. Steel, P. J. Waddell, and D. Penny. 1996. Evolution of chlorophyll and bacteriochlorophyll: The problem of invariant sites in sequence analysis. Proceedings of the National Academy of Sciences U.S.A. 93:1930–1934.

Lockhart, P. J., M. A. Steel, M. D. Hendy, and D. Penny. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Molecular Biology and Evolution 11:605-612.

Maddison, W. P., and D. R. Maddison. 2003. Mesquite: a modular system for evolutionary analysis. Version 0.994. http://mesquiteproject.org.

McGuire, G., M. C. Denham, and D. J. Balding. 2001. Models of sequence evolution for DNA sequences containing gaps. Molecular Biology and Evolution 18:481-490.

Minin, V., Z. Abdo, P. Joyce, and J. Sullivan. 2003. Performance-based selection of likelihood models for phylogeny estimation. Systematic Biology 52:674-683.

Moriarty, E. C., and D. C. Cannatella. 2004. Phylogenetic relationships of North American chorus frogs (*Pseudacris*). Molecular Phylogenetics and Evolution 30: 409-420.

Muse, S. V. 1995. Evolutionary analyses of DNA sequences subject to constraints on secondary structure. Genetics 139:1429-1439.

Muse, S. V., and B. S. Gaut. 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Molecular Biology and Evolution 11:715-724.

Nielsen, R., and Z. Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929-936.

Posada, D., and K. A. Crandall. 2001. Selecting the best-fit model of nucleotide substitution. Systematic Biology 50:580-601.

Rambaut, A., and N. C. Grassly. 1997. Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Computational Applied Biosciences 13:235-238.

Robinson, D. F., and L. R. Foulds. 1981. Comparison of phylogenetic trees. Mathematical Biosciences 53:131-147.

Rodríguez, F., J. L. Oliver, A. Marín, and J. R. Medina. 1990. The general stochastic model of nucleotide substitution. Journal of Theoretical Biology 142:485–501.

Sanderson, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Molecular Biology and Evolution 14:1218-1231.

Schwarz, G. 1974. Estimating the dimension of a model. Annuals of Statistics 6:461-464.

Steel, M., P. J. L. Székely, P. L. Erdös, and P. J. Waddell. 1993. A complete family of phylogenetic invariants for any number of taxa under Kimura's 3ST model. New Zealand Journal of Botany 31:289-296.

Sullivan, J., K. E. Holsinger, and C. Simon. 1995. Among-site rate variation and phylogenetic analysis of 12S rRNA in sigmodontine rodents. Molecular Biology and Evolution 12:988-1001.

Sullivan, J., and D. L. Swofford. 2001. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated? Systematic Biology 50:723-729.

Suzuki, Y., G. V. Glazko, and M. Nei. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proceedings of the National Academy of Sciences U.S.A. 99:16138-16143.

Swofford, D. L., P. J. Waddell, J. P. Huelsenbeck, P. G. Foster, P. O. Lewis, and J. S. Rogers. 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. Systematic Biology 50:525-539.

Tavaré, S. 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Lectures in Mathematical Life Science 17:57–86.

Thorne, J. L., H. Kishino, I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. Molecular Biology and Evolution 15:1647–1657.

Tillier, E. R. M., and R. A. Collins. 1998. High apparent rate of simultaneous compensatory base-pair substitutions in ribosomal RNA. Genetics. 148:1993-2002.

Waddell, P., and D. Penny. 1996. Evolutionary trees of apes and humans from DNA sequences. In*: Handbook of Symbolic Evolution* (A. J. Lock & C. R. Peters, eds.), pp. 53-73. Clarendon Press, Oxford.

Wilcox, T. P., D. J. Zwickl, T. A. Heath, and D. M. Hillis. 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. Molecular Phylogenetics and Evolution 25:361-371.

Yang, Z. 1993. Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Molecular Biology and Evolution 10:1396-1401.

Yang, Z., N. Goldman, and A. Friday. 1994. Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. Molecular Biology and Evolution 11:316-324.

Zar, J. H. 1999. *Biostatistical Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey.

Chapter 2

Reproductive Character Displacement Is Not the Only Possible Outcome of Reinforcement*

Abstract. We study the form of the clines in a female mating preference and male display trait using simulations of a hybrid zone. Allopatric populations of two species are connected by demes in a stepping stone arrangement. Results show that reproductive character displacement (a pattern of increased prezygotic isolation in sympatry compared to allopatry) may or may not result when there is reinforcement (defined here as the strengthening of prezygotic isolation as a result of selection against hybrids, relative to the amount of prezygotic isolation present when hybrids are not selected against). Further, reproductive character displacement of the preference may or may not occur when it occurs in the male display. We conclude that the absence of reproductive character displacement is not evidence against the operation of reinforcement.

*Significant portions of this chapter have been previously published as Lemmon, Smadja, & Kirkpatrick, 2004. Journal of Evolutionary Biology 17:177–183.

2.1 INTRODUCTION

Secondary contact between partially reproductively isolated populations may have several possible outcomes: extinction of one of the two populations, stable coexistence with hybridization, fusion of the two populations, or an increase in premating divergence and formation of distinct species (Barton & Hewitt, 1981; Liou & Price, 1994). In 1940, Dobzhansky postulated that if two divergent populations produce hybrids of low fitness where they come into contact, natural selection will enhance premating isolation (Dobzhansky, 1940; Howard, 1993). This process was termed reinforcement by Blair (1955).

Testing the plausibility of reinforcement has been a challenge for theoretical researchers. Reinforcement was long considered to be controversial because it was thought that very strong selection was required to compensate for the negative effect of recombination and gene flow (Paterson, 1978; Spencer et al., 1986). More recently, however, an increasing number of new models have moved the hypothesis of reinforcement into the foreground of speciation research by demonstrating the plausibility of this process under more realistic conditions (for a review, see Turelli et al., 2001 and Kirkpatrick & Ravigné, 2002).

Demonstrating reinforcement in nature is not trivial. As a result, empirical researchers have focused on one pattern that may result from reinforcement: reproductive character displacement (RCD), which is a pattern of greater divergence of an isolating trait in areas of sympatry between closely related taxa than in areas of

allopatry (Brown & Wilson 1956; Howard 1993). Howard emphasizes the distinction between reinforcement as a *process* and RCD as one potential *pattern* that can result from this process. With the difficulty of demonstrating reinforcement in mind, Howard focuses on RCD while omitting discussion of other patterns that may result from reinforcement.

The search for reinforcement in nature has become confused for five reasons. The first source of confusion stems from arguments regarding whether the term reinforcement should be used in cases where postzygotic isolation is already complete. Butlin (1987a, b), for example, has argued that the evolution of prezygotic isolation in this case should be termed reproductive character displacement. His reason is that *speciation* by reinforcement cannot occur because complete postzygotic isolation implies that the two entities are already distinct species. As Howard (1993) points out, however, the process of reinforcement itself may still occur since selection against hybridization can still lead to the evolution of increased prezygotic isolation. We agree with Howard's argument that the processes of speciation and reinforcement can operate independently and therefore use of the term reinforcement is appropriate when postzygotic isolation is already complete. The second reason for confusion is the failure to recognize that RCD can result from processes other than reinforcement, for example when there is interference between the mate recognition signals of taxa that do not hybridize. Howard (1993) addresses this issue and outlines the evidence needed to demonstrate that a pattern of RCD observed has resulted from the process of reinforcement.

The third reason for confusion is the failure to recognize that reinforcement can lead to patterns other than RCD. This confusion typically manifests itself in the false idea that in order for reinforcement to occur, prezygotic isolation in sympatry must be strengthened beyond the degree of prezygotic isolation seen in allopatry (Noor, 1999). We argue here that the evolutionary process by which prezygotic isolating barriers are strengthened by selection against hybridization is the same regardless of whether RCD is the resulting pattern. We therefore define reinforcement broadly as the strengthening of prezygotic isolation as a result of selection against hybrids, relative to the amount of prezygotic isolation present when hybrids are not selected against. By our definition, therefore, reinforcement can occur even when prezygotic isolation in sympatry is not strengthened beyond the degree of prezygotic isolation seen in allopatry. The reason is that selection against hybrids can result in fewer heterospecific matings (and thus increased prezygotic isolation) regardless of whether or not RCD is present.

A fourth reason for confusion is the failure to recognize that patterns of divergence in the female mating preference may differ from those in the male display trait. Finally, some workers have failed to recognize the role that sexual selection can play in generating selection against hybrids. If hybrids have decreased mating opportunities or decreased fertility, then reinforcement may occur even in the absence of viability selection against hybrids (Coyne & Orr, 1989; Liou & Price, 1994; Kirkpatrick & Servedio, 1999; Noor, 1999).

Here we present a simple model of a hybrid zone with the aim of clarifying two of these causes of confusion. We use simulations to make the following points: that reinforcement can lead to patterns other than RCD, and that the male display trait and female preference do not always evolve the same pattern of divergence. We wish to caution empiricists not to draw conclusions about the influence of selection against hybrids on prezygotic isolation based solely on presence or absence or RCD. We conclude with a discussion of some possible disparities among researchers in the interpretation of the definition of reinforcement and some problems that may arise as a result of this disparity.

2.2 MODEL

We begin by assuming a simple genetic system of three diploid loci with free recombination. The three loci are a male trait locus (T), a female preference locus (P), and a hybrid incompatibility locus (I). We denote the two alleles at the male trait as T_0 and T_1 (with frequencies t_0 and t_1), those at the preference locus as P_0 and P_1 (with frequencies p_0 and p_1), and those at the hybrid incompatibility locus as I_0 and I_1 (with frequencies i_0 and i_1). Allopatric divergence at the trait locus, denoted ΔT, is defined to be the difference between t_0 in the two allopatric populations. Allopatric divergence at the preference locus, ΔP, and incompatibility locus, ΔI, are defined in a similar fashion.

The hybrid zone is represented by ten populations arranged in a stepping stone model, with the left and right-most populations being identified as allopatric, and all others as sympatric. The order of events in each generation is migration, natural selection, and sexual selection (mating). Migration is one-way from allopatry to sympatry but two-way in sympatry. The rate of migration between adjacent populations is denoted *m*. The effective size of each population is infinite.

We assume that the trait locus is under both natural and sexual selection. A linear environmental gradient causes viability selection that favors allele T_1 in the left allopatric population and T_0 in the right allopatric population. The relative viabilities of the genotypes T_0T_0, T_0T_1, and T_1T_1 in the left allopatric population are $1 - s_T$, $1 - s_T/2$, and 1, respectively, while in the right allopatric population the viabilities of the T_0T_0 and T_1T_1 males are reversed. Viabilities of the three genotypes change in a linear fashion as one moves across the hybrid zone.

Females choose their mates based on their genotype at the preference locus P, with the P_0 allele conferring a preference for males that carry the T_0 allele and the P_1 allele a preference for the T_1 allele. The frequency of matings between different male and female genotypes is proportional to the product of the frequencies of those genotypes and the preference that the female has for the male. Table 2.1 shows these preferences.

We assume conditions that produce a stable polymorphism in the preference and trait loci in the allopatric populations, thus permitting the preference and

trait to evolve to more extreme values in sympatry (that is, the pattern of RCD). Over-dominant natural selection acts on the female preference locus with viabilities in the left allopatric population of 1/2, 1, and $1 - s_P$, for the genotypes P_0P_0, P_0P_1, and P_1P_1, respectively. Conversely, in the right allopatric population the viabilities of the P_0P_0 and P_1P_1 females are reversed. We chose the value for s_P in order to give the desired value of ΔP. Natural selection does not act on the female preference locus in sympatry.

Selection against hybrids provides the force that may lead to reinforcement. Here we assume that viability selection against hybrids is disruptive, acting against heterozygotes at the incompatibility locus. The relative viabilities of the genotypes I_0I_0, I_0I_1, and I_1I_1 are respectively 1, $1 - s_I$, and 1 in all populations, where s_I is the strength of viability selection against hybrids. The left allopatric population is fixed for allele I_0, and the right for allele I_1.

Using simulations, we track genotype frequencies through time. At the beginning of each simulation, allopatric populations are brought to equilibrium. We then allow secondary contact by allowing migration into the sympatric populations. We ran the simulations until all populations reached equilibrium.

2.3 QUANTIFYING CLINE SHAPE

Figure 2.1 gives the interpretation of our measures of RCD and reinforcement. To quantify the cline shape, we noted the difference between t_0 in first sympatric

population on the left and the adjacent allopatric population. This quantity, which we term b_T, indicates if RCD is present. A positive value of b_T means there is RCD, that is, the male trait has greater divergence in sympatry than in allopatry (a "reversed" cline). Furthermore, we can study how the intensity of selection against hybrids affects b_T to ask if increasing selection against hybrids leads to increasing premating isolation and thus reinforcement. We can ask similar questions about the evolution of the preference using the analogous quantity b_P.

To quantify the amount of reinforcement, we compare the shape of the cline under two conditions: when selection acts against hybrids ($s_I > 0$) and when selection does not act against hybrids ($s_I = 0$). This quantity, which we define as $\delta = b_{s>0} - b_{s=0}$ (the slope observed when selection is acting against hybrids minus the slope observed when selection is not acting against hybrids), tells us the effect of selection against hybrids on the shape of the cline. When prezygotic isolation is strengthened by selection against hybrids (reinforcement), we expect δ to be positive for either the male trait or female preference.

2.4 RESULTS

A first basic observation from the simulations is that reinforcement does not always produce RCD. This result can be seen in Figure 2.2, which presents the evolution of a cline through time under two different conditions. With no selection acting against

hybrids (Generation 0), a simple cline forms. After selection against hybrids is introduced, reinforcement occurs (Generations 50 and 100) and the cline steepens. Under moderate selection against hybrids ($s_I = 0.65$), the resulting cline at equilibrium is monotonic. Under more intense selection against hybrids ($s_I = 0.90$), the resulting cline is reversed. We say that reinforcement has occurred in both cases because selection against hybrids resulted in increased in prezygotic isolation, relative to the amount of prezygotic isolation present when hybrids are not selected against (compare Equilibrium and Generation 0). RCD, conversely, is only evident in one of the two cases.

Figure 2.3 and Table 2.2 also show cases in which there is reinforcement (that is, increasing intensity of selection against hybrids leads to increased divergence of the preference and/or male trait in sympatry) but not RCD (that is, greater divergence in sympatry than allopatry, indicated by positive values of b_T and/or b_P). Under a broad range of conditions, viability selection against hybrids (s_I) leads to reinforcement of both the preference and male trait, indicated by positive values of δ_T and δ_P.

A second basic observation is that the clines in the preference and male trait can be qualitatively different in shape. Under some conditions, both the preference and male trait show RCD, while under others only the male trait does. Examples are shown in Figure 2.3. Although we typically observed monotonic and reversed clines, we also observed a third type of cline, in which the pattern of divergence in sympatry was the reverse of that in allopatry. Figure 2.4 shows examples of different patterns that can be produced by reinforcement. We also observed that the male trait consistently diverges

more than the preference in our model. That may not be a general pattern, however, as the quantitative outcome may depend on behavioral and genetic details.

The simulations also suggest the roles that different parameters may have in promoting RCD and reinforcement. Table 2.2 shows that the degree of RCD in the male trait and female preference is enhanced by increased female mating preferences (larger α) and by larger differences between the average female mating preferences of the allopatric populations (larger ΔP). Increased divergence of the trait in allopatry (ΔT), however, generally leads to a decrease in the amount of RCD observed for the male trait and female preference (smaller b_T and b_P). As might be anticipated, increased migration (m) leads to decreased reinforcement of both the male trait and preference (smaller δ_T and δ_P). Our simulations do not allow us to say how these conclusions might generalize to other assumptions regarding the genetics, geography, behavior, etc.

2.5 DISCUSSION

The model presented here demonstrates two simple points. First, the absence of RCD does not imply the absence of reinforcement. This result makes the point that RCD is not the unique signature of reinforcement. Second, the preference and male trait can have qualitatively different clines. This result underlines the importance of studying the patterns of divergence in both the trait and the preference. While the quantitative results

we find depend on the detailed assumptions made in our model, we expect that these two qualitative conclusions may hold under a broad range of conditions.

We have focused here on the concept of reinforcement originated by Dobzhansky (1940), defined by Blair (1955), and further clarified by Howard (1993). This definition describes the process that selection against hybrids may lead to: the enhancement of prezygotic isolation. We have shown that prezygotic isolation can be enhanced by selection against hybrids without resulting in reproductive character displacement (RCD), defined as the pattern in which there is more divergence in sympatry than allopatry (Brown & Wilson, 1956).

There may be a tendency, because of empirical convenience, to equate reinforcement with RCD. This has led to a conception of reinforcement as an increase in prezygotic isolation in sympatry *relative to allopatry*, due to selection against hybrids. Use of this definition leads researchers to view RCD as a necessary outcome of reinforcement. This definition also downplays the importance of the evolution of prezygotic isolation in the stages where RCD is not present. During these stages the frequency of heterospecific matings is reduced and fewer hybrids zygotes are formed just as they are when RCD is present. Our simulations show that RCD is just one possible outcome that can result from selection against hybrids. We see no difference between the process that steepens a monotonic cline and the process that transforms a monotonic cline into a reversed cline. In both cases the propensity to mate with heterospecifics evolves by indirect selection generated by direct selection on less fit hybrids.

As Howard (1993) and Noor (1999) point out, RCD may result from evolutionary processes other than reinforcement. It is therefore critical to distinguish between the evolutionary process that enhances isolation, which may or may not be reinforcement, and the pattern of the outcome, which may or may not be RCD. This is not to say that the study of RCD in nature is not important; we may indeed learn volumes about speciation by concentrating on geographic regions of enhanced prezygotic isolation. The point is just that we cannot draw conclusions about whether reinforcement has occurred based solely on the presence or absence of RCD.

How then can we determine if reinforcement has occurred in nature? Ideally, we might want to compare the hybrid zone of interest with another zone that is identical but that has no selection against hybrids. One could, for example, compare the proportion of heterospecific matings in sympatry under the two cases, expecting to see fewer where there is selection against hybrids. Alternatively, we could compare the slope of the cline under the two cases. We would expect greater divergence (that is, larger values of b_T and/or b_p) when selection is acting against hybrids than when selection is not. While these approaches are practical for simulation studies, they clearly are not feasible in nature.

Perhaps asking whether reinforcement has occurred is not the most informative question. Existing analytical models of reinforcement show that prezygotic isolation is strengthened by selection acting against hybrids under very general conditions (Kirkpatrick & Servedio, 1999; Kirkpatrick, 2000, 2001), and our findings are consistent

with that conclusion. These results suggest we can expect there will be some reinforcement whenever there is some assortative mating and some selection against hybrids. If our aim is to understand the role of postzygotic isolation in the formation and maintenance of species, then perhaps we should not be asking whether or not reinforcement is occurring, but instead be trying to understand what effect reinforcement has on the patterns of divergence we see in nature.

Table 2.1. The frequency of matings between different male and female genotypes.

		Male genotype		
		T_0T_0	**T_0T_1**	**T_1T_1**
Female genotype	**P_0P_0**	$(1 + \alpha)^2$	$(1 + \alpha)$	1
	P_0P_1	1	1	1
	P_1P_1	1	$(1 + \alpha)$	$(1 + \alpha)^2$

Table 2.2. Cline shape (*b*) and the amount of reinforcement (δ) under different evolutionary assumptions. Subscripts T and P correspond to the male trait and female preference loci, respectively. Parameters include the strength of female preference (α), rate of migration (*m*), amount of allopatric divergence in the male trait (Δ*T*) and female preference (Δ*P*), and strength of viability selection against hybrids (s_I). Each row presents the parameter values assumed and the outcome of one simulation, where a dash (-) indicates a parameter value that is equal to the value in the top row.

α	*m*	Δ*T*	Δ*P*	s_I	b_T	b_P	δ_T	δ_P
0.4	0.05	0.4	0.4	0.4	0.0364	-0.0279	0.0611	0.0153
0.2	-	-	-	-	-0.1437	-0.042	0.0183	0.0081
0.8	-	-	-	-	0.1643	0.0012	0.0589	0.0263
1	-	-	-	-	0.1869	0.0128	0.0531	0.0303
-	0.025	-	-	-	0.0939	-0.0217	0.1013	0.0202
-	0.05	-	-	-	0.0364	-0.0279	0.0611	0.0153
-	0.075	-	-	-	0.0153	-0.0295	0.0464	0.0142
-	0.1	-	-	-	0.005	-0.0299	0.0393	0.0141
-	-	0.2	-	-	0.0748	-0.0269	0.0614	0.0152
-	-	0.4	-	-	0.0364	-0.0279	0.0611	0.0153
-	-	0.6	-	-	-0.0011	-0.0289	0.0612	0.0154
-	-	0.8	-	-	-0.0364	-0.0298	0.0616	0.0156
-	-	-	0.2	-	-0.1942	-0.0335	0.0002	0.0017
-	-	-	0.4	-	0.0364	-0.0279	0.0611	0.0153
-	-	-	0.6	-	0.1368	-0.0384	0.0454	0.0191
0.4	0.05	0.4	0.4	0.4	0.0364	-0.0279	0.0611	0.0153
-	-	-	-	0	-0.0248	-0.0432	0	0
-	-	-	-	0.2	0.0021	-0.0363	0.0269	0.0069
-	-	-	-	0.6	0.094	-0.0098	0.1188	0.0334
-	-	-	-	0.8	0.1366	0.0157	0.1614	0.0589
-	-	-	-	1	0.1594	0.0435	0.1842	0.0867

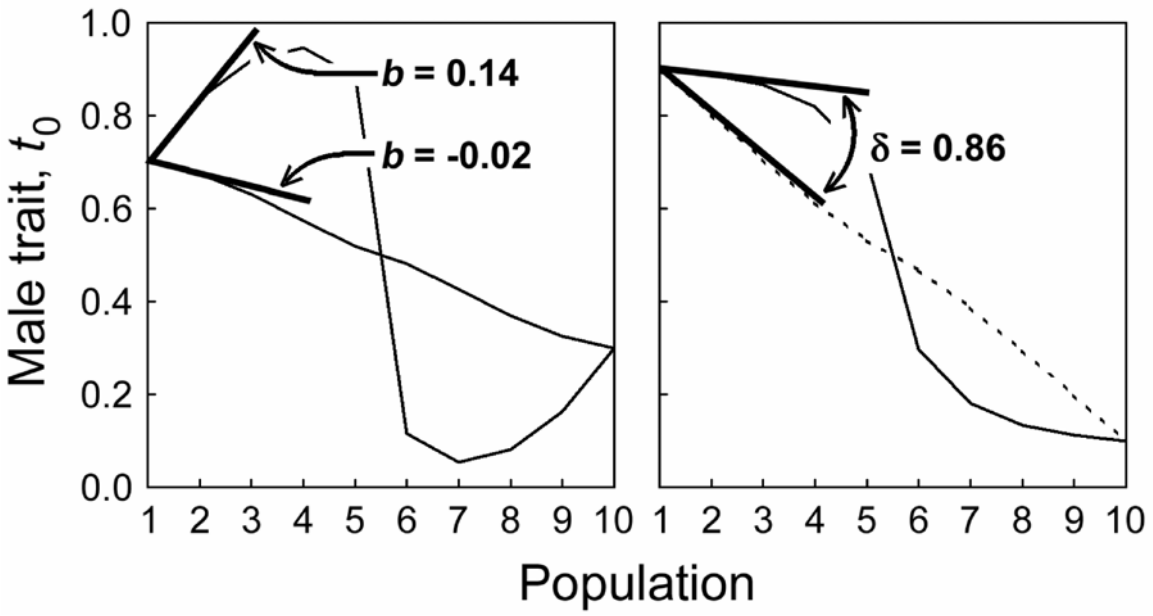

Figure 2.1. Measures of reproductive character displacement and reinforcement. The left panel shows the interpretations for b, our measure of cline shape, and the right panel shows the interpretations for δ, our measure of reinforcement. A dashed line (- -) indicates that no selection is acting against hybrids, whereas a solid line (–) indicates that hybrids have reduced fitness.

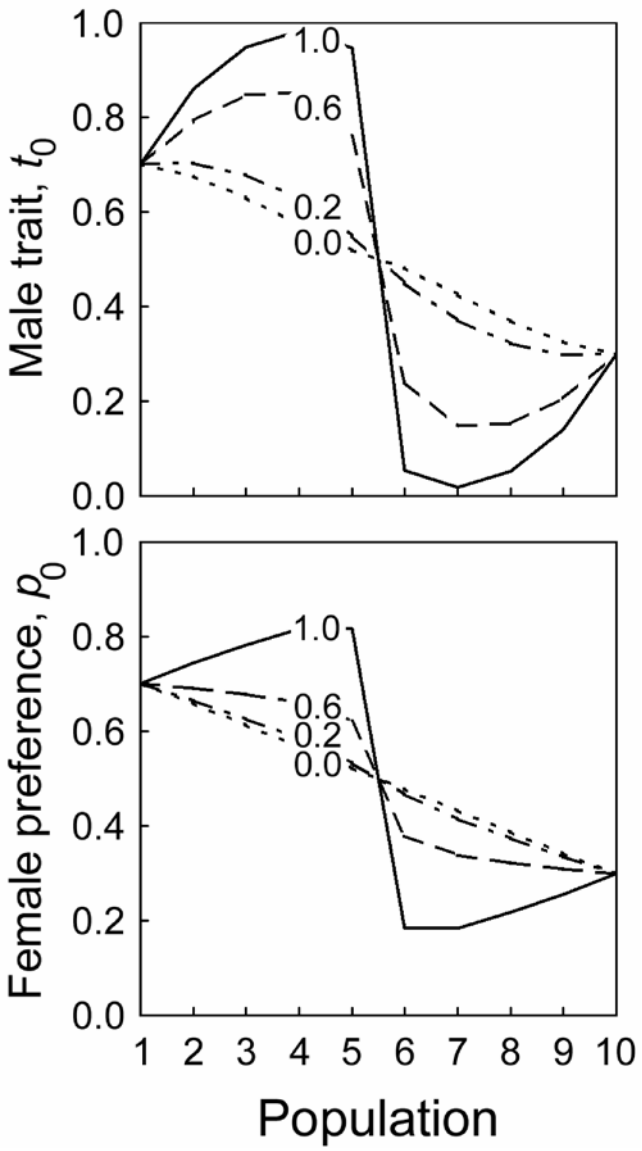

Figure 2.2. The evolution of a cline in response to selection against hybridization. (a) A simple cline in mating preferences forms when there is no selection against hybrids. (b & c) Selection against hybrids results in reinforcement of mating preferences in sympatry. (d) The resulting shape of the cline at equilibrium depends upon a number of conditions, including the strength of selection against hybrids (sI). Under some conditions (sI = 0.90), the resulting cline shows the pattern of reproductive character displacement whereas in others (sI = 0.65), the resulting cline is monotonic. The process leading to the strengthening of prezygotic isolation in sympatry is the same in both cases. Parameter values are $\alpha = 0.4$, $m = 0.05$, $\Delta T = 0.4$, and $\Delta P = 0.4$.

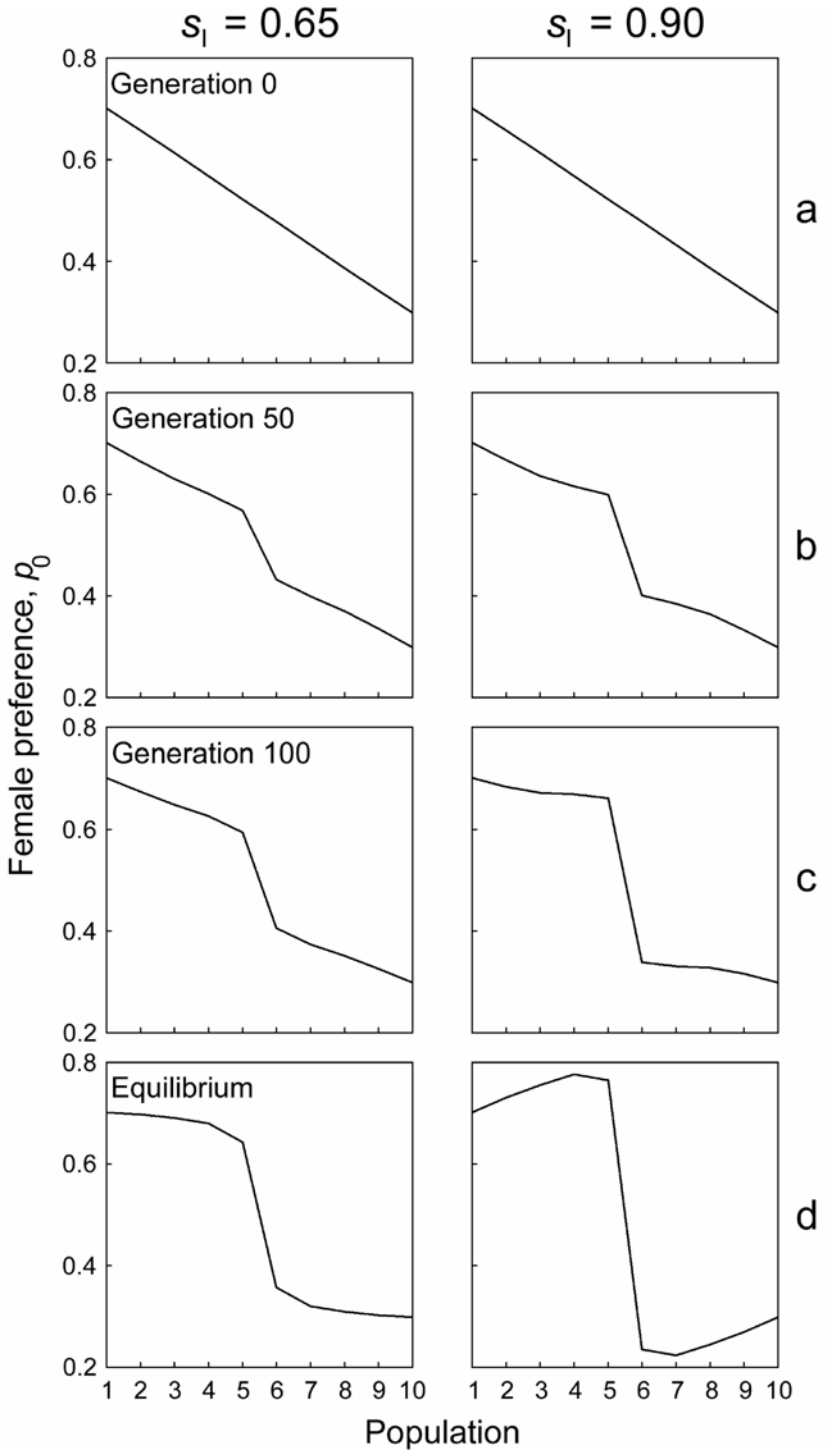

Figure 2.3. Patterns of divergence in the male trait (above) and female preference (below). The four clines in each panel correspond to different intensities of viability selection against hybrids (with the value for s_I shown on each curve). Parameter values are $\alpha = 0.4$, $m = 0.05$, $\Delta T = 0.4$, and $\Delta P = 0.4$.

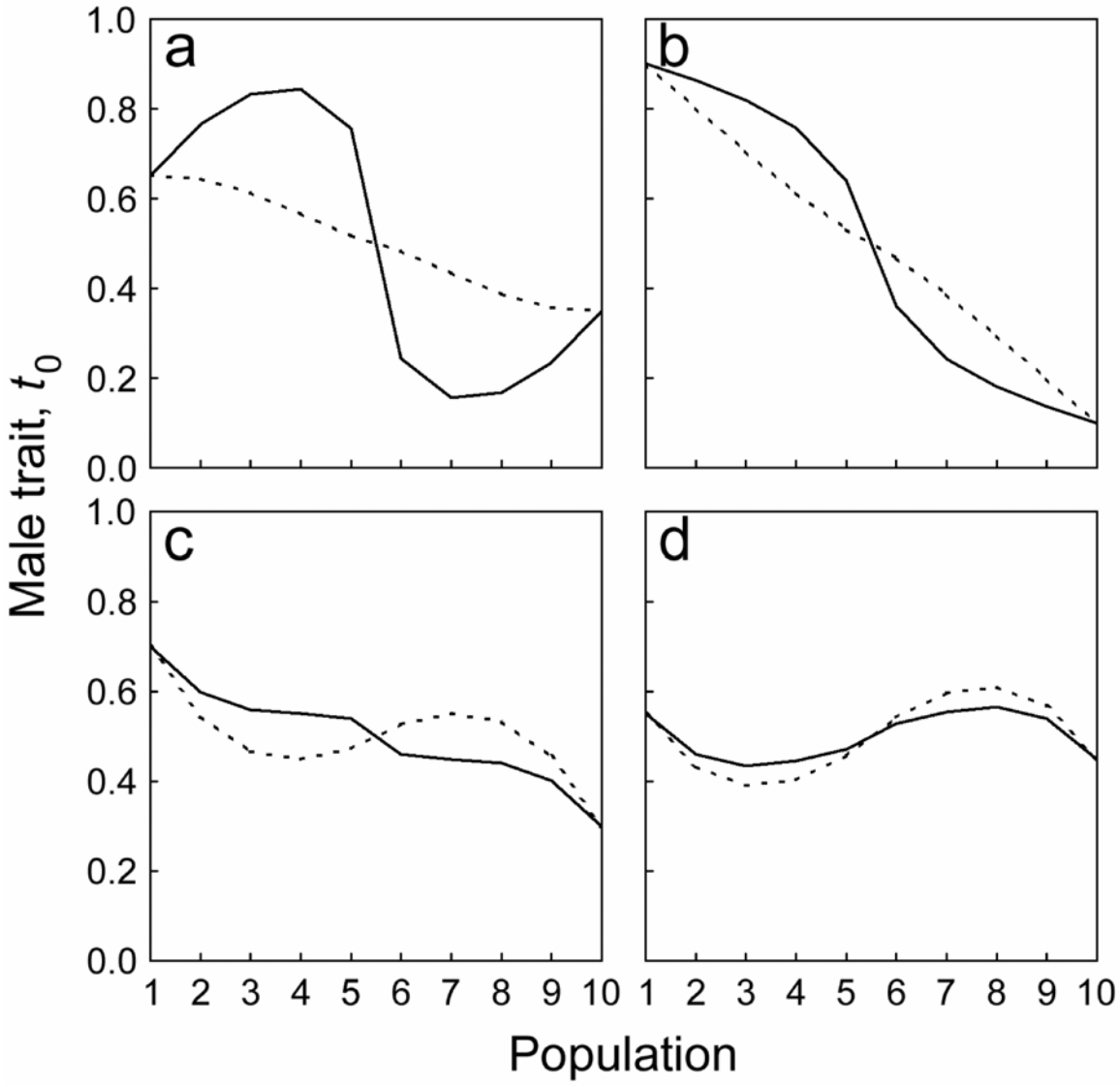

Figure 2.4. Equilibrium cline shapes resulting from reinforcement. A dashed line (- -) indicates that no selection is acting against hybrids, whereas a solid line (–) indicates that hybrids have reduced fitness. Reinforcement can result in (a) a reversed cline, (b) a monotonic cline, or (c & d) other types of clines. Note that although reinforcement occurred in all four of these cases, reproductive character displacement resulted in only one case (a). Parameter values are (a) $\alpha = 0.4$, m = 0.05, $\Delta T = 0.3$, $\Delta P = 0.4$, and sI = 0.6, (b) $\alpha = 0.4$, m = 0.05, $\Delta T = 0.8$, $\Delta P = 0.4$, and sI = 0.4, (c) $\alpha = 0.4$, m = 0.05, $\Delta T = 0.4$, $\Delta P = 0.3$, and sI = 0.6, and (d) $\alpha = 0.2$, m = 0.05, $\Delta T = 0.1$, $\Delta P = 0.4$, and sI = 0.6.

References

Akaike, N. H., and G. M. Hewitt. 1981. Hybrid zones and speciation. In: *Evolution and Speciation/Essays in Honor of M.J.D. White* (W.R. Atchley & D.S. Woodruff, eds), pp. 109-145. Cambridge University Press, Cambridge.

Blair, W. F. 1955. Mating call and stage of speciation in the *Microhyla olivacea-M. carolinensis* complex. Evolution 9:469-480.

Brown, W. L. Jr., and E. O. Wilson. 1956. Character displacement. Systematic Zoology 5:49-64.

Butlin, R. K. 1987a. Speciation by reinforcement. Trends in Ecology and Evolution 2:8-13.

Butlin, R. K. 1987b. Species, speciation, and reinforcement. American Naturalist 130:461-464.

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

Dobzhansky, T. 1940. Genetics and the origin of species. Columbia University Press, N.Y.

Howard, D. J. 1993. Reinforcement: Origin, dynamics, and fate of an evolutionary hypothesis. In: *Hybrid Zones and the Evolutionary Process* (R.G. Harrison, ed.), pp. 46-69. Oxford University Press, New-York.

Kirkpatrick, M. 2000. Reinforcement and divergence under assortative mating. Proceedings of the Royal Society of London B: Biology 267:1649-1655.

Kirkpatrick, M. 2001. Reinforcement during ecological speciation. Proceedings of the Royal Society of London B: Biology 268:1259-1263.

Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection:models and experiments. American Naturalist 159:S22-S35.

Kirkpatrick, M., and M. Servedio. 1999. The reinforcement of mating preferences on an island. Genetics 151:1764-1772.

Liou, L. W., and T. D. Price. 1994. Speciation by reinforcement of premating isolation. Evolution 48:1451-1459.

Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. Heredity 83:503-508.

Paterson, H. E. H. 1978. More evidence against speciation by reinforcement. South African Journal of Science 74:369-371.

Spencer, H. G., B. H. McArdle, and B. M. Lambert. 1986. A theoretical investigation of speciation by reinforcement. American Naturalist 128:241-262.

Turelli, M., N. H. Barton, and J. A. Coyne. 2001. Theory and speciation. Trends in Ecology and Evolution 16:330-343.

Chapter 3

Reinforcement and the Genetics of Hybrid Incompatibilities*

Abstract. Recent empirical studies suggest that genes involved in speciation are often sex-linked. We derive a general analytic model of reinforcement to study the effects of sex linkage on reinforcement under three forms of selection against hybrids: one-locus, two-locus, and ecological incompatibilities. We show that the pattern of sex linkage can have a large effect on the amount of reinforcement due to hybrid incompatibility. Sex linkage of genes involved in postzygotic isolation generally increases the strength of reinforcement, but only if genes involved in prezygotic isolation are also sex-linked. We use exact simulations to test the accuracy of the approximation and find that qualitative predictions made assuming weak selection can hold when selection is strong.

*Significant portions of this chapter have been previously published as Lemmon & Kirkpatrick, 2006. Genetics, 173: 1145-1155.

3.1 INTRODUCTION

Speciation is the evolution of prezygotic or postzygotic isolation. Postzygotic isolation is thought to evolve through the accumulation of genetic incompatibilities during allopatric separation. Prezygotic isolation can evolve through reinforcement, which is the evolution of increased prezygotic isolation as a result of selection against hybrids (Dobzhansky, 1940; Blair, 1955; Howard, 1993). Empirical evidence for reinforcement comes from studies of insects, birds, fish, amphibians, and other taxa (Howard, 1993; Coyne and Orr, 2004).

One empirical pattern that has recently emerged is that many genes involved in speciation are sex-linked (Grula and Taylor, 1980; Heisler, 1984; Reinhold, 1998; Ritchie, 2000; Iyengar et al., 2002; Lindholm and Breden, 2002; Sætre et al., 2003). How does sex linkage affect the potential for reinforcement?

Theoretical studies of reinforcement have explored several genetic and geographic situations, demonstrating that reinforcement is expected under general conditions (Kirkpatrick and Servedio, 1999; Kirkpatrick, 2000, 2001; Kirkpatrick and Ravigné 2002; Lemmon et al., 2004). Kirkpatrick (2001) studied reinforcement due to ecological incompatibilities in haploids. Theoretical studies have also demonstrated that increased linkage between incompatibility loci generally decreases the amount of

reinforcement (Kirkpatrick and Servedio, 1999; Servedio and Sætre, 2003). The only study considering the effects of sex linkage is a simulation study by Servedio and Sætre, (2003), who compared the amount of reinforcement expected when all loci are autosomal to the amount expected when all loci are Z-linked. They concluded that sex linkage enhances reinforcement. Servedio and Sætre did not, however, consider incompatibilities between autosomal and sex-linked genes, which are quite common in nature (Schartl, 1995; Presgraves, 2003; Barbash et al., 2004). What is missing from the theoretical literature is an analytic model of reinforcement that can accommodate any pattern of sex linkage and any number of incompatibility genes.

In this paper, we derive a general analytic model of reinforcement that allows for any form of pre- and postzygotic isolation. We study the effects of sex linkage by applying the general model to three specific types of postzygotic isolation (hybrid incompatibility): selection at a single locus, selection on two incompatible loci, and selection on an ecological intermediate. We focus on the evolution of prezygotic isolation (female preference) on an island population that hybridizes with migrants arriving from a continental population. There are several reasons why this situation is of interest. First, islands are a prolific source of new species (Mayr, 1963). Second, the assumption is not as restrictive as it first appears. This model will, for example, apply to cases where migration is two-way but selection or some other force constrains the

evolution of speciation genes in one population. Third and most importantly, the measure of reinforcement is clear. Reinforcement is simply the amount of divergence in the preference between the island and the continent. When selection does not act against hybridization, the preference on the island will match that of the continent because migration will reduce any initial divergence in the preference to zero. Any divergence in the preference that remains, then, is a result of reinforcement.

Our model suggests that sex linkage of hybrid incompatibility genes enhances reinforcement, but only when the female preference genes are also sex-linked. We also show that autosome-X (or autosome-Z) incompatibilities are favorable to reinforcement regardless of whether the preference genes are autosomal, or X-linked (or Z-linked). We test the accuracy of the approximations using exact four-locus simulations.

3.2 RESULTS

Notation and Assumptions. Our model describes the evolution of biallelic loci in a diploid, sexually reproducing population. The notation, which follows Kirkpatrick et al., (2002), is described in detail below and summarized in Table 3.1. We say that genes carried by an individual occupy *positions*. A position is defined to be the locus at which a gene resides, the sex of the individual carrying it, and the sex of the individual from

which it was inherited. Four positions exist for an autosomal locus, three exist for a X-linked (Z-linked) locus, and one exists for a Y-linked (W-linked) locus. At an autosomal locus *i*, for example, the gene carried by a female and inherited from a female (the individual's mother) is denoted i_{ff}, while the gene she inherited from a male (her father) is written i_{fm}. Segregating at each locus are two alleles, denoted 0 and 1. On the island, the frequency of allele 1 at position $\mathbb{i}$ (where $\mathbb{i}$ could represent i_{ff}, i_{fm}, i_{mf}, or i_{mm}) is denoted $p_{\mathbb{i}}$, with $q_{\mathbb{i}} = 1 - p_{\mathbb{i}}$. Loci may be linked or unlinked.

We divide the loci into three nonoverlapping sets: one containing the female preference loci, one containing the male trait loci, and one containing the hybrid incompatibility loci. These sets are denoted $\mathbb{P}$, $\mathbb{T}$, and $\mathbb{H}$, respectively. The set of all positions in females that affect the preference is written $\mathbb{P}_f$. With *n* autosomal loci, for example, there are 2*n* positions in this set. Each set may contain any number of loci. We make no assumptions about the type of natural and sexual selection on the male trait, which means that any type of mating system is applicable.

A female's mating preference phenotype *P* and a male's display trait phenotype *T* are allowed to be any aspect of their phenotypes that affects who they are likely to mate. For simplicity, we assume that the genes affecting the preference and trait have additive effects and that there is no imprinting (that is, alleles inherited from mothers and fathers are expressed equally). The female preference loci are assumed to be free of direct

selection. This allows us to isolate the effects of reinforcement that results from selection acting on the male trait and hybrid incompatibility loci. We make no assumptions about how the preference loci are inherited (autosomally or sex-linked). For simplicity, however, we do assume they all share the same mode of inheritance. The trait loci are autosomal. The effects of sex linkage of the male trait and female preference are studied in Hall and Kirkpatrick (2006).

All loci that affect the hybrids but not the preference or the male trait are designated as the hybrid incompatibility loci. These loci can have any pattern of additive or nonadditive gene action (that is, dominance and/or epistasis) involving any combination of the loci. We make no assumptions about the mode of inheritance of these loci: they may be autosomal, sex-linked, or cytoplasmic.

We study the evolution of female preference on an island that receives migrants from a continent that is at equilibrium. Migration occurs at a rate *m*, which is defined to be the proportion of newly arrived migrants on the island just after migration. The model applies equally to a pair of sympatric species between which there is one-way introgression. In that case, *m* represents the rate of hybridization between the focal species and the other species, and all descendants of hybrid matings are considered to belong to the focal population.

We assume a life cycle that begins with zygotes, followed by selection on the hybrid incompatibility loci, followed by migration, followed by natural and/or sexual selection on the male trait loci, followed by mating. The cycle ends with transmission and the generation of new zygotes. We also assume non-overlapping generations and that the effects of genetic drift and mutation are negligible.

The female preference on the island will evolve as an indirect response to direct selection on other loci, namely the male trait and hybrid incompatibility loci. Indirect selection depends on two things. The first is the strength of direct selection. We denote the strength of direct selection acting on a set of positions $\mathbb{U}$ by $a_{\mathbb{U}}$. Kirkpatrick et al. (2002) explain how $a_{\mathbb{U}}$ can be calculated for any pattern of selection, including arbitrary forms of epistasis and dominance. This means that results for reinforcement that are derived in terms of the $a_{\mathbb{U}}$ can be applied to any kind of hybrid incompatibility, as we will see shortly. The second thing on which indirect selection depends is the strength of associations (linkage disequilibria) among positions. We denote the associations among positions in the set $\mathbb{U}$ as $D_{\mathbb{U}}$. The algebraic definition for $D_{\mathbb{U}}$ is given in Kirkpatrick et al. (2002), which can be consulted for more details.

General Results. We begin by deriving a model of reinforcement that is general with regards to the form of selection against hybrids. We then apply the general model to

specific types of hybrid incompatibility in order to determine the effects of sex linkage on the potential for reinforcement. For purposes of clarity, only the main results are presented below; the detailed derivation of each result is given in the online materials.

Under our assumption that the preference genes have additive effects, the preference phenotype of a particular female can be written

$$P = \overline{P} + \sum_{i \in \mathbb{P}_f} b_i \zeta_i, \tag{3.1}$$

where $\overline{P}$ is the mean preference among female zygotes on the island, b_i is the difference in the preference caused by carrying allele 1 rather than allele 0 at position i, and $\zeta_i = q_i$ if the female carries allele 1 at position i and $-p_i$ otherwise. The summation includes one term for each of the positions affecting the preference in females.

Reinforcement of the preference on the island will result in divergence between the mean values of the preference on the continent (denoted P^C) and the island. The change in the mean preference in females from the start of one generation to the next is

$$\Delta \overline{P} \;=\; \sum_{i \in P_f} b_i \, \Delta p_i, \tag{3.2}$$

where Δp_i is the per-generation change in the allele frequency at position i.

In the online materials, we show that the per-generation change, Δp_i, can be written as a function of the change caused by selection and migration *within* a generation, which is:

$$p_i''' - p_i \;=\; \sum_{A \subseteq H} a_A \, D_{Ai} \;+\; m(p_i^C \;-\; p_i') \;+\; \sum_{A \subseteq T} a_A \, D_{Ai}'' \,. \tag{3.3}$$

We use primes to denote variables at different stages in the life cycle, with no primes denoting a value in zygotes, one prime after selection on hybrid incompatibility, two primes after migration, and three primes after natural and sexual selection on the male trait. A superscript C denotes a value in the continental population.

To make further progress, we need expressions for the associations (the *D*) that appear in (3.3), which change under the forces of selection, migration, nonrandom mating, and recombination. Here we will use the "quasi-linkage equilibrium", or QLE,

approximation (Barton and Turelli, 1991; Kirkpatrick et al., 2002). The key assumptions are that the selection coefficients (the a) and associations (the D) are much smaller than 1, which allows us to neglect terms where they appear as higher powers. The assumption regarding the associations will be met when the forces that generate associations (selection and migration) are weak relative to the forces breaking them down (recombination), so we require that recombination rates be not too small. We also assume that migration is weak relative to selection, such that m is of $O(a^2)$. The online materials show how we can then derive $O(a)$ approximations for the associations.

To simplify the analysis, we assume that the preference loci are unlinked to the incompatibility loci. We find the amount of reinforcement at equilibrium by combining (3.2) and (3.3), and setting the change in (3.2) to zero. In the online materials we find that the difference between the mean preference on the island and continent is

$$\hat{P} - P^{C} = M(1 + I) + O(a), \tag{3.4}$$

where a hat denotes a value at equilibrium and $O(a)$ indicates that terms proportional to the a have been neglected. M represents the effect of selection acting on the male trait loci:

$$M = \frac{F_P}{m} \sum_{i \in P} b_i \sum_{A \subseteq T} a_A \tilde{D}^{"}_{Ai}, \quad (3.5)$$

where the tilda represents a QLE approximation and F_P is the proportion of preference genes found in females (for example, F_P equals 1/2 when the preference loci are all autosomally inherited and 2/3 when they are all X-linked). Hall and Kirkpatrick (2006) use (3.5) to study in detail how sex linkage of the female preference and male display trait affect reinforcement.

The value of I quantifies the strength of hybrid incompatibility in terms of its effect on the reinforcement of the female preference. I is defined as:

$$I = -F_P \sum_{A \subseteq H_f} a_A d_A \phi^f_A - (1 - F_P) \sum_{A \subseteq H_m} a_A d_A \phi^m_A . \quad (3.6)$$

The left and right terms represent the effect of selection against hybrid incompatibilities in females and males, respectively. Here, $d_A = \prod_{i \in A} (p^C_i - \hat{p}_i)$. Expressions for ϕ^f_A and ϕ^m_A are derived in the online materials. Their values depend on three things: 1) the

type of hybrid incompatibility, 2) how the preference and incompatibility loci are inherited and 3) the recombination rates between hybrid incompatibility loci.

Equations 3.4 - 3.6 are our primary results. By specifying how the hybrid incompatibility genes are selected and inherited, we can use (3.6) to make quantitative predictions. Before doing that, however, we can draw three general conclusions. The first conclusion is that the amount of reinforcement depends on how the female preference and hybrid incompatibility genes are inherited. That follows from the fact that the F_{P} (the fraction of preference positions in females) and the $\phi_{\mathbb{A}}$ depend on the mode of inheritance.

The second conclusion is that the amount of reinforcement will increase with increasing strengths of selection against hybrids (reflected in larger values of $a_{\mathbb{A}}$) and increasing amounts of divergence at the hybrid incompatibility loci (reflected in larger absolute values of $d_{\mathbb{A}}$). Note, however, that reinforcement can occur when $I = 0$, implying that selection on the male trait alone can produce reinforcement (Kirkpatrick and Servedio, 1999). Selection on genes other than the male trait genes may enhance reinforcement, but only if $M > 0$ (implying that there is natural or sexual selection on the male trait favoring it to diverge from the continent).

The third conclusion is that the total effect of selection against hybrids on reinforcement can be separated into an effect due to selection on females (the first summation in (3.6)) and an effect due to selection on males (the second summation in (3.6)). These effects are weighted by F_P and $(1 - F_P)$, which are the proportion of time the preference genes spend in females and males, respectively. If all preference loci are autosomal, for example, then $F_P = 1/2$ and effects from the two sexes are weighted equally. If all preference genes are X-linked, however, then $F_P = 2/3$ and the female effect is given more weight. The mode of inheritance of the female preference genes, therefore, influences the relative effects of selection against male and female hybrids on reinforcement (see also Hall and Kirkpatrick, 2006).

To simplify the analyses, we now make three additional assumptions. The first assumption is that the island and continent are nearly fixed for alternative alleles (alleles 1 and 0 respectively) at the hybrid incompatibility loci. When this is true, $d_{\mathbb{A}} \approx (-1)^n$, where n is the number of positions in $\mathbb{A}$. The assumption is good as long as the strength of migration is weak relative to the strength of selection against hybrids and migrants.

The second assumption is that males are the heterogametic sex. We make this assumption only for clarity of presentation. For taxa in which females are the heterogametic sex, an expression for I can be obtained from (3.6) by simply

interchanging the notation for males and females. This means that an autosome-X incompatibility, for example, has the same effect as autosome-Z incompatibility.

Third, we assume that the quantity *M* defined in (3.5) is independent of how the preference loci are inherited. This approximation is based on the results of Hall and Kirkpatrick (2006), who find that changing the mode of inheritance of the preference from autosomal to X-linked alters *M* by less than 10% when the trait is autosomal. This simplification allows us to isolate the consequences of sex linkage on reinforcement through its effects on *I*, which represents the force of selection on hybrid incompatibility.

We will now use (3.6) to derive equations for *I* considering three types of incompatibilities: those due to selection on a single locus, those due to interactions between two loci, and those due to selection against an ecological intermediate. For one and two locus incompatibilities, we consider all possible combinations of three modes of inheritance (autosomal, X-linked, Y-linked) for the female preference and hybrid incompatibility loci. We then compare the cases and determine the combinations that are expected to produce the most reinforcement.

One-Locus Incompatibilities. In the previous section, we derived a general equation describing the effect of any form of selection against hybrids on reinforcement. Here we use that equation to study incompatibilities resulting from selection on a single locus, *k*.

This type of incompatibility can evolve when the fitness effects of the locus are environment-dependent, for example when one allele is favored on the continent and another on the island. For generality we allow females and males to have different fitnesses. Our notation for the strengths of selection against different genotypes is given in the upper half of Figure 3.1, where s_0 is the strength of selection against an individual homozygous/hemizygous for the continental allele, and s_1 is the strength of selection against a heterozygote. The preference loci may be autosomal or X-linked and the incompatibility locus may be autosomal, X-linked, or Y-linked.

In the online materials, we derive expressions for $a_{\mathbb{A}}$, $\phi_{\mathbb{A}}^{\mathrm{f}}$, and $\phi_{\mathbb{A}}^{\mathrm{m}}$. Plugging those results into (3.6) and simplifying, we obtain the expressions for the effect on reinforcement, *I*. Those expressions are summarized in Table 3.2, where f and m denote coefficients for females and males, respectively. The effect of two or more independent one-locus incompatibilities can be studied by simply summing up the appropriate values from Table 3.2.

Results in Table 3.2 show that under general conditions, reinforcement is expected to be strongest when the preference and the incompatibility have the same mode of inheritance (either autosomal, X-linked, or Z-linked). Other combinations are expected to produce smaller amounts of reinforcement. This conclusion is robust so

long as males and females with the same genotype have equal fitness, the island allele is not recessive, and there is no overdominance (i.e. $s_1 > s_0/2$).

We can also see from Table 3.2 that the amount of reinforcement generally increases with the strength of selection against hybrids. The one exception is when the preference and incompatibility loci reside on different sex chromosomes. No reinforcement is possible in this case because the two loci cannot be inherited together.

Two-Locus Incompatibilities. Two-locus hybrid incompatibilities, first discussed by Bateson (1909), Dobzhansky (1934, 1937), and Muller (1939, 1940, 1942), are thought to be common in nature and have been the focus of extensive empirical and theoretical studies (Coyne and Orr, 2004). Here we apply the general result (3.6) to two-locus hybrid incompatibilities.

We denote the loci involved as k and l, and the rate of recombination between them as r_{kl}. Again, we allow females and males to have different fitnesses. Our notation for the strengths of selection against different genotypes is given in the lower half of Figure 3.1. Five combinations of modes of inheritance for the two incompatibility loci exist: autosome-autosome (A-A), autosome-X (A-X), autosome-Y (A-Y), X-X, and X-Y. Since the preference may be autosomal or X-linked, there are a total of ten combinations when all three sets of loci are considered.

In the online materials, we derive expressions for $a_{\mathbb{A}}$, $\phi^{\mathrm{f}}_{\mathbb{A}}$, and $\phi^{\mathrm{m}}_{\mathbb{A}}$ for each of the ten combinations of modes of inheritance. Plugging these expressions into (3.6) yields the expressions for I given in Table 3.3. For simplicity of presentation, Table 3.3 presents the expressions for I assuming that males and females with the same genotype have the same fitness and that dosage compensation exists. More general expressions for I are given in the online materials.

Table 3.3 can be used to show that increased linkage between the incompatibility loci decreases the amount of reinforcement for commonly observed types of incompatibilities. These results are consistent with previous theory (Kirkpatrick and Servedio, 1999; Servedio and Sætre, 2003).

The general model derived above can accommodate any pattern of dominance and epistasis, as well as any pattern of sex linkage. Here we compare different types of incompatibilities in order to determine how sex linkage affects reinforcement. To make the analysis tractable, we reduce the number of parameters by making the following biologically realistic assumptions: 1) the free recombination between hybrid incompatibility, 2) dosage compensation, and 3) males and females with the same genotype have equal fitness. Note that the third assumption does not imply that males and females of the same hybrid class have equal fitness.

To determine the effect of sex linkage on reinforcement, we computed *I* for A-A, A-X, A-Y, X-X, and X-Y incompatibilities using Table 3.3, then compared the values in a pairwise fashion. We studied the five patterns of hybrid incompatibility presented in the top row of Figure 3.2. Patterns A and B are two simple ways in which pure types have high fitness and hybrids are selected against. Pattern C is consistent with the type of incompatibility studied by Servedio and Sætre (2003). Pattern D is compatible with the type studied by Turelli and Orr (2000), who assume that the incompatibility evolved by drift. Pattern E is inconsistent with that assumption. By comparing the results from the five patterns, we can identify how our conclusions may be affected by particular assumptions about the type of two-locus incompatibility.

Two conclusions can be drawn from Figure 3.2. First, it is clear that the type of incompatibility that is expected to contribute the most to reinforcement depends on whether or not the female preference is sex linked. This result holds for all five patterns of hybrid incompatibility studied (also see online materials for a more detailed analysis). When the preference is autosomally inherited (center row), A-A and A-X incompatibilities are expected to contribute the most. When the preference is X-linked (bottom row), however, X-X and A-X incompatibilities are expected to be more important. This result is consistent with our results for one-locus incompatibilities.

The second conclusion is that the mode of inheritance of the hybrid incompatibility has only a moderate effect when A-Y and X-Y incompatibilities are ignored. When A-Y and X-Y incompatibilities are considered, however, sex linkage can have a very large effect (see online materials), in some cases exceeding a 10-fold effect. The reason is that very little reinforcement is expected when one or more incompatibility loci are Y-linked and the preference locus is X-linked. Analogous results apply to taxa with Z-W sex determination.

Ecological incompatibilities. How does selection against ecologically-inferior hybrids favor reinforcement of prezygotic isolation? Here we apply the general model derived above to the situation in which genes contribute additively to a quantitative trait, such as body size or bill length. In this model, hybrids are selected against because they have intermediate phenotypes that are selected against. We assume that in the absence of migration, the island and continent populations would be fixed for alternative alleles at a set of loci. Below we present an equation for the amount of reinforcement assuming that all incompatibility loci are either autosomal, X-linked, or Y-linked. More general equations allowing for any combination of loci are given in the online materials.

For simplicity, we assume that the *n* loci influencing the ecological trait have equal allele frequencies and equal effects on the trait. We also assume that the

ecological trait is determined by a large number of loci, each with a small effect. The mean values of the ecological trait on the continent and island are given by Z^C and Z, respectively. Selection against the hybrids is a function of the strength of selection of the directional selection gradient β and the stabilizing selection gradient Γ acting on the ecological trait. (Negative values of Γ correspond to stabilizing selection, and positive values to disruptive selection.) The values of the selection gradients depend on the fitness function for the ecological trait and also on the distribution of that trait in the island population. That distribution evolves in response to selection and migration, which causes the values of the selection gradients to change. The β and Γ in the expressions that follow refer to the equilibrium values for the gradients. See Lande and Arnold (1993) and Kirkpatrick (2001) for more details.

In the online materials we show that the effect on reinforcement from selection on the ecological trait is

$$I \approx \phi\left(\tfrac{1}{2}\beta\left|\hat{Z} - Z^C\right| + \Gamma\left(\hat{Z} - Z^C\right)^2\right), \tag{3.7}$$

where $\hat{Z}$ is the equilibrium value of the ecological trait on the island and the value of ϕ depends on how the preference and ecological trait are inherited (see Table 3.4).

Table 3.4 shows that the type of incompatibility that is expected to contribute most to reinforcement depends on how the female preference is inherited. Autosomal incompatibilities contribute more when the preference also autosomal, whereas X-linked incompatibilities contribute more when the preference is X-linked. Y-linked incompatibilities are expected to contribute relatively little, regardless of how the female preference is inherited. Equation 3.7 also shows that the amount of reinforcement due to selection acting on an ecological trait increases linearly with the strength of directional selection acting on the island, linearly with the strength of disruptive selection acting on the island, and faster than linearly (quadratically) with the equilibrium amount of divergence in the ecological trait between the continent and island.

3.3 SIMULATIONS

The analytic model developed above utilizes a QLE approximation that is accurate when selection and migration are sufficiently weak. Here we use exact simulations in order to answer three questions: 1) How does the accuracy of the analytic approximation decrease with an increasing strength of selection? 2) Do qualitative results obtained from the weak selection approximation hold when selection is strong? 3) What happens when migration is not weak relative to selection? We find that the

analytic approximation is quite good when selection is weak, and that qualitative conclusions hold when selection is strong.

Notation and Assumptions. The analytic model allows any number of loci to contribute to the female preference, male trait, and hybrid incompatibility. Here we consider a special case of that model: when one locus contributes to the female preference, one locus contributes to the male trait, and two loci contribute to the hybrid incompatibility. We assume that females have a preference of $1 + \alpha$ for males with a trait genotype that matches her preference genotype, relative to males with other genotypes. Mating probabilities are calculated using preference values and genotype frequencies taken just before mating (Kirkpatrick 1982). We also assume natural selection disfavoring continental male trait alleles on the island, with heterozygotes taking a fitness of $1 - s_T / 2$, and homozygotes/hemizygotes taking a fitness of $1 - s_T$. To maximize the potential for reinforcement, we assume that the continental allele frequencies at the preference and trait loci are equal to 0.5. Initial allele frequencies on the island were 1.0 for the male trait and 0.5 for the female preference (initial frequencies on the island had a negligible effect on equilibrium values). The continent and island were initially fixed for alternate alleles at the hybrid incompatibility loci. We assumed a heritability of 1 and that all loci are unlinked. Without loss of generality, we choose the following phenotypic values for the preference and trait: 0 for individuals

homozygous/hemizygous for the continental allele, 1 for heterozygous individuals, and 2 for individuals homozygous/hemizygous for the island allele. Simulations were run until the change in each allele frequency between generations was less than 10^{-12}.

Accuracy of the Approximation. How does the accuracy of the analytic approximation decrease as the strength of selection is increased? We answered this question by comparing exact values for the amount of reinforcement (obtained from the simulations at equilibrium) with the corresponding approximate values (obtained from the analytic model given in Equation 3.4). To calculate the analytic approximation, we used the expression for (5) derived by Hall and Kirkpatrick (2006) in terms of measurable quantities (see their Equation 5). Error of the analytic approximation was calculated as $\left(\hat{P} - P\right) \big/ \left(P - P^{C}\right)$, where $\hat{P}$ and P are the approximate and exact values of the preference on the island, respectively. This is a measure of the accuracy of the entire model, which is a function of the accuracy of (3.5) (the male trait component) and of (3.6) (the hybrid incompatibility component).

We varied the strength of reinforcement by varying the values of the parameters describing migration and selection, while holding the relationship among those parameters constant. More specifically, we denoted the strength of selection as s, and set $m = s/80$, $s_T = s$, $\alpha = 2s$, $s_{00} = 0$, $s_{11} = s/4$, $s_{01} = s_{10} = s_{12} = s_{21} = s/2$, $s_{02} = s_{20} = s$. These assumptions correspond to the incompatibility type E from Figure 3.2.

Figure 3.3 shows how the error in the analytic approximation increases with the strength of selection for five different types of sex linkage (preference and trait are autosomal). Figure 3.3 suggests that when selection is weak the accuracy of the analytic approximation is quite good (the error is on the order of the strength of selection).

Servedio (2004) criticizes the QLE approximation, suggesting that conclusions drawn from models assuming weak selection can not be applied to situations involving strong selection. To test this claim, we extended the simulations to include conditions of very strong selection. Results from these simulations are presented in Figure 3.4. The main figure presents the results obtained when the preference and trait are autosomal and the inset figure presents results obtained when they are X-linked.

We can conclude from Figure 3.4 that qualitative patterns observed when selection is weak can hold when selection is very strong. These simulations present two patterns that Servedio found problematic. First, the simulations involving an autosomal preference and trait show a rank-order change in the relative amounts of reinforcement as selection becomes strong, indicated by the fact that the lines corresponding to A-Y and X-X incompatibilities cross when $s \approx 0.8$. This rank-order change, however, does not change the qualitative conclusions: incompatibilities that are especially favorable to reinforcement remain so. Second, the simulations involving an X-linked preference and

trait show a highly non-linear relationship between the strength of selection and the amount of reinforcement. Note that despite this non-linear relationship, the qualitative patterns observed when selection is weak are also observed when selection is strong. For example, the observation that when the preference and trait are X-linked, A-Y and X-Y incompatibilities are weaker than other types is consistent across the entire range of *s*.

Selection vs. Drift. Recent empirical studies suggest that hybrid incompatibilities evolve by selection instead of drift (Presgraves et al., 2003; Barbash et al., 2004; Wu and Ting, 2004). Authors of some recent theoretical studies of hybrid incompatibilities, however, have assumed that incompatibilities evolve by drift (Turelli and Orr, 2000; Gavrilets, 2003). Is the outcome of secondary contact the same for incompatibilities that evolve by selection as opposed to drift? To answer this question we focused on types D and E from Figure 3.2. Type D is consistent with the assumption that the incompatibility evolved by drift, whereas type E is consistent with the assumption that the incompatibility evolved by selection. For each of these two types, we conducted a series of simulations in which we independently varied the relative strength of selection and migration and recorded the amount of reinforcement at equilibrium. Simulation procedures were the same as described in the *Notation and assumptions* section above, with autosomal inheritance of all 4 loci, $\alpha = 2$, and $s_T = 1$.

The simulations revealed a phenomenon that is not currently appreciated. Postzygotic isolation can be lost through genetic swamping despite strong selection against hybrids, if selection against the ancestral genotype is weak. The reason for this result is that individuals of a particular hybrid class may have different genotypes. If one or more of these genotypes (such as the ancestral genotype) have high fitness, then individuals with those genotypes will survive, allowing introgression of alleles from one species to the other. This will occur even if selection against the hybrid class as a whole is strong. Our simulations demonstrated the result of this pattern. When incompatibility type D was assumed in the simulations, the incompatibility could only be maintained if all F1 hybrids died ($s = 1$) or there was no migration $m = 0$). This implies that for the incompatibility that evolved by drift, swamping could occur despite strong selection against the hybrid classes. When incompatibility type E was assumed in the simulations, in contrast, the incompatibility was maintained whenever $m < s/12$.

The simple conclusion that we can draw from these simulations is that the strength of selection against the hybrid classes may not be the best indicator of whether or not postzygotic isolation is likely to be maintained (or whether or not reinforcement will occur). Other factors, such as whether the postzygotic isolation evolved by selection or drift, may also determine the fate of incompatibilities in hybridizing

populations. This result has important implications for the relative importance of selection and drift in the evolution of hybrid incompatibility.

3.4 DISCUSSION

We have shown that sex linkage influences how hybrid incompatibility contributes to reinforcement. In general, sex-linked incompatibilities are expected to produce more reinforcement than autosomal incompatibilities when female preference is also sex-linked. When the preference is not sex-linked, however, autosomal (and autosome-X) incompatibilities produce more reinforcement. (Autosome-X incompatibilities are favorable to reinforcement regardless of how the preference is inherited.) These results hold for many types of hybrid incompatibility, including selection against heterozygotes at a single locus, selection against incompatible alleles at two loci, and selection against ecological intermediates. The strength of the effect, which depends on the genetic details, can be quite large but is typically less than 2-fold when Y-linkage is ignored.

The weak selection approximation we utilized has recently been criticized by Servedio (2004), who suggested that results from models assuming weak selection can not apply to situations involving strong selection. Our simulations suggest otherwise.

We found that even when reinforcement changes nonlinearly with increasing selection, qualitative patterns hold regardless of the strength of selection. Simulations have also shown the approximation to be more accurate when strong selection is a function of many genes each with small effect rather than a function of a few genes with large effect (data not shown). This suggests that the accuracy we report is a conservative estimate since we considered the worst case (a few genes with large effect). We conclude, therefore, that our qualitative conclusions may hold even when selection against hybrids is strong.

Our simulations also show that incompatibilities that evolve in allopatry by drift or weak selection are likely to be lost by swamping during secondary contact, even when the strength of selection against hybrids is strong. This observation suggests that the rate of hybridization during secondary contact may determine the types of incompatibilities that can persist. With very infrequent hybridization, incompatibilities that evolved by weak selection can be maintained, but with moderate to strong hybridization, only incompatibilities that evolved by very strong selection can be maintained. Incompatibilities that evolved by drift will be lost in the face of any degree of hybridization, assuming some introgression.

Secondary contact should act as a sieve, therefore, removing incompatibilities that evolved by drift or weak selection but retaining those that evolved by strong

selection. This *incompatibility sieve* should produce at least two empirical patterns. First, taxa that have not come into secondary contact since their divergence should have a greater proportion of incompatibilities that evolved by drift and weak selection than taxa that hybridize. The second pattern is that if incompatibilities sometimes evolve by drift in allopatry, then upon secondary contact the amount of intrinsic postzygotic isolation should decline until only incompatibilities that evolved by strong selection remain. If a sufficient amount of postzygotic isolation remains, the populations may remain distinct, otherwise, the populations will fuse into one.

The incompatibility sieve is also expected to produce the striking pattern that is observed in hybridizing taxa: incompatibilities between hybridizing taxa evolved by selection. The *OdsH* locus in *Drosophila*, for example, has been shown to be a result of gene duplication followed by positive selection (Ting et al., 1998, 2000, 2004; Sun et al., 2004; Wu and Ting, 2004). Both the *Nup96* gene and the candidate factor *Nup153* appear to be a product of recent positive selection (Presgraves et al., 2003; D. Presgraves, *pers. comm*.). Barbash et al. (2004) have shown that *Hmr*, a factor causing incompatibility between *D. melanogaster* and its sibling species, has also been under recent positive selection. While there may be other factors contributing to this pattern, such as observational or publication bias, the incompatibility sieve provides a plausible

hypothesis for why we may be observing so many incompatibilities that evolved by selection rather than by drift.

Table 3.1. Summary of notation.

Loci, contexts, positions, and sets of positions

i	A single locus i
$\mathbb{i}$	A single position at locus i
i_{fm}	The position at the i locus that is found in a female and inherited from her father
$\mathbb{A}$	A set of positions
$\mathbb{A}_{\mathrm{f}}$	A set of positions in females
U	A set of loci
$\mathbb{P}$	The set of positions that affect the female preference
$\mathbb{T}$	The set of positions that affect the male trait
$\mathbb{H}$	The set of positions that affect hybrid fitness but not the trait or preference

Summations and products

$\sum_{\mathbb{i} \in \mathbb{A}}$	A sum over all positions $\mathbb{i}$ in the set $\mathbb{A}$
$\sum_{\mathbb{U} \subseteq \mathbb{A}}$	A sum over all subsets of the set $\mathbb{A}$, including the full set $\mathbb{A}$ and empty set $\varnothing$
$\prod_{\mathbb{i} \in \mathbb{A}}$	A product over all positions $\mathbb{i}$ in the set $\mathbb{A}$

Allele frequencies

$p_{\mathbb{i}}$	The frequency of the 1 allele at the position $\mathbb{i}$ on the island
$\hat{p}$	The equilibrium frequency of the 1 allele on the island
p^{C}	The equilibrium frequency of the 1 allele on the continent

Δp The change in the allele frequency during one generation

$d_{\mathbb{A}}$ $\prod_{i \in \mathbb{A}} (\hat{p}_i - p_i^C)$, the allele frequency divergence between continent and island

Phenotypes, fitnesses and selection

P The preference of a particular female on the island

$\overline{P}$ The average preference in females on the island

$\hat{P}$ The equilibrium value of the preference in females on the island

P^C The average preference in females on the continent at equilibrium

Z The value of the ecological trait on the island

Z^C The value of the ecological trait on the continent

$\hat{Z}$ The equilibrium value of the ecological trait on the island

b_i The difference in the preference of carrying allele 1 rather than 0 at position i

ζ_i Equals q_i if the female carries allele 1 at position i and $-p_i$ otherwise

$a_{\mathbb{A}}$ The selection coefficient for the set of positions $\mathbb{A}$

s A traditional selection coefficient

β The strength of directional selection on the island

Γ The strength of disruptive selection on the island

Associations

$D_{\mathbb{A}i}$ The association among the positions in the set $\mathbb{A}i$ at the start of the generation

$D''_{\mathbb{A}i}$ The association among the positions in the set $\mathbb{A}i$ after migration

Miscellaneous

r_{kl}	The probability that recombination will break up the loci k and l
m	The proportion of newly arrived migrants on the island, just after migration
I	The effect on reinforcement due to selection on the hybrid incompatibility loci
M	The effect on reinforcement due to selection on the male trait loci
F_{P}	The proportion of female preference genes found in females

Table 3.2. Effect of one-locus incompatibilities on reinforcement (*I*).

Incompatibility	Autosomal Preference	X-Linked Preference
Autosomal	$2\,s_{1_f} + 2\,s_{1m}$	$\frac{8}{3}\,s_{1_f} + \frac{2}{3}\,s_{1_m}$
X-Linked	$2\,s_{1_f} + s_{0m}$	$4\,s_{1_f} + \frac{5}{6}\,s_{0_m}$
Y-Linked	s_{0m}	0

Table 3.3. Effect of two-locus incompatibilities on reinforcement (I)

Incompatibility	Autosomal preference	X-linked preference
A-A	$\frac{4s_{11} + 4(s_{12} + s_{21})r_{kl}}{r_{kl} + 1}$	$\frac{3s_{11} + 3(s_{12} + s_{21})r_{kl}}{r_{kl} + 1}$
A-X	$\frac{2}{3}s_{10} + \frac{4}{3}s_{11} + \frac{1}{3}s_{20} + 2s_{12} + \frac{2}{3}s_{21}$	$\frac{6}{13}s_{10} + \frac{80}{39}s_{11} + \frac{29}{78}s_{20} + \frac{32}{39}s_{12} + \frac{76}{39}s_{21}$
A-Y	$\frac{2}{3}s_{10} + \frac{1}{3}s_{20} + \frac{4}{3}s_{12}$	$\frac{2}{3}s_{12}$
X-X	$\frac{s_{00}(r_{kl} - 5) + 10s_{11} + 2(s_{02} + s_{20})(r_{kl} + 5) + 6(s_{12} + s_{21})r_{kl}}{3r_{kl} + 5}$	$\frac{\frac{1}{6}s_{00}(5 - r_{kl}) + 4s_{11} + (s_{02} + s_{20} + 4s_{12} + 4s_{21})r_{kl}}{r_{kl} + 1}$
X-Y	$s_{02} + s_{20}$	$\frac{5}{6}s_{02}$

Table 3.4. Effect of ecological incompatibilities on reinforcement (ϕ).

Incompatibility	Autosomal preference	X-linked preference
Autosomal	8	6
X-linked	6	8.5
Y-linked	2	0

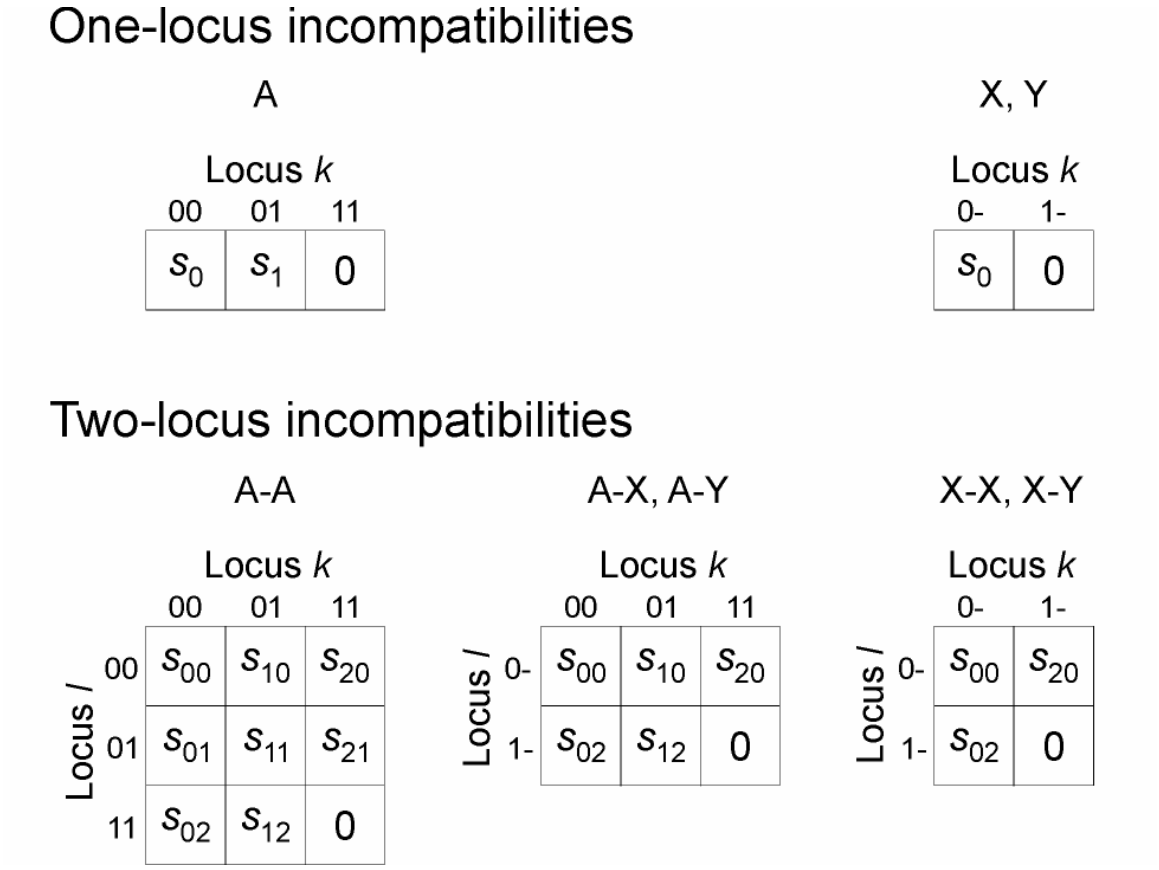

Figure 3.1. Notation denoting the strengths of selection against genotypes of the heterogametic sex. The homogametic sex maintains the genotypes (and notation) seen in the left-most column, regardless of sex linkage. Fitnesses are defined to be relative to the island genotype, which is fixed for the 1 allele at all incompatibility loci. Hybrid incompatibilities are a result of selection on either one or two loci, where each locus may be autosomal (A), X-linked (X), or Y-linked (Y). For two locus incompatibilities, the pattern of sex linkage is denoted using two letters joined by a dash. A-X, for example, indicates that the *k* locus is autosomal and the *l* locus is X-linked.

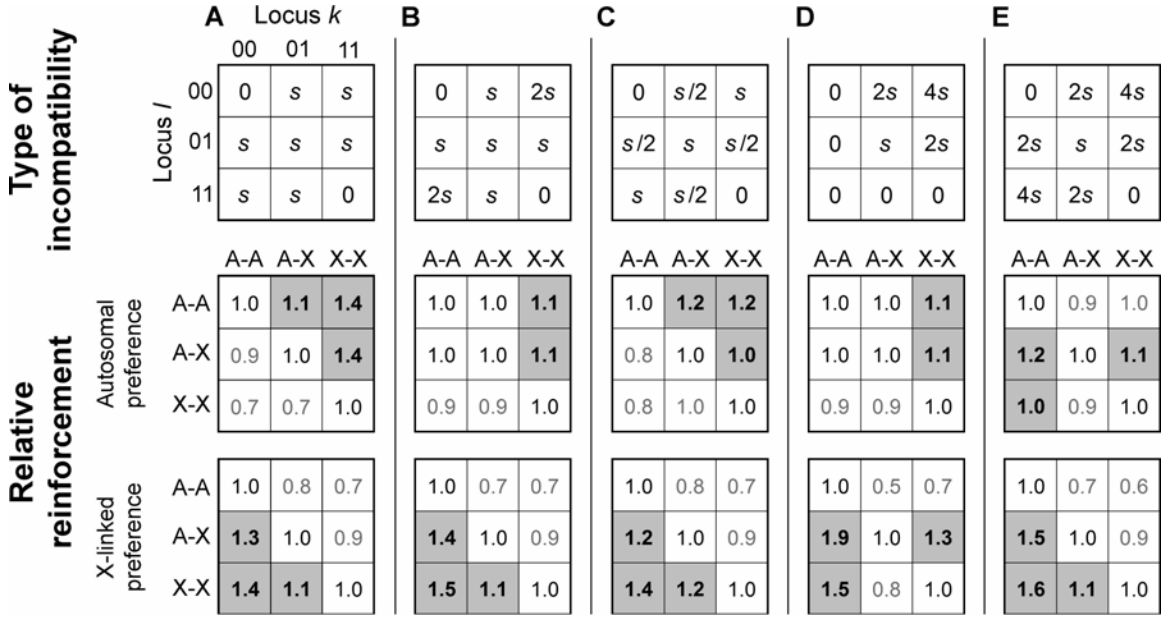

Figure 3.2. The effect of sex linkage on reinforcement for five types of hybrid incompatibility. The five matrices in the top row denote five types of hybrid incompatibility at two loci, *k* and *l*. Types D and E correspond to empirically observed patterns (Turelli and Orr 2000) and type C corresponds to that studied by Servedio and Sætre (2003). Values in the center and bottom rows indicate the relative amounts of reinforcement due to A-A, A-X and X-X incompatibilities. The amount of reinforcement for the pattern of sex linkage indicated to the left of the value is divided by that for the pattern indicated above the value. For example, the value 1.4 found in the lower left most square indicates that an X-X incompatibility is expected to produce 1.4 times more reinforcement than an A-A incompatibility. Values greater than one before rounding are shaded. For comparisons involving A-Y and X-Y incompatibilities, see the online materials.

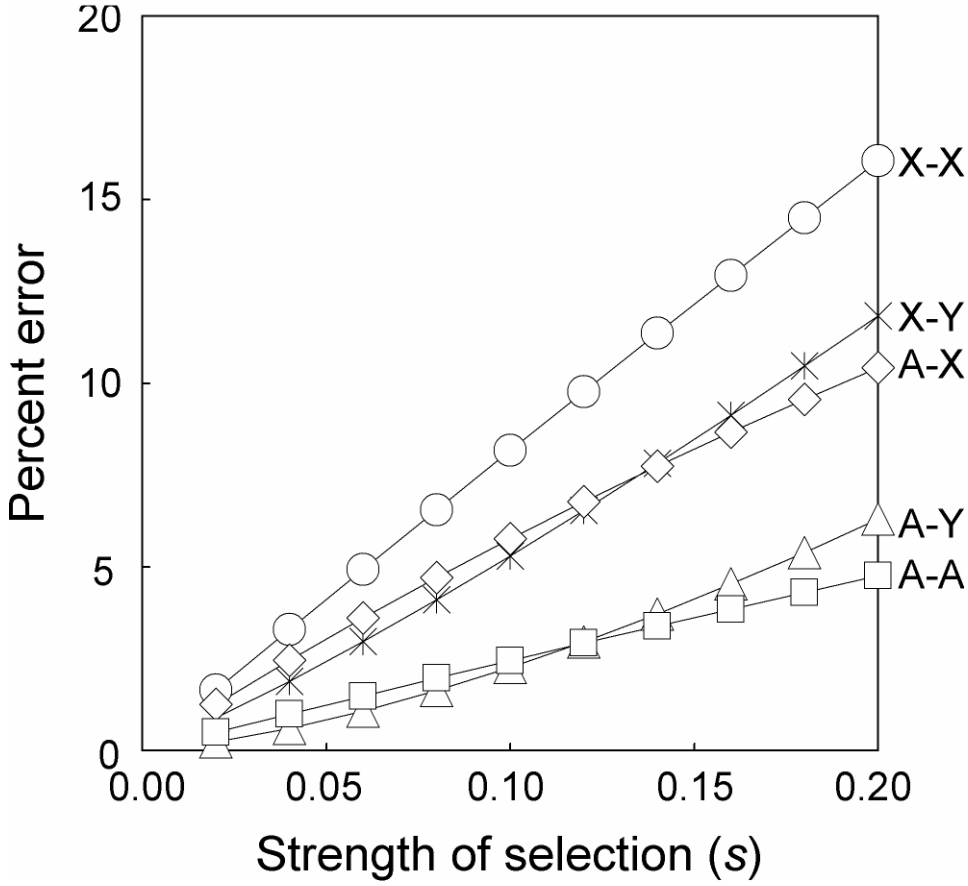

Figure 3.3. Accuracy of the analytic approximation for weak to moderate selection. Percent error, calculated as $(P_{approx} - P_{exact}) / (P_{exact} - P^{C})$, is plotted as a function of the strength of selection (s) for five types of hybrid incompatibility: Autosome-Autosome (A-A), Autosome-X (A-X), Autosome-Y (A-Y), X-X, and X-Y. Exact values were obtained from four-locus simulations. The strengths of migration (m), natural selection on the male trait (s_T) and sexual selection (α) were varied as a function of s, where $m = s / 80$, $s_T = s$, $\alpha = 2s$. The strength of selection on hybrid incompatibility also varied as a function of s, where $s_{00} = 0$, $s_{11} = s / 4$, $s_{01} = s_{10} = s_{12} = s_{21} = s / 2$, $s_{02} = s_{20} = s$.

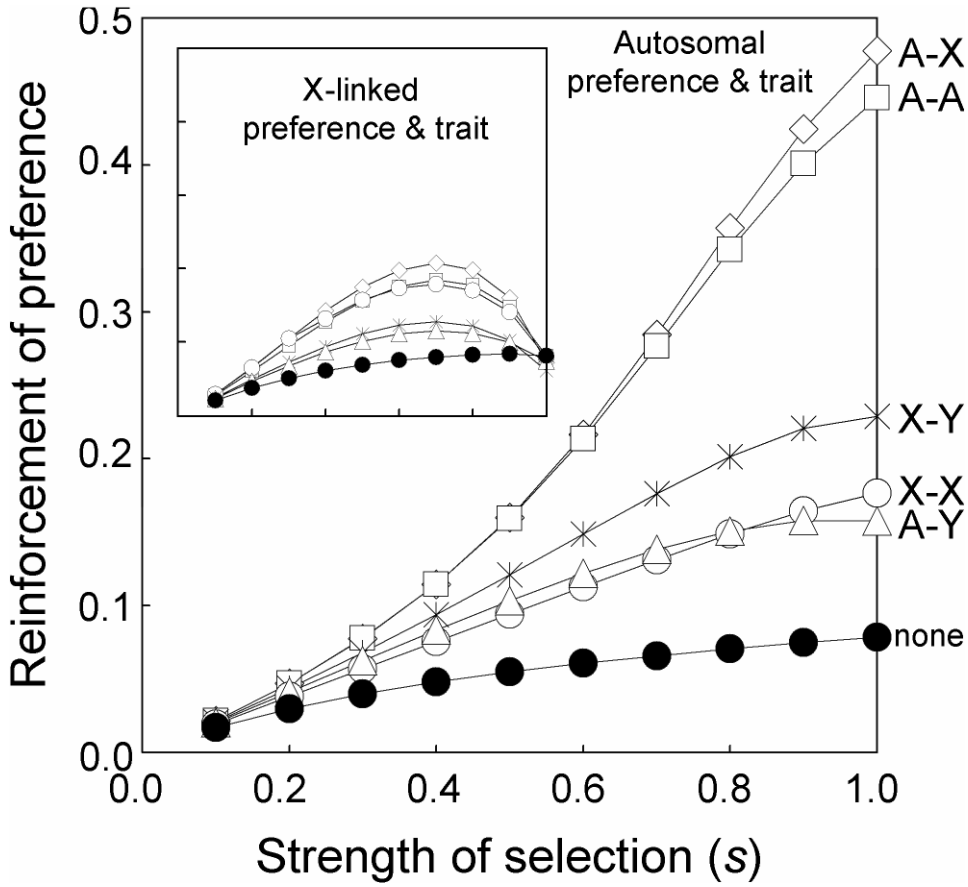

Figure 3.4. Reinforcement of female preference when selection against hybrids is strong. The amount of reinforcement ($P_{exact} - P^C$) is plotted as a function of the strength of selection. Exact values were obtained from four-locus simulations. The strengths of migration, natural selection on the male trait, sexual selection, and selection on the hybrid incompatibility were varied as a function of s (see Figure 3.3 caption). The main figure presents results obtained when the preference and trait are autosomal, whereas the inset figure presents results obtained when the preference and trait are X-linked. The curve labeled none corresponds to the situation in which there is no selection against hybrid incompatibility.

References

Barbash, D. A., P. Awadalla, and A. M. Tarone. 2004. Functional divergence caused by ancient positive selection of a *Drosophila* hybrid incompatibility locus. Plos Biology 2:839-848.

Barton, N. H., and M. Turelli. 1991. Natural and sexual selection on many loci. Genetics 127:229-255.

Bateson, W. 1909. Heredity and variation in modern lights. In: *Darwin and Modern Science* (A. C. Seward, ed.), pp. 85-101. Cambridge University Press, Cambridge.

Blair, W. F. 1955. Mating call and stage of speciation in the *Microhyla olivacea-M. carolinensis* complex. Evolution 9:469-480.

Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Inc., Sunderland.

Dobzhansky, T. 1934. Studies on hybrid sterility. I. Spermatogenesis in pure and hybrid *Drosophila pseudoobscura*. Zeitschrift Fuer Zellforschung Und Microskopische Anatomie 21:169-221.

Dobzhansky, T. 1937. *Genetics and the Origin of Species*. Columbia University Press, New York.

Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. American Naturalist 74:312-321.

Gavrilets, S. 2003. Perspective: Models of speciation: What have we learned in 40 years? Evolution 57:2197-2215.

Grula, J. W., and O. R. Taylor. 1980. The effect of X-chromosome inheritance on mate-selection behavior in the sulfur butterflies, *Colias eurytheme* and *C. philodice*. Evolution 34:688-695.

Hall, D. W., and M. Kirkpatrick. 2006. Reinforcement and sex linkage. Evolution 60:908-921.

Heisler, I. L. 1984. Inheritance of female mating propensities for yellow locus genotypes in *Drosophila melanogaster*. Genetical Research 44:133-149.

Howard, D. J. 1993. Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. In: *Hybrid Zones and the Evolutionary Process* (R. G. Harrison, ed.), pp. 46-69. Oxford University Press, New York.

Iyengar, V. K., H. K. Reeve, and T. Eisner. 2002. Paternal inheritance of a female moth's mating preference. Nature 419:830-832.

Kirkpatrick, M. 1982. Sexual selection and the evolution of female choice. Evolution 36:1-12.

Kirkpatrick, M. 2000. Reinforcement and divergence under assortative mating. Proceedings of the Royal Society of London B: Biology 267:1649-1655.

Kirkpatrick, M. 2001. Reinforcement during ecological speciation. Proceedings of the Royal Society of London B: Biology 268:1259-1263.

Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. Genetics 161:1727-1750.

Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection: models and experiments. American Naturalist 159:S22-S35.

Kirkpatrick, M., and M. R. Servedio. 1999. The reinforcement of mating preferences on an island. Genetics 151:865-884.

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

Lemmon, A. R., C. Smadja and M. Kirkpatrick. 2004. Reproductive character displacement is not the only possible outcome of reinforcement. Journal of Evolutionary Biology 17:117-183.

Lindholm, A., and F. Breden. 2002. Sex chromosomes and sexual selection in poeciliid fishes. American Naturalist 160:S214-S224.

Mayr, E. 1963. *Animal Species and Evolution*. Belknap Press, Cambridge.

Muller, H. J. 1939. Reversibility in evolution considered from the standpoint of genetics. Biological Reviews 14:261-280.

Muller, H. J. 1940. Bearing of the *Drosophila* work on systematics. In *The New Systematics (*J. S. Huxley, ed.), pp. 185-268. Clarendon Press, Oxford.

Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. Biological Symposia 6:71-125.

Presgraves, D. C. 2003. A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. Genetics 163:955-972.

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.

Reinhold, K. 1998. Sex linkage among genes controlling sexually selected traits. Behavioral Ecology Sociobiology 44:1-7.

Ritchie, M. G. 2000. The inheritance of female preference functions in a mate recognition system. Proceedings of the Royal Society of London B: Biology 267:327-332.

Sætre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon et al. 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. Proceedings of the Royal Society of London B: Biology 270:53-59.

Schartl, M. 1995. Platyfish and swordtails - a genetic system for the analysis of molecular mechanisms in tumor-formation. Trends in Genetics 11:185-189.

Servedio, M. R. 2004. The evolution of premating isolation: Local adaptation and natural and sexual selection against hybrids. Evolution 58:913-924.

Servedio, M. R., and G. P. Sætre. 2003. Speciation as a positive feedback loop between postzygotic and prezygotic barriers to gene flow. Proceedings of the Royal Society of London B: Biology 270:1473-1479.

Sun, S., C. T. Ting and C. I. Wu. 2004. The normal function of a speciation gene, Odysseus, and its hybrid sterility effect. Science 305:81-83.

Ting, C. T., S. C. Tsaur, M. L. Wu and C. I. Wu. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. Science 282:1501-1504.

Ting, C. T., S. C. Tsaur and C. I. Wu. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, Odysseus. Proceedings of the National Academy of Sciences U.S.A. 97:5313-5316.

Ting, C.-T., S.-C. Tsaur, S. Sun, W. E. Browne, Y.-C. Chen et al. 2004. Gene duplication and speciation in *Drosophila*:evidence from the Odysseus locus. Proceedings of the National Academy of Sciences U.S.A. 101:12232-12235.

Turelli, M., and H. A. Orr. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154:1663-1679.

Wu, C. I., and C. T. Ting. 2004. Genes and speciation. Nature Reviews Genetics 5:114-122.

Chapter 4

The Utility of Diagnostic Markers for Estimating Hybrid Fitness Components in Natural Hybrid Zones

Abstract. The majority of data on postzygotic isolation come from laboratory studies of intrinsic genetic incompatibilities. A thorough understanding of speciation, however, will require quantification of extrinsic postzygotic isolation in natural hybrid zones. I present a new statistical approach to estimating hybrid fitness in the wild, using multilocus genotype data taken from individuals sampled at two or more stages of the life cycle. After developing the statistical model, I apply maximum likelihood and Bayesian estimators to simulated data in order to determine the feasibility of obtaining precise viability estimates. I find that accurate estimates of hybrid viability can be obtained with a few hundred individuals and four or more diagnostic loci, as long as the degree of pre- and postzygotic isolation is not substantial and the number of hybrid classes is not large. By applying this new method to a mussel data set from the *Mytilus edulis-M. trossulus* hybrid zone, I provide new insight into the forces maintaining isolation of these species.

4.1 INTRODUCTION

A thorough understanding of speciation requires knowledge of how pre- and postzygotic isolation evolve in natural systems. Although intrinsic postzygotic isolation can be quantified using lab experiments (Coyne and Orr, 2004), extrinsic isolation is much more difficult to study, and therefore has only been quantified in a small number of systems (e.g. Helbig, 1991; Wang et al., 1997; Rundle, 2002; Höbel and Gerhardt, 2003). One reason extrinsic postzygotic isolation is difficult to quantify is that estimates are context dependent; extrinsic isolation can vary seasonally (Grant and Grant, 1993) and geographically (Fitzpatrick and Shaffer, 2004). Estimating both forms of isolation in a large number of species pairs will be required to answer the longstanding debate regarding the relative importance of extrinsic and intrinsic isolation in the process of speciation.

The increasing availability of diagnostic nuclear markers offers promise for researchers interesting in quantifying isolation in natural systems. In addition to their application to studying cline shape (Barton and Hewitt 1989), nuclear markers have been used to estimate the hybrid class identity of individuals sampled in hybrid zones (Pritchard et al., 2000; Anderson and Thompson, 2002). This information has been combined with traditional statistical approaches to determine whether selection is

operating against hybrids (Peterson et al., 2005), if it varies geographically (Fitzpatrick and Shaffer, 2004), and if it varies throughout the life cycle (Toro et al., 2004). Unfortunately traditional methods do not provide estimates of relative fitness for individuals from different hybrid classes (e.g. F1, F2, and backcross hybrids). Yet, hybrid fitness is an important measure isolation between populations in the process of speciation and can be used to elucidate the potential factors that maintain this isolation.

The most straightforward way to estimate fitness in natural systems is to employ a censusing approach. An estimate of the relative viability of F1 hybrids between two stages of the life cycle, for example, could be obtained by sampling individuals at the two stages, estimating the frequency of F1 hybrids in each sample, and taking the ratio of the frequencies at the two points in time. One difficulty with this approach is that the hybrid class to which a particular individual belongs is rarely known. Diagnostic nuclear markers, however, can be used to resolve hybrid class identity. Some studies suggest that as few as four diagnostic markers are needed to accurately identify F1 hybrids, although identification of advanced hybrids is considerably more difficult (Boecklen and Howard, 1997; Vähä and Primmer, 2006). Methods for estimating hybrid fitness from frequency information should take this uncertainty into account.

Here, I present a statistical model that can be used to estimate hybrid fitness components from the multilocus genotypes of individuals collected at two or more stages of a life cycle. The approach employed is to estimate the hybrid class frequencies at each stage, then to compute fitnesses from the frequency estimates. I use two likelihood-based estimators and simulated data to study the effect of five factors on the accuracy of viability estimates: the degree of prezygotic isolation, the degree of postzygotic isolation, the number of individuals sampled, the number of loci sampled, and hybrid class. I then apply the method to a published, 1401-sample mussel data set and estimate the strength of selection against hybrids during two periods of their life cycle. I conclude by discussing possible limitations of the method and how the method can be extended to alleviate these limitations.

4.2 METHODS

Assumptions. Three components are developed in this section: 1) the statistical model used to calculate likelihood, 2) the equilibrium equations used to simulate data, 3) the computational approaches (implementation) used to estimate hybrid class frequencies. The assumptions made for each component are outlined in Table 4.1 and detailed in the appropriate section. Three general assumptions are also made: First, the sample is

assumed to be representative of the population of interest (both geographically and temporally). Second, individuals are assumed to be sampled randomly with respect to their hybrid class identity. Third, I assume that migration (between the time points sampled) does not affect the relative frequencies of individuals in the different hybrid classes.

Likelihood Functions. Here, I derive equations describing the likelihood of obtaining a set of multilocus genotypes, given the hybrid class frequencies (the parameters of interest). I make several simplifying assumptions. The data are the genotypes of n_I individuals that have been sampled randomly from a hybrid zone and genotyped at n_L loci. Each locus is assumed to segregate two codominant alleles, 0 and 1, which are assumed to be fixed in the A and B parental species, respectively. The genotype at each locus can take the values 0, 1, or 2, where 0 denotes an individual homozygous for the 0 allele, 1 denotes a heterozygous individual, and 2 denotes an individual homozygous for the 1 allele. The markers are assumed to be unlinked and free of direct selection. Let **G** indicate the full genotype data matrix (which has dimension n_I x n_L) and $\mathbf{g}_i$ denote the genotype of individual *i*. The eight-locus genotype for the 6th individual sampled, for example, might take the form $\mathbf{g}_6 = \{0, 2, 1, 2, 1, 0, 2, 1\}$. Also, let **p** denote a vector of n_H hybrid class frequencies that sum to one. For example, p_1 might indicate the proportion of pure individuals from species A when the sample was taken. Lastly, let **a**

represent a vector of allele frequencies, where a_i indicates the expected frequency of the 1 allele in the *i*th hybrid class.

Under these assumptions, the likelihood of obtaining the data matrix, given the hybrid class frequencies, is

$$L(\mathbf{G} \mid \mathbf{p}, \mathbf{h}, \mathbf{a}) = \prod_i \sum_j p_j P(\mathbf{g}_i \mid h_i = j), \tag{4.1}$$

where h_i denotes the hybrid class identity of individual *i*. The maximum likelihood estimate of the vector of hybrid class frequencies, $\hat{\mathbf{h}}$, contains the values that maximize the likelihood of observing **G**. The product in Equation 4.1 is taken over the n_{I} sampled individuals and the sum is taken over the n_{H} hybrid classes to which each individual could potentially belong. Since the loci are assumed to be unlinked, the conditional probability can be calculated by taking the product of the probabilities for each locus:

$$P(\mathbf{g}_i \mid h_i = j) = \prod_l P(g_{il} \mid a_{\mathrm{f}_j}, a_{\mathrm{m}_j}) \tag{4.2}$$

where a_{f_j} and a_{m_j} are the frequencies of the 1 allele in the father and mother of an individual from the *j*th hybrid class, respectively. Note that Equation 4.2 may hold even

when linkage disequilibrium exists in the population, so long as linkage disequilibrium does not exist within a hybrid class. This can be seen in the conditional probability on the left side of the equation. Since the probability of observing the genotype $\mathbf{g}_i$ is conditioned on the hybrid class identity of individual *i*, linkage disequilibrium can be generated through differential fitness of individuals from different hybrid classes, so long as individuals with the same hybrid class (but potentially different genotypes) have equal fitness. Finally, one can compute the probability of observing a genotype at locus *l* in individual *i* by assuming Mendelian inheritance:

$$P(g_{il} \mid a_{f_j}, a_{m_j}) = \begin{cases} (1 - a_{f_j})(1 - a_{m_j}) & \text{if } g_{il} = 0, \\ a_{f_j}(1 - a_{m_j}) + (1 - a_{f_j})\, a_{m_j} & \text{if } g_{il} = 1, \\ a_{f_j}\, a_{m_j} & \text{if } g_{il} = 2. \end{cases} \tag{4.3}$$

Analysis of Simulated Data. Simulated data were used to identify the conditions in which accurate estimation of hybrid class viability is feasible. Data sets were simulated under twenty five combinations of pre- and postzygotic isolation. Prezygotic isolation was assumed to operate through preference of pure individuals for conspecifics. A proportion of pure individuals, C_P, were assumed to have a preference for conspecifics and thus only mated with individuals from the same species. All other individuals mated randomly, except for second generation hybrids, which did not mate. The degree of

prezygotic isolation (C_P) took the value 0.0, 0.2, 0.4, 0.6, or 0.8. Postzygotic isolation was controlled through a viability parameter, V_h, which determined the viability of hybrids relative to pure individuals. Individuals from different hybrid classes were assumed to have the same viability. The degree of postzygotic isolation ($1 - V_h$) took the value 0.0, 0.2, 0.4, 0.6, or 0.8.

The statistical approach developed above can be used to estimate the frequencies of individuals in different hybrid classes (see below). To estimate viability of individuals in a particular class, frequencies estimated from samples taken at two time points are needed. Accordingly, for each combination of pre- and postzygotic isolation, I simulated two data matrices, one assuming equilibrium frequencies before viability selection (calculated using Equation F5 derived in the Appendix F) and one assuming equilibrium frequencies after viability selection (calculated using Equation F7). Each data matrix contained 1024 individuals and 32 loci. Individuals were first randomly assigned a hybrid class identity based on the equilibrium frequencies then randomly assigned genotypes using expected genotype frequencies for that hybrid class. Again, no assumptions were made regarding linkage equilibrium of the population as a whole. The above procedure was repeated to produce fifty replicate pairs of data matrices.

In order to understand the effects of sample size on the accuracy of fitness estimates, each data set was subsampled in twenty-five different ways. Each subsample

included 64, 128, 256, 512, or 1024 individuals and 2, 4, 8, 16, or 32 loci. In all, a total of 62,500 data sets were analyzed using the maximum likelihood and Bayesian approached outlined below.

Maximum Likelihood Estimation. Hybrid class frequencies ($\hat{\mathbf{h}}$) can be estimated by finding the frequencies that maximize the likelihood of observing the genotypes in the data set. Maximum likelihood frequencies were estimated using the following algorithm: 1) choose starting frequencies for the hybrid classes, 2) compute the likelihood of observing the data given the current hybrid class frequencies, 3) randomly choose two hybrid classes, 4) hold the other frequencies constant and find the optimum frequencies for the chosen classes using the Newton-Raphson optimization method (Ypma 1995), 5) repeat steps 2-4 until the change in the likelihood score between iterations is less than 10^{-12} for thirty consecutive iterations.

To increase the chance of finding the global solution, each maximum likelihood analysis was performed twice, once using random starting frequencies drawn from a flat Dirichlet distribution, and once using rough approximates for the frequencies. Rough approximates of the hybrid class frequencies were obtained by computing the frequency of sampled individuals that were homozygous at all loci, assigning one half of this value to each of the two pure classes, and distributing the remaining frequency evenly across the other hybrid classes. The rough approximation method is one way of quickly

producing starting points that tend to be closer to the maximum likelihood estimate. After both analyses were performed, final frequency estimates were taken from the analysis that produced the highest likelihood score.

Bayesian Estimation. Hybrid class frequencies can also be estimated using a Bayesian approach. Because analytic solutions are not available, I used a Markov chain Monte Carlo (MCMC) approach to estimate the posterior probability distribution. A flat Dirichlet distribution was assumed as a prior on hybrid class frequencies for all analyses. This prior reflects ignorance of the hybrid class frequencies. The Markov chain was initiated with equal starting frequencies for the six classes. At each generation, two hybrid classes were randomly chosen and new values for the frequencies of those classes were proposed in the following way: 1) a random number was drawn from a Beta distribution with $\alpha = 1/5$ and $\beta = 5$, 2) this value was then added to the frequency of one of the two chosen hybrid classes and subtracted from the frequency of the other chosen class. Proposals were rejected if the value of either frequency did not fall between zero and one. Proposals were also rejected if a uniform random number between zero and one was greater than the ratio of the likelihoods of the proposed and current states.

The posterior distribution was estimated using samples taken from the Markov chain every 50 generations. After the chain was run for 50,000 generations, the

likelihood scores of the samples were analyzed to identify the first sample in which the likelihood score was greater than the median likelihood of the following samples. Samples taken before this point were discarded as burn-in because they were likely biased by the starting condition. A point estimate of the hybrid class frequencies was then obtained from the post-burnin samples by computing the median of the marginal distribution for each class and normalizing the median frequencies.

Analysis of Empirical Data. I applied the Bayesian estimator described above to a data set that was previously analyzed by Toro et al. (2004). These workers sampled 1401 mussels from a hybrid zone between *Mytilus edulis* and *M. trossulus* in Trinity Bay on the east coast of Newfoundland and genotyped each individual at two diagnostic nuclear markers: *ITS* and *Glu-5'*. Individuals were assigned to three age classes, (larvae, juveniles, and adults), based on shell length. Because samples were not collected by following a single cohort, I assume that hybrid class frequencies are at equilibrium. I address two questions: 1) do viabilities of pure and hybrid individuals vary during the life cycle? and 2) are hybrids disfavored, relative to pure individuals?

Hybrid class frequencies, in the form of posterior probability distributions, were estimated separately for larval, juvenile, and adult samples. Markov chains were run for 10^6 generations and sampled every 50 generations. Viabilities for early and late periods of the life cycle were then estimated from the frequency posterior distributions. The

viability posterior distribution was obtained from two frequency posterior distributions by: 1) randomly selecting, without replacement, one sample from each of the two frequency distributions, 2) computing, for each hybrid class, the ratio of frequencies found in the two samples, and 3) repeating steps 1-2 until all samples had been drawn. The early viabilities were calculated by dividing juvenile frequencies by larval frequencies. Likewise, the late viabilities were calculated by dividing adult frequencies by the juvenile frequencies.

4.3 RESULTS

Accuracy of Fitness Estimates. The fitness of individuals from different hybrid classes can be accurately estimated provided that a sufficiently large number of individuals and loci are sampled. Figure 4.1 presents the error in the fitness estimates as a function of the hybrid class of interest, the degree of prezygotic isolation, the degree of postzygotic isolation, the number of individuals sampled, and the number of loci sampled. Here, error is computed as the absolute difference between the true and estimated viabilities divided by the true viability. Error decreased for pure individuals (Figure 4.1a) as the degree of pre- and postzygotic isolation increased, and as the number of individuals sampled increased. The number of loci genotyped had relatively small effect on the

amount of error. Error rates for pure individuals were quite low, on the order of 10% or less, when the number in individuals sampled was 256 or larger.

The fitness of F1 hybrids can also be accurately estimated (Figure 4.1b). In this case, however, the error decreased as the degree of prezygotic isolation *increased* or the degree of postzygotic isolation *decreased*. The number of loci and individuals sampled had a larger effect on the error of the fitness estimates of F1 hybrids than that of pure individuals. Nonetheless, error in F1 fitness estimates was quite low (on the order of 10%) when the degree of postzygotic isolation was less than 0.6 and at least 256 individuals and 4 loci were sampled.

The fitness of backcross and F2 hybrids is more difficult to estimate (Figure 4.1c-d). Error in these cases increased substantially as the degree of pre- and postzygotic isolation increased. The number of loci and individuals sampled also had substantial effect on the error. For both types of hybrids, error was generally quite high (> %30) unless the number of loci exceeded 2 and the number of individuals sampled exceeded 256. Estimates of advanced hybrid fitness, however, is still quite accurate for a large portion of parameter space. Error for backcross fitness estimates ranged between 10% and 20% when the degrees of pre- and postzygotic isolation were less than 0.8, the number of loci was greater than 2, and the number of individuals was greater than 256. Error rates for F2 hybrids, however, generally exceeded 30% unless the degrees of pre-

and postzygotic isolation were less than 0.6, the number of loci was greater than 2 and the number of individuals was 512 or greater.

Estimator Efficiency. The ability of both maximum likelihood and Bayesian methods to accurately estimate observed hybrid frequencies depends strongly on the number of loci sampled but not on the number of individuals sampled. Figures 4.2a-d present the relationship between the frequencies estimated using maximum likelihood (Figure 4.2a) or Bayesian (Figure 4.2b-d) methods and the frequencies actually sampled (observed) for all of the data sets analyzed. Comparing the estimated frequencies to the observed frequencies, instead of the true frequencies (i.e. the frequencies used to simulate the data), allows estimator error to be separated by sampling error. Both methods performed quite well, except when the number of loci fell below 8 (Figure 4.2a-b). When the number of loci was 2 or 4, both methods produced substantial error. In general, the error was larger for the maximum likelihood estimator (2a) than for the Bayesian estimator (2b). As indicated by Figure 4.2c, estimates of F2 frequency were positively biased at the expense of backcross frequencies, which were negatively biased. This bias disappeared when more than 2 loci were available. This result suggests that when few loci are available, combining F2 and backcross frequencies into a single "advanced" hybrid class may be more appropriate than interpreting them separately. Lastly, Figure

4.2d suggests that sampling more individuals, though it may reduce sampling error, does not improve the efficiency of the estimators.

Maximum Likelihood vs. Bayesian Estimators. Maximum likelihood and Bayesian estimates of hybrid frequency were very similar when a sufficient number of individuals or loci were available. In Figure 4.3, the Bayesian frequency estimates are plotted against the maximum likelihood estimates. Two interesting patterns are worth noting. The first pattern is that when a small amount of information is available to estimate the hybrid frequencies (loci = 2, individuals = 64), the Bayesian estimates of F2 and backcross frequencies appear to be influenced by the flat prior assumed in the analysis. This result is suggested by the fact that frequencies less than 1/6 are positively biased and those greater than 1/6 are negatively biased, relative to maximum likelihood estimates. This bias disappears when *either* the number of loci *or* the number of individuals is increased. The second interesting pattern is that when the number of loci is small (2), and the number of individuals is fairly large, the two estimators produce very similar estimates despite producing positively biased estimates of F2 hybrid frequency (compare Figures 4.2c and 4.3b). This bias is absent when the number of loci is large (regardless of the number of individuals).

Empirical Example. Estimates of hybrid class frequencies in the *Mytilus edulis-M. trossulus* hybrid zone indicate that F1 hybrids are very rare but that advanced hybrids are

relatively more common. This result supports the findings of Toro et al. (2004). Marginal posterior density estimates are presented in Figure 4.4 for the three time points. Because mussels were only genotyped at two loci, and estimates of F2 and backcross frequencies are known to be biased when few loci are available (see Estimator Efficiency), frequencies of second generation hybrids were summed as an estimate of "advanced" hybrid frequency.

Viabilities varied significantly during the life cycle for pure *M. edulis* (posterior probability (pp) = 0.99966), *M. trossulus* (pp = 1.00000), and advanced hybrids (pp = 0.99750), but not for F1 hybrids (pp = 0.6609848). Marginal density estimates for early and late viability are given in Figure 4.5. Pure *M. trossulus* were favored early in the life cycle (pp = 0.97599) but disfavored late in the life cycle (pp = 1.00000). In contrast, advanced hybrids were favored early in the life cycle (pp = 0.98997) but disfavored late in the life cycle (pp = 0.96159). There was no evidence for selection favoring *M. edulis* early in the life cycle (pp = 0.508635), but results suggest that pure *M. edulis* individuals were favored late in the life cycle (pp = 1.00000). There was also no evidence for selection against F1 hybrids (early pp = 0.36891; late pp = 0.61539). This result is not surprising in light of the high degree of uncertainty in the viability of F1 hybrids, which is a result of the low frequency of this class.

Advanced hybrids had lower viability than pure individuals early in the life cycle (vs. pure *M. trossulus*, pp = 0.99734; vs. pure *M. edulis*, pp = 0.88792). Advanced hybrid viability late in the life cycle exceeded that of pure *M. trossulus* (pp = 1.00000) but was less than that of pure *M. edulis* (pp = 0.99740).

4.4 DISCUSSION

I have presented a statistical method of estimating hybrid fitness components using diagnostic markers. Four results emerged from analysis of the simulated data. First, accurate estimates of pure and early hybrid viability can be accurately estimated with as few as 256 individuals and 4 loci per sampled stage. More individuals are needed if the hybrid class of interest is in low frequency. Second, when fewer than four loci are available, estimates of F2 frequency tend to be positively biased, whereas estimates of backcross frequency tend to be negatively biased. Third, increasing the degree of pre- and postzygotic isolation (which reduces the abundance of hybrids and increases the abundance of pure individuals) results in less precise viability estimates for hybrids (especially advanced hybrids), but more precise viability estimates for pure individuals. This suggests that hybrid fitness components may be most accurately estimated during early stages of speciation. Lastly, maximum likelihood and Bayesian

estimates of hybrid class frequency are similar, though Bayesian estimates can be influenced by the prior if insufficient data exist.

Analysis of the 1401-sample mussel data set provides new insight into the forces maintaining *Mytilus edulis* and *M. trossulus* species identity. First, although Toro et al. (2004) were not able to detect selection on hybrids, I was able to find significant evidence for selection against advanced hybrids early in the life cycle and selection favoring advanced hybrids late in the life cycle. Second, Toro et al. concluded that hybrid viability may be intermediate between the two parental species. Though I found early viability of advanced hybrid to be intermediate, I found late viability of advanced hybrids to be lower than that of either species (although only significantly lower than *M. trossulus*). This suggests that selection against advanced hybrids may contribute to postzygotic isolation in this system.

The method developed here could be extended in a number of ways. One way is to extend the estimators to accommodate more hybrid classes. A reversible-jump MCMC approach, for example, could be used to allow flexibility in the number of hybrid classes. A second way to improve the method is to relax the assumption that all loci are diagnostic following the approach taken by Anderson and Thompson (2002), in which allele frequencies in allopatry are estimated as nuisance parameters. Finally, the method could be improved by relaxing the assumption that all the loci are unlinked,

which may not be appropriate for some data sets. In this case one could add additional parameters to the model that could account for the rate of recombination between two or more loci.

For some organisms, such as anurans, mated pairs can be readily obtained. Information from the genotypes of mated pairs could be used in combination with adults (just prior to mating) to study prezygotic isolation and sexual selection against hybrids. Samples of mated pairs and early zygotes could be used to estimate the degree of postmating prezygotic isolation, such as sterility or gametic incompatibility. In either case, the method developed here would need to be extended such that the frequencies of hybrid class pairings could be estimated.

Comparing viability estimates from samples taken at different geographic locations may also allow researchers to study geographic variation in pre- and postzygotic isolation. These estimates, in combination with ecological data, could then provide important insight into the extrinsic factors that contribute to speciation. In addition, comparing viability estimates from samples taken during different points in time may allow researchers to study temporal variation in isolation resulting from seasonal climate change, for example.

Given the growing availability of diagnostic nuclear markers, and the paucity of studies estimating extrinsic postzygotic isolation, it is likely that future speciation research will increasingly focus on analysis of natural hybrid zones. The method developed here offers a new approach to studying temporal and geographic variation in pre- and postzygotic isolation. Studying reproductive isolation in this way is a necessary step towards a comprehensive understanding of the process of speciation.

Table 4.1. Assumptions made during the development of the three components of the study. Note that many of the restrictive assumptions only pertain to the simulated data.

Type of Assumption	Statistical Model	Simulated Data	Implementation
Hybrid Classes	--	6	6
Migration	--	frequencies unchanged	--
Selection on Markers	none	none	--
Recombination	free	free	free
Sampling	--	random	--
Type of Markers	diagnostic codominant	diagnostic codominant	diagnostic codominant
Mating	--	symmetric hybridization advanced hybrids do not mate	--
Viability	--	equal for pure individuals equal for hybrids	--

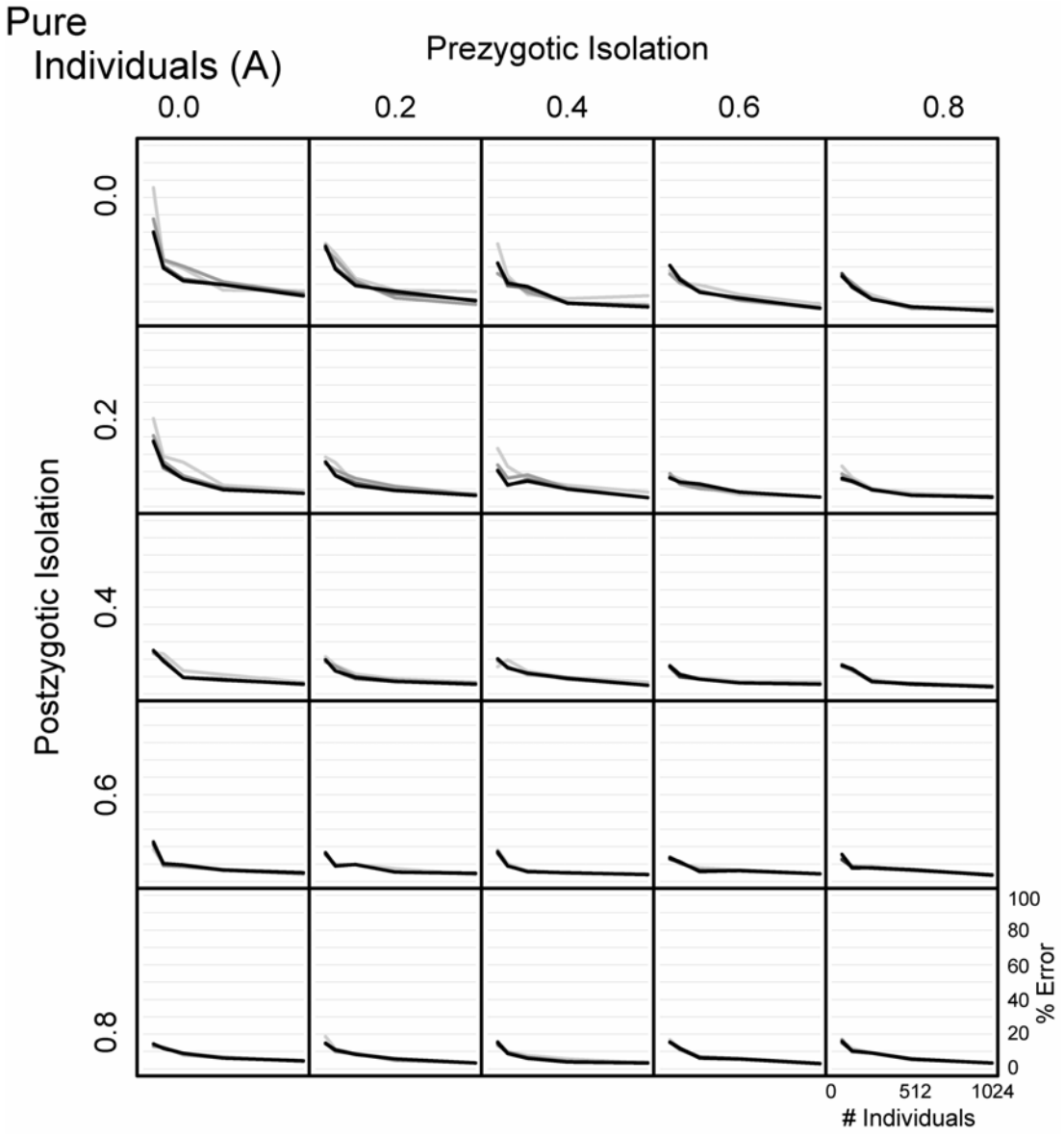

Figure 4.1. Total percent error in estimates of viability as a function of the hybrid class (panel), amount of prezygotic isolation (column), amount of postzygotic isolation (row), number of individuals (x-axis), and the number of loci (shade of curve). Results shown for 2, 4, 8, 16, and 32 loci are shaded from light grey to black, respectively. Error is calculated as (true viability – estimated viability)/(true viability). The curves represent the median error across 50 replicates.

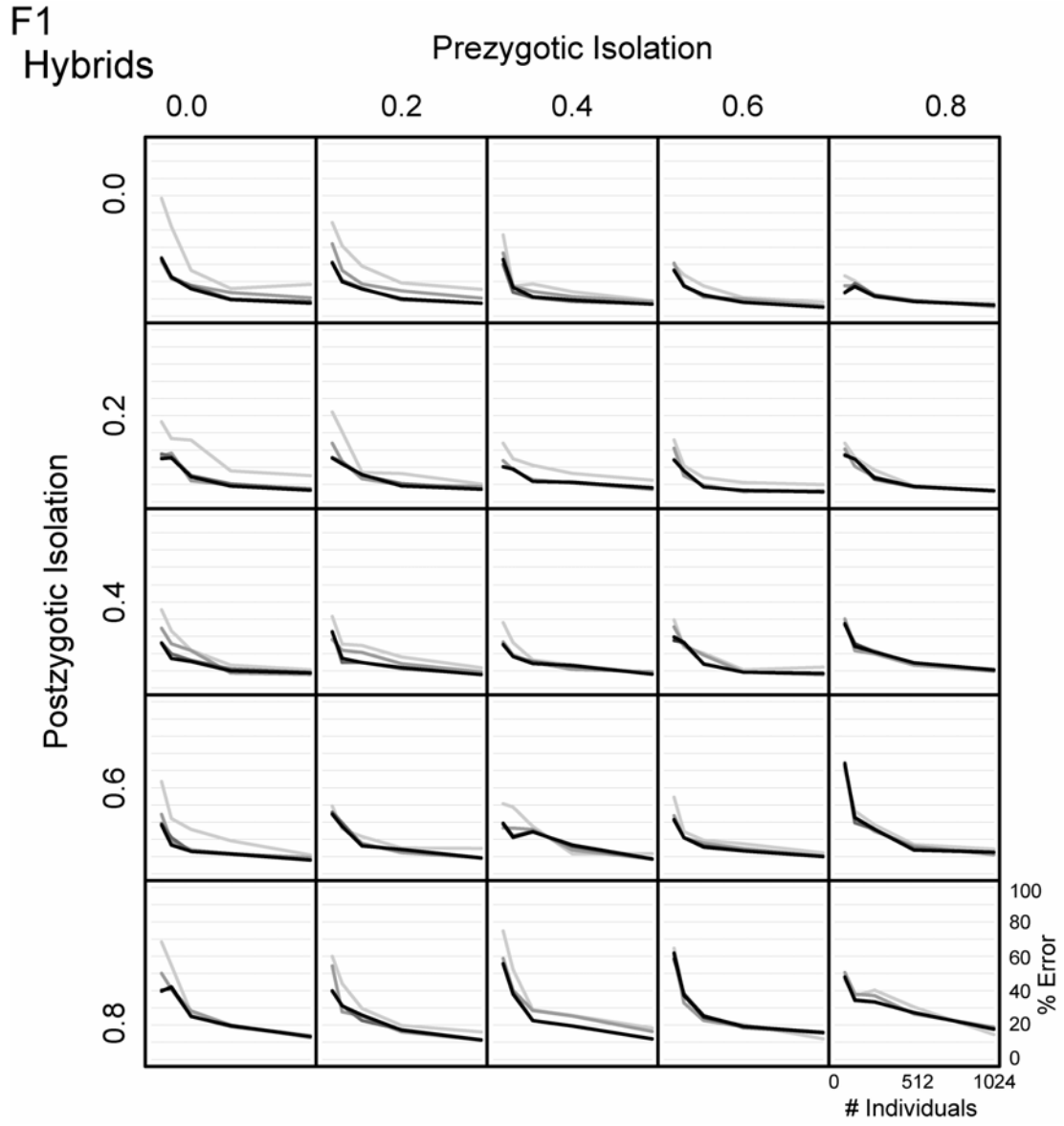
F1
Hybrids
Prezygotic Isolation
0.0
0.2
0.4
0.6
0.8
Postzygotic Isolation
0.0
0.2
0.4
0.6
0.8
100
80
60
40
20
0
% Error
0
512
1024
Individuals

Figure 4.1b.

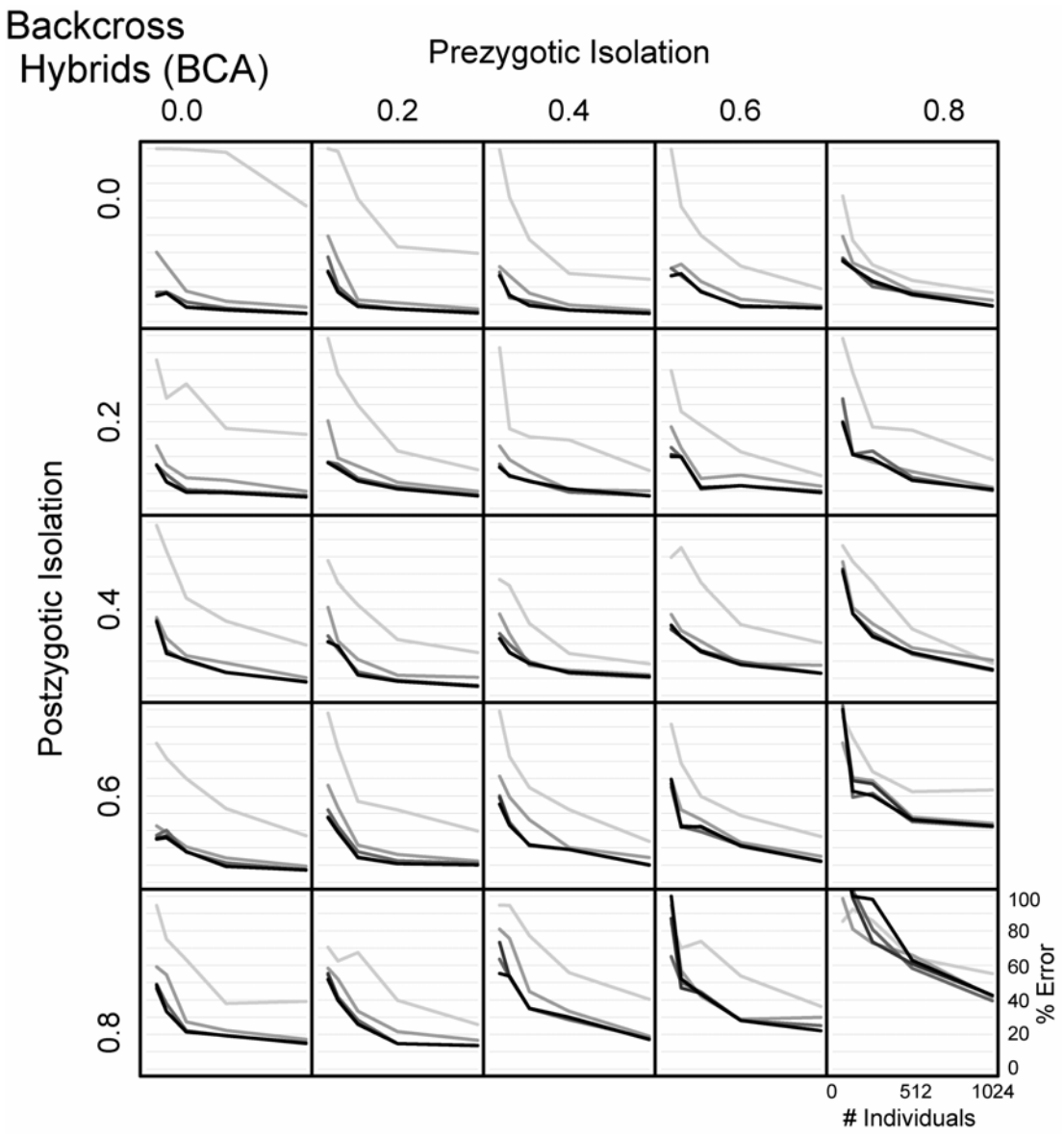
Backcross
Hybrids (BCA)
Prezygotic Isolation
0.0
0.2
0.4
0.6
0.8
Postzygotic Isolation
0.0
0.2
0.4
0.6
0.8
100
80
60
40
20
0
% Error
0
512
1024
Individuals

Figure 4.1c.

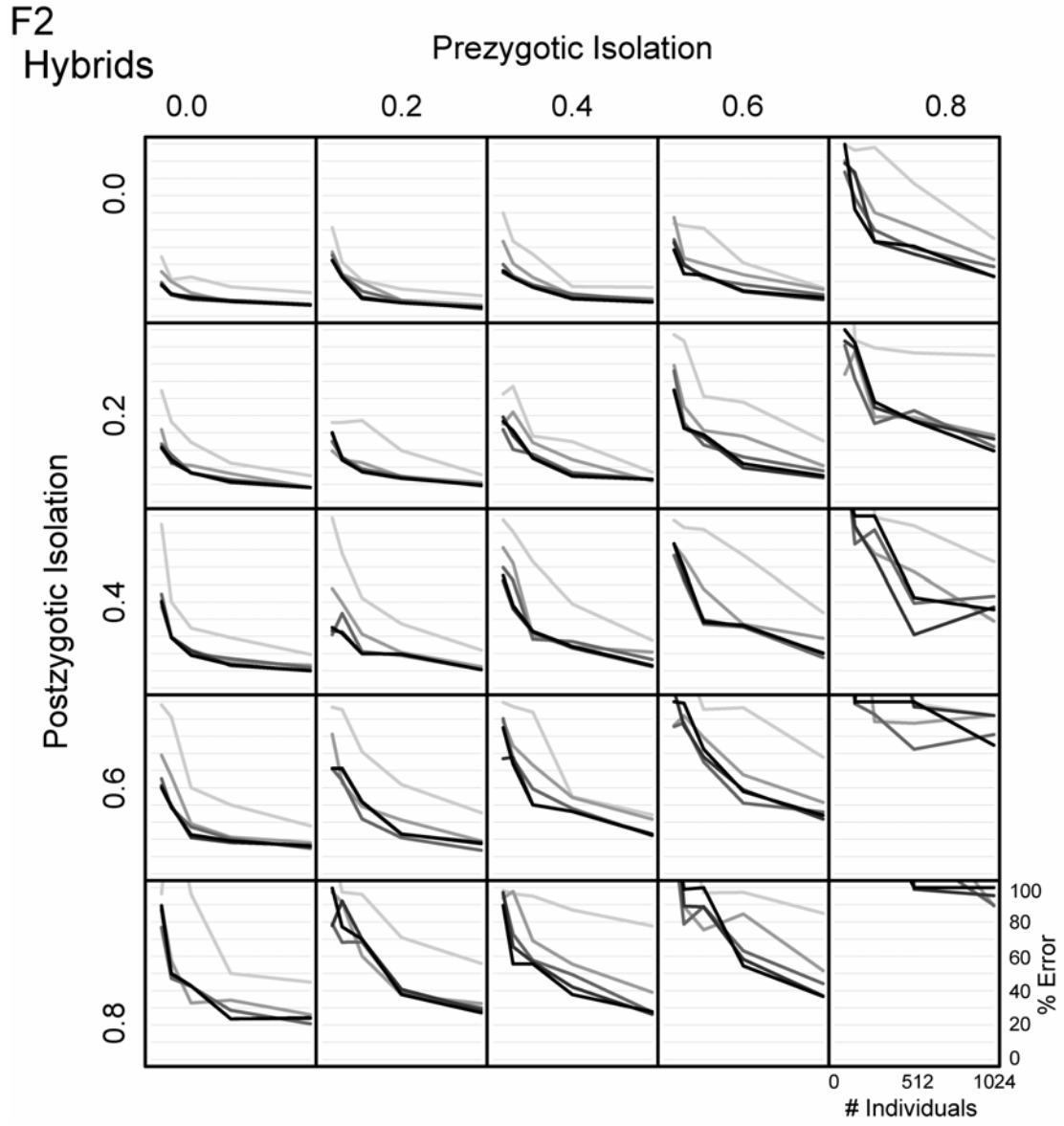
F2
Hybrids
Prezygotic Isolation
0.0
0.2
0.4
0.6
0.8
Postzygotic Isolation
0.0
0.2
0.4
0.6
0.8
100
80
60
40
20
0
% Error
0
512
1024
Individuals

Figure 4.1d.

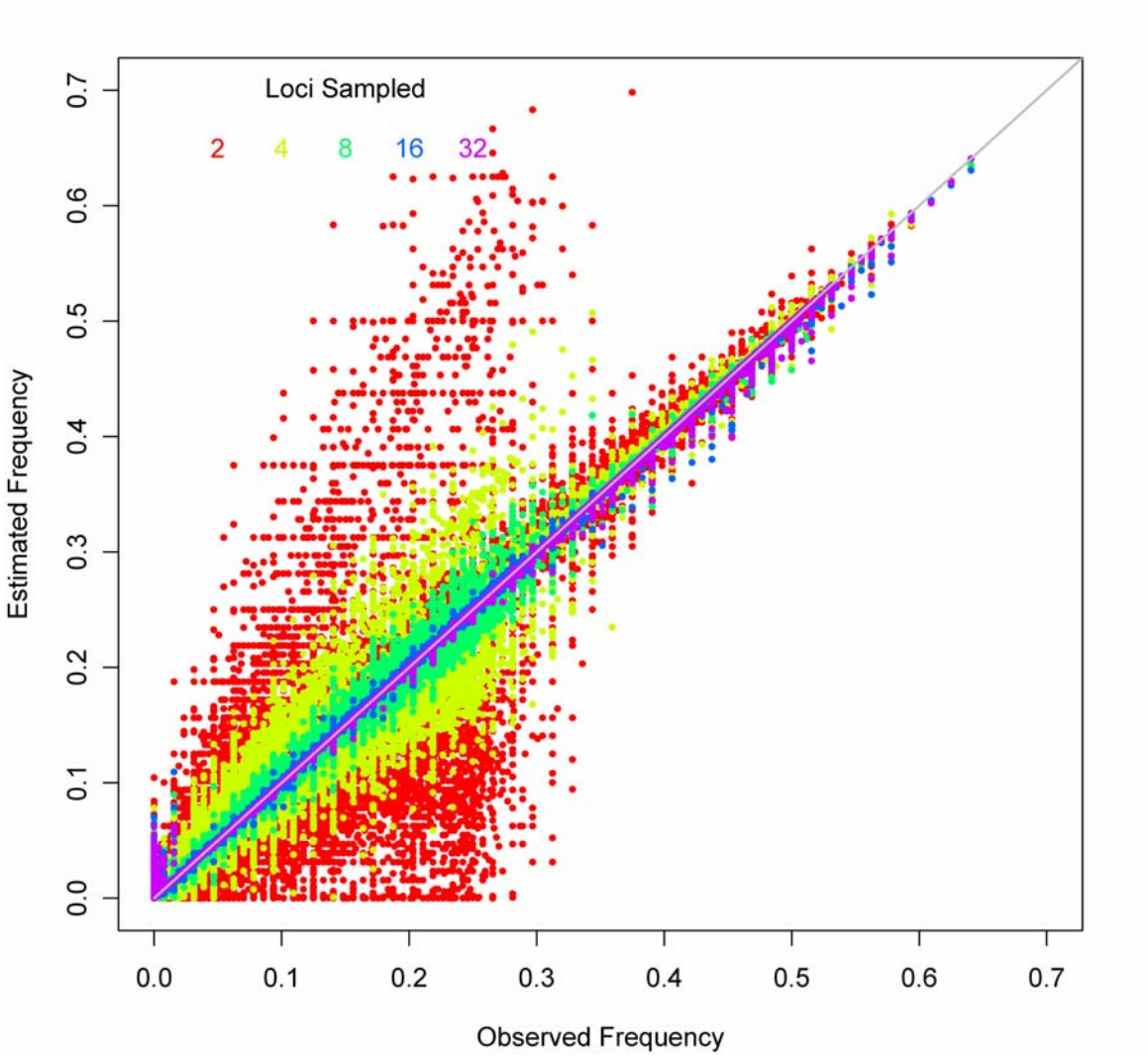

Figure 4.2. Estimator error as a function of the type of estimator, the number of loci, the hybrid class, and the number of individuals. In each graph, estimated frequencies are plotted against observed (sampled) frequencies for all analyses performed. In (a) and (b), which present results for the maximum likelihood and Bayesian estimators, respectively, the color of each point corresponds to the number of loci used in the analysis. In (c), each point is colored according to the hybrid class corresponding to that point. In (d), points are colored by the number of individuals used in the analysis. Note that the number of loci has the largest effect on estimator error whereas the number of individuals has the smallest effect. Also note that when the number of loci is small, F2 frequency estimates are positively biased at the expense of backcross frequency estimates (see panel c). Panels b-d present results from the Bayesian estimator.

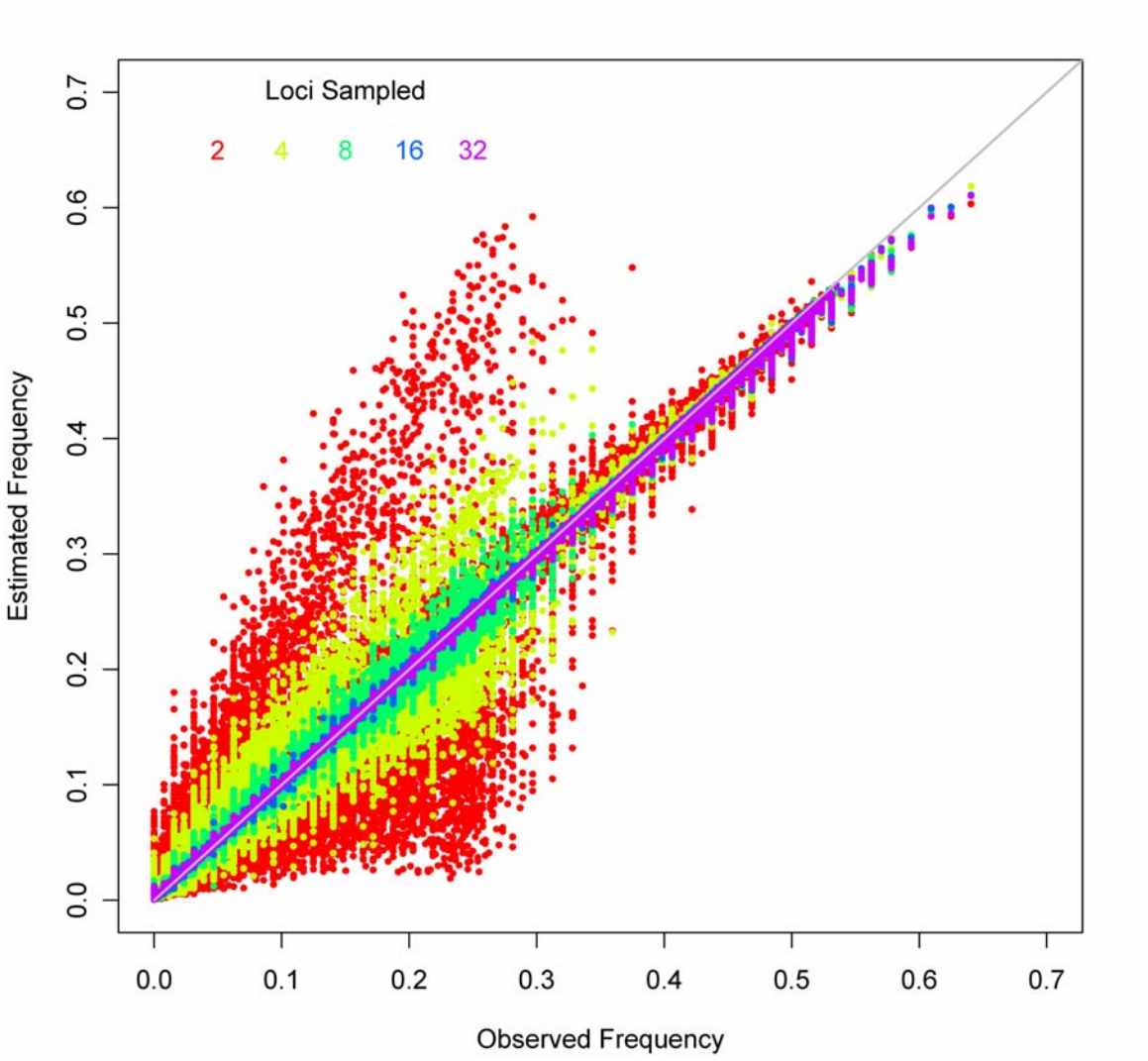
Loci Sampled
2
4
8
16
32
Estimated Frequency
Observed Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

Figure 4.2b.

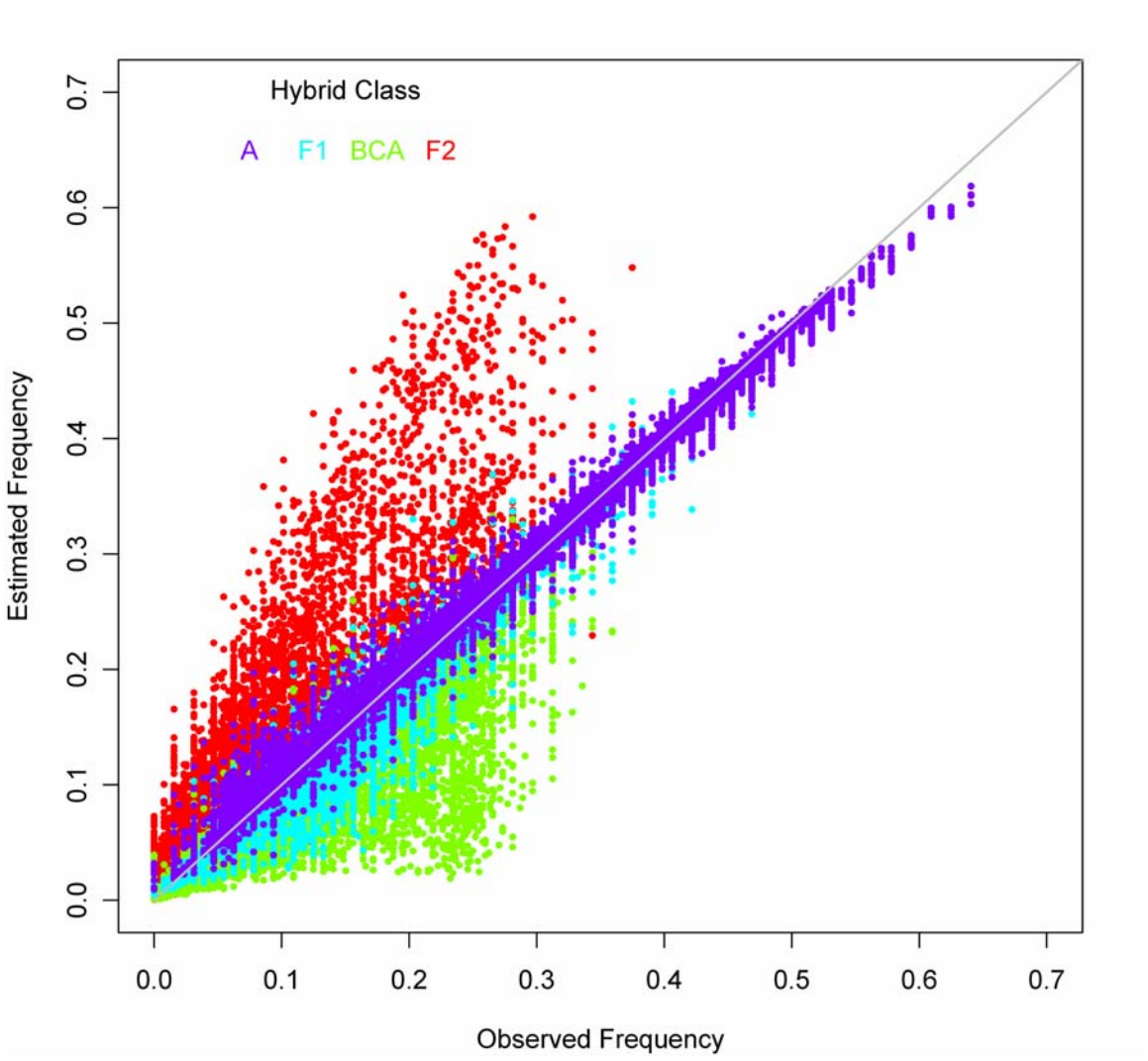
Hybrid Class
A
F1
BCA
F2
Estimated Frequency
Observed Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

Figure 4.2c.

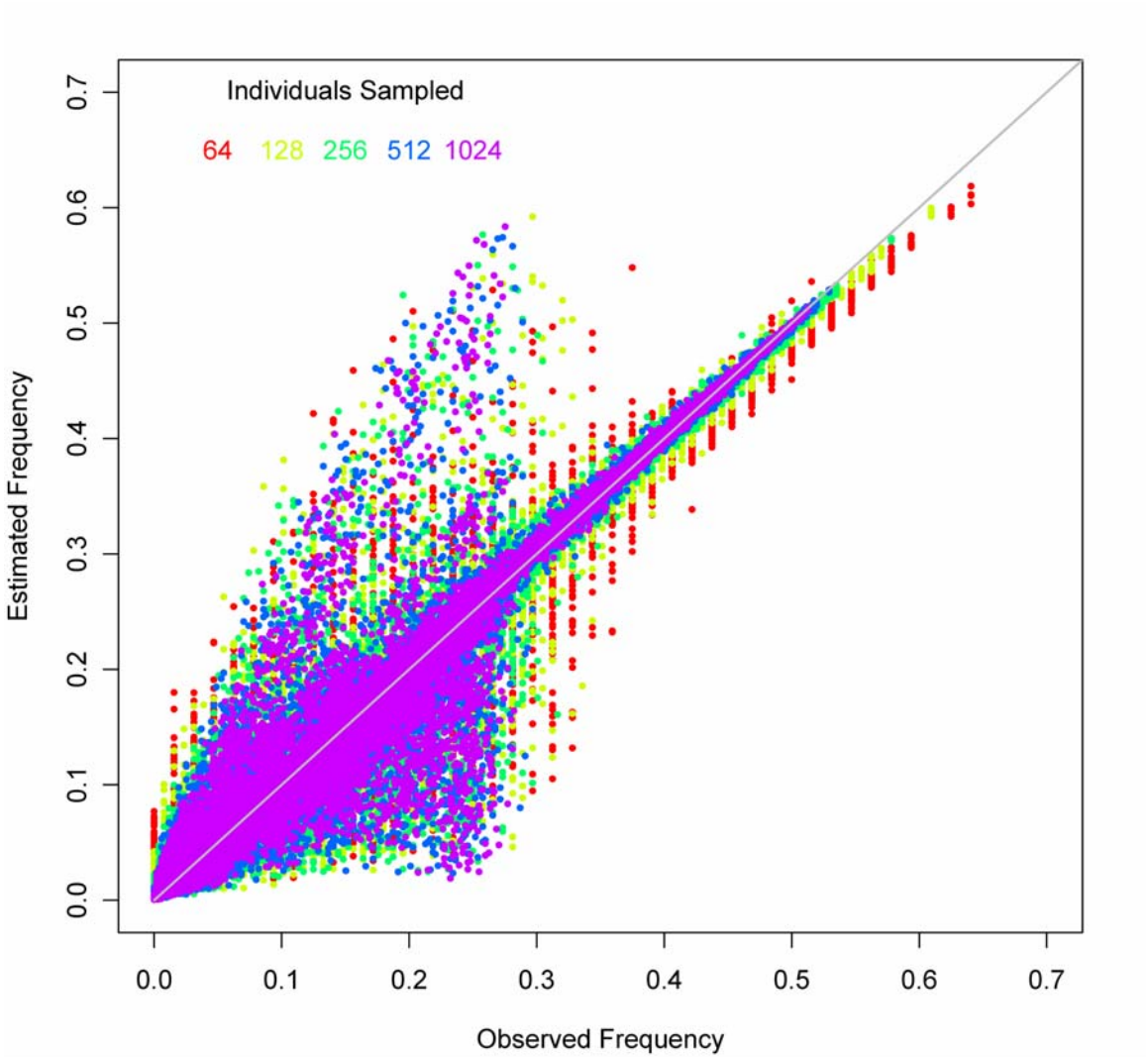
Individuals Sampled
64 128 256 512 1024
Estimated Frequency
Observed Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

Figure 4.2d.

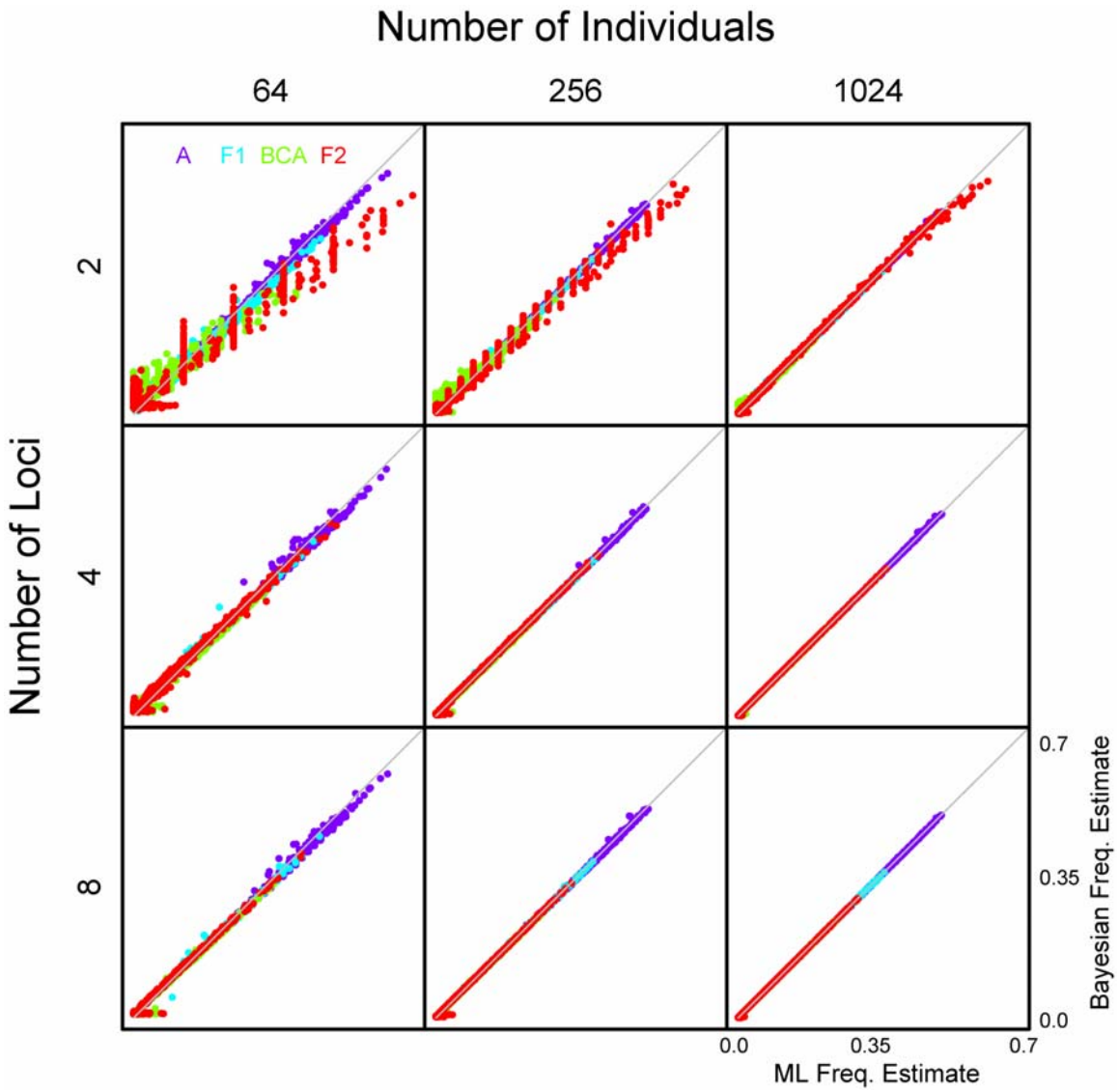

Figure 4.3. Maximum likelihood vs. Bayesian estimates of hybrid class frequencies. The nine panels present different combinations of the number of individuals and the number of loci used in the analysis. In each panel, the frequency estimates obtained from Bayesian analyses are plotted against the corresponding estimates obtained from maximum likelihood analyses. Note that while increasing the number of loci or individuals improves agreement between the two estimators, both estimators produce positively biased estimates of F2 hybrid frequency when the number of loci is small (compare this figure to Figure 4.2c). Also note that when the number of loci and individuals is small, Bayesian estimates of F2 frequency appear to be biased by the prior.

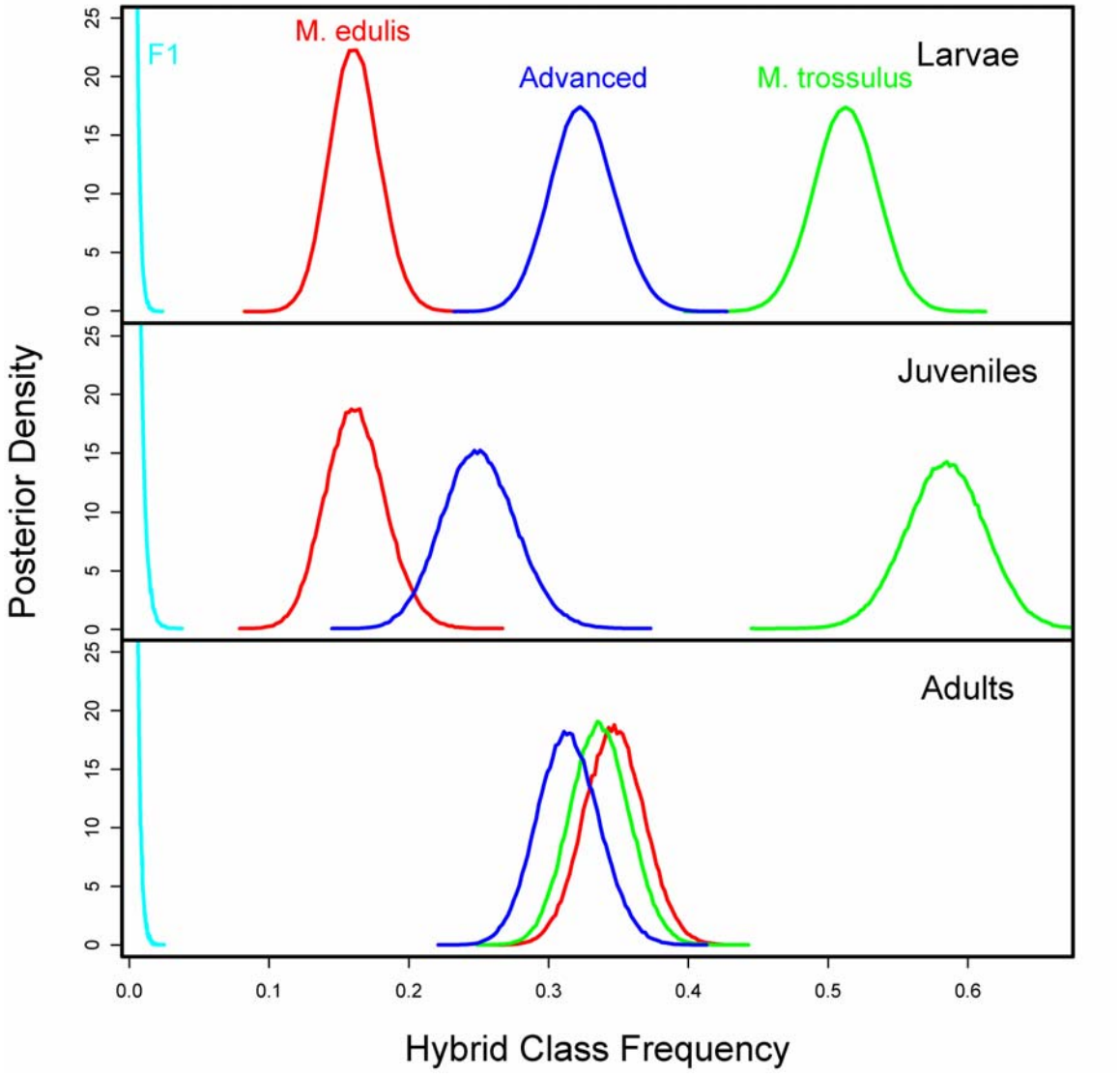

Figure 4.4. Estimates of pure *Mytilus edulis*, pure *M. trossulus*, F1, and advanced hybrid frequency in larvae, juvenile, and adults. At each of the 3 stages, the posterior probability distribution of hybrid class frequencies was estimated using MCMC. Each curve represents the marginal density for a particular class. Note that the class labeled Advanced includes three types of hybrids: F2, BCA and BCB.

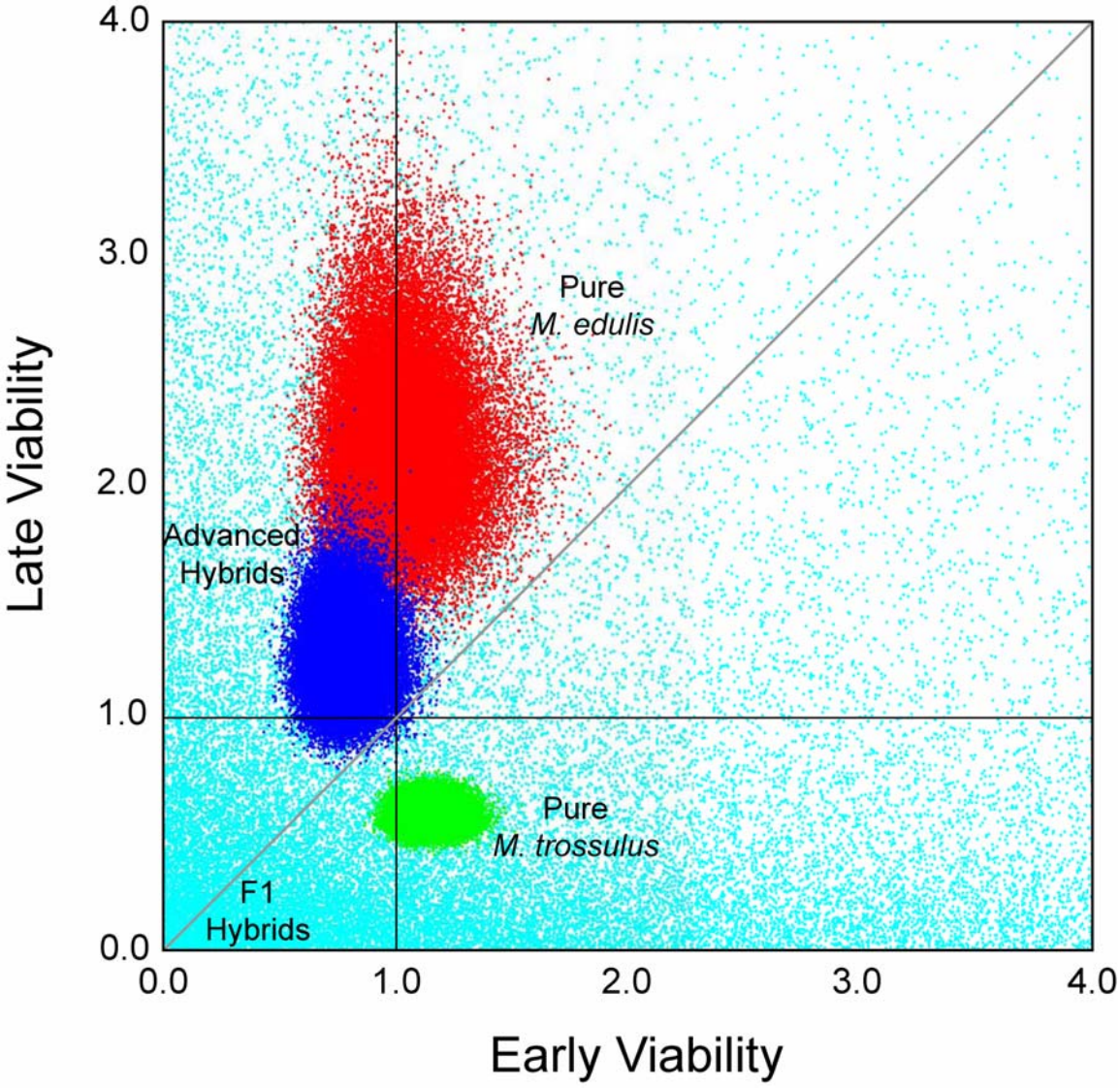

Figure 4.5. Estimates of pure *Mytilus edulis*, pure *M. trossulus*, F1, and advanced hybrid viability during two life history stages. Early viability was calculated as juvenile (2mm-15mm) frequency divided by larval frequency and late viability was calculated as adult (>15mm) frequency divided by juvenile frequency. Points of a particular color represent the marginal density estimated using MCMC. Points to the right of the vertical line indicate support for an increase in frequency during the early part of the life cycle whereas points above the horizontal line indicate support for an increase in frequency during the late part of the life cycle. Points below the diagonal indicate support for greater viability during the early part of the life cycle, whereas points below the diagonal line indicate support for greater viability during the late part of the life cycle. Note that the class labeled Advanced includes three types of hybrids: F2, BCA and BCB.

References

Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217-1229.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Annual Review of Ecology and Systematics 16:113-148.

Boecklen, W. J., and D. J. Howard. 1997. Genetic analysis of hybrid zones: numbers of markers and power of resolution. Ecology 78:2611-2616.

Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.

Fitzpatrick, B. M., and H. B. Shaffer. 2004. Environment-dependent admixture dynamics in a tiger salamander hybrid zone. Evolution 58:1282-1293.

Grant, B. R., and P. R. Grant. 1993. Evolution of Darwin's finches caused by a rare climatic event. Proceedings of the Royal Society of London B: Biology 251:111-117.

Helbig, A. J. 1991. Inheritance of migratory direction in a bird species:A cross-breeding experiment with SE and SW migrating blackcaps (*Sylvia atricapilla*). Behavioral Ecology and Sociobiology 28:9-12.

Höbel, G., and H. C. Gerhardt. 2003. Reproductive character displacement in the acoustic communication system of green tree frogs (*Hyla cinerea*). Evolution 57:894-904.

Peterson, M. A., K. J. Monsen, H. Pedersen, T. McFarland, and J. Bearden. 2005. Direct and indirect analysis of the fitness of *Chrysochus* (Coleoptera:Chrysomelidae) hybrids. Biological Journal of the Linnean Society 84:273-286.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Rundle, H. D. 2002. A test of ecologically dependent postmating isolation between sympatric sticklebacks. Evolution 56:322-329.

Toro, J., D. J. Innes, R. J. Thompson. 2004. Genetic variation among life-history stages of mussels in a *Mytilus edulis-M. trossulus* hybrid zone. Marine Biology 145:713-725.

Vähä, J-P., and C. R. Primmer. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Molecular Ecology 15:63-72.

Wang, H., E. D. McArthur, S. C. Sanderson, J. H. Graham, and D. C. Freeman. 1997. Narrow hybrid zone between two subspecies of big sagebrush (*Artemisia tridentata*:Asteraceae). IV. Reciprocal transplant experiments. Evolution 51:95-102.

Ypma, T. J. 1995. Historical development of the Newton-Raphson method. SIAM Review 37:531-551.

Appendix A

Model of Reinforcement

Introduction. We being by deriving a model of reinforcement that is general with regards to the form of selection against hybrids. We then apply the general model to specific types of hybrid incompatibility in order to determine how the pattern of sex linkage of female preference and hybrid incompatibility genes affects the potential for reinforcement.

The methods and notation follow Kirkpatrick et al. (2002; hereafter KJB), which can be consulted for further details. We say that genes carried by an individual occupy positions, with two positions at each locus. At locus *i*, for example, the gene carried by a female and inherited from a female (the individual's mother) is denoted i_{ff}, while the gene she inherited from a male (her father) is written i_{fm}. The sets of all positions that affect the female preference, male trait, and hybrid incompatibility are denoted $\mathbb{P}$, $\mathbb{T}$, and $\mathbb{H}$, respectively. The entire genome is denoted $\mathbb{W}$. The set of all positions in females that affect the preference is written $\mathbb{P}_f$; with *n* autosomal loci, there are 2*n* positions in this set. On the island, the frequency of allele 1 at position $\mathbb{i}$ (where $\mathbb{i}$ could represent i_{ff}, i_{fm}, i_{mf}, or i_{mm}) is denoted $p_{\mathbb{i}}$, with $q_{\mathbb{i}} = 1 - p_{\mathbb{i}}$. The notation is summarized in Table 3.1.

Central to our analysis is the "quasi-linkage equilibrium", or QLE, approximation that was introduced by Barton and Turelli (1991). When selection acting on individual loci and groups of loci is weak relative to rates of recombination, the associations (linkage disequilibrium) between loci will rapidly evolve to values that are small relative to their maximum possible values.

The General Model. We make the assumption that preference genes have additive effects (that is, that they show no dominance or epistasis) and no imprinting (that is, alleles inherited from mothers and fathers are expressed equally). Without making any further assumptions, we can then write the preference for an individual as

$$P = \overline{P} + \sum_{i \in \mathbb{P}_f} b_i \zeta_i, \qquad \text{(A1)}$$

where $\overline{P}$ is the mean preference among female zygotes, b_i is the difference in the preference caused by carrying allele 1 rather than allele 0 at locus i, and $\zeta_i = q_i$ if the female carries allele 1 at position i and $-p_i$ otherwise. The summation includes one term for each of the positions affecting the preference in females.

The change in the mean preference in females from the start of one generation to the next is

$$\Delta \overline{P} = \sum_{i \in P_f} b_i \Delta p_i \quad \text{(A2)}$$

where Δp_i is the change in allele frequency at position i from the start of one generation to the start of the next. This change is caused by forces acting within a generation, i.e. migration and selection, and by transmission between generations. Note that transmission can cause a change even at equilibrium. For example, when there is no selection acting on one sex, allele frequencies after selection are different in males and females, and thus they change during transmission.

We assume a life cycle that begins with selection on hybrid incompatibility, followed by migration, followed by selection on the male trait. Assuming selection on hybrid incompatibility precedes migration is biologically plausible and produces simpler equations than assuming selection on hybrid incompatibility follows migration. Qualitative results are similar under the second assumption. In what follows, we use primes to denote the stage of the life cycle, with no primes denoting the start of the generation, one denoting just after selection on hybrid incompatibility, two denoting just after migration, three denoting just after natural and sexual selection on the male trait, and four denoting just after transmission. We assume no direct selection on the female preference loci.

We are seeking the equilibrium for the mean preference. We first note that assuming no meiotic drive, one can show that the allele frequency change is asymptotically equal for all the positions at a locus. Further, the rate of change is equal to the change within a generation averaged over all the positions at that locus. Thus for locus *j*,

$$\Delta \tilde{p}_i = \frac{1}{n_i} \sum_{j:\, j = i} (p'''_j - p_j), \tag{A3}$$

where n_i is the number of positions at locus i (e.g. $n_i = 4$ for an autosomal locus in a diploid dioecious population) and tildas denote values at QLE. The summation is over all positions j that are at the same locus as position i. The proof of (A3) is given in Appendix B below. The change in allele frequencies due to migration and selection within a generation is

$$p'''_j - p_j = \sum_{\mathbb{A} \subseteq \mathbb{H}} a_{\mathbb{A}} D_{\mathbb{A}j} + m(p^C_j - p'_j) + \sum_{\mathbb{A} \subseteq \mathbb{T}} a_{\mathbb{A}} D''_{\mathbb{A}j}, \tag{A4}$$

where m is proportion of newly arrived migrants, p^C_j is the allele frequency among migrants from the continent, $a_{\mathbb{A}}$ is the selection coefficient on positions in set $\mathbb{A}$, and $D_{\mathbb{A}j}$ is the association (linkage disequilibrium) among positions in the set $\mathbb{A}j$.

The $a_{\mathbb{A}}$ are coefficients for selection (both natural and sexual) as defined by KJB (their equation 7). Key to our approach is the fact that these coefficients can describe any form of selection, including arbitrary patterns of dominance and multilocus epistasis. We can use this result, therefore, to study any specific type of hybrid incompatibility, and in a later section we calculate their values for several specific types of hybrid incompatibility. To make further progress on the general model we assume at this point that the a are much smaller than 1, meaning that the force of selection on individual loci and sets of loci is weak relative to rates of recombination. Under these conditions, the associations rapidly evolve to values that are of order a. Further, it is possible to derive simple approximations for them that are accurate to order a, and we will do that below. We therefore replace $D_{\mathbb{A}\mathbb{j}}$ in (A4) by $\tilde{D}_{\mathbb{A}\mathbb{j}} + O(a^2)$, where a tilda denotes a QLE approximation.

We can now find the asymptotic rate of change in the mean preference by substituting these results into (A2), which with some minor rearrangement gives

$$\Delta \overline{P} \;=\; m \sum_{i \in \mathbb{P}_f} \sum_{j : j = i} \frac{1}{n_i} b_i (p_j^C - p_j') + \sum_{i \in \mathbb{P}_f} \sum_{j : j = i} \sum_{\mathbb{A} \subseteq \mathbb{H}} \frac{1}{n_i} b_i a_{\mathbb{A}} \tilde{D}_{\mathbb{A}j} + \sum_{i \in \mathbb{P}_f} \sum_{j : j = i} \sum_{\mathbb{A} \subseteq \mathbb{T}} \frac{1}{n_i} b_i a_{\mathbb{A}} \tilde{D}''_{\mathbb{A}j} + O(a^3) . \qquad \text{(A5)}$$

The three terms on the right represent, respectively, migration, the effects of selection on hybrid incompatibility, and selection on the male trait. Consider the first term. Differences between the allele frequencies at positions at the same locus are $O(a)$, and so we can write $p_j' = p_j' + O(a)$, where p_j' is the average of the allele frequencies at locus j. Likewise, since the changes in allele frequencies due to selection against hybridization are $O(a)$, we can write $p_j' = p_j + O(a)$. The inner summation then causes the term corresponding to a given preference position in a female to be summed n_i times. Considering the second and third terms, we can replace the two outer summations by a single summation over all positions in set $\mathbb{P}$ (including those in males and females) if each term in the sum is multiplied by n_i^f, the number of positions at locus i carried by females.

We now assume that all preference positions have the same mode of inheritance, and that migration is much weaker than selection (specifically, that m is $O(a^2)$). This gives

$$\begin{aligned}\Delta \overline{P} &= m \textstyle\sum_{\mathbb{i} \in \mathbb{P}_f} b_i (p_i^C - p_i) + \sum_{\mathbb{i} \in \mathbb{P}} \frac{n_i^f}{n_i} b_i \sum_{\mathbb{A} \subseteq \mathbb{H}} a_{\mathbb{A}} \tilde{D}_{\mathbb{A}\mathbb{i}} \\ &\quad + \sum_{\mathbb{i} \in \mathbb{P}} \frac{n_i^f}{n_i} b_i \sum_{\mathbb{A} \subseteq \mathbb{T}} a_{\mathbb{A}} \tilde{D}''_{\mathbb{A}\mathbb{i}} + O(a^3) \\ &= m (P^C - \overline{P}) + F_P \textstyle\sum_{\mathbb{i} \in \mathbb{P}} b_i \sum_{\mathbb{A} \subseteq \mathbb{H}} a_{\mathbb{A}} \tilde{D}_{\mathbb{A}\mathbb{i}} \\ &\quad + F_P \sum_{\mathbb{i} \in \mathbb{P}} b_i \sum_{\mathbb{A} \subseteq \mathbb{T}} a_{\mathbb{A}} \tilde{D}''_{\mathbb{A}\mathbb{i}} + O(a^3),\end{aligned} \tag{A6}$$

where P^C is the mean female preference on the continent, and F_P is the fraction of preference positions carried by females (e.g., 1/2 for autosomal preferences, 2/3 for X-linked preferences).

Setting $\Delta \overline{P}$ to zero and rearranging, we find at equilibrium that the amount of reinforcement is

$$\hat{P} - P^{C} = \frac{F_P}{m}\left(\sum_{\mathbb{i} \in \mathbb{P}_f} b_i \sum_{\mathbb{A} \subseteq \mathbb{H}_f} a_{\mathbb{A}} \tilde{D}_{\mathbb{A}\mathbb{i}} + \sum_{\mathbb{i} \in \mathbb{P}_m} b_i \sum_{\mathbb{A} \subseteq \mathbb{H}_m} a_{\mathbb{A}} \tilde{D}_{\mathbb{A}\mathbb{i}}\right) + \frac{F_P}{m} \sum_{\mathbb{i} \in \mathbb{P}} b_i \sum_{\mathbb{A} \subseteq \mathbb{T}} a_{\mathbb{A}} \tilde{D}^{"}_{\mathbb{A}\mathbb{i}} + O(a)\,. \tag{A7}$$

where $\hat{P}$ is the mean preference at equilibrium. Here we have split up the sum corresponding to selection on hybrid incompatibility into two terms, one reflecting selection on females and the second selection on males. We are able split the sum because $\tilde{D}_{\mathbb{A}\mathbb{i}} = 0$ whenever $\mathbb{A}\mathbb{i}$ contains more than one sex of carrier.

Hybrid incompatibility drives evolution of the preference through the genetic associations between incompatibility alleles and preference alleles, which appear as $\tilde{D}_{\mathbb{A}\mathbb{i}}$ in equation (A7). These associations are generated by admixture (that is, migration or hybridization). As we show in Appendix C, when there is a lack of pleiotropy between the male trait and the hybrid incompatibility the QLE values for the associations among positions are of the form

$$\tilde{D}_{\mathbb{A}\mathbb{i}} = m\, d_{Ai}\, \phi_{\mathbb{A}\mathbb{i}}, \tag{A8}$$

where

$$d_{Ai} = (p_i^C - p_i) \prod_{j \in A} (p_j^C - p_j), \quad \text{(A9)}$$

and $\phi_{\mathbb{A}i}$ is a constant factor that depends on the mode of inheritance for the preference position i and the incompatibility positions in the set $\mathbb{A}$, as well as the probabilities of recombination in males and females breaking up loci in the set *Ai*. Appendix C shows how to compute $\phi_{\mathbb{A}i}$ for any mode of inheritance for the preference and incompatibility loci.

In order to simplify (A7) further, however, we need an expression for $\phi_{\mathbb{A}i}$ that is independent of the preference position i. $\phi_{\mathbb{A}i}$ is independent of i when all loci affecting the preference have the same mode of inheritance and the probability of recombination breaking up the set *Ai* in each sex is the same for all preference loci. This condition is met, for example, when there is only one loci affects the preference or when all of the preference loci are unlinked to the incompatibility loci when possible. We assume that the latter case is true. This means that when the preference is autosomal, $(1 - r_{Ai}^f) = \frac{1}{2}(1 - r_A^f)$ and $(1 - r_{Ai}^m) = \frac{1}{2}(1 - r_A^m)$, where the superscript denotes whether recombination is occuring in females (f) or males (m). When the preference and one or more of the incompatibility loci are X-linked, however, $(1 - r_{Ai}^f) = \frac{1}{2}(1 - r_A^f)$ and $(1 - r_{Ai}^m) = (1 - r_A^m)$. With this assumption we can write (A8) as

$$\tilde{D}_{(\mathbb{A}\mathfrak{i})_{\mathrm{f}}} \;=\; m\, d_{Ai}\, \phi^{\mathrm{f}}_{\mathbb{A}} \quad \text{and} \quad \tilde{D}_{(\mathbb{A}\mathfrak{i})_{\mathrm{m}}} \;=\; m\, d_{Ai}\, \phi^{\mathrm{m}}_{\mathbb{A}}. \tag{A10}$$

The values of the $\phi^{\mathrm{f}}_{\mathbb{A}}$ and $\phi^{\mathrm{m}}_{\mathbb{A}}$ are given in Table C1. Equation (A10) shows that the associations between the preference and incompatibility loci are generated by introgression, and are proportional to its rate (represented by m). These are the associations that link selection against hybrids to reinforcement of the preference. Now we can write

$$\begin{aligned} \hat{P} - P^{\mathrm{C}} \;=\; & \\ & F_{\mathrm{P}}\Big(\textstyle\sum_{\mathfrak{i} \in \mathbb{P}_{\mathrm{f}}} b_i\, d_i \sum_{\mathbb{A} \subseteq \mathbb{H}_{\mathrm{f}}} a_{\mathbb{A}}\, d_A\, \phi^{\mathrm{f}}_{\mathbb{A}} \;+\; \sum_{\mathfrak{i} \in \mathbb{P}_{\mathrm{m}}} b_i\, d_i \sum_{\mathbb{A} \subseteq \mathbb{H}_{\mathrm{m}}} d_A\, \phi^{\mathrm{m}}_{\mathbb{A}}\Big) \\ & \qquad + \tfrac{1}{m}\, F_{\mathrm{P}} \sum\nolimits_{\mathfrak{i} \in \mathbb{P}} b_i \sum\nolimits_{\mathbb{A} \subseteq \mathbb{T}} a_{\mathbb{A}}\, \tilde{D}'_{\mathbb{A}\mathfrak{i}} \;+\; O(a) \\ \;=\; & \left(\hat{P} - P^{\mathrm{C}}\right)\left[F_{\mathrm{P}}\, I_{\mathrm{f}} \;+\; (1 - F_{\mathrm{P}})\, I_{\mathrm{m}}\right] \;+\; M \;+\; O(a)\,. \end{aligned} \tag{A11}$$

Here,

$$I_{\mathrm{f}} \;=\; -\textstyle\sum_{\mathbb{A} \subseteq \mathbb{H}_{\mathrm{f}}} a_{\mathbb{A}}\, d_A\, \phi^{\mathrm{f}}_{\mathbb{A}} \quad \text{and} \quad I_{\mathrm{m}} \;=\; -\sum_{\mathbb{A} \subseteq \mathbb{H}_{\mathrm{m}}} a_{\mathbb{A}}\, d_A\, \phi^{\mathrm{m}}_{\mathbb{A}} \tag{A12}$$

represent the effects on the preference of selection on hybrid incompatibility acting on females and males, respectively, and

$$M = \frac{F_{\mathrm{P}}}{m} \sum_{i \in \mathbb{P}} b_i \sum_{\mathbb{A} \subseteq \mathbb{T}} a_{\mathbb{A}} \tilde{D}''_{\mathbb{A}i} \tag{A13}$$

represents the effects on the preference of natural and sexual selection on the male display.

Rearranging finally gives us an expression for the amount of reinforcement that is general to all forms of inheritance for the hybrid incompatibility:

$$\hat{P} - P^{\mathrm{C}} = \frac{M}{1-I} + O(a), \tag{A14}$$

where

$$I = F_{\mathrm{P}} I_{\mathrm{f}} + (1 - F_{\mathrm{P}}) I_{\mathrm{m}}. \tag{A15}$$

Written this way, we have partly isolated the effects of the modes of inheritance of the display trait loci from the those of the hybrid incompatibility loci. The modes of

inheritance of hybrid incompatibility affect I_f and I_m, but it is plausible that (at least to a first approximation) they will not change M. Likewise, the mode of inheritance of the display trait loci will only affect M. Since our focus of this paper is on how inheritance of hybrid incompatibility affects speciation, we will simplify our analysis by treating M as a fixed entity. It is possible that reinforcement will cause the male trait to evolve in a way that causes M to vary depending on the strength of selection on hybrid incompatibility. Lastly, note that the mode of inheritance of the preference will affect M, F_P, I_f, and I_m.

We can see from (A14) that the approximation becomes less accurate as the force of selection against hybrids increases (the denominator approaches zero as I approaches 1). A second approximation that will be more accurate for values of I close to 1 can be obtained using a Taylor approximation. Multiplying the numerator and denomerator of (A14) by $1 + I$ and neglecting the I^2 that appears in the denominator, we obtain

$$\hat{P} - P^C = M(1 + I) + O(a). \qquad \text{(A16)}$$

Equation (A16) has a simple and intuitive interpretation. Selection on the male trait is sufficient to cause reinforcement, even without additional selection on the incompatibility loci (in which case $I = 0$). Selection against hybrid incompatibilities

further amplifies the amount of reinforcement. We will now see how different forms of hybrid incompatibility affect reinforcement.

One- and Two-Locus Incompatibilities. In the previous section we derived equations describing the effect of selection against hybrid incompatibility on the amount of reinforcement. As (A12) shows, this effect depends on the values of three terms: $a_{\mathbb{A}}$, d_A, and $\phi_{\mathbb{A}}$. Values for $\phi_{\mathbb{A}}$, derived in Appendix C, depend on how the preference and incompatibility loci are inherited and on the rates of recombination breaking them up.

The value of $a_{\mathbb{A}}$ depends on the intensity and pattern of selection on the hybrid incompatibility loci. Here we study one- and two-locus incompatibilities. We allow each incompatibility locus to be either autosomal, X-linked, Y-linked, Z-linked, or W-linked. Our notation for the strengths of selection against different genotypes at the hybrid incompatibility loci are given in Figure 3.1. We assign the island genotype, defined to be homozygous/hemizygous for the 1 allele at all loci, a fitness of one and other genotypes a fitness relative to that of the island genotype. Using this notation we show in Appendix D how to derive expressions for the $a_{\mathbb{A}}$ found in (A12) in terms of the selection coefficients given in Figure 3.1. Table D1 presents those expressions for all types of one- and two-locus incompatibilities.

The value of $d_{\mathbb{A}}$ depends on the amount of divergence between the island and continent at the hybrid incompatibility loci. The value of $d_{\mathbb{A}}$ turns out to be quite simple

since we are assuming that migration is much weaker than selection (i.e. $m = O(s^2)$). This assumption is reasonable because when migration is not relatively weak, the continent and island cannot maintain a polymorphism at the incompatibility loci and reinforcement cannot occur. One can easily show that under this assumption, $p_j^C - p_j = -1 + O(s)$ for any locus j. This fact allows us to use the following $O(s)$ approximation for $d_{\mathbb{A}}$:

$$d_{\mathbb{A}} = (-1)^n + O(s), \tag{A17}$$

where n is the number of elements in $\mathbb{A}$. This result holds regardless of the number of loci in $\mathbb{A}$ or the pattern of inheritance of those loci.

Now that we have expressions for $a_{\mathbb{A}}$, d_A, and $\phi_{\mathbb{A}}$ we can compute I_{f} and I_{m} for any type of one- or two-locus incompatibility by plugging values from Table C1, Table D1, and equation (A17) into (A12). Before we give the results, we summarize the assumptions required to arrive at the results, which are:

1) QLE (selection is weak relative to rates of recombination)

2) preference genes have additive effects

3) no imprinting

4) two alleles per locus

5) selection against hybrid incompatibility happens before migration

6) selection on the male trait happens after migration

7) no direct selection on the female preference

8) no meiotic drive

9) all preference positions have the same mode of inheritance

10) migration is weak relative to the strength of selection

11) no pleiotropy (each locus affects only the incompatibility, trait, or preference)

12) preference positions are unlinked to the incompatibility loci when possible

The effects of selection against hybrids on the amount of reinforcement (I_f and I_m) are presented in the following table. A, X, and Y indicate that all loci are autosomally inherited, X-linked, and Y-linked, respectively. A-A, A-X, A-Y, X-X, and X-Y denote autosome-autosome, autosome-X, autosome-Y, X-X, and X-Y incompatibilities, respectively. Equations pertaining to a Z-W system of sex-determination are identical, except with f and m subscripts interchanged.

Table A1. The effect of selection against male (I_f) and female (I_m) hybrids on reinforcement.

Autosomal preference

	I_f	I_m
A	$4\,s_{1_f}$	$4\,s_{1_m}$
X	$4\,s_{1_f}$	$2\,s_{0_m}$
Y	0	$2\,s_{0_m}$
A-A	$4\,(s_{12_f} + s_{21_f}) + 8\,\frac{s_{11_f} - s_{12_f} - s_{21_f}}{2 + r_{kl}^f + r_{kl}^m}$	$4\,(s_{12_m} + s_{21_m}) + 8\,\frac{s_{11_m} - s_{12_m} - s_{21_m}}{2 + r_{kl}^f + r_{kl}^m}$
A-X	$\frac{8}{3}\,s_{11_f} + \frac{4}{3}\,s_{12_f} + \frac{4}{3}\,s_{21_f}$	$\frac{4}{3}\,s_{10_m} + \frac{8}{3}\,s_{12_m} + \frac{2}{3}\,s_{20_m}$
A-Y	0	$\frac{4}{3}\,s_{10_m} + \frac{8}{3}\,s_{12_m} + \frac{2}{3}\,s_{20_m}$
X-X	$4\,(s_{12_f} + s_{21_f}) + 20\,\frac{s_{11_f} - s_{12_f} - s_{21_f}}{5 + 3\,r_{kl}^f}$	$2\,(s_{02_m} + s_{20_m}) + 2\,\frac{(5 - r_{kl}^f)\,(s_{00_m} - s_{02_m} - s_{20_m})}{5 + 3\,r_{kl}^f}$
X-Y	0	$2\,(s_{02_m} + s_{20_m})$

X-linked preference

	I_f	I_m
A	$4\,s_{1_\mathrm{f}}$	$2\,s_{1_\mathrm{m}}$
X	$6\,s_{1_\mathrm{f}}$	$\frac{5}{2}\,s_{0_\mathrm{m}}$
Y	0	0
A-A	$4\,(s_{12_\mathrm{f}} + s_{21_\mathrm{f}}) + 4\,\frac{(5-r_{kl}^\mathrm{m})\,(s_{11_\mathrm{f}} - s_{12_\mathrm{f}} - s_{21_\mathrm{f}})}{5 + r_{kl}^\mathrm{f}\,(3 - r_{kl}^\mathrm{m}) + r_{kl}^\mathrm{m}}$	$2\,(s_{12_\mathrm{m}} + s_{21_\mathrm{m}}) + 2\,\frac{(5-r_{kl}^\mathrm{f})\,(s_{11_\mathrm{m}} - s_{12_\mathrm{m}} - s_{21_\mathrm{m}})}{5 + r_{kl}^\mathrm{f}\,(3 - r_{kl}^\mathrm{m}) + r_{kl}^\mathrm{m}}$
A-X	$\frac{40}{13}\,s_{11_\mathrm{f}} + \frac{12}{13}\,s_{12_\mathrm{f}} + \frac{38}{13}\,s_{21_\mathrm{f}}$	$\frac{18}{13}\,s_{10_\mathrm{m}} + \frac{8}{13}\,s_{12_\mathrm{m}} + \frac{29}{26}\,s_{20_\mathrm{m}}$
A-Y	0	$2\,s_{12_\mathrm{m}}$
X-X	$6\,(s_{12_\mathrm{f}} + s_{21_\mathrm{f}}) + 6\,\frac{s_{11_\mathrm{f}} - s_{12_\mathrm{f}} - s_{21_\mathrm{f}}}{1 + r_{kl}^\mathrm{f}}$	$\frac{5}{2}\,(s_{02_\mathrm{m}} + s_{20_\mathrm{m}}) + \frac{(5-r_{kl}^\mathrm{f})\,(s_{00_\mathrm{m}} - s_{02_\mathrm{m}} - s_{20_\mathrm{m}})}{2 + 2\,r_{kl}^\mathrm{f}}$
X-Y	0	$\frac{5}{2}\,s_{02_\mathrm{m}}$

To compute the total contribution to reinforcement, I, contributions from females and males are weighted by the proportion of preference genes found in females (F_{P}) and males ($1 - F_{\mathrm{P}}$) as follows:

$$I = F_{\mathrm{P}} I_{\mathrm{f}} + (1 - F_{\mathrm{P}}) I_{\mathrm{m}} \tag{A18}$$

where $F_{\mathrm{P}} = \frac{1}{2}$ with an autosomal preference and $F_{\mathrm{P}} = \frac{2}{3}$ with an X-linked preference. The following table provides expressions for the total contribution to reinforcement.

Table A2. The total effect of selection against hybrids (I) on reinforcement.

Autosomal preference

	I
A	$2\,s_{1_{\mathrm{f}}} + 2\,s_{1_{\mathrm{m}}}$
X	$2\,s_{1_{\mathrm{f}}} + s_{0_{\mathrm{m}}}$
Y	$s_{0_{\mathrm{m}}}$
A-A	$2\,(s_{12_{\mathrm{f}}} + s_{21_{\mathrm{f}}}) + 4\,\frac{s_{11_{\mathrm{f}}} - s_{12_{\mathrm{f}}} - s_{21_{\mathrm{f}}}}{2 + r_{kl}^{\mathrm{f}} + r_{kl}^{\mathrm{m}}} + 2\,(s_{12_{\mathrm{m}}} + s_{21_{\mathrm{m}}}) + 4\,\frac{s_{11_{\mathrm{m}}} - s_{12_{\mathrm{m}}} - s_{21_{\mathrm{m}}}}{2 + r_{kl}^{\mathrm{f}} + r_{kl}^{\mathrm{m}}}$
A-X	$\frac{4}{3}\,s_{11_{\mathrm{f}}} + \frac{2}{3}\,s_{12_{\mathrm{f}}} + \frac{2}{3}\,s_{21_{\mathrm{f}}} + \frac{2}{3}\,s_{10_{\mathrm{m}}} + \frac{4}{3}\,s_{12_{\mathrm{m}}} + \frac{1}{3}\,s_{20_{\mathrm{m}}}$
A-Y	$\frac{2}{3}\,s_{10_{\mathrm{m}}} + \frac{4}{3}\,s_{12_{\mathrm{m}}} + \frac{1}{3}\,s_{20_{\mathrm{m}}}$
X-X	$2\,(s_{12_{\mathrm{f}}} + s_{21_{\mathrm{f}}}) + 10\,\frac{s_{11_{\mathrm{f}}} - s_{12_{\mathrm{f}}} - s_{21_{\mathrm{f}}}}{5 + 3\,r_{kl}^{\mathrm{f}}} + (s_{02_{\mathrm{m}}} + s_{20_{\mathrm{m}}}) + \frac{(5 - r_{kl}^{\mathrm{f}})\,(s_{00_{\mathrm{m}}} - s_{02_{\mathrm{m}}} - s_{20_{\mathrm{m}}})}{5 + 3\,r_{kl}^{\mathrm{f}}}$
X-Y	$s_{02_{\mathrm{m}}} + s_{20_{\mathrm{m}}}$

X-linked preference

	I
A	$\frac{8}{3}\, s_{1_\mathrm{f}} + \frac{2}{3}\, s_{1_\mathrm{m}}$
X	$4\, s_{1_\mathrm{f}} + \frac{5}{6}\, s_{0_\mathrm{m}}$
Y	0
A-A	$\frac{8}{3}\,(s_{12_\mathrm{f}} + s_{21_\mathrm{f}}) + \frac{8}{3}\,\frac{(5-r_{kl}^{\mathrm{m}})\,(s_{11_\mathrm{f}} - s_{12_\mathrm{f}} - s_{21_\mathrm{f}})}{5 + r_{kl}^{\mathrm{f}}\,(3 - r_{kl}^{\mathrm{m}}) + r_{kl}^{\mathrm{m}}}$ $+ \frac{2}{3}\,(s_{12_\mathrm{m}} + s_{21_\mathrm{m}}) + \frac{2}{3}\,\frac{(5-r_{kl}^{\mathrm{f}})\,(s_{11_\mathrm{m}} - s_{12_\mathrm{m}} - s_{21_\mathrm{m}})}{5 + r_{kl}^{\mathrm{f}}\,(3 - r_{kl}^{\mathrm{m}}) + r_{kl}^{\mathrm{m}}}$
A-X	$\frac{80}{39}\, s_{11_\mathrm{f}} + \frac{24}{39}\, s_{12_\mathrm{f}} + \frac{76}{39}\, s_{21_\mathrm{f}} + \frac{18}{39}\, s_{10_\mathrm{m}} + \frac{8}{39}\, s_{12_\mathrm{m}} + \frac{29}{78}\, s_{20_\mathrm{m}}$
A-Y	$\frac{2}{3}\, s_{12_\mathrm{m}}$
X-X	$4\,(s_{12_\mathrm{f}} + s_{21_\mathrm{f}}) + 4\,\frac{s_{11_\mathrm{f}} - s_{12_\mathrm{f}} - s_{21_\mathrm{f}}}{1 + r_{kl}^{\mathrm{f}}} + \frac{5}{6}\,(s_{02_\mathrm{m}} + s_{20_\mathrm{m}}) + \frac{1}{6}\,\frac{(5-r_{kl}^{\mathrm{f}})\,(s_{00_\mathrm{m}} - s_{02_\mathrm{m}} - s_{20_\mathrm{m}})}{1 + r_{kl}^{\mathrm{f}}}$
X-Y	$\frac{5}{6}\, s_{02_\mathrm{m}}$

Ecological Incompatibilities. Here we answer the question: how does selection against ecologically-inferior hybrids favor reinforcement of prezygotic isolation? Here we study a model in which genes contribute additively to a quantitative trait, such as body size or bill length. In this model, hybrids are unfit because they have intermediate phenotypes that are selected against. We'll assume that in the absence of hybridization, the island and the continent species would be fixed for a set of loci that have additive effects on the ecological trait. It turns out to simplify things substantially if we assume that the n diploid loci influencing the trait are interchangeable: then they have equal allele frequencies and equal effects on the trait. Without loss of generality, we can define the scale of measurement for the ecological trait such that $Z = 0$ when all the loci are fixed for the favored allele (allele 0). The mean of the continent is denoted Z^C.

Selection coefficients: We can approximate the fitness function of an individual in the vicinity of $Z = 0$ by a quadratic:

$$W = 1 + \beta Z + \Gamma Z^2. \qquad \text{(A19)}$$

Here, fitness is a function of the directional selection gradient β and the stabilizing selection gradient Γ acting on the ecological trait. (Negative values of Γ correspond to stabilizing selection, and positive values to disruptive selection.) The values of the

selection gradients depend on the fitness function for the ecological trait and also on the distribution of that trait in the island population. That distribution evolves in response to selection and migration, which causes the values of the selection gradients to change. The β and Γ in the expressions that follow refer to the equilibrium values for the gradients. See Lande and Arnold (1993) and Kirkpatrick (2001) for more details.

Using our usual notation, the phenotype of an individual can be written

$$Z = \overline{Z} + \sum_{i \in \mathbb{I}} b_i \zeta_i, \tag{A20}$$

where mean phenotype is

$$\overline{Z} = 2 n p_i b_i. \tag{A21}$$

Recall that p_i is the frequency of the 1 allele. Under our assumptions, when $p_i = 1$ the mean phenotype is Z^{C}. That means $b_i = \frac{Z^{\mathrm{C}}}{2n}$, and so

$$\overline{Z} = 2 n p_i \left(\frac{Z^{\mathrm{C}}}{2n}\right) = p_i Z^{\mathrm{C}}. \tag{A22}$$

The fitness function needed to calculate the *a*s is

$$\begin{aligned} \frac{W}{\overline{W}} &\approx 1 + \beta\,(p_i\,Z^{\mathrm{C}} + \textstyle\sum_{\mathbb{i}\in\mathbb{I}}\, b_{\mathbb{i}}\,\zeta_{\mathbb{i}}) + \Gamma\,(p_i\,Z^{\mathrm{C}} + \textstyle\sum_{\mathbb{i}\in\mathbb{I}}\, b_{\mathbb{i}}\,\zeta_{\mathbb{i}})^2 \\ &\approx 1 + \beta\left(p_i\,Z^{\mathrm{C}} + \frac{Z^{\mathrm{C}}}{2n}\textstyle\sum_{\mathbb{i}\in\mathbb{I}}\,\zeta_{\mathbb{i}}\right) + \Gamma\left(p_i\,Z^{\mathrm{C}} + \frac{Z^{\mathrm{C}}}{2n}\textstyle\sum_{\mathbb{i}\in\mathbb{I}}\,\zeta_{\mathbb{i}}\right)^2 \end{aligned} \tag{A23}$$

The approximation assumes that all individuals in the population are near $Z = 0$. This is consistent with our assumption that the hybridization rate is very small, and so the frequencies of alleles from the continent are very small. Picking out the appropriate coefficients of the ζs shows that the selection coefficients are

$$a_i = \frac{1}{2n}\,\beta\,Z^{\mathrm{C}} + \frac{1}{n}\,\Gamma\,Z^{\mathrm{C}\,2}\,p_i, \tag{A24}$$

and

$$a_{ij} = \frac{1}{4n^2}\,\Gamma\,Z^{\mathrm{C}\,2}. \tag{A25}$$

Assuming selection is much stronger than hybridization, the equilibrium value for the allele frequencies for the ecological trait are:

$$p_i \approx \frac{m}{\beta b_i} \approx \frac{2mn}{\beta Z^{C}}, \tag{A26}$$

which completes the calculation for the selection coefficients:

$$a_i = \frac{1}{2n}\beta Z^{C} + \frac{2}{\beta} m \Gamma Z^{C}. \tag{A27}$$

The effect of ecological selection: Now we derive an expression for I using the selection coefficients derived above. We will assume that females and males with the same value of Z have equal fitness. Under this assumption, $I_f = I_m$. From the derivation of the as above, we know that there are two types of as, those that involve one position and those that involve two positions. Separating the sum into two parts yields

$$I = -\sum_{i \in H} a_i d_i \phi_i - \sum_{i \in H} \sum_{\substack{j \in H, \\ j \neq i}} a_{ij} d_{ij} \phi_{ij}. \tag{A28}$$

Since we are assuming that the continent and island are nearly fixed for alternative alleles, $d_i \approx -1$ and $d_{ij} \approx 1$. Now we have

$$I = \sum_{i \in H} a_i \phi_i - \sum_{i \in H} \sum_{\substack{j \in H, \\ j \neq i}} a_{ij} \phi_{ij}. \tag{A29}$$

Plugging in the values for the *a*s yields

$$I = \sum_{i \in H} \left(\frac{1}{2n} \beta Z^C + \frac{2}{\beta} m \Gamma Z^C \right) \phi_i - \sum_{i \in H} \sum_{\substack{j \in H, \\ j \neq i}} \left(\frac{1}{4n^2} \Gamma Z^C \right.$$

$$= \left(\frac{1}{2n} \beta Z^C + \frac{2}{\beta} m \Gamma Z^C \right) \sum_{i \in H} \phi_i - \left(\frac{1}{4n^2} \Gamma Z^{C\,2} \right) \sum_{i \in H} \sum_{\substack{j \in H, \\ j \neq i}} \tag{A30}$$

If we assume that n is large, then n^2 is very large and the right term is nearly zero so

$$I \approx \left(\frac{1}{2n} \beta Z^C + \frac{2}{\beta} m \Gamma Z^C \right) \sum_{i \in H} \phi_i. \tag{A31}$$

The summation in equation (A31) adds up the contributions to reinforcement for all positions affecting the ecological trait. Using $n\,\overline{\phi} = \sum_{i \in \mathbb{H}} \phi_i$, where $\overline{\phi}$ is the average contribution of a position to reinforcement, we have

$$\begin{aligned} I &\approx \left(\tfrac{1}{2n}\,\beta\,Z^{\mathrm{C}} \;+\; \tfrac{2}{\beta}\,m\,\Gamma\,Z^{\mathrm{C}}\right) n\,\overline{\phi} \\ &= \left(\tfrac{1}{2}\,\beta\,Z^{\mathrm{C}} \;+\; \tfrac{2n}{\beta}\,m\,\Gamma\,Z^{\mathrm{C}}\right)\overline{\phi}. \end{aligned} \tag{A32}$$

Using the fact that $m\,n/\beta \;=\; Z^{\mathrm{C}}\,p_i/2$, we substitute to obtain

$$I \approx \left(\tfrac{1}{2}\,\beta\,Z^{\mathrm{C}} \;+\; \Gamma\,{Z^{\mathrm{C}}}^{2}\,p_i\right)\overline{\phi}. \tag{A33}$$

Since we are assuming that $p_i \approx 1$ for all i,

$$I \approx \left(\tfrac{1}{2}\,\beta\,Z^{\mathrm{C}} \;+\; \Gamma\,{Z^{\mathrm{C}}}^{2}\right)\overline{\phi}. \tag{A34}$$

Since we have defined Z to be in the vicinity of zero at equilibrium, we can replace $Z^{C} = Z^{C} - \hat{Z} = \left| \hat{Z} - Z^{C} \right|$. Our final result for the effect of ecological selection on the amount of reinforcement is

$$I \approx \overline{\phi} \left(\tfrac{1}{2} \beta \left| \hat{Z} - Z^{C} \right| + \Gamma \left(\hat{Z} - Z^{C} \right)^{2} \right). \qquad \text{(A35)}$$

To study the affect of sex linkage, we define π_A, π_X, and π_Y to be the proportions of the ecological trait loci that are autosomally inherited, X-linked, and Y-linked, respectively. Using this notation, results from Table C1, and the fact that $r_{kl} = 0$ when ϕ_A contains one position, one can show that $\overline{\phi} = 8\pi_A + 6\pi_X + 2\pi_Y$ when the preference is autosomally inherited and $\overline{\phi} = 6\pi_A + \frac{17}{2}\pi_X$ when the preference is X-linked.

References

Barton, N. H., and M. Turelli. 1991. Natural and sexual selection on many loci. Genetics 127:229-255.

Kirkpatrick, M.. 2001. Reinforcement during ecological speciation. Proceedings of the Royal Society of London B: Biology 268:1259-1263.

Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. Genetics 161:1727-1750.

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

Appendix B

Asymptotic Change in Allele Frequencies Among Positions at a Locus

Here we show that after a short period, the allele frequencies of all the positions at a locus evolve at the same rate. This rate is equal to the change within a generation averaged over all the positions at that locus. Begin by writing the initial frequencies of the alleles at a locus as the vector $\mathbf{p}_0$. After one generation,

$$\mathbf{p}_1 = \mathbf{T}(\mathbf{p}_0 + \boldsymbol{\Delta}), \tag{B1}$$

where $\mathbf{T}$ is a matrix of transmission coefficients and $\boldsymbol{\Delta}$ is the vector of changes in allele frequencies within a generation caused by selection and migration. At QLE, $\boldsymbol{\Delta}$ is constant. After two generations of QLE, we have

$$\mathbf{p}_2 = \mathbf{T}(\mathbf{p}_1 + \boldsymbol{\Delta}) = \mathbf{T}^2\mathbf{p}_0 + \mathbf{T}^2\boldsymbol{\Delta} + \mathbf{T}\boldsymbol{\Delta}, \tag{B2}$$

and after *t* generations,

$$\mathbf{p}_t = \mathbf{T}^t\mathbf{p}_0 + (\sum_{i=1}^{t} \mathbf{T}^t)\,\boldsymbol{\Delta}. \qquad \text{(B3)}$$

Thus the vector of changes in allele frequencies at generation t is

$$\Delta\mathbf{p}_t = \mathbf{p}_{t+1} - \mathbf{p}_t = (\mathbf{T}^{t+1} - \mathbf{T}^t)\,\mathbf{p}_0 + \mathbf{T}^t\,\boldsymbol{\Delta}. \qquad \text{(B4)}$$

Since $\mathbf{T}$ is a stochastic matrix, for t large we have $\mathbf{T}^{t+1} - \mathbf{T}^t \to 0$ as $t \to \infty$. Thus asymptotically the allele frequency change is

$$\Delta\,\tilde{\mathbf{p}} = \mathbf{T}^t\,\Delta. \qquad \text{(B5)}$$

This implies that, for large t, the vector of allele frequency change $\Delta\tilde{\mathbf{p}}$ is proportional to the leading eigenvector of $\mathbf{T}$.

We assume that meiosis is normal in the sense that there is no meiotic drive. In that case, transmission is *conservative*, meaning that all positions at a locus are transmitted with equal probability. This implies that the rows as well as the columns of

T sum to unity, that is, the matrix is "doubly-stochastic". In that case, the leading eigenvector is $\{1/n,\ 1/n,\ ...\}^{\mathbf{T}}$ (Karlin and Taylor 1975), implying that

$$\Delta \tilde{\mathbf{p}} \quad = \quad \{1/n,\ 1/n,\ ...\}^{\mathbf{T}} \bar{\mathbf{\Delta}}, \tag{B6}$$

where $\bar{\Delta}$ is simply the mean allele frequency change within a generation among all the positions at the locus (that is, the mean of the elements of Δ).

To sum up, at QLE and in the absence of meiotic drive, the change in allele frequency at position i is equal to the average change at all the positions at that locus within a generation:

$$\Delta \tilde{p}_i \quad = \quad \frac{1}{n_i} \sum_{j:\, j \,=\, i} (p'_j - p_j), \tag{B7}$$

where n_i is the number of positions at locus *i* and the summation is over all positions at locus *i*.

References

Karlin, S., and H. M. Taylor. 1975 *A first course in stochastic processes*. Academic Press, New York.

Appendix C

Associations Between Preference and Incompatibility Positions at QLE

Here we find the QLE approximations for the $\tilde{D}_{\mathbb{A}\mathbb{i}}$, the associations between preference and incompatibility positions that appear in (A7). To do so, we derive a recursion for their evolution. Because we are interested in the affect of hybrid incompatibility on reinforcement, we assume that there is a lack of pleiotropy between the hybrid incompatibility and the male trait, that is $\mathbb{H} \cap \mathbb{T} = \emptyset$. Under this assumption, selection on the male trait loci will not affect the associations between the preference and incompatibility positions.

It is convenient to work backwards, starting with the associations in zygotes at the beginning of the next generation, which we denote as $D''''_{\mathbb{A}\mathbb{i}}$. We assume that the life cycle begins with selection against incompatibilities, followed by migration, followed by mating and transmission. The associations in zygotes can be written in terms of the associations among mated adults of the current generation:

$$D''''_{\mathbb{A}\mathbb{i}} = \sum_{\mathbb{j}: j = i} \sum_{\mathbb{U}: U = A} t_{\mathbb{A}\mathbb{i} \leftarrow \mathbb{U}\mathbb{j}} D'''_{\mathbb{U}\mathbb{j}}, \tag{C1}$$

where $t_{\mathbb{A}i \leftarrow \mathbb{U}j}$ is a transmission coefficient, defined as the probability that the positions in set $\mathbb{A}i$ were inherited from those in set $\mathbb{U}j$, and $D'''_{\mathbb{U}j}$ is the association between positions in set $\mathbb{U}j$ in the current generation after selection and migration (see KJB equation 12). Since selection on the male trait will not affect the associations between the preference and incompatibility loci, $D'''_{\mathbb{U}j} = D''_{\mathbb{U}j}$, so we can rewrite (C1) as

$$D''''_{\mathbb{A}i} = \sum_{j: j=i} \sum_{\mathbb{U}: U=A} t_{\mathbb{A}i \leftarrow \mathbb{U}j} D''_{\mathbb{U}j} . \qquad \text{(C2)}$$

The associations after migration, in turn, can be written in terms of the associations before migration but after selection on the hybrid incompatibility. We assume that migration is weak ($m << 1$), an assumption required both by our QLE approximation and to prevent swamping of the locally-adapted allele in the island population. Then when all positions in the set $\mathbb{U}j$ have the same sex-of-carrier,

$$D''_{\mathbb{U}j} = D'_{\mathbb{U}j} + m\, d_{\mathbb{U}j} , \qquad \text{(C3)}$$

where

$$d_{\mathbb{U}\mathbb{j}} = (p_{\mathbb{j}}^{C} - p_{\mathbb{j}}) \prod_{\mathbb{i} \in \mathbb{U}} (p_{\mathbb{i}}^{C} - p_{\mathbb{i}}) \quad \text{(C4)}$$

and a single prime denotes a value after selection but before migration (see KJB equation 34). When $\mathbb{U}\mathbb{j}$ involves positions with both sexes-of-carrier, $D''_{\mathbb{U}\mathbb{j}} = D'_{\mathbb{U}\mathbb{j}}$. At QLE, the differences in allele frequencies among different positions at the same locus are $O(a)$. This fact allows us to rewrite $d_{\mathbb{U}\mathbb{j}}$ as

$$d_{\mathbb{U}\mathbb{j}} = (p_j^{C} - p_j) \prod_{i \in U} (p_i^{C} - p_i)^{n_{\mathbb{U}i}} + O(a) \quad \text{(C5)}$$

where $n_{\mathbb{U}i}$ is defined to be the number of positions in $\mathbb{U}$ that are at the locus i and p_i is the allele frequency at locus i, averaged across all positions at that locus. This expression simplifies nicely when no two positions in $\mathbb{U}$ are at the same locus (i.e. $n_{\mathbb{U}i} = 1$ for all $\mathbb{U}i$), which is true when all positions in $\mathbb{U}$ have the same sex of carrier and the same sex of origin. When this is true we can replace $d_{\mathbb{U}\mathbb{j}}$ with $d_{Uj} + O(a)$, where d_{Uj} is defined as

$$d_{Uj} = (p_j^{C} - p_j) \prod_{i \in U} (p_i^{C} - p_i). \quad \text{(C6)}$$

To complete the life cycle, we write the associations after selection in terms of those in zygotes. Since $\mathbb{j}$ represents a preference position, and since we assume that the preference is free of direct selection, the change in associations between $\mathbb{j}$ and incompatibility positions caused by selection is $O(a^2)$. Thus we can write

$$D'_{\mathbb{U}\mathbb{j}} \;=\; D_{\mathbb{U}\mathbb{j}} \,+\, O(a^2). \tag{C7}$$

We can now find the QLE approximation for the associations. We set the values of the associations at the beginning of successive generations equal, $D'''_{\mathbb{A}\mathbb{i}} = D_{\mathbb{A}\mathbb{i}} \equiv \tilde{D}_{\mathbb{A}\mathbb{i}}$, and drop terms $O(a^2)$. Substituting equations (C3) and (C7) into (C2) shows that when the positions in set $\mathbb{A}\mathbb{i}$ include both sexes of origin, then $\tilde{D}_{\mathbb{A}\mathbb{i}} = 0$. When they all have a single sex-of-origin,

$$\begin{aligned} \tilde{D}_{\mathbb{A}\mathbb{i}} \;&=\; \sum\nolimits_{\mathbb{j}:\, j = i} \sum\nolimits_{\mathbb{U}:\, U = A} t_{\mathbb{A}\mathbb{i}\leftarrow\mathbb{U}\mathbb{j}} (\tilde{D}_{\mathbb{U}\mathbb{j}} \,+\, m\, d_{\mathbb{U}\mathbb{j}}) \\ &=\; m\, d_{Ai} \,+\, \textstyle\sum_{\mathbb{j}:\, j = i} \sum_{\mathbb{U}:\, U = A} t_{\mathbb{A}\mathbb{i}\leftarrow\mathbb{U}\mathbb{j}}\, \tilde{D}_{\mathbb{U}\mathbb{j}} \end{aligned} \tag{C8}$$

Expression (C8) represents a linear system of equations that can always be solved to give the QLE approximations for the associations under any kind of

inheritance, which is specified by the transmission coefficients. We could, for example, study a case in which some preference genes are autosomal while others are cytoplasmic, and hybrid incompatibilities involving autosome-X epistatic interactions. The number of associations described by equation (C8) depends on the mode of inheritance. For example, if preference locus i and all the hybrid incompatibility loci A are autosomal, then there are four kinds of associations: $D_{(Ai)_{\mathrm{ff}}}$, $D_{(Ai)_{\mathrm{fm}}}$, $D_{(Ai)_{\mathrm{mf}}}$, and $D_{(Ai)_{\mathrm{mm}}}$. On the other hand, if A and i are all X-linked, then there are only three kinds of these associations: $D_{(Ai)_{\mathrm{ff}}}$, $D_{(Ai)_{\mathrm{fm}}}$, and $D_{(Ai)_{\mathrm{mf}}}$.

Given any particular form of inheritance, we can solve for these associations by first writing them as a vector $\tilde{\mathbf{D}}$ (where the order of the elements is arbitrary). We can then write equation (C8) in matrix form as

$$\tilde{\mathbf{D}} = \mathbf{T}\tilde{\mathbf{D}} + m\, d_{Ai}\, \mathbf{1}, \tag{C9}$$

where $\mathbf{T}$ is a matrix of transmission coefficients and $\mathbf{1}$ is a vector of 1s. The solution for the associations is therefore

$$\tilde{\mathbf{D}} = m\, d_{Ai}\, (\mathbf{I} - \mathbf{T})^{-1}\, \mathbf{1}, \tag{C10}$$

where **I** is the identity matrix.

To make the calculation clear, take the example where set *Ai* consists of a mixture of autosomal and X-linked loci. Then the QLE values for the associations given by equation (C8) are

$$\begin{pmatrix} \tilde{D}_{(Ai)_{\mathrm{ff}}} \\ \tilde{D}_{(Ai)_{\mathrm{fm}}} \\ \tilde{D}_{(Ai)_{\mathrm{mf}}} \end{pmatrix}$$

$$= m\, d_{i\,k} \begin{pmatrix} 1 - t_{(Ai)_{\mathrm{ff}} \leftarrow (Ai)_{\mathrm{ff}}} & - t_{(Ai)_{\mathrm{ff}} \leftarrow (Ai)_{\mathrm{fm}}} & - t_{(Ai)_{\mathrm{ff}} \leftarrow (Ai)_{\mathrm{mf}}} \\ - t_{(Ai)_{\mathrm{fm}} \leftarrow (Ai)_{\mathrm{ff}}} & 1 - t_{(Ai)_{\mathrm{fm}} \leftarrow (Ai)_{\mathrm{fm}}} & - t_{(Ai)_{\mathrm{fm}} \leftarrow (Ai)_{\mathrm{mf}}} \\ - t_{(Ai)_{\mathrm{mf}} \leftarrow (Ai)_{\mathrm{ff}}} & - t_{(Ai)_{\mathrm{mf}} \leftarrow (Ai)_{\mathrm{fm}}} & 1 - t_{(Ai)_{\mathrm{mf}} \leftarrow (Ai)_{\mathrm{mf}}} \end{pmatrix}$$

$$= m\, d_{i\,k} \begin{pmatrix} 1 - \frac{1}{2}(1 - r^{\mathrm{f}}_{Ai}) & - \frac{1}{2}(1 - r^{\mathrm{f}}_{Ai}) & 0 \\ 0 & 1 & - \frac{1}{2}(1 - r^{\mathrm{m}}_{Ai}) \\ - \frac{1}{2}(1 - r^{\mathrm{f}}_{Ai}) & - \frac{1}{2}(1 - r^{\mathrm{f}}_{Ai}) & 1 \end{pmatrix}^{-1} \tag{C11}$$

$$= m\, d_{Ai} \begin{pmatrix} \frac{3 - r^{\mathrm{f}}_{Ai}}{r^{\mathrm{f}}_{Ai}(2 - r^{\mathrm{m}}_{Ai}) + r^{\mathrm{m}}_{Ai}} \\ \frac{3 + r^{\mathrm{f}}_{Ai} - 2\, r^{\mathrm{m}}_{Ai}}{r^{\mathrm{f}}_{Ai}(2 - r^{\mathrm{m}}_{Ai}) + r^{\mathrm{m}}_{Ai}} \\ \frac{3 - r^{\mathrm{f}}_{Ai}}{r^{\mathrm{f}}_{Ai}(2 - r^{\mathrm{m}}_{Ai}) + r^{\mathrm{m}}_{Ai}} \end{pmatrix},$$

where r^{f}_{Ai} is the probability that a recombination event occurs somewhere among the loci in set *Ai* in females, and r^{m}_{Ai} is the corresponding rate in males. These results show that for loose linkage (which is consistent with the assumptions of the approximation), the magnitudes of the associations are greater when recombination rates between the preference locus *i* and the incompatibility loci *A* are smaller.

These results allow us to work out the associations needed to calculate the strength of reinforcement for any mixture of modes of inheritance for the preference and incompatibility loci. The results simplify greatly, however, if all the preference loci have the same recombination rates with all the incompatibility loci. This happens in two situations: when there is just a single preference locus, or when the recombination rates r^{f}_{Ai} and r^{m}_{Ai} are the same for all preference loci *i*. That happens when the preference loci are autosomally inherited, and unlinked to any incompatibility loci. This is a biologically plausible situation, and so we will make that assumption in what follows. In this case, a transmission coefficient involving a preference locus *i* can be written

$$t_{\mathbb{A}i_{xy}\leftarrow\mathbb{B}i_{yz}} = \tfrac{1}{2}\, t_{\mathbb{A}\leftarrow\mathbb{B}}, \qquad \text{(C12)}$$

where *x*, *y*, and *z* can take the values m and f. Substituting that into equation (C8) gives the associations when the preference is autosomal and unlinked to the incompatibility loci, which themselves are X-linked or a mixture of X-linked and autosomal:

$$\begin{pmatrix} \tilde{D}_{(Ai)_{\mathrm{ff}}} \\ \tilde{D}_{(Ai)_{\mathrm{fm}}} \\ \tilde{D}_{(Ai)_{\mathrm{mf}}} \end{pmatrix} = 2\,m\,d_{Ai} \begin{pmatrix} \frac{5 - r_A^{\mathrm{f}}}{5 + r_A^{\mathrm{f}}(3 - r_A^{\mathrm{m}}) + r_A^{\mathrm{m}}} \\ \frac{5 + r_A^{\mathrm{f}} - 2\,r_A^{\mathrm{m}}}{5 + r_A^{\mathrm{f}}(3 - r_A^{\mathrm{m}}) + r_A^{\mathrm{m}}} \\ \frac{5 - r_A^{\mathrm{f}}}{5 + r_A^{\mathrm{f}}(3 - r_A^{\mathrm{m}}) + r_A^{\mathrm{m}}} \end{pmatrix}. \qquad \text{(C13)}$$

All other forms of inheritance can be worked out in the same way.

We can finally write the associations needed in (A7) for the case of autosomal preferences. For a preference position i and a set of incompatibility positions $\mathbb{A}$,

$$\tilde{D}_{\mathbb{A}i} = m\,d_{Ai}\,\phi_{\mathbb{A}}, \qquad \text{(C14)}$$

where $\phi_{\mathbb{A}} = 0$ if $\mathbb{A}$ contains both sexes of origin or both sexes of carrier. When $\mathbb{A}i$ contains only one sex of carrier and only one sex of origin, the value of $\phi_{\mathbb{A}}$ depends on the mode of inheritance of the preference and the incompatibility loci. Values for $\phi_{\mathbb{A}}$

assuming autosomal preference and values assuming X-linked preference are given in Table C1 below. In the table, A indicates that all loci are autosomally inherited, X indicates that all loci are X-linked, Y indicates that all loci are Y-linked, A-X indicates that one or more locus is autosomally inherited and one or more locus is X-linked, and A-Y indicates that one or more locus is autosomally inherited and one or more locus is Y-linked.

Table C1. The effect of recombination and sex linkage ($\phi_{\mathbb{A}}$) on reinforcement.

Autosomal preference

	$\phi_{(A)_{\mathrm{ff}}}$	$\phi_{(A)_{\mathrm{fm}}}$
A	$\frac{4 - r^{\mathrm{f}}_{A_{\mathrm{A}}} + r^{\mathrm{m}}_{A_{\mathrm{A}}}}{2 + r^{\mathrm{f}}_{A_{\mathrm{A}}} + r^{\mathrm{m}}_{A_{\mathrm{A}}}}$	$\frac{4 + r^{\mathrm{f}}_{A_{\mathrm{A}}} - r^{\mathrm{m}}_{A_{\mathrm{A}}}}{2 + r^{\mathrm{f}}_{A_{\mathrm{A}}} + r^{\mathrm{m}}_{A_{\mathrm{A}}}}$
X	$\frac{2\,(5 - r^{\mathrm{f}}_{A_{\mathrm{X}}})}{5 + 3\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}$	$\frac{2\,(5 - r^{\mathrm{f}}_{A_{\mathrm{X}}})}{5 + 3\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}$
Y	0	0
A-X	$\frac{36 - 4\,r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(1 - r^{\mathrm{f}}_{A_{\mathrm{X}}}) - 4\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}{27 + r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(1 - r^{\mathrm{f}}_{A_{\mathrm{X}}})\,(5 - r^{\mathrm{m}}_{A_{\mathrm{A}}}) + r^{\mathrm{f}}_{A_{\mathrm{X}}}\,(5 - r^{\mathrm{m}}_{A_{\mathrm{A}}}) + r^{\mathrm{m}}_{A_{\mathrm{A}}}}$	$\frac{4\,(9 + r^{\mathrm{f}}_{A_{\mathrm{A}}} + r^{\mathrm{f}}_{A_{\mathrm{X}}} - r^{\mathrm{f}}_{A_{\mathrm{A}}}\,r^{\mathrm{f}}_{A_{\mathrm{X}}} - 2\,r^{\mathrm{m}}_{A_{\mathrm{A}}})}{27 + r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(1 - r^{\mathrm{f}}_{A_{\mathrm{X}}})\,(5 - r^{\mathrm{m}}_{A_{\mathrm{A}}}) + r^{\mathrm{f}}_{A_{\mathrm{X}}}\,(5 - r^{\mathrm{m}}_{A_{\mathrm{A}}}) + r^{\mathrm{m}}_{A_{\mathrm{A}}}}$
A-Y	0	0
X-Y	0	0

	$\phi_{(A)_{\mathrm{mf}}}$	$\phi_{(A)_{\mathrm{mm}}}$
A	$\frac{4 - r^{\mathrm{f}}_{A_{\mathrm{A}}} + r^{\mathrm{m}}_{A_{\mathrm{A}}}}{2 + r^{\mathrm{f}}_{A_{\mathrm{A}}} + r^{\mathrm{m}}_{A_{\mathrm{A}}}}$	$\frac{4 + r^{\mathrm{f}}_{A_{\mathrm{A}}} - r^{\mathrm{m}}_{A_{\mathrm{A}}}}{2 + r^{\mathrm{f}}_{A_{\mathrm{A}}} + r^{\mathrm{m}}_{A_{\mathrm{A}}}}$
X	$\frac{2\,(5 - r^{\mathrm{f}}_{A_{\mathrm{X}}})}{5 + 3\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}$	0
Y	0	2
A-X	$\frac{36 - 4\,r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(1 - r^{\mathrm{f}}_{A_{\mathrm{X}}}) - 4\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}{27 + r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(1 - r^{\mathrm{f}}_{A_{\mathrm{X}}})\,(5 - r^{\mathrm{m}}_{A_{\mathrm{A}}}) + r^{\mathrm{f}}_{A_{\mathrm{X}}}\,(5 - r^{\mathrm{m}}_{A_{\mathrm{A}}}) + r^{\mathrm{m}}_{A_{\mathrm{A}}}}$	0
A-Y	0	$\frac{4}{3 + r^{\mathrm{m}}_{A_{\mathrm{A}}}}$
X-Y	0	0

X-linked preference

	$\phi_{(A)_{\mathrm{ff}}}$	$\phi_{(A)_{\mathrm{fm}}}$
A	$\frac{2\,(5-r^{\mathrm{f}}_{A_{\mathrm{A}}})}{5-r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(-3+r^{\mathrm{m}}_{A_{\mathrm{A}}})+r^{\mathrm{m}}_{A_{\mathrm{A}}}}$	$\frac{2\,(5+r^{\mathrm{f}}_{A_{\mathrm{A}}}-2\,r^{\mathrm{m}}_{A_{\mathrm{A}}})}{5+r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(3-r^{\mathrm{m}}_{A_{\mathrm{A}}})+r^{\mathrm{m}}_{A_{\mathrm{A}}}}$
X	$\frac{5-r^{\mathrm{f}}_{A_{\mathrm{X}}}}{2+2\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}$	$\frac{7+r^{\mathrm{f}}_{A_{\mathrm{X}}}}{2+2\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}$
Y	0	0
A-X	$\frac{18-2\,r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(1-r^{\mathrm{f}}_{A_{\mathrm{X}}})-2\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}{13+r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(1-r^{\mathrm{f}}_{A_{\mathrm{X}}})\,(3-r^{\mathrm{m}}_{A_{\mathrm{A}}})+r^{\mathrm{f}}_{A_{\mathrm{X}}}\,(3-r^{\mathrm{m}}_{A_{\mathrm{A}}})+r^{\mathrm{m}}_{A_{\mathrm{A}}}}$	$\frac{2\,(11+r^{\mathrm{f}}_{A_{\mathrm{A}}}+r^{\mathrm{f}}_{A_{\mathrm{X}}}-r^{\mathrm{f}}_{A_{\mathrm{A}}}\,r^{\mathrm{f}}_{A_{\mathrm{X}}}-4\,r^{\mathrm{m}}_{A_{\mathrm{A}}})}{13+r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(1-r^{\mathrm{f}}_{A_{\mathrm{X}}})\,(3-r^{\mathrm{m}}_{A_{\mathrm{A}}})+r^{\mathrm{f}}_{A_{\mathrm{X}}}\,(3-r^{\mathrm{m}}_{A_{\mathrm{A}}})+r^{\mathrm{m}}_{A_{\mathrm{A}}}}$
A-Y	0	0
X-Y	0	0

	$\phi_{(A)_{\mathrm{mf}}}$	$\phi_{(A)_{\mathrm{mm}}}$
A	$\frac{2\,(5-r^{\mathrm{f}}_{A_{\mathrm{A}}})}{5+r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(3-r^{\mathrm{m}}_{A_{\mathrm{A}}})+r^{\mathrm{m}}_{A_{\mathrm{A}}}}$	0
X	$\frac{5-r^{\mathrm{f}}_{A_{\mathrm{X}}}}{2+2\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}$	0
Y	0	0
A-X	$\frac{18-2\,r^{\mathrm{f}}_{A_{\mathrm{A}}}(1-r^{\mathrm{f}}_{A_{\mathrm{X}}})-2\,r^{\mathrm{f}}_{A_{\mathrm{X}}}}{13+r^{\mathrm{f}}_{A_{\mathrm{A}}}\,(1-r^{\mathrm{f}}_{A_{\mathrm{X}}})\,(3-r^{\mathrm{m}}_{A_{\mathrm{A}}})+r^{\mathrm{f}}_{A_{\mathrm{X}}}\,(3-r^{\mathrm{m}}_{A_{\mathrm{A}}})+r^{\mathrm{m}}_{A_{\mathrm{A}}}}$	0
A-Y	0	0
X-Y	0	0

Appendix D

Selection Coefficients for One- and Two-Locus Incompatibilities

Here we present $O(a)$ approximations for the selection coefficients (the $a_{\mathbb{A}}$) that appear in equation (A12) under a variety of assumptions about how the hybrid incompatibility genes are inherited. The general procedure for deriving the selection coefficients is given in KJB. In the next section we give, as an example, the derivation for the case when hybrid incompatibilities are caused by two autosomal loci. The results for the selection coefficients for other types of hybrid incompatibilities, derived in a similar fashion, are summarized below in Table D1.

Assuming that interactions between alleles at two autosomal loci, *k* and *l*, contribute to decreased hybrid fitness, our notation for the selection coefficients is given in the bottom left corner of Figure 3.1. We allow females and males to have different fitnesses, so we use the subscripts f and m to denote selection coefficients pertaining to females and males, respectively. Given this notation we can write the general fitness functions as

$$
\begin{aligned}
W_{\mathrm{f}} \;=\; & 1 - s_{00_{\mathrm{f}}}((1 - X_{k_{\mathrm{ff}}})(1 - X_{k_{\mathrm{fm}}})(1 - X_{l_{\mathrm{ff}}})(1 - X_{l_{\mathrm{fm}}})) - \\
& s_{01_{\mathrm{f}}}((1 - X_{k_{\mathrm{ff}}})(1 - X_{k_{\mathrm{fm}}})(1 - X_{l_{\mathrm{ff}}})\, X_{l_{\mathrm{fm}}}) - \\
& s_{01_{\mathrm{f}}}((1 - X_{k_{\mathrm{ff}}})(1 - X_{k_{\mathrm{fm}}})\, X_{l_{\mathrm{ff}}}(1 - X_{l_{\mathrm{fm}}})) - \\
& s_{02_{\mathrm{f}}}((1 - X_{k_{\mathrm{ff}}})(1 - X_{k_{\mathrm{fm}}})\, X_{l_{\mathrm{ff}}}\, X_{l_{\mathrm{fm}}}) - \\
& s_{10_{\mathrm{f}}}((1 - X_{k_{\mathrm{ff}}})\, X_{k_{\mathrm{fm}}}(1 - X_{l_{\mathrm{ff}}})(1 - X_{l_{\mathrm{fm}}})) - \\
& s_{11_{\mathrm{f}}}((1 - X_{k_{\mathrm{ff}}})\, X_{k_{\mathrm{fm}}}(1 - X_{l_{\mathrm{ff}}})\, X_{l_{\mathrm{fm}}}) - \\
& s_{11_{\mathrm{f}}}((1 - X_{k_{\mathrm{ff}}})\, X_{k_{\mathrm{fm}}}\, X_{l_{\mathrm{ff}}}(1 - X_{l_{\mathrm{fm}}})) - \\
& s_{12_{\mathrm{f}}}((1 - X_{k_{\mathrm{ff}}})\, X_{k_{\mathrm{fm}}}\, X_{l_{\mathrm{ff}}}\, X_{l_{\mathrm{fm}}}) - \\
& s_{10_{\mathrm{f}}}(X_{k_{\mathrm{ff}}}(1 - X_{k_{\mathrm{fm}}})(1 - X_{l_{\mathrm{ff}}})(1 - X_{l_{\mathrm{fm}}})) - \\
& s_{11_{\mathrm{f}}}(X_{k_{\mathrm{ff}}}(1 - X_{k_{\mathrm{fm}}})(1 - X_{l_{\mathrm{ff}}})\, X_{l_{\mathrm{fm}}}) - \\
& s_{11_{\mathrm{f}}}(X_{k_{\mathrm{ff}}}(1 - X_{k_{\mathrm{fm}}})\, X_{l_{\mathrm{ff}}}(1 - X_{l_{\mathrm{fm}}})) - \\
& s_{12_{\mathrm{f}}}(X_{k_{\mathrm{ff}}}(1 - X_{k_{\mathrm{fm}}})\, X_{l_{\mathrm{ff}}}\, X_{l_{\mathrm{fm}}}) - \\
& s_{20_{\mathrm{f}}}(X_{k_{\mathrm{ff}}}\, X_{k_{\mathrm{fm}}}(1 - X_{l_{\mathrm{ff}}})(1 - X_{l_{\mathrm{fm}}})) - \\
& s_{21_{\mathrm{f}}}(X_{k_{\mathrm{ff}}}\, X_{k_{\mathrm{fm}}}(1 - X_{l_{\mathrm{ff}}})\, X_{l_{\mathrm{fm}}}) - s_{21_{\mathrm{f}}}(X_{k_{\mathrm{ff}}}\, X_{k_{\mathrm{f\,m}}}\, X_{l_{\mathrm{ff}}}(1 - X_{l_{\mathrm{fm}}}))
\end{aligned}
\tag{D1}
$$

and

$$
\begin{aligned}
W_{\mathrm{m}} \;=\; & 1 - s_{00_{\mathrm{m}}}((1 - X_{k_{\mathrm{mf}}})(1 - X_{k_{\mathrm{mm}}})(1 - X_{l_{\mathrm{mf}}})(1 - X_{l_{\mathrm{mm}}})) - \\
& s_{01_{\mathrm{m}}}((1 - X_{k_{\mathrm{mf}}})(1 - X_{k_{\mathrm{mm}}})(1 - X_{l_{\mathrm{mf}}})\, X_{l_{\mathrm{mm}}}) - \\
& s_{01_{\mathrm{m}}}((1 - X_{k_{\mathrm{mf}}})(1 - X_{k_{\mathrm{mm}}})\, X_{l_{\mathrm{mf}}}(1 - X_{l_{\mathrm{mm}}})) - \\
& s_{02_{\mathrm{m}}}((1 - X_{k_{\mathrm{mf}}})(1 - X_{k_{\mathrm{mm}}})\, X_{l_{\mathrm{mf}}}\, X_{l_{\mathrm{mm}}}) - \\
& s_{10_{\mathrm{m}}}((1 - X_{k_{\mathrm{mf}}})\, X_{k_{\mathrm{mm}}}(1 - X_{l_{\mathrm{mf}}})(1 - X_{l_{\mathrm{mm}}})) - \\
& s_{11_{\mathrm{m}}}((1 - X_{k_{\mathrm{mf}}})\, X_{k_{\mathrm{mm}}}(1 - X_{l_{\mathrm{mf}}})\, X_{l_{\mathrm{mm}}}) - \\
& s_{11_{\mathrm{m}}}((1 - X_{k_{\mathrm{mf}}})\, X_{k_{\mathrm{mm}}}\, X_{l_{\mathrm{mf}}}(1 - X_{l_{\mathrm{mm}}})) - \\
& s_{12_{\mathrm{m}}}((1 - X_{k_{\mathrm{mf}}})\, X_{k_{\mathrm{mm}}}\, X_{l_{\mathrm{mf}}}\, X_{l_{\mathrm{mm}}}) - \\
& s_{10_{\mathrm{m}}}(X_{k_{\mathrm{mf}}}(1 - X_{k_{\mathrm{mm}}})(1 - X_{l_{\mathrm{mf}}})(1 - X_{l_{\mathrm{mm}}})) - \\
& s_{11_{\mathrm{m}}}(X_{k_{\mathrm{mf}}}(1 - X_{k_{\mathrm{mm}}})(1 - X_{l_{\mathrm{mf}}})\, X_{l_{\mathrm{mm}}}) - \\
& s_{11_{\mathrm{m}}}(X_{k_{\mathrm{mf}}}(1 - X_{k_{\mathrm{mm}}})\, X_{l_{\mathrm{mf}}}(1 - X_{l_{\mathrm{mm}}})) - \\
& s_{12_{\mathrm{m}}}(X_{k_{\mathrm{mf}}}(1 - X_{k_{\mathrm{mm}}})\, X_{l_{\mathrm{mf}}}\, X_{l_{\mathrm{mm}}}) - \\
& s_{20_{\mathrm{m}}}(X_{k_{\mathrm{mf}}}\, X_{k_{\mathrm{mm}}}(1 - X_{l_{\mathrm{mf}}})(1 - X_{l_{\mathrm{mm}}})) - \\
& s_{21_{\mathrm{m}}}(X_{k_{\mathrm{mf}}}\, X_{k_{\mathrm{mm}}}(1 - X_{l_{\mathrm{mf}}})\, X_{l_{\mathrm{mm}}}) - s_{21_{\mathrm{m}}}(X_{k_{\mathrm{mf}}}\, X_{k_{\mathrm{m\,m}}}\, X_{l_{\mathrm{mf}}}(1 - X_{l_{\mathrm{mm}}}))
\end{aligned}
\tag{D2}
$$

where W_f is the fitness function for the females and W_m is the fitness function for the males. X_i takes the value 1 when the genotype of interest contains the position i and takes the value 0 when it does not. Replacing the X's using

$\zeta_k = X_k - p_k \approx X_k - 1 + O(s) \approx X_k - 1$ yields

$$
\begin{aligned}
W_f = {} & 1 - s_{00_f}\,\zeta_{k_{ff}}\,\zeta_{k_{fm}}\,\zeta_{l_{ff}}\,\zeta_{l_{fm}} + \\
& s_{10_f}\,(1+\zeta_{k_{ff}})\,\zeta_{k_{fm}}\,\zeta_{l_{ff}}\,\zeta_{l_{fm}} + s_{10_f}\,\zeta_{k_{ff}}\,(1+\zeta_{k_{fm}})\,\zeta_{l_{ff}}\,\zeta_{l_{fm}} - \\
& s_{20_f}\,(1+\zeta_{k_{ff}})\,(1+\zeta_{k_{fm}})\,\zeta_{l_{ff}}\,\zeta_{l_{fm}} + s_{01_f}\,\zeta_{k_{ff}}\,\zeta_{k_{fm}}\,(1+\zeta_{l_{ff}})\,\zeta_{l_{fm}} - \\
& s_{11_f}\,(1+\zeta_{k_{ff}})\,\zeta_{k_{fm}}\,(1+\zeta_{l_{ff}})\,\zeta_{l_{fm}} - s_{11_f}\,\zeta_{k_{ff}}\,(1+\zeta_{k_{fm}})\,(1+\zeta_{l_{ff}})\,\zeta \\
& s_{21_f}\,(1+\zeta_{k_{ff}})\,(1+\zeta_{k_{fm}})\,(1+\zeta_{l_{ff}})\,\zeta_{l_{fm}} + s_{01_f}\,\zeta_{k_{ff}}\,\zeta_{k_{fm}}\,\zeta_{l_{ff}}\,(1+\zeta_l \\
& s_{11_f}\,(1+\zeta_{k_{ff}})\,\zeta_{k_{fm}}\,\zeta_{l_{ff}}\,(1+\zeta_{l_{fm}}) - s_{11_f}\,\zeta_{k_{ff}}\,(1+\zeta_{k_{fm}})\,\zeta_{l_{ff}}\,(1+\zeta_l \\
& s_{21_f}\,(1+\zeta_{k_{ff}})\,(1+\zeta_{k_{fm}})\,\zeta_{l_{ff}}\,(1+\zeta_{l_{fm}}) - \\
& s_{02_f}\,\zeta_{k_{ff}}\,\zeta_{k_{fm}}\,(1+\zeta_{l_{ff}})\,(1+\zeta_{l_{fm}}) + \\
& s_{12_f}\,(1+\zeta_{k_{ff}})\,\zeta_{k_{fm}}\,(1+\zeta_{l_{ff}})\,(1+\zeta_{l_{fm}}) + \\
& s_{12_f}\,\zeta_{k_{ff}}\,(1+\zeta_{k_{fm}})\,(1+\zeta_{l_{ff}})\,(1+\zeta_{l_{fm}})
\end{aligned}
\tag{D3}
$$

and

$$
\begin{aligned}
W_{\mathrm{m}} \;=\; & 1 - s_{00_{\mathrm{m}}}\,\zeta_{k_{\mathrm{mf}}}\,\zeta_{k_{\mathrm{mm}}}\,\zeta_{l_{\mathrm{mf}}}\,\zeta_{l_{\mathrm{mm}}} + s_{10_{\mathrm{m}}}\,(1+\zeta_{k_{\mathrm{mf}}})\,\zeta_{k_{\mathrm{mm}}}\,\zeta_{l_{\mathrm{mf}}}\,\zeta_{l_{\mathrm{mm}}} + \\
& s_{10_{\mathrm{m}}}\,\zeta_{k_{\mathrm{mf}}}\,(1+\zeta_{k_{\mathrm{mm}}})\,\zeta_{l_{\mathrm{mf}}}\,\zeta_{l_{\mathrm{mm}}} - s_{20_{\mathrm{m}}}\,(1+\zeta_{k_{\mathrm{mf}}})\,(1+\zeta_{k_{\mathrm{mm}}})\,\zeta_{l_{\mathrm{mf}}}\,\zeta_{l_{\mathrm{mm}}} + \\
& s_{01_{\mathrm{m}}}\,\zeta_{k_{\mathrm{mf}}}\,\zeta_{k_{\mathrm{mm}}}\,(1+\zeta_{l_{\mathrm{mf}}})\,\zeta_{l_{\mathrm{mm}}} - s_{11_{\mathrm{m}}}\,(1+\zeta_{k_{\mathrm{mf}}})\,\zeta_{k_{\mathrm{mm}}}\,(1+\zeta_{l_{\mathrm{mf}}})\,\zeta_{l_{\mathrm{mm}}} - \\
& s_{11_{\mathrm{m}}}\,\zeta_{k_{\mathrm{mf}}}\,(1+\zeta_{k_{\mathrm{mm}}})\,(1+\zeta_{l_{\mathrm{mf}}})\,\zeta_{l_{\mathrm{mm}}} + \\
& s_{21_{\mathrm{m}}}\,(1+\zeta_{k_{\mathrm{mf}}})\,(1+\zeta_{k_{\mathrm{mm}}})\,(1+\zeta_{l_{\mathrm{mf}}})\,\zeta_{l_{\mathrm{mm}}} + \\
& s_{01_{\mathrm{m}}}\,\zeta_{k_{\mathrm{mf}}}\,\zeta_{k_{\mathrm{mm}}}\,\zeta_{l_{\mathrm{mf}}}\,(1+\zeta_{l_{\mathrm{mm}}}) - s_{11_{\mathrm{m}}}\,(1+\zeta_{k_{\mathrm{mf}}})\,\zeta_{k_{\mathrm{mm}}}\,\zeta_{l_{\mathrm{mf}}}\,(1+\zeta_{l_{\mathrm{mm}}}) - \\
& s_{11_{\mathrm{m}}}\,\zeta_{k_{\mathrm{mf}}}\,(1+\zeta_{k_{\mathrm{mm}}})\,\zeta_{l_{\mathrm{mf}}}\,(1+\zeta_{l_{\mathrm{mm}}}) + \\
& s_{21_{\mathrm{m}}}\,(1+\zeta_{k_{\mathrm{mf}}})\,(1+\zeta_{k_{\mathrm{mm}}})\,\zeta_{l_{\mathrm{mf}}}\,(1+\zeta_{l_{\mathrm{mm}}}) - \\
& s_{02_{\mathrm{m}}}\,\zeta_{k_{\mathrm{mf}}}\,\zeta_{k_{\mathrm{mm}}}\,(1+\zeta_{l_{\mathrm{mf}}})\,(1+\zeta_{l_{\mathrm{mm}}}) + \\
& s_{12_{\mathrm{m}}}\,(1+\zeta_{k_{\mathrm{mf}}})\,\zeta_{k_{\mathrm{mm}}}\,(1+\zeta_{l_{\mathrm{mf}}})\,(1+\zeta_{l_{\mathrm{mm}}}) + \\
& s_{12_{\mathrm{m}}}\,\zeta_{k_{\mathrm{mf}}}\,(1+\zeta_{k_{\mathrm{mm}}})\,(1+\zeta_{l_{\mathrm{mf}}})\,(1+\zeta_{l_{\mathrm{mm}}})
\end{aligned}
\qquad \text{(D4)}
$$

Lastly, we determine the selection coefficients by looking for the coefficients of the ζ terms who's subscript matches that of the a that we desire, which are the a containing one position or two positions with the same sex-of-carrier and sex-of-origin. The a's are given in Table D1 for the case of two autosomal loci, as well as for all other combinations assuming X-Y or Z-W sex determination.

Table D1. Selection coefficients ($a_{\mathbb{A}}$).

	$a_{k_{\mathrm{ff}}}$	$a_{k_{\mathrm{fm}}}$	$a_{k_{\mathrm{mf}}}$	$a_{k_{\mathrm{mm}}}$	$a_{l_{\mathrm{ff}}}$	$a_{l_{\mathrm{fm}}}$	$a_{l_{\mathrm{mf}}}$	$a_{l_{\mathrm{mm}}}$
A	$s_{1_{\mathrm{f}}}$	$s_{1_{\mathrm{f}}}$	$s_{1_{\mathrm{m}}}$	$s_{1_{\mathrm{m}}}$	0	0	0	0
X	$s_{1_{\mathrm{f}}}$	$s_{1_{\mathrm{f}}}$	$s_{0_{\mathrm{m}}}$	0	0	0	0	0
Y	0	0	0	$s_{0_{\mathrm{m}}}$	0	0	0	0
Z	0	$s_{0_{\mathrm{f}}}$	$s_{1_{\mathrm{m}}}$	$s_{1_{\mathrm{m}}}$	0	0	0	0
W	$s_{0_{\mathrm{f}}}$	0	0	0	0	0	0	0

	$a_{k_{\mathrm{ff}}}$	$a_{k_{\mathrm{fm}}}$	$a_{k_{\mathrm{mf}}}$	$a_{k_{\mathrm{mm}}}$	$a_{l_{\mathrm{ff}}}$	$a_{l_{\mathrm{fm}}}$	$a_{l_{\mathrm{mf}}}$	$a_{l_{\mathrm{mm}}}$
A-A	$s_{12_{\mathrm{f}}}$	$s_{12_{\mathrm{f}}}$	$s_{12_{\mathrm{m}}}$	$s_{12_{\mathrm{m}}}$	$s_{21_{\mathrm{f}}}$	$s_{21_{\mathrm{f}}}$	$s_{21_{\mathrm{m}}}$	$s_{21_{\mathrm{m}}}$
X-X	$s_{12_{\mathrm{f}}}$	$s_{12_{\mathrm{f}}}$	$s_{02_{\mathrm{m}}}$	0	$s_{21_{\mathrm{f}}}$	$s_{21_{\mathrm{f}}}$	$s_{20_{\mathrm{m}}}$	0
A-X	$s_{12_{\mathrm{f}}}$	$s_{12_{\mathrm{f}}}$	$s_{12_{\mathrm{m}}}$	$s_{12_{\mathrm{m}}}$	$s_{21_{\mathrm{f}}}$	$s_{21_{\mathrm{f}}}$	$s_{20_{\mathrm{m}}}$	0
A-Y	0	0	$s_{12_{\mathrm{m}}}$	$s_{12_{\mathrm{m}}}$	0	0	0	$s_{20_{\mathrm{m}}}$
X-Y	0	0	$s_{02_{\mathrm{m}}}$	0	0	0	0	$s_{20_{\mathrm{m}}}$
Z-Z	0	$s_{02_{\mathrm{f}}}$	$s_{12_{\mathrm{m}}}$	$s_{12_{\mathrm{m}}}$	0	$s_{20_{\mathrm{f}}}$	$s_{21_{\mathrm{m}}}$	$s_{21_{\mathrm{m}}}$
A-Z	$s_{12_{\mathrm{f}}}$	$s_{12_{\mathrm{f}}}$	$s_{12_{\mathrm{m}}}$	$s_{12_{\mathrm{m}}}$	0	$s_{20_{\mathrm{f}}}$	$s_{21_{\mathrm{m}}}$	$s_{21_{\mathrm{m}}}$
A-W	$s_{12_{\mathrm{f}}}$	$s_{12_{\mathrm{f}}}$	0	0	$s_{20_{\mathrm{f}}}$	0	0	0
Z-W	0	$s_{02_{\mathrm{f}}}$	0	0	$s_{20_{\mathrm{f}}}$	0	0	0

	$a_{k_{\mathrm{ff}}\, l_{\mathrm{ff}}}$	$a_{k_{\mathrm{fm}}\, l_{\mathrm{fm}}}$
A-A	$-s_{11_{\mathrm{f}}} + s_{12_{\mathrm{f}}} + s_{21_{\mathrm{f}}}$	$-s_{11_{\mathrm{f}}} + s_{12_{\mathrm{f}}} + s_{21_{\mathrm{f}}}$
X-X	$-s_{11_{\mathrm{f}}} + s_{12_{\mathrm{f}}} + s_{21_{\mathrm{f}}}$	$-s_{11_{\mathrm{f}}} + s_{12_{\mathrm{f}}} + s_{21_{\mathrm{f}}}$
A-X	$-s_{11_{\mathrm{f}}} + s_{12_{\mathrm{f}}} + s_{21_{\mathrm{f}}}$	$-s_{11_{\mathrm{f}}} + s_{12_{\mathrm{f}}} + s_{21_{\mathrm{f}}}$
A-Y	0	0
X-Y	0	0
Z-Z	0	$-s_{00_{\mathrm{f}}} + s_{02_{\mathrm{f}}} + s_{20_{\mathrm{f}}}$
A-Z	0	$-s_{10_{\mathrm{f}}} + s_{12_{\mathrm{f}}} + s_{20_{\mathrm{f}}}$
A-W	$-s_{10_{\mathrm{f}}} + s_{12_{\mathrm{f}}} + s_{20_{\mathrm{f}}}$	0
Z-W	0	0

	$a_{k_{\mathrm{mf}}\, l_{\mathrm{mf}}}$	$a_{k_{\mathrm{mm}}\, l_{\mathrm{mm}}}$
A-A	$-s_{11_{\mathrm{m}}} + s_{12_{\mathrm{m}}} + s_{21_{\mathrm{m}}}$	$-s_{11_{\mathrm{m}}} + s_{12_{\mathrm{m}}} + s_{21_{\mathrm{m}}}$
X-X	$-s_{00_{\mathrm{m}}} + s_{02_{\mathrm{m}}} + s_{20_{\mathrm{m}}}$	0
A-X	$-s_{10_{\mathrm{m}}} + s_{12_{\mathrm{m}}} + s_{20_{\mathrm{m}}}$	0
A-Y	0	$-s_{10_{\mathrm{m}}} + s_{12_{\mathrm{m}}} + s_{20_{\mathrm{m}}}$
X-Y	0	0
Z-Z	$-s_{11_{\mathrm{m}}} + s_{12_{\mathrm{m}}} + s_{21_{\mathrm{m}}}$	$-s_{11_{\mathrm{m}}} + s_{12_{\mathrm{m}}} + s_{21_{\mathrm{m}}}$
A-Z	$-s_{11_{\mathrm{m}}} + s_{12_{\mathrm{m}}} + s_{21_{\mathrm{m}}}$	$-s_{11_{\mathrm{m}}} + s_{12_{\mathrm{m}}} + s_{21_{\mathrm{m}}}$
A-W	0	0
Z-W	0	0

Appendix E

The Effect of Sex Linkage of Two-Locus Incompatibilities on Reinforcement

In Figure 3.2, we compare the amount of reinforcement for A-A, A-X and X-X incompatibilities. The following figures also present the comparisons involving A-Y and X-Y incompatibilities. The main conclusion that we can draw from these figures is that when the preference is X-linked, reinforcement is expected to much stronger when the incompatibility is a function interactions involving autosomal and/or X-linked genes only, rather than interactions involving one or more Y-linked genes.

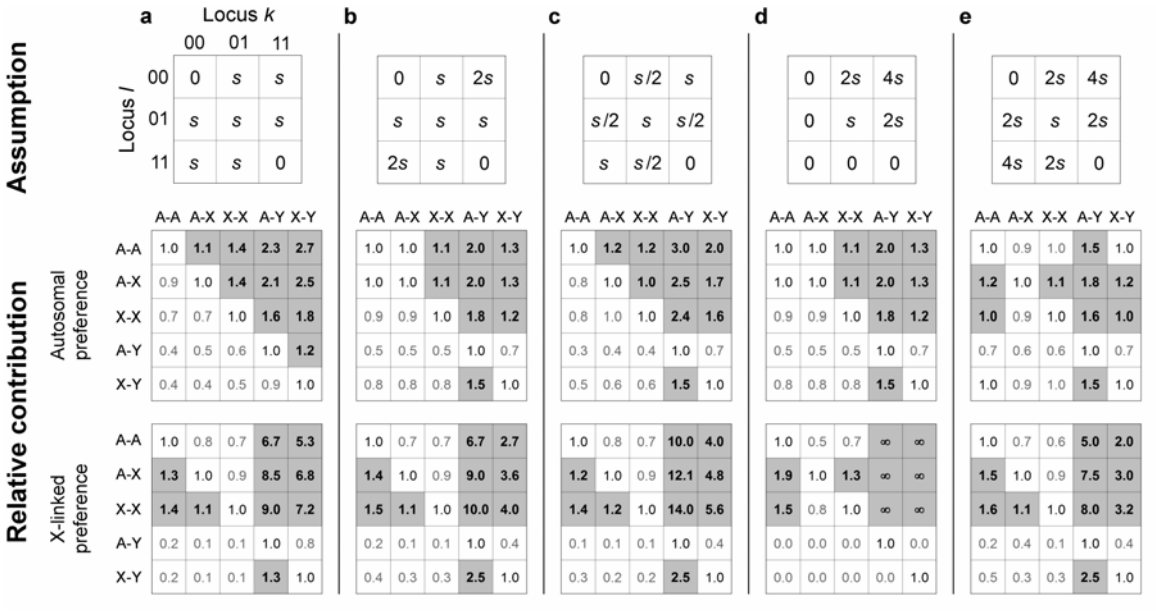

Figure E1. The effect of sex-linkage on reinforcement for five types of hybrid incompatibility. The five matrices in the top row denote five types of hybrid incompatibility at two loci, k and l. Values in the center and bottom rows indicate the relative amounts of reinforcement due to A-A, A-X, A-Y, X-X, and X-Y incompatibilities. The amount of reinforcement for the pattern of sex linkage indicated to the left of the value is divided by that for the pattern indicated above the value. For example, the value 0.2 found in the lower left most square indicates that an A-A incompatibility is expected to produce 5 times more reinforcement than an X-Y incompatibility when the preference is X-linked. Values greater than one before rounding are shaded.

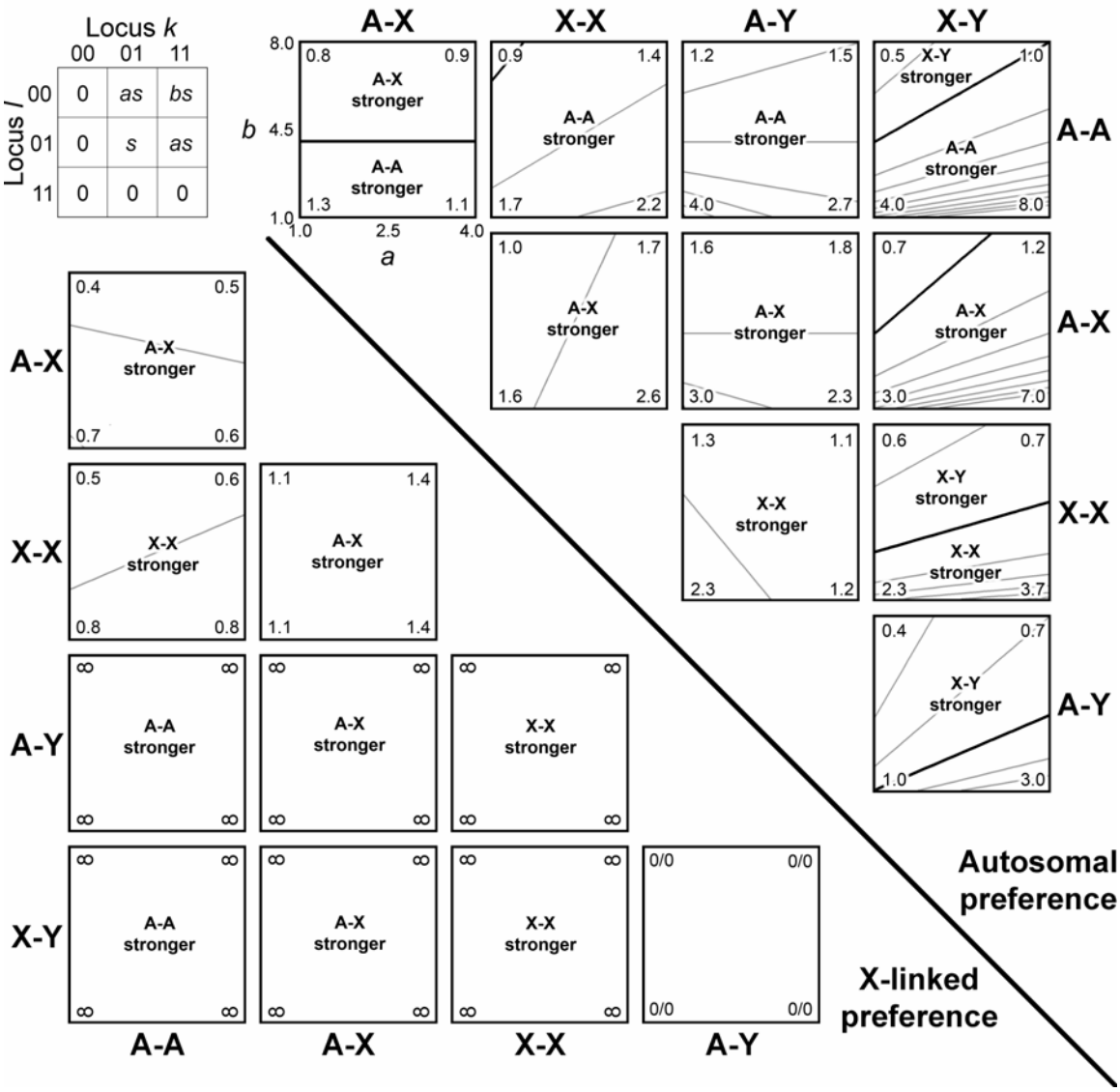

Figure E2. The effect of dominance and epistasis on the contribution to reinforcement for five two-locus incompatibilities that may have evolved by drift. The type of selection assumed is indicated in the matrix located in the upper left corner of the figure, where the pattern of dominance and epistasis is a function of *a* and *b*. Each graph is a contour plot showing the expected amount of reinforcement for the type of incompatibility labeled to the left/right of the graph relative to the amount for the type labeled above/below the graph. A bold line indicates that the amount of reinforcement is equal for the two types of incompatibility being compared in the graph. Values in the corners of each graph indicate the relative amounts of reinforcement at those points in parameter space. For example, the value 1.0 seen in the upper right corner of the figure indicates that A-A and X-Y incompatibilities are expected to produce equal amounts of reinforcement when $a = 4$ and $b = 8$. The presence of many contour lines indicates that conclusions about the relative amount of reinforcement are sensitive to the specific pattern of dominance and epistasis, whereas the presence of fewer lines indicates that conclusions are more robust.

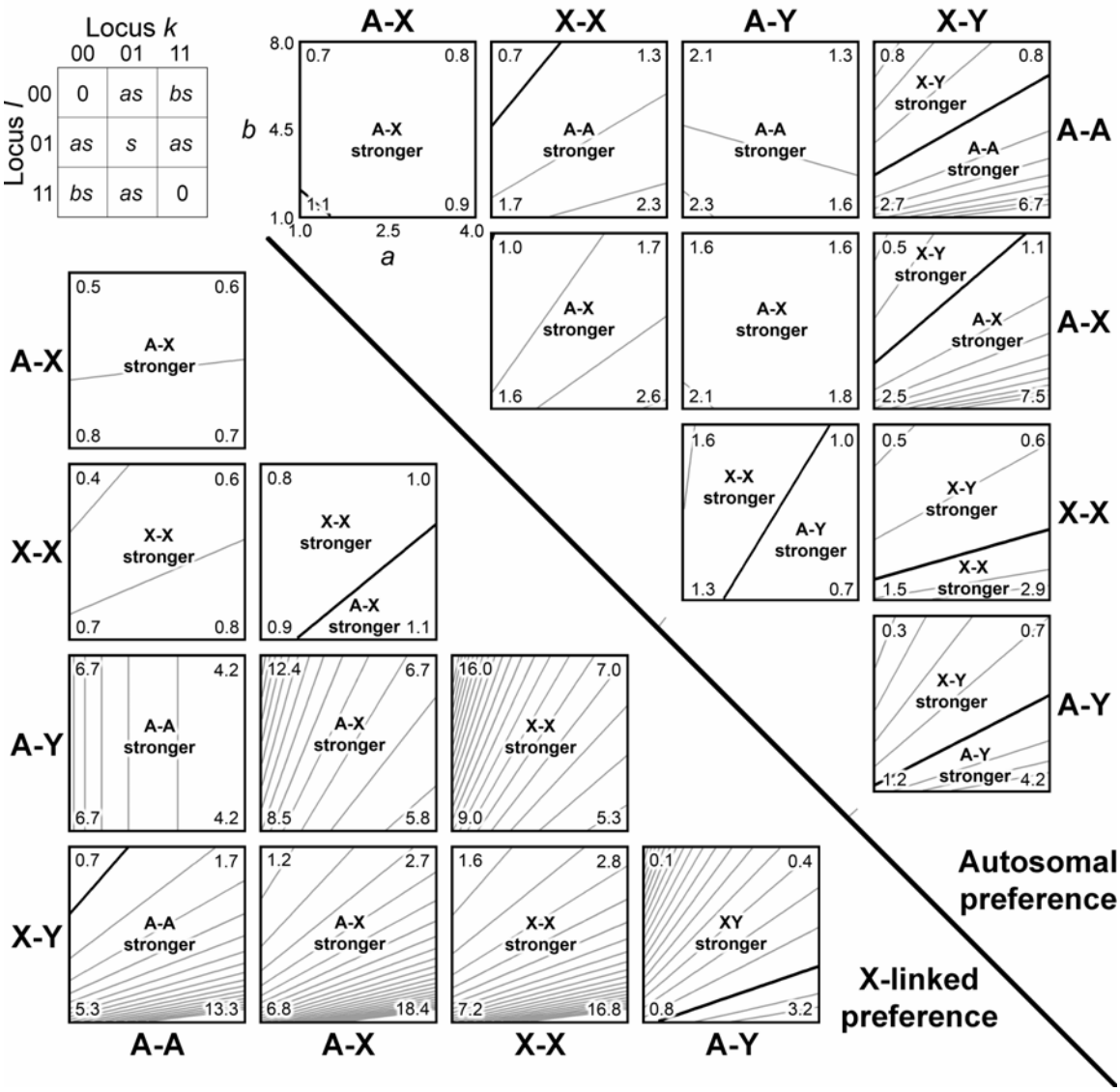

Figure E3. The effect of dominance and epistasis on the contribution to reinforcement for five two-locus incompatibilities that may have evolved by selection. This figure is analogous to Figure E2.

Appendix F

Equilibrium Hybrid Class Frequencies

Here I derive equations describing the equilibrium frequencies of 6 different hybrid classes, denoted A, B, F1, BCA, BCB, and F2. Classes A and B contain pure individuals from species A and B respectively. The F1 class contains first generation hybrids, produced through the mating of class A and B individuals. The BCA and BCB classes contain backcross hybrids, produced through the mating of an F1 hybrid and either a class A or class B individual, respectively. The F2 class contains individuals produced through mating two individuals from the F1 class. Let P_h denote the frequency of hybrid class h in zygotes. Also, let V_{h} denote the viability of hybrids, relative to pure individuals. Note that for simplicity, the viabilities of F1, BCA, BCB, and F2 hybrids are assumed to be equal. This allows us the degree of postzygotic isolation to be determined by a single parameter. When $V_{\mathrm{h}} = 0$, postzygotic isolation is complete, whereas when $V_{\mathrm{h}} = 1$, there is no postzygotic isolation. The frequencies of the hybrid classes after viability selection can be expressed as

$$P'_{A} = \frac{P_{A}}{\bar{V}}$$

$$P'_{B} = \frac{P_{B}}{\bar{V}}$$

$$P'_{F1} = \frac{P_{F1} V_{h}}{\bar{V}}$$

$$P'_{BCA} = \frac{P_{BCA} V_{h}}{\bar{V}} \qquad \text{(F1)}$$

$$P'_{BCB} = \frac{P_{BCB} V_{h}}{\bar{V}}$$

$$P'_{F2} = \frac{P_{F2} V_{h}}{\bar{V}},$$

where $\bar{V}$ is the average viability.

The second assumption is that second generation hybrids (BCA, BCB and F2) do not mate. Therefore, it will be convenient to define P''_{h}, as the frequency of the hybrid class *h* in the mating population. These frequencies can be obtained by normalizing as follows:

$$P''_{A} = \frac{P'_{A}}{P'_{A}+P'_{B}+P'_{F1}}$$

$$P''_{B} = \frac{P'_{B}}{P'_{A}+P'_{B}+P'_{F1}}$$

$$P''_{F1} = \frac{P'_{F1}}{P'_{A}+P'_{B}+P'_{F1}}$$

$$P''_{BCA} = \frac{0}{P'_{A}+P'_{B}+P'_{F1}} = 0 \qquad \text{(F2)}$$

$$P''_{BCB} = \frac{0}{P'_{A}+P'_{B}+P'_{F1}} = 0$$

$$P''_{F2} = \frac{0}{P'_{A}+P'_{B}+P'_{F1}} = 0$$

Let C_P, denote the degree of conspecific preference, a variable specifying the degree of prezygotic isolation. When $C_P = 0$, all individuals mate randomly and there is no prezygotic isolation. When $C_P > 0$, however, a proportion C_P of the pure individuals mate with conspecifics and the remaining proportion, $(1 - C_P)$, mate randomly. All F1 hybrids are assumed to mate randomly, regardless of the degree of prezygotic isolation. Second generation hybrids were assumed to not mate. Under this form of mating, the frequencies of the hybrid classes in zygotes in the next generation are

$$P'''_{\mathrm{A}} = \frac{(P''_{\mathrm{A}}(1-C_{\mathrm{p}}))^2}{(P''_{\mathrm{A}}+P''_{\mathrm{B}})(1-C_{\mathrm{p}})+P''_{\mathrm{F1}}} + P''_{\mathrm{A}} C_{\mathrm{p}}$$

$$= \frac{\left(\frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}(1-C_{\mathrm{p}})\right)^2}{\left(\frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}+\frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}\right)(1-C_{\mathrm{p}})+\frac{P_{\mathrm{F1}} V_{\mathrm{h}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}} + \frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}} C_{\mathrm{p}}$$

$$= \frac{P_A((C_p-1)(P_A+C_p P_B)-C_p P_{\mathrm{F1}} V_h)}{(C_p-1)(P_A+P_B)^2+(C_p-2) P_{\mathrm{F1}} V_h (P_A+P_B)-P_{\mathrm{F1}}^2 V_h^2}$$

$$P'''_{\mathrm{B}} = \frac{(P''_{\mathrm{B}}(1-C_{\mathrm{p}}))^2}{(P''_{\mathrm{A}}+P''_{\mathrm{B}})(1-C_{\mathrm{p}})+P''_{\mathrm{F1}}} + P''_{\mathrm{B}} C_{\mathrm{p}} \qquad \text{(F3)}$$

$$= \frac{\left(\frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}(1-C_{\mathrm{p}})\right)^2}{\left(\frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}+\frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}\right)(1-C_{\mathrm{p}})+\frac{P_{\mathrm{F1}} V_{\mathrm{h}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}} + \frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}} C_{\mathrm{p}}$$

$$= \frac{P_B((C_p-1)(C_p P_A+P_B)-C_p P_{\mathrm{F1}} V_h)}{(C_p-1)(P_A+P_B)^2+(C_p-2) P_{\mathrm{F1}} V_h (P_A+P_B)-P_{\mathrm{F1}}^2 V_h^2}$$

$$P'''_{\mathrm{F1}} = \frac{2 P''_{\mathrm{A}} P''_{\mathrm{B}}(1-C_{\mathrm{p}})^2}{(P''_{\mathrm{A}}+P''_{\mathrm{B}})(1-C_{\mathrm{p}})+P''_{\mathrm{F1}}}$$

$$= \frac{2\frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}\frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}(1-C_{\mathrm{p}})^2}{\left(\frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}+\frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}\right)(1-C_{\mathrm{p}})+\frac{P_{\mathrm{F1}} V_{\mathrm{h}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}} V_{\mathrm{h}}}}$$

$$= -\frac{2(C_p-1)^2 P_A P_B}{((C_p-1)(P_A+P_B)-P_{\mathrm{F1}} V_h)(P_A+P_B+P_{\mathrm{F1}} V_h)}$$

$$P'''_{\mathrm{BCA}} = \frac{2\,P''_{\mathrm{A}}\,P''_{\mathrm{F1}}(1-C_{\mathrm{p}})}{(P''_{\mathrm{A}}+P''_{\mathrm{B}})\,(1-C_{\mathrm{p}})+P''_{\mathrm{F1}}}$$

$$= \frac{2\,\frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}\,\frac{P_{\mathrm{F1}}\,V_{\mathrm{h}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}\,(1-C_{\mathrm{p}})}{\left(\frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}+\frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}\right)(1-C_{\mathrm{p}})+\frac{P_{\mathrm{F1}}\,V_{\mathrm{h}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}}$$

$$= \frac{2\,(C_p-1)\,P_A\,P_{\mathrm{F1}}\,V_h}{((C_p-1)\,(P_A+P_B)-P_{\mathrm{F1}}\,V_h)\,(P_A+P_B+P_{\mathrm{F1}}\,V_h)}$$

$$P'''_{\mathrm{BCB}} = \frac{2\,P''_{\mathrm{B}}\,P''_{\mathrm{F1}}(1-C_{\mathrm{p}})}{(P''_{\mathrm{A}}+P''_{\mathrm{B}})\,(1-C_{\mathrm{p}})+P''_{\mathrm{F1}}}$$ (F3, cont)

$$= \frac{2\,\frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}\,\frac{P_{\mathrm{F1}}\,V_{\mathrm{h}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}\,(1-C_{\mathrm{p}})}{\left(\frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}+\frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}\right)(1-C_{\mathrm{p}})+\frac{P_{\mathrm{F1}}\,V_{\mathrm{h}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}}$$

$$= \frac{2\,(C_p-1)\,P_B\,P_{\mathrm{F1}}\,V_h}{((C_p-1)\,(P_A+P_B)-P_{\mathrm{F1}}\,V_h)\,(P_A+P_B+P_{\mathrm{F1}}\,V_h)}$$

$$P'''_{\mathrm{F2}} = \frac{{P''_{\mathrm{F1}}}^{2}}{(P''_{\mathrm{A}}+P''_{\mathrm{B}})\,(1-C_{\mathrm{p}})+P''_{\mathrm{F1}}}$$

$$= \frac{\left(\frac{P_{\mathrm{F1}}\,V_{\mathrm{h}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}\right)^2}{\left(\frac{P_{\mathrm{A}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}+\frac{P_{\mathrm{B}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}\right)(1-C_{\mathrm{p}})+\frac{P_{\mathrm{F1}}\,V_{\mathrm{h}}}{P_{\mathrm{A}}+P_{\mathrm{B}}+P_{\mathrm{F1}}\,V_{\mathrm{h}}}}$$

$$= -\frac{P_{\mathrm{F1}}^2\,V_h^2}{((C_p-1)\,(P_A+P_B)-P_{\mathrm{F1}}\,V_h)\,(P_A+P_B+P_{\mathrm{F1}}\,V_h)}$$

The changes in the hybrid frequencies are

$$\Delta P_{\mathrm{A}} = P'''_{\mathrm{A}} - P_{\mathrm{A}} = \frac{P_A\left((C_p-1)\left(P_A+C_p P_B\right)-C_p P_{\mathrm{F1}} V_h\right)}{(C_p-1)\left(P_A+P_B\right)^2+(C_p-2) P_{\mathrm{F1}} V_h \left(P_A+P_B\right)-P_{\mathrm{F1}}^2 V_h^2} - P_{\mathrm{A}}$$

$$\Delta P_{\mathrm{B}} = P'''_{\mathrm{B}} - P_{\mathrm{B}} = \frac{P_B\left((C_p-1)\left(C_p P_A+P_B\right)-C_p P_{\mathrm{F1}} V_h\right)}{(C_p-1)\left(P_A+P_B\right)^2+(C_p-2) P_{\mathrm{F1}} V_h \left(P_A+P_B\right)-P_{\mathrm{F1}}^2 V_h^2} - P_{\mathrm{B}}$$

$$\Delta P_{\mathrm{F1}} = P'''_{\mathrm{F1}} - P_{\mathrm{F1}} = -\frac{2\left(C_p-1\right)^2 P_A P_B}{\left((C_p-1)\left(P_A+P_B\right)-P_{\mathrm{F1}} V_h\right)\left(P_A+P_B+P_{\mathrm{F1}} V_h\right)} - P_{\mathrm{F1}}$$

$$\Delta P_{\mathrm{BCA}} = P'''_{\mathrm{BCA}} - P_{\mathrm{BCA}} = \frac{2\left(C_p-1\right) P_A P_{\mathrm{F1}} V_h}{\left((C_p-1)\left(P_A+P_B\right)-P_{\mathrm{F1}} V_h\right)\left(P_A+P_B+P_{\mathrm{F1}} V_h\right)} - P_{\mathrm{BCA}} \quad \text{(F4)}$$

$$\Delta P_{\mathrm{BCB}} = P'''_{\mathrm{BCB}} - P_{\mathrm{BCB}} = \frac{2\left(C_p-1\right) P_B P_{\mathrm{F1}} V_h}{\left((C_p-1)\left(P_A+P_B\right)-P_{\mathrm{F1}} V_h\right)\left(P_A+P_B+P_{\mathrm{F1}} V_h\right)} - P_{\mathrm{BCB}}$$

$$\Delta P_{\mathrm{F2}} = P'''_{\mathrm{F2}} - P_{\mathrm{F2}} = -\frac{P_{\mathrm{F1}}^2 V_h^2}{\left((C_p-1)\left(P_A+P_B\right)-P_{\mathrm{F1}} V_h\right)\left(P_A+P_B+P_{\mathrm{F1}} V_h\right)} - P_{\mathrm{F2}}$$

To obtain the equilibrium frequencies in zygotes, $\tilde{P}_h$, one can solve for P_{A}, P_{B}, and P_{F1} after setting (F4) equal to zero. After some simplification, the equilibrium frequencies can be written as

$$\tilde{P}_{\mathrm{A}} = \frac{-(C_p\,(C_p^3+(2\,V_h+X+2)\,C_p^2+(-2\,V_h\,(X+8)+X-2)\,C_p+2\,V_h\,(X+3)-3\,X-2)+X+1)}{(8\,(C_p-V_h-1)\,((V_h-1)\,C_p^2-2\,(V_h+1)\,C_p+V_h+1))}$$

$$\tilde{P}_{\mathrm{B}} = \frac{-(C_p\,(C_p^3+(2\,V_h+X+2)\,C_p^2+(-2\,V_h\,(X+8)+X-2)\,C_p+2\,V_h\,(X+3)-3\,X-2)+X+1)}{(8\,(C_p-V_h-1)\,((V_h-1)\,C_p^2-2\,(V_h+1)\,C_p+V_h+1))}$$

$$\tilde{P}_{\mathrm{F1}} = \frac{(C_p-1)\,(C_p\,(4\,V_h+C_p\,(C_p-4\,V_h+X+5)-4\,X-3)+X+1)}{4\,(C_p-V_h-1)\,((V_h-1)\,C_p^2-2\,(V_h+1)\,C_p+V_h+1)}$$

$$\tilde{P}_{\mathrm{BCA}} = \frac{4\,(C_p-1)\,C_p\,V_h}{(3\,V_h-2)\,C_p^2+(V_h\,(X-8)-2\,(X+1))\,C_p-V_h\,(X-1)} \tag{F5}$$

$$\tilde{P}_{\mathrm{BCB}} = \frac{4\,(C_p-1)\,C_p\,V_h}{(3\,V_h-2)\,C_p^2+(V_h\,(X-8)-2\,(X+1))\,C_p-V_h\,(X-1)}$$

$$\tilde{P}_{\mathrm{F2}} = \frac{2\,(C_p-1)\,V_h^2}{(V_h-1)\,C_p^2+(V_h\,(2\,V_h+X-5)-X-2)\,C_p-2\,V_h\,(V_h+X+1)-X-1},$$

where

$$X = \sqrt{(C_p+1)^2+8\,C_p\,V_h}\,. \tag{F6}$$

Equations F5, which were checked for accuracy using simulations, indicate the equilibrium frequencies of individuals before viability selection. The frequencies after viability selection can be obtained by simply multiplying the (F5) by the viability and normalizing, yielding

$$\tilde{P}'_{\mathrm{A}} = \frac{\tilde{P}_{\mathrm{A}}}{\bar{V}}$$

$$\tilde{P}'_{\mathrm{B}} = \frac{P_{\mathrm{B}}}{\bar{V}}$$

$$\tilde{P}'_{\mathrm{F1}} = \frac{\tilde{P}_{\mathrm{F1}} V_{\mathrm{h}}}{\bar{V}}$$

$$\tilde{P}'_{\mathrm{BCA}} = \frac{\tilde{P}_{\mathrm{BCA}} V_{\mathrm{h}}}{\bar{V}} \qquad \text{(F7)}$$

$$\tilde{P}'_{\mathrm{BCB}} = \frac{\tilde{P}_{\mathrm{BCB}} V_{\mathrm{h}}}{\bar{V}}$$

$$\tilde{P}'_{\mathrm{F2}} = \frac{\tilde{P}_{\mathrm{F2}} V_{\mathrm{h}}}{\bar{V}}.$$

Consolidated References

Akaike, H. 1974. A new look at statistical model identification. IEEE Trans. Autom. Contr. 19:716-723.

Alfaro, M. E., S. Zoller, and F. Lutzoni. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Molecular Biology and Evolution 20:255-266.

Amenta, N., and J. Klinger. 2002. Case Study: Visualizing sets of evolutionary trees. IEEE Information Visualization. Pages 71-74.

Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217-1229.

Barbash, D. A., P. Awadalla, and A. M. Tarone. 2004. Functional divergence caused by ancient positive selection of a *Drosophila* hybrid incompatibility locus. Plos Biology 2:839-848.

Barton, N. H., and G. M. Hewitt. 1981. Hybrid zones and speciation. In: *Evolution and Speciation/Essays in Honor of M.J.D. White* (W.R. Atchley & D.S. Woodruff, eds), pp. 109-145. Cambridge University Press, Cambridge.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Annual Review of Ecology and Systematics 16:113-148.

Barton, N. H., and M. Turelli. 1991. Natural and sexual selection on many loci. Genetics 127:229-255.

Bateson, W. 1909. Heredity and variation in modern lights. In: *Darwin and Modern Science* (A. C. Seward, ed.), pp. 85-101. Cambridge University Press, Cambridge.

Blair, W. F. 1955. Mating call and stage of speciation in the *Microhyla olivacea-M. carolinensis* complex. Evolution 9:469-480.

Boecklen, W. J., and D. J. Howard. 1997. Genetic analysis of hybrid zones: numbers of markers and power of resolution. Ecology 78:2611-2616.

Bollback, J. P. 2002. Bayesian model adequacy and choice in phylogenetics. Molecular Biology and Evolution 19:1171-1180.

Brown, W. L. Jr., and E. O. Wilson. 1956. Character displacement. Systematic Zoology 5:49-64.

Butlin, R. K. 1987a. Speciation by reinforcement. Trends in Ecology and Evolution 2:8-13.

Butlin, R. K. 1987b. Species, speciation, and reinforcement. American Naturalist 130:461-464.

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

Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.

Cummings, M. P., S. A. Handley, D. S. Myers, D. L. Reed, A. Rokas, and K. Winka. 2003. Comparing bootstrap and posterior probability values in the four-taxon case. Systematic Biology 52:477-487.

Cunningham, C. W., H. Zhu, and D. M. Hillis. 1998. Best-fit maximum-likelihood models for phylogenetic inference: Empirical tests with known phylogenies. Evolution 52:978-987.

Dobzhansky, T. 1934. Studies on hybrid sterility. I. Spermatogenesis in pure and hybrid *Drosophila pseudoobscura*. Zeitschrift Fuer Zellforschung Und Microskopische Anatomie 21:169-221.

Dobzhansky, T. 1937. *Genetics and the Origin of Species*. Columbia University Press, New York.

Dobzhansky, T. 1940. Genetics and the origin of species. Columbia University Press, N.Y.

Dobzhansky, T. 1940. Speciation as a stage in evolutionary divergence. American Naturalist 74:312-321.

Douady, C. J., F. Delsuc, Y. Boucher, W. F. Doolittle, and E. J. P. Douzery. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. Molecular Biology and Evolution 20:248-254.

Erixon, P., B. Svennblad, T. Britton, and B. Oxelman. 2003. The reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Systematic Biology 52:665-673.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Systematic Zoology 27:401-410.

Felsenstein, J., and G. A. Churchill. 1996. A hidden Markov model approach to variation among sites in rate of evolution. Molecular Biology and Evolution 13:93–104.

Fitzpatrick, B. M., and H. B. Shaffer. 2004. Environment-dependent admixture dynamics in a tiger salamander hybrid zone. Evolution 58:1282-1293.

Galtier, N., and M. Gouy. 1998. Inferring pattern and process: Maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. Molecular Biology and Evolution 15:871–879.

Gaut, B. S., and P. O. Lewis. 1995. Success of maximum likelihood phylogeny inference in the four-taxon case. Molecular Biology and Evolution 12:152–162.

Gavrilets, S. 2003. Perspective: Models of speciation: What have we learned in 40 years? Evolution 57:2197-2215.

Golding, G. B. 1983. Estimates of DNA and protein sequence divergence: An examination of some assumptions. Molecular Biology and Evolution 1: 125-142.

Goldman, N, and Z. H. Yang. 1994. Codon-based model of nucleotide substitution for protein-coding DNA sequences. Molecular Biology and Evolution 11:725-736.

Goldman, N. 1993. Statistical tests of models of DNA substitution. Journal of Molecular Evolution 36:182-198.

Grant, B. R., and P. R. Grant. 1993. Evolution of Darwin's finches caused by a rare climatic event. Proceedings of the Royal Society of London B: Biology 251:111-117.

Grula, J. W., and O. R. Taylor. 1980. The effect of X-chromosome inheritance on mate-selection behavior in the sulfur butterflies, *Colias eurytheme* and *C. philodice*. Evolution 34:688-695.

Gu, X., Y-X. Fu, and W-H. Li. 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. Molecular Biology and Evolution 12: 546-557.

Hall, D. W., and M. Kirkpatrick. 2006. Reinforcement and sex linkage. Evolution 60:908-921.

Halpern, A. L., and W. J. Bruno. 1998. Evolutionary distances for protein-coding sequences: Modeling site-specific residue frequencies. Molecular Biology and Evolution 15: 910-917.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22:160–174.

Heisler, I. L. 1984. Inheritance of female mating propensities for yellow locus genotypes in *Drosophila melanogaster*. Genetical Research 44:133-149.

Helbig, A. J. 1991. Inheritance of migratory direction in a bird species:A cross-breeding experiment with SE and SW migrating blackcaps (*Sylvia atricapilla*). Behavioral Ecology and Sociobiology 28:9-12.

Höbel, G., and H. C. Gerhardt. 2003. Reproductive character displacement in the acoustic communication system of green tree frogs (*Hyla cinerea*). Evolution 57:894-904.

Howard, D. J. 1993. Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. In: *Hybrid Zones and the Evolutionary Process* (R. G. Harrison, ed.), pp. 46-69. Oxford University Press, New York.

Huelsenbeck, J. P. 2001. MrBayes 3.0b3: Bayesian inference of phylogeny. Distributed by the author. Department of Biology, Univ. Rochester, Rochester, New York.

Huelsenbeck, J. P., and D. M. Hillis. 1993. Success of phylogenetic methods in the four-taxon case. Systematic Biology 42:247–264.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Huelsenbeck, J. P., B. Larget, and D. Swofford. 2000. A compound Poisson process for relaxing the molecular clock. Genetics 154:1879-1892.

Huelsenbeck, J. P., B. Larget, R. E. Miller, and F. Ronquist. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Systematic Biology 51:673-688.

Huelsenbeck, J. P., F. Ronquist, R. Nielsen, and J. P. Bollback. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294:2310-2314.

Iyengar, V. K., H. K. Reeve, and T. Eisner. 2002. Paternal inheritance of a female moth's mating preference. Nature 419:830-832.

Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. In: *Mammalian Protein Metabolism* (H. N. Munro, ed.), pp. 21–132. Academic Press, New York.

Karlin, S., and H. M. Taylor. 1975. *A First Course in Stochastic Processes*. Academic Press, New York.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16:111-120.

Kirkpatrick, M. 1982. Sexual selection and the evolution of female choice. Evolution 36:1-12.

Kirkpatrick, M. 2000. Reinforcement and divergence under assortative mating. Proceedings of the Royal Society of London B: Biology 267:1649-1655.

Kirkpatrick, M. 2001. Reinforcement during ecological speciation. Proceedings of the Royal Society of London B: Biology 268:1259-1263.

Kirkpatrick, M., and M. R. Servedio. 1999. The reinforcement of mating preferences on an island. Genetics 151:865-884.

Kirkpatrick, M., and M. Servedio. 1999. The reinforcement of mating preferences on an island. Genetics 151:1764-1772.

Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection: models and experiments. American Naturalist 159:S22-S35.

Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. Genetics 161:1727-1750.

Kishino, H., J. L. Thorne, and W. J. Bruno. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Molecular Biology and Evolution 18:352-361.

Kuhner, M. K., and J. Felsenstein. 1994. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. Molecular Biology and Evolution 11:459–468.

Lanave, C., G. Preparata, C. Saccone, and G. Serio. 1984. A new method for calculating evolutionary substitution rates. Journal of Molecular Evolution 20:86–93.

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

Lemmon, A. R., and M. C. Milinkovitch. 2002. The metapopulation genetic algorithm: an efficient solution for the problem of large phylogeny estimation. Proceedings of the National Academy of Sciences U.S.A. 99:10516-10521.

Lemmon, A. R., C. Smadja and M. Kirkpatrick. 2004. Reproductive character displacement is not the only possible outcome of reinforcement. Journal of Evolutionary Biology 17:117-183.

Lindholm, A., and F. Breden. 2002. Sex chromosomes and sexual selection in poeciliid fishes. American Naturalist 160:S214-S224.

Liou, L. W., and T. D. Price. 1994. Speciation by reinforcement of premating isolation. Evolution 48:1451-1459.

Lockhart, P. J., A. W. D. Larkum, M. A. Steel, P. J. Waddell, and D. Penny. 1996. Evolution of chlorophyll and bacteriochlorophyll: The problem of invariant sites in sequence analysis. Proceedings of the National Academy of Sciences U.S.A. 93:1930–1934.

Lockhart, P. J., M. A. Steel, M. D. Hendy, and D. Penny. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Molecular Biology and Evolution 11:605-612.

Maddison, W. P., and D. R. Maddison. 2003. Mesquite: a modular system for evolutionary analysis. Version 0.994. http://mesquiteproject.org.

Mayr, E. 1963. *Animal Species and Evolution*. Belknap Press, Cambridge.

McGuire, G., M. C. Denham, and D. J. Balding. 2001. Models of sequence evolution for DNA sequences containing gaps. Molecular Biology and Evolution 18:481-490.

Minin, V., Z. Abdo, P. Joyce, and J. Sullivan. 2003. Performance-based selection of likelihood models for phylogeny estimation. Systematic Biology 52:674-683.

Moriarty, E. C., and D. C. Cannatella. 2004. Phylogenetic relationships of North American chorus frogs (*Pseudacris*). Molecular Phylogenetics and Evolution 30: 409-420.

Muller, H. J. 1939. Reversibility in evolution considered from the standpoint of genetics. Biological Reviews 14:261-280.

Muller, H. J. 1940. Bearing of the *Drosophila* work on systematics. In *The New Systematics (*J. S. Huxley, ed.), pp. 185-268. Clarendon Press, Oxford.

Muller, H. J. 1942. Isolating mechanisms, evolution and temperature. Biological Symposia 6:71-125.

Muse, S. V. 1995. Evolutionary analyses of DNA sequences subject to constraints on secondary structure. Genetics 139:1429-1439.

Muse, S. V., and B. S. Gaut. 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. Molecular Biology and Evolution 11:715-724.

Nielsen, R., and Z. Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929-936.

Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. Heredity 83:503-508.

Paterson, H. E. H. 1978. More evidence against speciation by reinforcement. South African Journal of Science 74:369-371.

Peterson, M. A., K. J. Monsen, H. Pedersen, T. McFarland, and J. Bearden. 2005. Direct and indirect analysis of the fitness of *Chrysochus* (Coleoptera:Chrysomelidae) hybrids. Biological Journal of the Linnean Society 84:273-286.

Posada, D., and K. A. Crandall. 2001. Selecting the best-fit model of nucleotide substitution. Systematic Biology 50:580-601.

Presgraves, D. C. 2003. A fine-scale genetic analysis of hybrid incompatibilities in *Drosophila*. Genetics 163:955-972.

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.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Rambaut, A., and N. C. Grassly. 1997. Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Computational Applied Biosciences 13:235-238.

Reinhold, K. 1998. Sex linkage among genes controlling sexually selected traits. Behavioral Ecology Sociobiology 44:1-7.

Ritchie, M. G. 2000. The inheritance of female preference functions in a mate recognition system. Proceedings of the Royal Society of London B: Biology 267:327-332.

Robinson, D. F., and L. R. Foulds. 1981. Comparison of phylogenetic trees. Mathematical Biosciences 53:131-147.

Rodríguez, F., J. L. Oliver, A. Marín, and J. R. Medina. 1990. The general stochastic model of nucleotide substitution. Journal of Theoretical Biology 142:485–501.

Rundle, H. D. 2002. A test of ecologically dependent postmating isolation between sympatric sticklebacks. Evolution 56:322-329.

Sætre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon et al. 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. Proceedings of the Royal Society of London B: Biology 270:53-59.

Sanderson, M. J. 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Molecular Biology and Evolution 14:1218-1231.

Schartl, M. 1995. Platyfish and swordtails - a genetic system for the analysis of molecular mechanisms in tumor-formation. Trends in Genetics 11:185-189.

Schwarz, G. 1974. Estimating the dimension of a model. Annuals of Statistics 6:461-464.

Servedio, M. R. 2004. The evolution of premating isolation: Local adaptation and natural and sexual selection against hybrids. Evolution 58:913-924.

Servedio, M. R., and G. P. Sætre. 2003. Speciation as a positive feedback loop between postzygotic and prezygotic barriers to gene flow. Proceedings of the Royal Society of London B: Biology 270:1473-1479.

Spencer, H. G., B. H. McArdle, and B. M. Lambert. 1986. A theoretical investigation of speciation by reinforcement. American Naturalist 128:241-262.

Steel, M., P. J. L. Székely, P. L. Erdös, and P. J. Waddell. 1993. A complete family of phylogenetic invariants for any number of taxa under Kimura's 3ST model. New Zealand Journal of Botany 31:289-296.

Sullivan, J., and D. L. Swofford. 2001. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated? Systematic Biology 50:723-729.

Sullivan, J., K. E. Holsinger, and C. Simon. 1995. Among-site rate variation and phylogenetic analysis of 12S rRNA in sigmodontine rodents. Molecular Biology and Evolution 12:988-1001.

Sun, S., C. T. Ting and C. I. Wu. 2004. The normal function of a speciation gene, Odysseus, and its hybrid sterility effect. Science 305:81-83.

Suzuki, Y., G. V. Glazko, and M. Nei. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proceedings of the National Academy of Sciences U.S.A. 99:16138-16143.

Swofford, D. L., P. J. Waddell, J. P. Huelsenbeck, P. G. Foster, P. O. Lewis, and J. S. Rogers. 2001. Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. Systematic Biology 50:525-539.

Tavaré, S. 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. Lectures in Mathematical Life Science 17:57–86.

Thorne, J. L., H. Kishino, I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. Molecular Biology and Evolution 15:1647–1657.

Tillier, E. R. M., and R. A. Collins. 1998. High apparent rate of simultaneous compensatory base-pair substitutions in ribosomal RNA. Genetics. 148:1993-2002.

Ting, C. T., S. C. Tsaur and C. I. Wu. 2000. The phylogeny of closely related species as revealed by the genealogy of a speciation gene, Odysseus. Proceedings of the National Academy of Sciences U.S.A. 97:5313-5316.

Ting, C. T., S. C. Tsaur, M. L. Wu and C. I. Wu. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. Science 282:1501-1504.

Ting, C.-T., S.-C. Tsaur, S. Sun, W. E. Browne, Y.-C. Chen et al. 2004. Gene duplication and speciation in *Drosophila*:evidence from the Odysseus locus. Proceedings of the National Academy of Sciences U.S.A. 101:12232-12235.

Toro, J., D. J. Innes, R. J. Thompson. 2004. Genetic variation among life-history stages of mussels in a *Mytilus edulis-M. trossulus* hybrid zone. Marine Biology 145:713-725.

Turelli, M., and H. A. Orr. 2000. Dominance, epistasis and the genetics of postzygotic isolation. Genetics 154:1663-1679.

Turelli, M., N. H. Barton, and J. A. Coyne. 2001. Theory and speciation. Trends in Ecology and Evolution 16:330-343.

Vähä, J-P., and C. R. Primmer. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Molecular Ecology 15:63-72.

Waddell, P., and D. Penny. 1996. Evolutionary trees of apes and humans from DNA sequences. In*: Handbook of Symbolic Evolution* (A. J. Lock & C. R. Peters, eds.), pp. 53-73. Clarendon Press, Oxford.

Wang, H., E. D. McArthur, S. C. Sanderson, J. H. Graham, and D. C. Freeman. 1997. Narrow hybrid zone between two subspecies of big sagebrush (*Artemisia tridentata*:Asteraceae). IV. Reciprocal transplant experiments. Evolution 51:95-102.

Wilcox, T. P., D. J. Zwickl, T. A. Heath, and D. M. Hillis. 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. Molecular Phylogenetics and Evolution 25:361-371.

Wu, C. I., and C. T. Ting. 2004. Genes and speciation. Nature Reviews Genetics 5:114-122.

Yang, Z. 1993. Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Molecular Biology and Evolution 10:1396-1401.

Yang, Z., N. Goldman, and A. Friday. 1994. Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. Molecular Biology and Evolution 11:316-324.

Ypma, T. J. 1995. Historical development of the Newton-Raphson method. SIAM Review 37:531-551.

Zar, J. H. 1999. *Biostatistical Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey.

Vita

Alan Richard Lemmon, son of Andrew and Mary Lemmon, was born in Provo Utah on April 21, 1976. After growing up in Broomfield, Colorado, he attended The University of Colorado at Boulder where he earned his Bachelor of Arts Degree with a double major in Environmental, Population & Organismal Biology and Molecular, Cellular & Developmental Biology. After graduating in December 2000, Alan spent six months in the lab of Michel Milinkovitch at the Free University of Brussels in Belgium, where he worked as a Research Associate. In August of 2001, he entered graduate school at the University of Texas at Austin in the Ecology, Evolution & Behavior graduate program. On June 21, 2003, Alan married the love of his life, Emily Claire Moriarty Lemmon.

Permanent address: 1 University Station #C0930, Austin, TX 78712.

This dissertation was typed by the author.