doi:10.1111/j.1558-5646.2009.00786.x

CHROMOSOMAL INVERSIONS AND SPECIES DIFFERENCES: WHEN ARE GENES AFFECTING ADAPTIVE DIVERGENCE AND REPRODUCTIVE ISOLATION EXPECTED TO RESIDE WITHIN INVERSIONS?

Jeffrey L. Feder[1,2,3] and Patrik Nosil[2,4,5]

[1]*Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556*

[2]*Institute for Advanced Study, Wissenschaftskolleg, Berlin, 14193, Germany*

[3]*E-mail: feder.2@nd.edu*

[4]*Department of Ecology and Evolutionary Biology, University of Colorado at Boulder, Colorado 80309*

[5]*E-mail: pnosil@zoology.ubc.ca*

Received March 17, 2009
Accepted June 23, 2009

Many factors can promote speciation, and one which has received much attention is chromosomal inversions. A number of models propose that the recombination suppressing effects of inversions facilitate the maintenance of differences between interbreeding populations in genes affecting adaptive divergence and reproductive isolation. These models predict that such genes will disproportionately reside within inversions, rather than in collinear regions. This hypothesis has received some support, but exceptions exist. Additionally, the effects of known low levels of recombination within inversions on these models are uninvestigated. Here, simulations are used to compare the maintenance of genetic differences between populations following secondary contact and hybridization in different inversion models. We compare regions with no recombination within them to regions with low recombination and to collinear regions with free recombination. Our most general finding is that the low levels of recombination within an inversion often result in the loss of accentuated divergence in inverted regions compared to collinear ones. We conclude that inversions can facilitate the maintenance of species differences under some conditions, but that large or qualitative differences between inverted and collinear regions need not occur. We also find that strong selection facilitates maintenance of divergence in a manner analogous to inversions.

KEY WORDS: **Epistasis, fitness trade-offs, gene flow, genomic islands, local adaptation, speciation.**

The factors both promoting and constraining the speciation process are a central topic in evolutionary biology (Coyne and Orr 2004; Gavrilets 2004; Nosil et al. 2009a). In this context, the role of chromosomal inversions in promoting speciation has received much attention (Coyne and Orr 2004; Hoffman and Rieseberg 2008 for review). For example, it has long been realized that structural chromosomal changes associated with inversions can result in reduced hybrid fitness (White 1978; King 1993; Coyne and Orr 2004, pp. 256–267 for review). The last decade has also seen the development of a number of "genic" models that propose a role for inversions in speciation which focuses on the genes within inversions, rather than structural changes, as being the causal agents for reproductive isolation (Noor et al. 2001, 2007; Rieseberg 2001; Ortiz-Barrientos et al. 2002; Butlin

2005; Machado et al. 2007; Hoffman and Rieseberg 2008; Santos 2009).

In these verbal genic models, inversions are initially established by either selection or drift, usually in allopatry. Following secondary contact and gene flow, inversions are predicted to facilitate the maintenance of genetic differences between populations in genes that affect adaptive divergence and reproductive isolation, thereby preventing the populations from being homogenized by gene flow. The basic premise is that inversions reduce introgression for large regions of the genome and protect favorable genotypic combinations within these regions from being broken up by recombination (Butlin 2005; Hoffman and Rieseberg 2008 for review). More specifically, favorable genotypic combinations at loci affecting adaptation to different environments are maintained within inversions, resulting in each gene involved exhibiting a greater composite selective differential between environments than it would individually (Rieseberg 2001). Additionally, inversions might tie up genes that confer divergent adaptation to those which affect assortative mating, thereby doubly promoting the maintenance of genetic differences between hybridizing populations (Butlin 2005). Finally, genotypic combinations within inversions that contribute to intrinsic genetic incompatibilities in hybrids, a form of reproductive isolation, can be protected from invasion by ancestral, compatible genotypic combinations (Noor et al. 2001). Extensions to these initial models on the maintenance of differences have now examined the role of inversions in the initial build up of genetic differences between populations (Navarro and Barton 2003; Gavrilets 2004) and on the factors driving the actual spread of inversions (Kirkpatrick and Barton 2006).

The genic inversion models make some explicit predictions. One is that inversions will be more common in sympatric versus allopatric taxon pairs. The underlying logic is that allopatric taxa without inversions can diverge because they need not counter the homogenizing effects of gene flow, and thus allopatric taxon pairs lacking inversions will occur. In contrast, secondary contact in sympatry between populations lacking inversions results in the fusion and homogenization of the two populations. Thus, distinct populations that are observed in sympatry will tend to harbor inversions. This prediction has been supported in *Drosophila* (Noor et al. 2001). Another prediction is that genes affecting adaptive divergence and reproductive isolation will reside within inversions, but that this tendency will be accentuated in sympatric taxa (relative to allopatric ones). A number of studies of hybridizing taxa, for example using quantitative trait locus (QTL) mapping, have now shown such genes to reside within inversions (Rieseberg et al. 1999; Noor et al. 2001; Feder et al. 2003a,b; Manoukis et al. 2008; see Hoffman and Rieseberg 2008 for full review). In an explicit test of this second prediction, Brown et al. (2004) compared the genetic basis of hybrid male sterility in a sympatric species pair of *Drosophila* (*D. pseudoobscura pseudoobscura* and *D. persimilis*) to that of an allopatric pair (*D. pseudoobscura bogotana* and *D. persimilis*). The two species pairs considered differ from one another in the same three inversions. Brown et al. (2004) report that virtually all of the sterility factors in the sympatric pair are associated with the three inverted regions, whereas sterility factors are present in the collinear regions in the allopatric pair.

However, recent population genomic studies (i.e., genome scans) sometimes contradict the importance of inversions. In this literature, genomic regions harboring genes affecting adaptive divergence or reproductive isolation are inferred indirectly via the accentuated genetic differentiation between populations displayed by such regions (i.e., greater differentiation than expected under neutrality, Beaumont 2005; Nielsen 2005; Storz 2005; Noor and Feder 2006; Stinchcombe and Hoekstra 2008). Studies using such approaches often report that regions of exceptionally high differentiation (outlier loci) are widely distributed across the genome, for example residing on numerous different chromosomes (Scotti-Saintagne et al. 2004; Achere et al. 2005; Grahame et al. 2006; Rogers and Bernatchez 2007; Egan et al. 2008; Nosil et al. 2008; Wood et al. 2008; Nosil et al. 2009a for review). These studies suggest that genes affecting adaptation and reproductive isolation are not always clustered within an inversion, although it is possible that each of these regions lies within a different inversion (with such inversions themselves being genomically dispersed). In other cases, inversions themselves were studied and no evidence for selection on them was revealed, suggesting the inversions do not harbor genes affecting adaptive divergence (Cohuet et al. 2004). In short, inversion models have some, but not overwhelming, empirical support. As noted by Hoffman and Rieseberg (2008) in their recent review of the evolutionary significance of inversions "the relative importance of inversions in the adaptive evolution of traits has rarely been addressed. It is therefore usually not clear if inversions play a critical role in adaptive shifts or if they only have a minor effect." (pp. 31).

Although the verbal models are intuitive, the efficacy of different genic inversions models has not been compared within a standardized theoretical framework. Such undertaking could be important, because the verbal models differ in whether fitness trade-offs, selection, and epistasis are involved (Noor et al. 2001; Rieseberg 2001; Gavrilets 2004), and thus a comparison of different scenarios is warranted. Additionally, the verbal models assume no recombination within inversions. However, some recombination does occur within inversions (Navarro et al. 1997; Jaarola et al. 1998; Hoffman and Rieseberg 2008 for review) and this has known theoretical significance because gene flow near a locus that contributes to reproductive isolation should be inversely proportional to the selection: recombination ratio (Barton 1979). Thus, the effectiveness of an inversion in limiting gene flow will depend

on how much recombination occurs within it, with even moderate levels of recombination potentially limiting the maintenance of genetic divergence (Felsenstein 1981; Ortiz-Barrientos et al. 2002). Indeed, there are empirical data indicating that recombination might be important for determining the efficacy of inversion models: genes exhibited accentuated divergence between species have been observed to reside near chromosomal breakpoints (*Helianthus* sunflowers, Rieseberg et al. 1999; Yatabe et al. 2007) or near centromeres (e.g., *Anopheles* mosquitoes, Turner et al. 2005; *Oryctolagus cuniculus* rabbits, Geraldes et al. 2006), where recombination is extensively reduced. Other studies showed that levels of genetic differentiation between the *Drosophila* species, *D. pseudoobscura,* and *D. persimilis* are somewhat elevated just outside of inversions, but drop off markedly even just a few megabases outside the inversion (Machado et al. 2007; Noor et al. 2007). Thus, any recombination might reduce the efficacy of the inversion models proposed by Noor and Rieseberg, but this has not been explicitly investigated.

Here, we use computer simulations to address theoretical issues pertaining to which particular inversion models are most effective at facilitating the maintenance of genetic differences in the face of gene flow. We vary selection strength, migration rates, recombination rates, and the extent of negative epistatic interactions between loci, thereby exploring the general conditions under which each model results in differences in genetic divergence between collinear versus inverted regions. In turn, these theoretical results might inform empirical discrepancies—to what extent might the differences among empirical results be explained by which model is acting or whether there is some recombination? Finally, by varying both the strength of selection and genetic architecture, we address issues concerning the role of each of these two factors in the maintenance of genetic divergence (Nosil et al. 2009b).

To formalize the verbal genic inversion models that were first proposed (e.g., Noor et al. 2001; Rieseberg 2001), we focus on the maintenance of existing differences between populations upon secondary contact. This is a useful and clear starting point, and future work could examine the initial build up of differences within inversions and the spread of inversions (cf. Navarro and Barton 2003; Kirkpatrick and Barton 2006, respectively). The genes we model in the simulations affect adaptive divergence and reproductive isolation, and thus are analogous to "speciation genes" (Coyne and Orr 2004; Wu and Ting 2004). Our study differs most explicitly from previous and related work on inversions (e.g., Noor et al. 2001; Rieseberg 2001) by considering the effects of variation in recombination rate and our most general finding is that the low levels of recombination within an inversion often result in the loss of accentuated divergence in inverted regions compared to collinear ones. We report on the time course of these patterns and discuss the findings in light of the expectation that genes affecting adaptive divergence and reproductive isolation will reside within (i.e., map to) inversions in natural populations.

Material and Methods

THE FIVE DIFFERENT INVERSION MODELS

We investigated the evolutionary dynamics of five different computer simulation models that could potentially generate disparities in genetic differentiation between populations in regions of high (collinear) versus low (inversion) recombination in the genome. A summary of these models is depicted in Figure 1. Models 1 and 2 examine the verbal argument of Rieseberg (2001) concerning ecological adaptation to different habitats. Model 1 involved fitness trade-offs in which alternate alleles at two loci (A and B) conferred higher fitness in one habitat and were disfavored in the other habitat. In this case, blocks of favored linked loci held together in inversions may be less prone to swamping by gene flow, as they would behave collectively as a "supergene" with a greater composite selective differential between habitats than each individual gene considered alone. In the second model, alleles at loci A and B were alternately favored in one of the two habitats and neutral in the other. In this instance, the equivalent of a Hill–Robertson effect (Hill and Robertson 1966; Santiago and Caballero 1998) may impede the homogenization of rearranged chromosomal regions by restricting the formation of gametes possessing the favored alleles at both loci. These first two models do not involve epistasis, but the following three models do.

Model 3 represents the hypothesis of Noor et al. (2001) involving negative epistasis and intrinsic postzygotic incompatibilities in hybrids. In model 3, the derived and neutral substitutions at loci A and B were assumed to be alternately fixed by genetic drift in populations 1 and 2. Thus, this model does not involve selection or adaptation to different ecological environments. However, hybrids possessing the two derived alleles suffered reduced viability due to a Dobzhansky–Muller (D–M) incompatibility (Bateson 1909; Dobzhansky 1937; Muller 1940, 1942; Orr 1995; Coyne and Orr 2004 for review). When this is the case, free recombination between the negatively interacting derived mutations could allow selection to act independently on the genes and eliminate them from populations. In contrast, low recombination, by restricting the formation of chromosomes possessing neutral ancestral alleles at both loci A and B, could impede the effective removal of the derived substitutions from the populations.

Models 4 and 5 were variations on the hypothesis of Noor et al. (2001) in that they involve negative epistasis but also further involve selection. Model 4 was similar to model 3, but instead of being neutral, the two derived substitutions were considered to be universally favored (beneficial) across the two populations. Model 5 was also based on intrinsic postzygotic isolation in hybrids and, to our knowledge, has not been previously proposed in

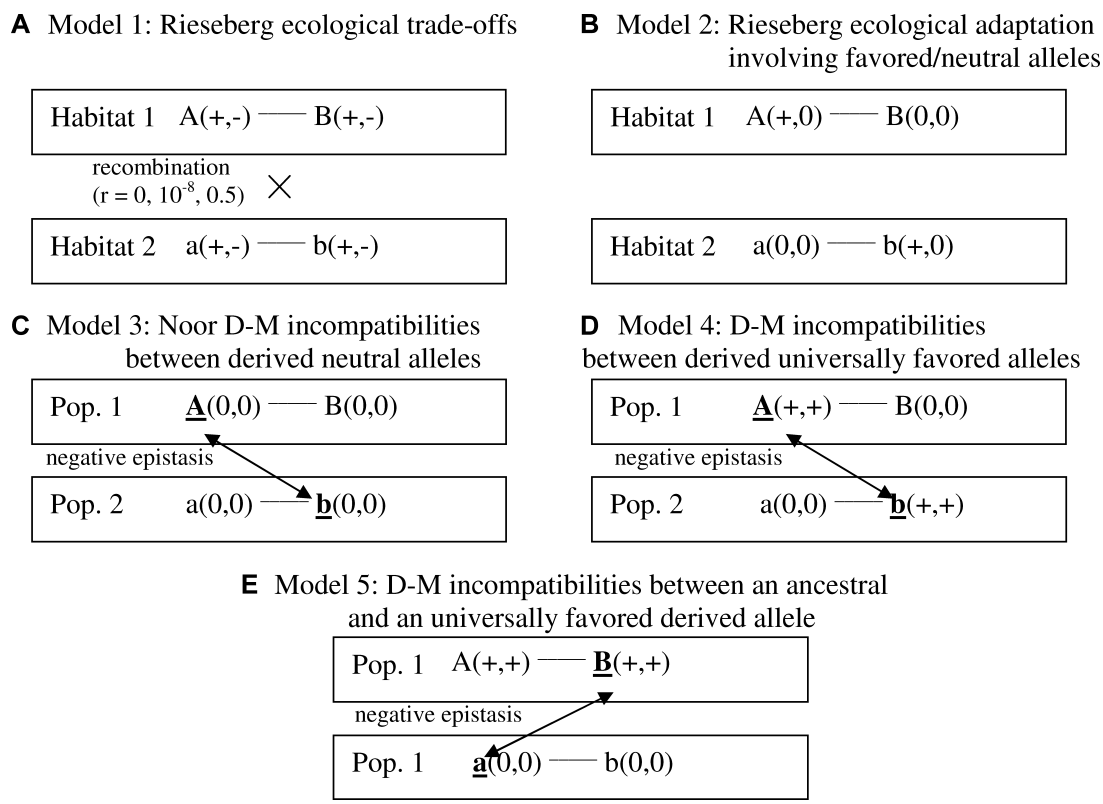

Figure 1. Diagrams of five models predicting that chromosomal rearrangements facilitate the maintenance of genetic differences between populations for loci involved in adaptive divergence or reproductive isolation. Models 1 and 2 are the ecological models of Rieseberg (2001) in which either (A) alleles responsible for fitness trade-offs or (B) alleles favored in one habitat and neutral in the other are disproportionately retained in inverted regions. Models 3 and 4 are the D–M incompatibility model of Noor et al. (2001) and a variation of it, respectively, in which negative epistasis between (C) derived, neutral alleles alternately fixed between populations or (D) derived alleles universally favored across taxa underlie intrinsic postzygotic isolation. (E) Model 5 concerns the hypothesis presented in the current study in which D–M incompatibilities exist between a derived favorable mutation at locus B in one population and the ancestral allele retained at an alternate locus A in the other taxon. In the diagrams, we show the allele states for the two loci A and B and, in parentheses, whether they are favored (+), neutral (0), or selected against (–) in a given habitat/population under each of the models. The first symbol provides the relative fitness for an allele in its parental habitat/population and the second symbol its relative fitness in the alternate habitat/population. Upper case allele designations A and B were used for habitat/population 1 and lower case a and b for habitat/population 2. The arrows in models 3–5 depict negative epistatic interactions between the designated alleles in individuals of mixed hybrid ancestry. Bold underlined alleles in models 3–5 indicate the derived substitutions in the population. The computer simulations varied recombination rates between loci A and B in the simulations ($r = 0$, 10^{-8}, and 0.5) under differing parameter values for migration, selection, and negative epistasis. This was done to investigate the potential for differential retention of these genes in inverted versus collinear regions.

the literature. Instead of derived–derived incompatibilities, model 5 examined the consequences of negative interactions between derived and ancestral allelic states. In this case, derived and universally favored substitutions were envisioned to first fix at locus A, and then at locus B, in population 1. Hybrids of mixed ancestry would be inviable due to D–M incompatibility between the ancestral allele at locus A still present in population 2 and the new derived allele at locus B in population 1. Introgression would be hampered for a rearranged region because the universally favored derived mutation at locus A would have difficulty disentangling itself from the D–M incompatibility causing derived mutation at locus B. However, in collinear regions, the derived A substitution could freely move and increase in frequency in population 2, setting the stage for the subsequent introgression of the second derived allele at locus B.

SIMULATION PROCEDURES

We used the same discrete generation, two-island population computer simulations of introgression following secondary contact to investigate the dynamics of the five inversion models. We considered the two populations to initially be geographically isolated during which time genetic differences and an inversion became alternately fixed in the two demes. After secondary contact, migration occurred at a rate m per generation between the two island

demes. Population densities were assumed to be equal and independently regulated in the two demes. Selection, when present in the model, was assumed to be soft, with both infinitely sized demes contributing equally to the migrant pool. We considered a life cycle with selection following migration and preceding mating (newborns > dispersal > selection > meiosis/recombination > mating > zygotes). We followed the fate of allelic variation segregating at two loci, designated A and B. We also explored the consequences of a more polygenic four-locus system for model 2.

Three different linkage relationships were investigated between loci. In the first case, loci were separated by 50 cM, corresponding to free recombination and a collinear arrangement in the genome. For the other two cases, either no recombination was allowed or only a very low level of recombination (1×10^{-8}) was permitted between loci. These latter two simulations represented loci residing in rearrangements.

When recombination was free, we modeled exchange in the context of a standard two-locus system with a random assortment occurring between segregating alleles at the two genes A and B during meiosis. In contrast, exchange in the low recombination case was modeled from the context of two alternative gene orders ("standard" and "inverted") segregating in populations. Exchange in homokaryotes (individuals with standard/standard or inverted/inverted chromosomes) occurred in the same manner as for the free recombination case except that the distance between loci was set at 10 instead of 50 cM. In contrast, the total genetic flux (exchange) rate of alleles between alternate rearrangements in heterokaryotype individuals was set at 1×10^{-8}. Genetic flux was apportioned at a ratio of 70:30 between double exchange and gene conversion events (all successful recombination events were assumed to involve double exchange in the simulations). The proportion of double recombination events involving a particular gene or combination of genes was calculated as the product of the recombination distances flanking the loci (as estimated in collinear homokaryotypes) divided by the sum total of recombination products for all possible gene combinations. Distances from the most proximate genes to the inversion breakpoints were both set at 10 cM, the same as that between loci. Thus, for the two-locus model, chromosomal segments containing locus A, loci AB, and locus B were each involved in one third of the double recombination events between rearrangements. For gene conversion, loci A and B were considered to be equally likely involved in a conversion event. We assumed gene conversion to be unbiased such that a conversion event in a heterokaryote resulted in a transformed (changed) allelic state in the standard and inverted arrangement two thirds of the time. We observed almost identical results when inversions were simply modeled as two loci with 10^{-8} recombination between them (not inverted versus collinear regions per se). Tables of the full results are available from the authors.

Three different intensities of selection were considered in the simulations ($s = 0.01$ [weak], 0.1 [moderate], and 1.0 [strong]). Three levels of migration were considered ($m = 0.001$ [low], 0.01 [moderate], and 0.1 [high]). Four degrees of negative epistasis were considered ($ep = 0.01$ [weak], 0.1 [moderate], 0.5 [strong], and 0.95 [very strong]). Selection was modeled to affect viability between juvenile and adult life stages with segregating alleles interacting in a partially dominant manner such that the relative fitness of alternate homozygotes and the heterozygote at a locus were 1, $1 + s$ and $1 + s/2$, respectively. Fitness interactions were multiplicative between loci. Hybrids that possessed alleles causing D–M incompatibilities were assigned a relative fitness value of $1 - ep$ regardless of the number of incompatible alleles they possessed or their genotype for other loci. Migration occurred independent of genotype.

We measured genetic differentiation between populations at different time intervals in regions of no, low, and free recombination. As a metric, we followed through time the allele frequency difference at the B locus for the first 20,000 generations after secondary contact. This time frame is useful for assessing patterns of introgression following relatively recent secondary contact events, such as those involving the end of the last Pleistocene glaciation 10–15,000 years ago. Additionally, differences between models were almost always observed within this time frame (see Results).

Results

The overarching pattern that emerged from the computer simulations was that unless recombination was completely eliminated and coupled with strong selection or epistasis and low migration rates, there was no large difference in genetic divergence maintained between rearranged versus collinear regions of the genome following secondary contact (Figs. 2–6 and Supporting information for graphical results and Table 1 for a summary of the general results from each model). In essence, low-level recombination in inverted regions on the order of 10^{-8}, likely a modest rate for inversions (Hoffman and Rieseberg 2008 for review), was often sufficient to result in similar levels of differentiation for genes involved in reproductive isolation/ecological adaptation across the genome. Moreover, observed quantitative differences between rearranged versus collinear regions were usually time dependent. Given a little recombination, even when parameter values were ideal for disproportionately retaining differentiation for reproductive isolation/ecological adaptation genes in rearranged regions for models 2, 3, and 5 (strong selection/epistasis and low migration), the difference dissipated or began to dissipate within the 20,000 generations modeled in the simulations (Figs. 3, 4, and 6). Thus, following recent secondary contact events there may be a time window when "speciation genes"

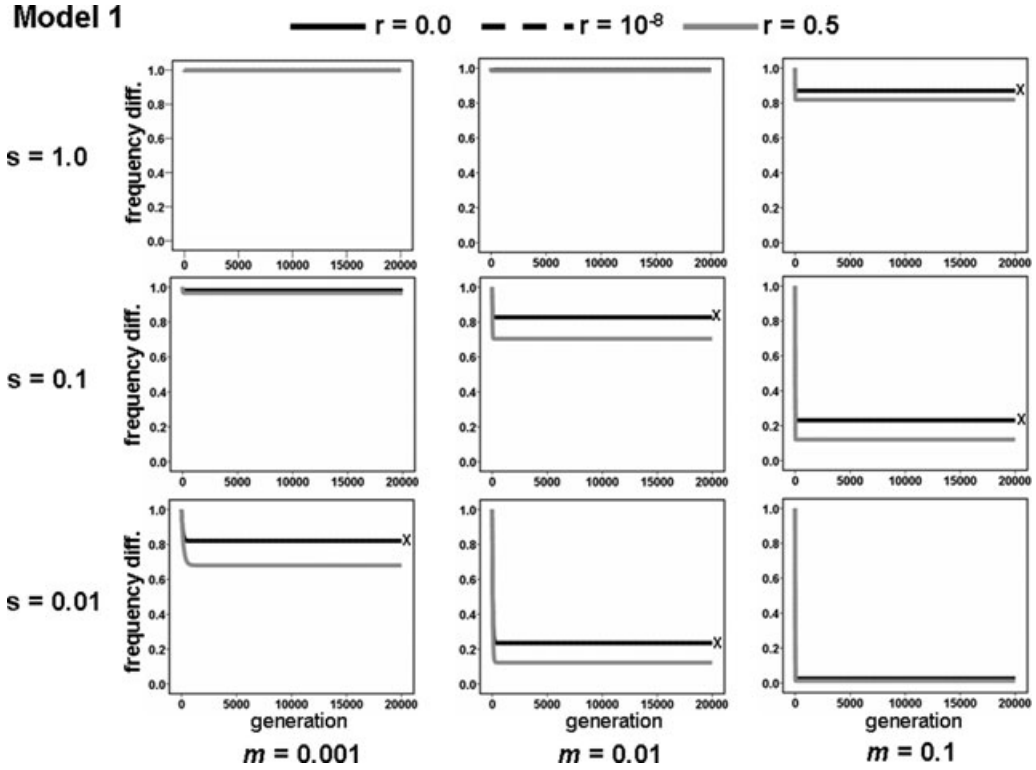

Figure 2. Simulation results from model 1, in which alternate alleles at two loci (A and B) conferred higher fitness in one habitat and were disfavored in the other habitat (i.e., "fitness trade-offs"). Shown is the difference in allele frequency between populations at locus B (frequency diff.) as a function of the number of generations since secondary contact between populations, for different selection (*s*) and migration (*m*) regimes. In each panel, allele frequencies differences are contrasted between scenarios in which the two loci reside in genomic regions with no recombination between loci ($r = 0.0$ = inversion with no recombination), low recombination between loci ($r = 10^{-8}$ = inversion with low recombination) and free recombination between loci ($r = 0.5$ = collinear genomic region). In this model, large frequency differences between populations were maintained at all genomic regions when selection was strong relative to migration, and frequency differences between populations were lost at all genomic regions when migration was high relative to selection. Thus, this model does not predict large differences among genomic regions in levels of genetic differentiation in genes affecting adaptive divergence. In some cases, the results for two of the three recombination rates were essentially identical, denoted by the small 'X' in the relevant panels.

appear to predominantly map to rearrangements, but this phenomenon may not (usually will not) persist indefinitely and is transitory.

The five models investigated in the study differed quantitatively, however, in their potential to maintain divergence between collinear and rearranged regions. For the ecologically based models proposed by Rieseberg (2001), inversion polymorphism had a much greater effect when alleles were favored in one habitat and neutral in the other (model 2; Fig. 3) compared to when they constituted true fitness trade-offs (model 1; Fig. 2). For model 1, fitness trade-offs between habitats generally resulted in the retention of detectable allele frequency differences between populations even for collinear regions, except for when selection was weak ($s = 0.01$) and migration high ($m = 0.1$; Fig. 2). For model 2 in contrast, free recombination allowed favored alleles in collinear regions to become disassociated from their disfavored genetic backgrounds in the alternate habitat. As a result, these favored/neutral alleles could flow more easily between populations when unlinked compared to when tightly linked in inversions (Fig. 3). But this difference was dependent on the absence of recombination. When recombination was low in model 2, differences between collinear regions and inverted regions disappeared relatively quickly. Indeed, unless selection was very strong ($s = 1.0$) and migration low ($m = 0.001$) the differences disappeared within the 20,000 generation time frame of the simulations. Increasing the number of loci from two to four delayed the period of decline and accentuated the difference between inverted and collinear regions, especially when the migration rate was low (Fig. 7). However, adding two additional loci did not prevent the ultimate dissipation of genetic differentiation for moderate migration rates unless selection was strong (Fig. 7). These results imply that if rearrangements contain large numbers of alternately interspersed favored/neutral alleles and migration rates are low following secondary contact, then genetic differentiation

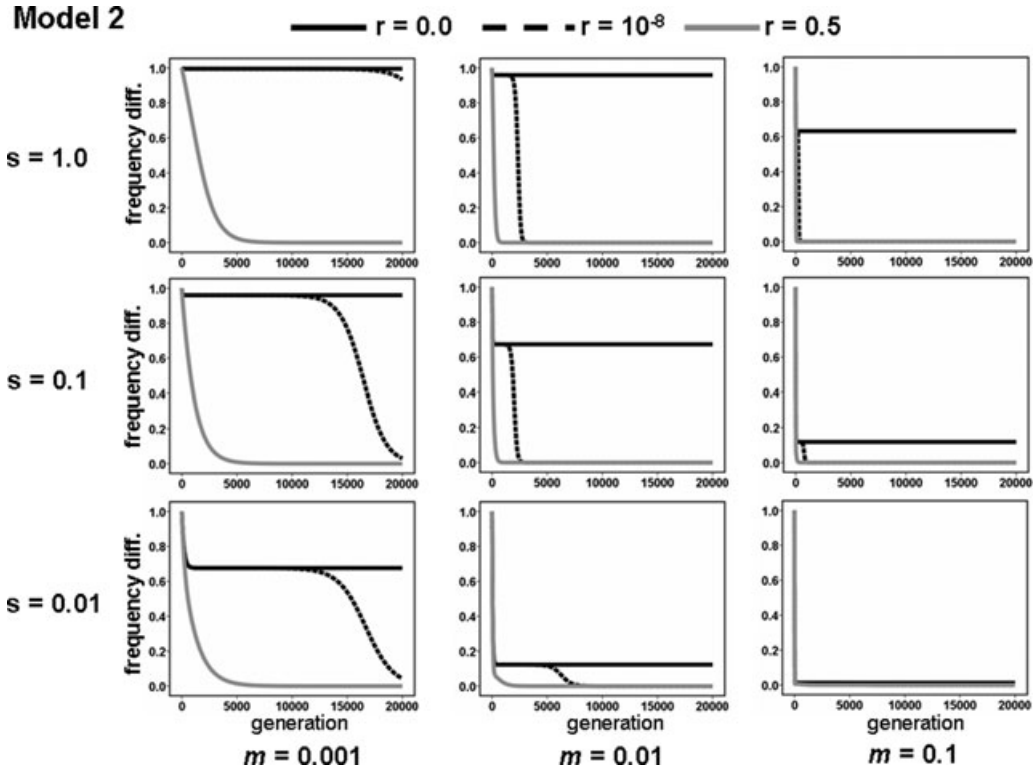

Figure 3. **Simulation results from model 2, in which alleles at loci A and B were alternately favored in one of the two habitats and neutral in the other. Shown is the difference in allele frequency between populations at locus B (frequency diff.) as a function of the number of generations since secondary contact between populations, for different selection (*s*) and migration (*m*) regimes. In each panel, allele frequencies differences are contrasted between scenarios in which the two loci reside in genomic regions with no recombination between loci ($r = 0.0$ = inversion with no recombination), low recombination between loci ($r = 10^{-8}$ = inversion with low recombination), and free recombination between loci ($r = 0.5$ = collinear genomic region). In this model, differences between regions with low versus free recombination were sometimes observed, but were transient. The differences between regions with low versus free recombination tended to decay sooner as selection was weakened or migration strengthened. See Figure 7 for results considering four, rather than two, loci.**

may be more strongly maintained within inverted versus collinear genomic regions for comparatively long periods of time.

The D–M incompatibility model 3 of Noor et al. (2001) generated a persistent difference between collinear and rearranged regions only when recombination was absent, migration low, and the alleles selectively neutral in their parental populations (Fig. 4). Relaxing these conditions resulted in dramatic reductions in the extent of differentiation seen for the inversions and, as was the case for ecological model 2, transformed the nature of the divergence from being long term to being ephemeral. As noted above, although recombination is reduced between alternative rearrangements in the genome, it is usually not eliminated altogether and rates of gene flux on the order of 10^{-8} that were evaluated in the simulations are likely modest and not uncommon (Hoffman and Rieseberg 2008 for review). Therefore, the assumption of no recombination is unlikely to be met in most instances.

Under the scenario of model 4 in which alleles where initially fixed by selection and are universally favored, essentially no difference was observed in the genetic differences maintained in derived-derived D–M incompatibilities between inverted and collinear regions. The only exceptions were for certain combinations of parameter values such as for moderate selection ($s = 0.1$), when negative epistasis was strong ($ep = 0.95$) and migration high ($m = 0.1$; see Fig. 5 for results with moderate selection and Supporting Information Figs. S1 and S2 for full results with strong and weak selection). However, even this exception occurred only in the absence of recombination. Thus, as was often the case for the other models, high levels of migration coupled with weak selection/epistasis led to the essential homogenization of allele frequency differences between populations in all regions of the genome (Figs. 2–6).

Model 5 proposed in the current article is based on derived-ancestral D–M incompatibilities. In this model, differences between rearranged and collinear regions of the genome were even less common and pronounced than in the other models. Indeed, model 5 was particularly susceptible to the effects of low-level recombination (10^{-8}) in negating any potential differences in the number of speciation genes retained between inverted versus collinear regions (see Fig. 6 for results with moderate selection,

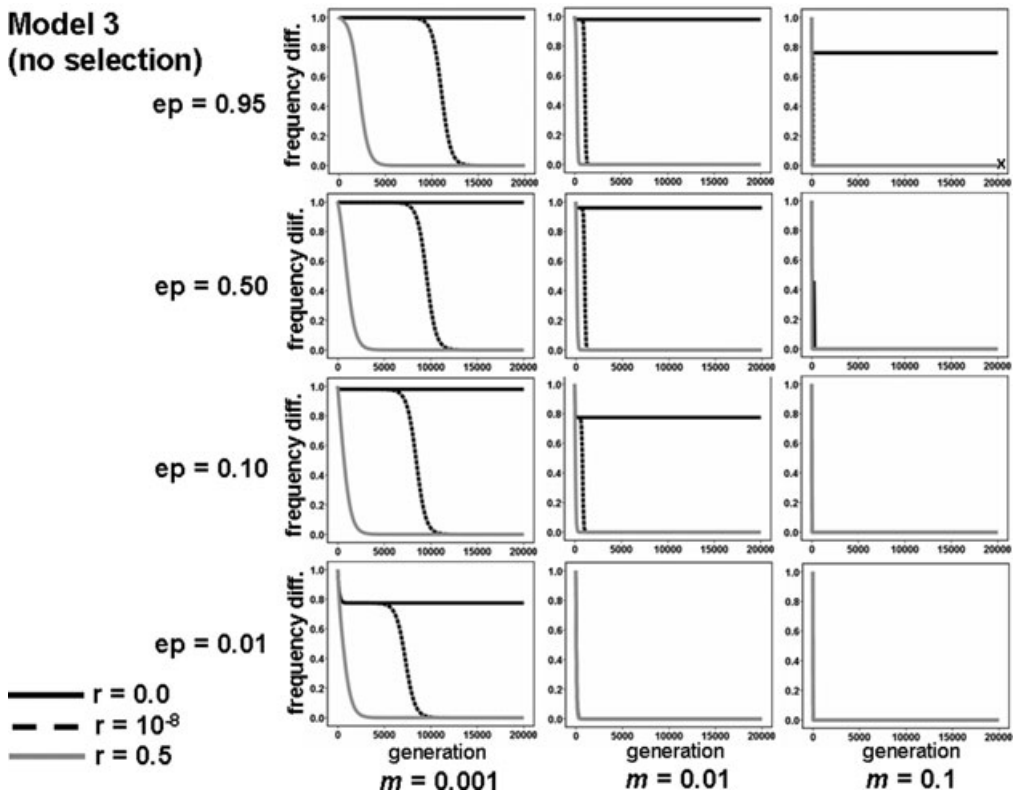

Figure 4. Simulation results from model 3, in which the derived and neutral substitutions at loci A and B were assumed to be alternately fixed by genetic drift in populations 1 and 2. Shown is the difference in allele frequency between populations at locus B (frequency diff.) as a function of the number of generations since secondary contact between populations, for different epistasis (*ep*) and migration (*m*) regimes. In each panel, allele frequencies differences are contrasted between scenarios in which the two loci reside in genomic regions with no recombination between loci ($r = 0.0$ = inversion with no recombination), low recombination between loci ($r = 10^{-8}$ = inversion with low recombination), and free recombination between loci ($r = 0.5$ = collinear genomic region). In this model, transient differences between regions with low versus free recombination were observed under some conditions, but these differences decayed relatively rapidly through time secondary contact. In some cases, the results for two of the three recombination rates were essentially identical, denoted by the small 'X' in the relevant panels.

see Supporting Information Figs. S3 and S4 for full results with strong and weak selection).

Discussion

GENERAL PATTERNS AND DIFFERENCES AMONG MODELS

We analyzed several simulation models that examined the effect of chromosomal inversions on the maintenance of genetic differences between hybridizing populations upon secondary contact. Our models formalize the verbal arguments of Rieseberg (2001) and Noor et al. (2001) and differ from several other mathematical inversion models by examining the maintenance of genetic differences, rather than their initial build up (e.g., Navarro and Barton 2003; Kirkpatrick and Barton 2006).

The results from our computer simulations imply that we might not always expect to observe stark, qualitative differences between the number of "speciation genes" mapping to rearranged versus collinear regions of the genome following secondary contact and introgression between populations. Certain models (2, 3, and to a limited extent 4 and 5) could produce windows of time in which a quantitative difference could potentially be detected between inverted versus collinear regions. But for the D–M incompatibility models 4 and 5, this difference depends on the complete absence of recombination for an inversion polymorphism, a perhaps unrealistic assumption. Moreover, model 3 will sometimes not apply, because it assumes that derived mutations contributing to D–M incompatibilities are neutral. Several genes affecting D–M incompatibilities have been shown to exhibit a history of evolving via positive selection, demonstrating that the assumption of neutrality will not always be met (e.g., *Hmr*, Barbash et al. 2003, 2004; *Nup96*, Presgraves et al. 2003; *Lhr*, Brideau et al. 2007; *Nup160*, Tang and Presgraves 2009; see Orr et al. 2004; Noor and Feder 2006 for reviews). Thus, model 2 may provide the most probable scenario for any significant time window of greater differentiation for rearranged than collinear regions. This is especially true if a large number of interspersed

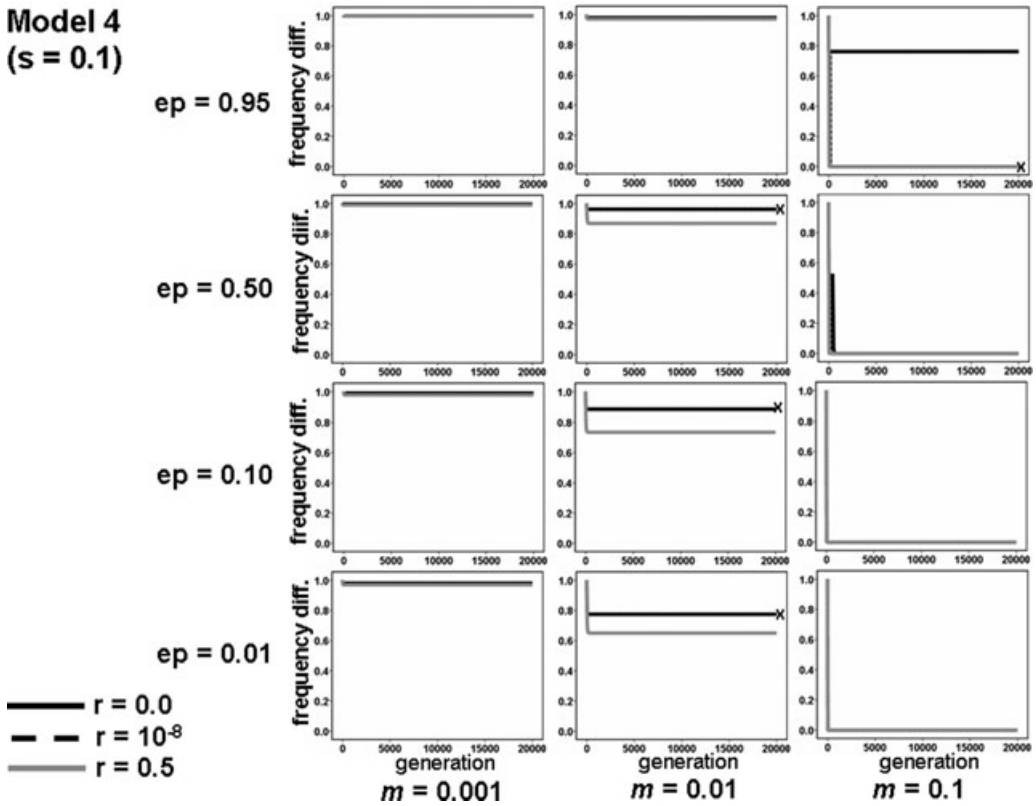

Figure 5. **Simulation results from model 4, a model similar to model 3, but in which instead of being neutral, the two derived substitutions were considered to be universally favored across the two populations (i.e., universally beneficial). Shown is the difference in allele frequency between populations at locus B (frequency diff.) as a function of the number of generations since secondary contact between populations, for different epistasis (*ep*) and migration (*m*) regimes (here moderate selection is considered, see Supporting Information for full results with strong and weak selection). In each panel, allele frequencies differences are contrasted between scenarios in which the two loci reside in genomic regions with no recombination between loci ($r = 0.0$ = inversion with no recombination), low recombination between loci ($r = 10^{-8}$ = inversion with low recombination), and free recombination between loci ($r = 0.5$ = collinear genomic region). In this model, differences were maintained in all genomic regions when migration was low or modest and tended to be lost in all genomic regions when migration was high. Thus, this model does not predict differences between inverted versus collinear regions in the level of genetic divergence maintained upon secondary contact. In some cases, the results for two of the three recombination rates were essentially identical, denoted by the small 'X' in the relevant panels.**

favored/neutral alleles reside in the rearrangements. Although the difference may not be permanent unless recombination is zero, a pronounced pattern could be observed if migration rates were low and/or secondary contact occurred in the not too distant past. However, as model 2 is based on alleles being favored in one habitat (population) and neutral in the other, the loci do not constitute genes affecting adaptive divergence or reproductive isolation in the strict sense of the terms. Consequently, none of the models appear to predict a dramatic and permanent difference in the number of loci causing reproductive isolation between rearranged versus collinear regions, unless certain specific circumstances are met.

We stress that our models do not address the possibility for inversions to serve as foci on which additional differences can be built. Instead we focused on the maintenance of divergence following secondary contact and introgression, a scenario many would argue is the common situation for many taxa in present-day sympatry (Coyne and Orr 2004). However, as shown by Navarro and Barton (2003), the possibility of additional divergence following contact (at least for universally favored alleles causing D–M incompatibilities) is again limited in time and will not necessarily produce an "all or none" qualitative effect. The take-home message from the simulation models is therefore that observed instances of dramatic differences between rearranged versus collinear regions may often reflect specific circumstances, as discussed below.

CONDITIONS UNDER WHICH ACCENTUATED GENETIC DIVERGENCE IS EXPECTED WITHIN INVERSIONS: EXTENSIVELY REDUCED RECOMBINATION, MULTIPLE LOCI AND RECENT SECONDARY CONTACT

Empirical examples of hybridizing taxa where genes underlying divergent adaptation or reproductive isolation reside within

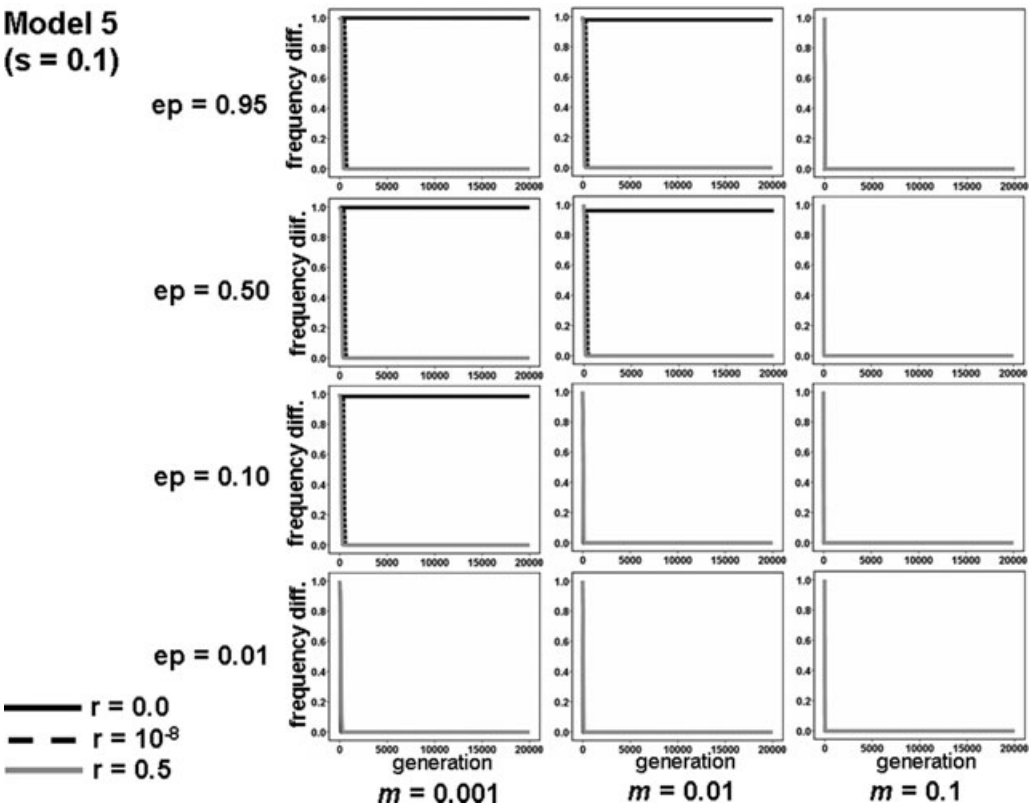

Figure 6. Simulation results from model 5, which examined the consequences of negative interactions between derived and ancestral allelic states. In this case, derived and universally favored substitutions were envisioned to first fix at locus A, and then at locus B, in population 1. Shown is the difference in allele frequency between populations at locus B (frequency diff.) as a function of the number of generations since secondary contact between populations, for different epistasis and migration regimes (here moderate selection is considered, see Supporting Information for full results with strong and weak selection). In each panel, allele frequencies differences are contrasted between scenarios in which the two loci reside in genomic regions with no recombination between loci ($r = 0.0$ = inversion with no recombination), low recombination between loci ($r = 10^{-8}$ = inversion with low recombination), and free recombination between loci ($r = 0.5$ = collinear genomic region). In this model, genetic differences after secondary contact were maintained only in regions of no recombination, and only when epistasis was strong relative to migration. Thus, differences in the level of genetic divergence maintained in regions with low versus free recombination are not expected.

inversions do exist (Rieseberg et al. 1999; Noor et al. 2001; Feder et al. 2003a,b; Manoukis et al. 2008; see Hoffman and Rieseberg 2008 for full review). At least four factors could explain these empirical observations, thereby reconciling our general simulation results with these empirical findings.

First, inversions may facilitate the maintenance of genetic differences when recombination within them is extensively reduced, for example when loci affecting divergence adaptation or reproductive isolation lie near a chromosomal breakpoint. Indeed, there are empirical data indicating that the degree of recombination might be important for determining the efficacy of inversion models. For example, a recent study combining QTL mapping and population genomics reported that accentuated divergence between hybridizing sunflower species (*Helianthus annuus* and *H. petiolaris*) occurred primarily at chromosomal breakpoints rather than throughout inversions (Rieseberg et al. 1999; Yatabe et al. 2007). Other studies showed that levels of genetic differentiation between *Drosophila* species drop off markedly even just a few megabases outside the inversion (Machado et al. 2007; Noor et al. 2007). Thus, the efficacy of inversion models should increase as genes underlying species differences reside in regions of increasingly reduced recombination. Notably, this argument extends beyond inversion models to any factors that reduce recombination. For example, genetic differences between hybridizing taxa might most easily persist near centromeres where, as in the case of chromosomal breakpoints, recombination is heavily reduced. Indeed, accentuated divergence near centromeres has been reported in genetic studies of mosquitoes (*Anopheles gambiae*, Turner et al. 2005) and the European rabbit (*O. cuniculus*, Geraldes et al. 2006).

Second, inversions may be more effective at facilitating the maintenance of genetic divergence when multiple (many)

Table 1. General summary of the simulation results from the five different inversion models. The three types of genomic regions considered are inverted regions with no recombination within them, inverted regions with low (10^{-8}) recombination with them, and collinear regions of free recombination (0.5). See text for details, Figure. 1 for a graphical description of each model and Figures. 2–6 for detailed results.

Model	General Results
Model 1: Rieseberg ecological trade-offs	Very similar degree of genetic differentiation maintained across all three types of genomic regions, dependent on the balance between the strength of divergent selection and rates of migration
Model 2: Rieseberg ecological adaptation involving favored/neutral alleles	Inverted regions with no recombination maintain the greatest degree of genetic differentiation Inverted regions with low recombination sometimes exhibit greater genetic differentiation than collinear regions for a time period (e.g., a few thousand generations following secondary contact), but equilibrium levels of genetic differentiation maintained for such regions are very similar to those observed for collinear regions (two locus model) When four, rather than two, loci are considered, inverted regions with low recombination maintain greater differentiation than collinear regions under some conditions (e.g., when migration is very low, m=0.001)
Model 3: Noor D–M incompatibilities between derived neutral alleles	Inverted regions with no recombination often maintain the greatest degree of genetic differentiation Inverted regions with low recombination exhibit greater genetic differentiation than collinear regions for a time period, but only when migration is very low (m=0.001) Equilibrium levels of genetic differentiation for inverted regions with low recombination are very similar to levels observed for collinear regions
Model 4: D–M incompatibilities between derived universally favored alleles	Equilibrium levels of genetic differentiation for inverted regions with low recombination very similar to levels observed for collinear regions or very similar degree of genetic differentiation is maintained across all three types of genomic regions
Model 5: D–M incompatibilities between an ancestral and an universally favored derived allele	As for model 4

loci affecting divergent adaptation and reproductive isolation are tightly clustered within them, as exemplified by our inversion model 2 being more effective when four instead of only two loci were modeled. Two types of data suggest that such clustering of genes occurs. Population genomic studies have revealed some instances in which multiple different outlier loci are clustered within one or a few genomic regions (e.g., *A. gambiae* mosquitoes, Turner et al. 2005; *Zeiraphera diniana* larch budmoth, Emelianov et al. 2004; Nosil et al. 2009a for review). Likewise, some QTL studies have demonstrated that multiple different adaptive traits map to a single or similar genomic regions (e.g., *Acyrthosiphon* pea aphids, Hawthorne and Via 2001; *Heliconius* mimetic butterflies, Joron et al. 2006; Kronforst et al. 2006; *Coregonus* whitefish ecotypes, Rogers and Bernatchez 2007; *Gasterosteus* sticklebacks, Albert et al. 2007), although pleiotropy could also contribute to these results. These examples do not necessarily concern inversions per se, but they nonetheless suggest that genomic clustering of genes affecting adaptation and reproductive isolation might occur, with such clustering potentially facilitating the maintenance of genetic differences between hybridizing populations.

Third, differentiation is expected to be accentuated within inversions when secondary contact and gene flow is recent (Figs. 2–6), as might be the case for many cases of recent postglacial secondary contact (e.g., *Rhagoletis pomonella* flies, Feder et al. 2003a,b).

The fourth and final point concerns the ability or power to actually detect genes affecting adaptive divergence and reproductive isolation in collinear versus rearranged genomic regions. Specifically, there may be a detection bias leading to a perceived, rather than real, difference in how often such genes reside within the different types of genomic regions. Genes in rearrangements could potentially be easier to detect through mapping studies and population genome scans because (1) they will have a larger effect size due to the collectively higher selection coefficient favoring them when tightly physically linked compared to when alone, and (2) they will generate greater hitchhiking effects for associated neutral variation. Consequently, given this consideration and our results suggesting differences between collinear and rearranged regions will not always be large, any observed difference in the number of speciation genes in collinear versus rearranged regions could be due to ascertainment bias.

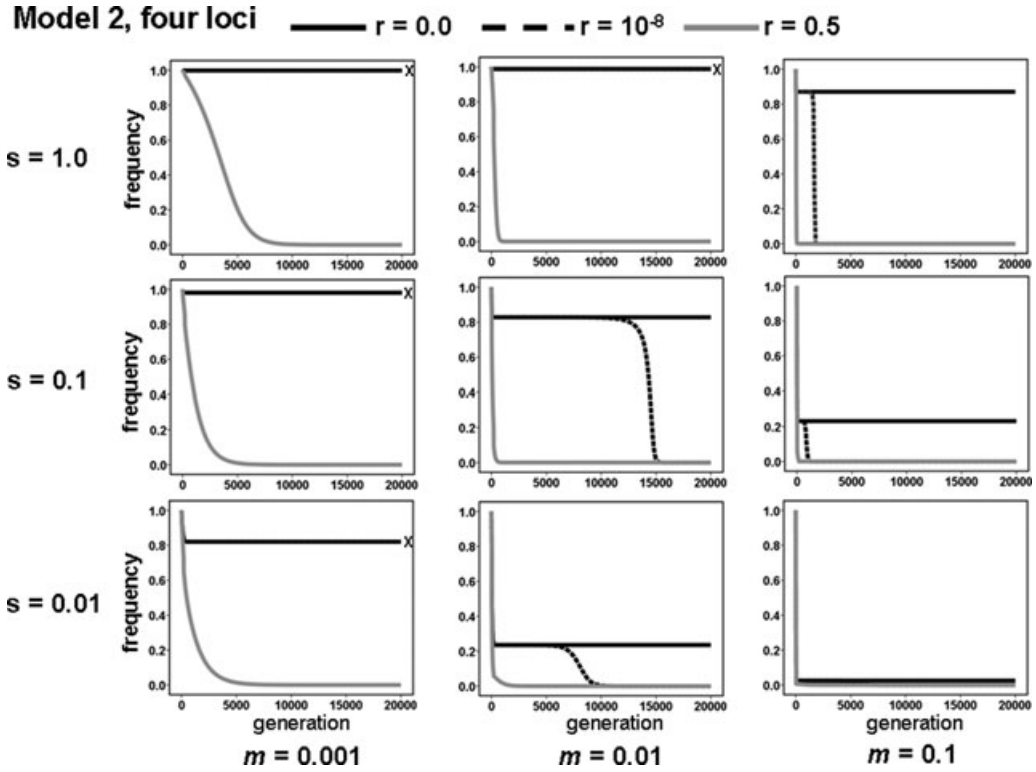

Figure 7. Simulation results from model 2 when four, rather than two, loci were considered (see Fig. 3 for the results from the two locus model). Shown is the difference in allele frequency between populations at locus B (frequency diff.) as a function of the number of generations since secondary contact between populations, for different selection (*s*) and migration (*m*) regimes. In each panel, allele frequencies differences are contrasted between scenarios in which the loci reside in genomic regions with no recombination between loci ($r = 0.0$ = inversion with no recombination), low recombination between loci ($r = 10^{-8}$ = inversion with low recombination), and free recombination between loci ($r = 0.5$ = collinear genomic region). In this model, inversions with no recombination between loci are effective at maintaining genetic differences upon secondary contact. For inversions with low recombination between loci, genetic differences can be maintained for extended periods of time, so long as selection is strong relative to migration. Thus, of all the models considered here, this one has the greatest potential to generate differences between inverted and collinear regions in the maintenance of genetic divergence upon secondary contact. In some cases, the results for two of the three recombination rates were essentially identical, denoted by the small 'X' in the relevant panels.

MAGNITUDE OF DIFFERENTIATION: SELECTION STRENGTH VERSUS GENETIC ARCHITECTURE

As quantitative data on speciation accumulate, it is becoming more evident that divergence in the speciation process often varies continuously (even if the end point is the development of a discontinuity) (Funk 1998; Jiggins and Mallet 2000; Funk et al. 2002, 2006; Mallet et al. 2007; Seehausen et al. 2008; Peccoud et al. 2009). For example, the completeness of reproductive isolation can vary, as can the degree of genotypic clustering in molecular markers or phenotypic traits, the sharpness of geographic clines in gene frequencies, and the extent of lineage sorting (reviewed in Nosil et al. 2009b).

A number of explanations have been proposed for variability in the degree of divergence that originates, and is maintained, during the speciation process (Nosil et al. 2009b). Increased time since population divergence and a lack of gene flow (i.e., allopatric divergence) are obvious factors promoting divergence. Other explanations concern genetic architecture and the nature of divergent natural selection. For example, speciation is predicted to be promoted by genetic factors that reduce recombination between genes under selection and those conferring reproductive isolation (Coyne and Orr 2004; Gavrilets 2004), by increased strength of divergent selection on any given gene (trait), and by divergent selection on a greater number of genes (traits) (i.e., 'multifarious selection' cf. Rice and Hostert 1993).

Our results inform debates about the importance of genetic versus selective factors in promoting the maintenance of genetic divergence, as well as providing some information on the interaction of the two. In general, we observed that strong selection facilitates maintenance of divergence in a manner analogous to reduced recombination. In some cases, strong selection alone was sufficient for strong genetic divergence to be maintained across all genomic regions, including collinear ones with free recombination (e.g., model 1 in Fig. 2, see also model 4 with strong selection

in Supporting information). In other cases, the only manner in which to maintain genetic divergence was to involve inversions with no recombination and large numbers of loci (e.g., model 2 in Fig. 3 and model 3 in Fig. 4). However, the maintenance of genetic divergence was often dependent on an interaction between selection strength and recombination rate, for example with differentiation being maintained in regions with low recombination when selection was strong. Thus, both selection and recombination were important factors contributing to the maintenance of genetic differences, and neither factor should be overlooked.

IMPLICATIONS FOR GENOMIC ISLANDS OF DIVERGENCE

A final topic concerns the size and distribution of genomic regions of accentuated divergence in the genome. Such regions have been referred to as "genomic islands of divergence" (Turner et al. 2005; Harr 2006) and were recently explicitly defined as "a region of the genome, of any size, whose divergence exceeds neutral expectations" (Nosil et al. 2009a, pp. 395). The accentuated divergence of islands often arises due to genes affecting adaptive divergence and reproductive isolation residing within such islands, with the hitchhiking effects associated with such genes resulting in elevated differentiation of physically linked loci (even if such loci are neutral, Charlesworth et al. 1997). In empirical studies, there appears to be much variability in the size and distribution of genomic islands. A number of factors, such as geographic mode of divergence, selection strength, and structural features of the genome such as chromosomal inversions, might affect this variability (reviewed in Nosil et al. 2009a, see also Via and West 2008). For example, it has been suggested that factors that reduce introgression, such as inversions or regions under divergent selection, will facilitate the "growth" of these islands of differentiation and might also result in their being clustered within the genome, rather than genomically dispersed (Gavrilets 2004; Via and West 2008). We observed that even small amounts of recombination can strongly reduce the role of inversions in maintaining genetic differences in the face of gene flow, raising questions about the efficacy of inversions, or even strong selection for that matter, in generating large genomic islands of divergence. Such questions represent an interesting avenue for further research.

CONCLUSIONS AND FUTURE DIRECTIONS

Our most general finding is that the low levels of recombination within an inversion often result in the loss of accentuated divergence in inverted regions compared to collinear ones. We conclude that inversions can facilitate the maintenance of species differences under some conditions, but that large or qualitative differences between inverted and collinear regions need not always occur. We also find that strong selection facilitates maintenance of divergence in a manner analogous to inversions. We considered here genes involved in adaptive divergence and those causing intrinsic genetic incompatibilities. Future work could examine genes affecting mate or habitat choice, because such genes can contribute to assortative mating and thus speciation (Coyne and Orr 2004 for review). It might be particularly relevant to examine the evolution of loci involved in mate or habitat choice in the context of their associations with genes involved in adaptive divergence or intrinsic genetic incompatibilities, for example because such associations could drive reinforcement upon secondary contract (Servedio and Noor 2003; Butlin 2005), as well as sympatric speciation.

An obvious extension to our study is to examine the factors studied here (e.g., different models and variation in recombination rate) but in the context of building up, rather than maintaining, differences within inversions (as in Navarro and Barton 2003). Likewise the effects of different models and recombination rates could be considered in the context of the factors facilitating the initial spread of an inversion in the first place (Kirkpatrick and Barton 2006). In all these cases, including for the scenarios examined in the current study, analytical solutions in the future are desirable. The collective results of such work will likely increase our understanding of not only the role of inversions in the origin and maintenance of genetic divergence, but more generally of the factors driving versus constraining adaptive divergence and speciation.

ACKNOWLEDGMENTS

We thank M. Kirkpatrick, M. Noor, J. Mallet, A. Meyer, and J. Galindo for discussions pertaining to inversions and speciation. During the preparation of the manuscript, the authors were hosted by the Institute for Advanced Study, Wissenschaftskolleg, Berlin. This work was supported by grants to JLF from the National Science Foundation and the United States Department of Agriculture.

LITERATURE CITED

Achere, V., J. M. Favre, G. Besnard, and S. Jeandroz. 2005. Genomic organization of molecular differentiation in Norway spruce (*Picea abies*). Mol. Ecol. 14:3191–3201.

Albert, A. Y. K., S. Sawaya, T. H. Vines, A. K. Knecht, C. T. Miller, B. R. Summers, S. Balabhadra, D. M. Kingsley, and D. Schluter. 2007. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. Evolution 62:76–85.

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

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

Barton, N. H. 1979. Gene flow past a cline. Heredity 43:333–339.

Bateson, W. 1909. Heredity and variation in modern lights. Pp. 85–101 *in* A. C. Seward, ed. Darwin and modern science. Cambridge Univ. Press, Cambridge.

Beaumont, M. A. 2005. Adaptation and speciation: what can F_{st} tell us? Trends Ecol. Evol. 20:435–440.

Brideau, N. J., H. A. Flores, J. Wang, S. Maheshwari, X. Wang, and D. A. Barbash. 2007. Two Dobzhansky–Muller genes interact to cause hybrid lethality in *Drosophila*. Science 314:1292–1295.

Brown, K. M., L. M. Burk, L. M. Henagan, and M. A. Noor. 2004. A test of the chromosomal rearrangement model of speciation in *Drosophila pseudoobscura*. Evolution 58:1856–1860.

Butlin, R. K. 2005. Recombination and speciation. Mol. Ecol. 14:2621–2635.

Charlesworth, B., M. Nordborg, and D. Charlesworth. 1997. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet. Res. 70:155–174.

Cohuet, A., I. Dia, F. Simard, M. Raymond, and D. Fontenille. 2004. Population structure of the malaria vector *Anopheles funestus* in Senegal based on microsatellite and cytogenetic data. Insect Mol. Biol. 13:251–258.

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

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

Egan, S. P., P. Nosil, and D. J. Funk. 2008. Selection and genomic differentiation during ecological speciation: isolating the contributions of host-association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. Evolution 62:1162–1181.

Emelianov, I., F. Marec, and J. Mallet. 2004. Genomic evidence for divergence with gene flow in host races of the larch budmoth. Proc. R. Soc. Lond. B 271:97–105.

Feder, J. L., S. H. Berlocher, J. B. Roethele, H. Dambroski, J. J. Smith, W. L. Perry, V. Gavrilovic, K. E. Filchak, J. Rull, and M. Aluja. 2003a. Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. Proc. Natl. Acad. Sci. USA 100:10314–10319.

Feder, J. L., F. B. Roethele, K. Filchak, J. Niedbalski, and J. Romero-Severson. 2003b. Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. Genetics 163:939–953.

Felsenstein, J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? Evolution 35:124–138.

Funk, D. J. 1998. Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. Evolution 52:1744–1759.

Funk, D. J., K. E. Filchak, and J. L. Feder. 2002. Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica 116:251–267.

Funk, D. J., P. Nosil, and W. Etges. 2006. Ecological divergence exhibits consistently positive associations with reproductive isolation across disparate taxa. Proc. Natl. Acad. Sci. USA 103:3209–3213.

Gavrilets, S. 2004. Fitness landscapes and the origin of species. Princeton Univ. Press, Princeton, NJ.

Geraldes, A., N. Ferrand, and N. W. Nachman. 2006. Contrasting patterns of introgression at X-linked loci across the hybrid zone between subspecies of the European rabbit (*Oryctolagus cuniculus*). Genetics 173:919–933.

Grahame, J. W., C. S. Wilding, and R. K. Butlin. 2006. Adaptation to a steep environmental gradient and an associated barrier to gene exchange in *Littorina saxatilis*. Evolution 60:268–278.

Harr, B. 2006. Genomic islands of differentiation between house mouse subspecies. Genome Res. 16:730–737.

Hawthorne, D. J., and S. Via. 2001. Genetic linkage of ecological specialization and reproductive isolation in pea aphids. Nature 412:904–907.

Hill, W. G., and A. Robertson. 1966. The effect of linkage on limits to artificial selection. Genet. Res. 8:269–294.

Hoffman, A. A., and L. H. Rieseberg. 2008. Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? Ann. Rev. Ecol. Evol. Syst. 39:21–42.

Jaarola, M., R. H. Martin, and T. Ashley. 1998. Direct evidence for suppression of recombination within two pericentric inversions in humans: a new sperm-FISH technique. Am. J. Hum. Genet. 63:218–224.

Jiggins, C. D., and J. Mallet. 2000. Bimodal hybrid zones and speciation. Trends Ecol. Evol. 15:250–255.

Joron, M., R. Papa, M. Beltrán, N. Chamberlain, J. Mavárez, S. Baxter, E. Bermingham, S. Humphray, J. Rogers, H. Beasley, et al. 2006. A conserved supergene locus controls colour pattern diversity in *Heliconius* butterflies. PLOS Biol. 4:e303.

King, M. 1993. Species evolution: the role of chromosomal change. Cambridge Univ. Press. Cambridge.

Kirkpatrick, M., and N. H. Barton. 2006. Chromosome inversions, local adaptation and speciation. Genetics 173:419–434.

Kronforst, M. R., L. G. Young, D. D. Kapan, C. McNeely, R. J. O'Neill, and L. E. Gilbert. 2006. Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. Proc. Natl. Acad. Sci. USA 103:6575–6580.

Machado, C. A., T. S. Haselkorn, and M. A. F. Noor. 2007. Evaluation of the genomic extent of effects of fixed inversion differences on intraspecific variation and interspecific gene flow in *Drosophila pseudoobscura* and *D. persimilis*. Genetics 175:1289–1306.

Mallet, J., M. Beltrán, W. Neukirchen, and M. Linares. 2007. Natural hybridization in heliconiine butterflies: the species boundary as a continuum. BMC Evol. Biol. 7:28.

Manoukis, N. C., J. R. Powell, M. B. Toure, A. Sacko, F. E. Edillo, M. B. Coulibaly, S. F. Traore, C. E. Taylor, and N. J. Besansky. 2008. A test of the chromosomal theory of ecotypic speciation in *Anopheles gambiae*. Proc. Natl. Acad. Sci. USA 105:2940–2945.

Muller, H. J. 1940. Bearings of the *Drosophila* work on systematics. Pp. 185–268 *in* J. S. Huxley, ed. The new systematics. Clarendon Press, Oxford.

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

Navarro, A., and N. H. Barton. 2003. Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. Evolution 57:447–459.

Navarro, A., E. Betran, A. Barbadilla, and A. Ruiz. 1997. Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. Genetics 146:695–709.

Nielsen, R. 2005. Molecular signatures of natural selection. Ann. Rev. Genet. 39:197–218.

Noor, M. A. F., and J. L. Feder. 2006. Speciation genetics: evolving approaches. Nat. Rev. Genet. 7:851–861.

Noor, M. A. F., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. Proc. Natl. Acad. Sci. USA 98:12084–12088.

Noor, M. A. F., D. A. Garfield, S. W. Schaeffer, and C. A. Machado. 2007. Divergent between the *Drosophila pseudoobscura* and *D. persimilis* genome sequences in relation to chromosomal inversions. Genetics 177:1417–1428.

Nosil, P., S. P. Egan, D. J. Funk. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: 'isolation-by-adaptation' and multiple roles for divergent selection. Evolution 62:316–336.

Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009a. Divergent selection and heterogeneous genomic divergence. Mol. Ecol. 18:375–402.

Nosil, P., L. Harmon, and O. Seehausen. 2009b. Ecological explanations for (incomplete) speciation. Trends Ecol. Evol. 24:145–156.

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

Orr, H. A., J. P. Masly, and D. C. Presgraves. 2004. Speciation genes. Curr. Opin. Genet. Dev. 14:675–679.

Ortiz-Barrientos, D., J. Reiland, J. Hey, and M. A. F. Noor. 2002. Recombination and the divergence of hybridizing species. Genetica 116:167–178.

Peccoud, J., A. Ollivier, M. Plantegenest, and J. C. Simon. 2009. A continuum of genetic divergence from sympatric host races to species in the pea aphid complex. Proc. Natl. Acad. Sci. USA 106:7495–7500.

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.

Rice, W. R., and E. E. Hostert. 1993. Laboratory experiments in speciation: what have we learned in 40 years? Evolution 47:1637–1653.

Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. Trends Ecol. Evol. 16:351–358.

Rieseberg, L. H., J. Whitton, and K. Gardner. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two wild sunflower species. Genetics 152:713–727.

Rogers, S. M., and L. Bernatchez. 2007. The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp. Salmonidae). Mol. Biol. Evol. 24:1423–1438.

Santiago, E., and A. Caballero. 1998. Effective size and polymorphism of linked neutral loci in populations under directional selection. Genetics 149:2105–2117.

Santos, M. 2009. Recombination load in a chromosomal inversion polymorphism of *Drosophila subobscura*. Genetics 181:803–809.

Scotti-Saintagne, C., S. Mariette, I. Porth, P. G. Goicoechea, T. Barreneche, K. Bodenes, K. Burg, and A. Kremer. 2004. Genome scanning for interspecific differentiation between two closely related oak species [*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.]. Genetics 168:1615–1626.

Seehausen, O., Y. Terai, I. S. Magalhaes, K. L. Carleton, H. D. J. Mrosso, R. Miyagi, I. Van der Sluijs, M. V. Schneider, M. Maan, H. Tachida, et al. 2008. Speciation through sensory drive in cichlid fish. Nature 455:620–627.

Servedio, M., and M. Noor. 2003. The role of reinforcement in speciation: theory and data. Ann. Rev. Ecol. Syst. 34:339–364.

Stinchcombe, J. R., and H. E. Hoekstra. 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity 100:158–170.

Storz, J. F. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. Mol. Ecol. 14:671–688.

Tang, S., and D. C. Presgraves. 2009. Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities. Science 323:779–782.

Turner, T. L., M. W. Hahn, and S. V. Nuzhdin. 2005. Genomic islands of speciation in *Anopheles gambiae*. PLOS Biol. 3:1572–1578.

Via, S., and J. West. 2008. The genetic mosaic suggests a new role for hitchhiking in ecological speciation. Mol. Ecol. 17:4334–4345.

White, M. J. D. 1978. Modes of speciation. W.H. Freeman and Company, San Francisco, CA.

Wood, H. M., J. W. Grahame, S. Humphray, J. Rogers, and R. K. Butlin. 2008. Sequence differentiation in regions identified by a genome scan for local adaptation. Mol. Ecol. 17:3123–3135.

Wu, I., and C.-T. Ting. 2004. Genes and speciation. Nat. Rev. Genet. 5:114–122.

Yatabe, Y., N. C. Kane, C. Scotti-Saintagne, and L. H. Rieseberg. 2007. Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. Genetics 175:1883–1893.

Associate Editor: M. Doebeli

Supporting Information

The following supporting information is available for this article:

Figure S1: Simulation results from model 4 with strong election (s = 1.0), a model similar to model 3, but in which instead of being neutral, the two derived substitutions were considered to be universally favored across the two populations (i.e., universally beneficial).

Figure S2: Simulation results from model 4 with weak selection (s = 0.01), a model similar to model 3, but in which instead of being neutral, the two derived substitutions were considered to be universally favored across the two populations (i.e., universally beneficial).

Figure S3: Simulation results from model 5 with strong election (s = 1.0), which examined the consequences of negative interactions between derived and ancestral allelic states.

Figure S4: Simulation results from model 5 with weak selection (s = 0.01), which examined the consequences of negative interactions between derived and ancestral allelic states.

Supporting Information may be found in the online version of this article.
(This link will take you to the article abstract).

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.