Geographic Mode of Speciation and Genomic Divergence

Jeffrey L. Feder,[1,2,3] Samuel M. Flaxman,[4] Scott P. Egan,[1,3] Aaron A. Comeault,[5] and Patrik Nosil[5]

[1]Department of Biological Sciences, [2]Environmental Change Initiative, and [3]Advanced Diagnostics and Therapeutics, University of Notre Dame, Notre Dame, Indiana 46556; email: feder.2@nd.edu, Scott.P.Egan.28@nd.edu

[4]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, Colorado 80309; email: Samuel.Flaxman@Colorado.EDU

[5]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, S102TN, United Kingdom; email: aacomeault@gmail.com, p.nosil@sheffield.ac.uk

Annu. Rev. Ecol. Evol. Syst. 2013. 44:73–97

First published online as a Review in Advance on August 26, 2013

The *Annual Review of Ecology, Evolution, and Systematics* is online at ecolsys.annualreviews.org

This article's doi:
10.1146/annurev-ecolsys-110512-135825

Keywords

allopatry, divergent selection, gene flow, genetic hitchhiking, population genomics, secondary contact

Abstract

Understanding speciation requires determining how inherent barriers to gene flow (reproductive isolation, RI) evolve between populations. The field of population genomics attempts to address this question by characterizing genome-wide patterns of divergence between taxa, often utilizing next-generation sequencing. Here, we focus on a central assumption of such "genome scans": regions displaying high levels of differentiation contain loci contributing to RI. Three major issues are discussed concerning the relationship between gene flow, genomic divergence, and speciation: (*a*) patterns expected in the presence versus absence of gene flow; (*b*) processes, such as direct selection and genetic hitchhiking, allowing for divergence with gene flow; and (*c*) the consequences of the timing of when gene flow occurs during speciation (e.g., continuous gene flow versus gene flow following secondary contact after a period of initial allopatric divergence). Theory and existing data are presented for each issue, and avenues for future work are highlighted.

1. INTRODUCTION: AN OVERVIEW OF POPULATION GENOMICS

RI: reproductive isolation

NGS: next-generation sequencing

F_{ST}: fixation index, a measure of genetic differentiation between populations

Divergent selection: selection acting in different directions between populations, including when selection favors two extremes within a single population (disruptive selection)

QTL: quantitative trait locus

Ever since Darwin, evolutionary biologists have been on a quest to understand how sexual populations diverge to become new species (Darwin 1859, Mayr 1963, Coyne & Orr 2004, Nosil 2012). A critical component of this task involves discerning how genetic barriers to gene flow evolve to reproductively isolate taxa (Coyne & Orr 2004, Gavrilets 2004). Historically, this enterprise has focused on uncovering individual genes contributing to reproductive isolation (RI). However, the emerging field of population genomics attempts to address this question by characterizing patterns of genome-wide divergence during speciation, now often utilizing next-generation sequencing (NGS) (Hudson 2008; Ellegren et al. 2012; Jones et al. 2012a,b). Such "genome scans" allow gene regions displaying exceptionally high levels of differentiation to be identified (e.g., high-F_{ST} statistical "outlier loci"); if populations are in geographic contact, it can be inferred that these regions experience reduced gene flow and contain loci under divergent selection that contribute to RI (Luikart et al. 2003, Turner et al. 2005, Li et al. 2008, Stinchcombe & Hoekstra 2008, Noor & Bennett 2009, Nosil et al. 2009, Butlin 2010, Turner & Hahn 2010).

Genome scans are therefore useful in several regards. First, they aid in identifying and mapping individual candidate loci responsible for RI. Second, they can help assess the number, size, and distribution of gene regions contributing to RI (i.e., the genome architecture of speciation). By genome architecture, we refer to general features of the organization of the genome, including gene order, gene density (number of loci per physical distance and recombination rate), and gene distribution along chromosomes, as well as structural features, such as chromosome number and size, centromere and telomere positions, and the presence of chromosomal inversions and translocations. As we discuss below, genome architecture can affect speciation through its effects on different forms of genetic hitchhiking and rates of recombination. Third, genome scans help discern the evolutionary processes driving and constraining divergence.

Several recent reviews have discussed the mapping and identification of specific genes associated with adaptation and RI (Orr et al. 2004, Presgraves 2007, Rieseberg & Blackman 2010, Barrett & Hoekstra 2011, Nosil & Schluter 2011). We refer readers to the sidebar, The Identity and Nature of Speciation Genes, and **Table 1** for a brief overview concerning our current understanding of such speciation genes. NGS has perhaps had its biggest impact on the issue of genome architecture, transforming our understanding of speciation from an individual gene to a

THE IDENTITY AND NATURE OF SPECIATION GENES

Resolution of the number and location of genes causing RI bears on the role genome architecture plays in speciation. Top-down and bottom-up approaches have been used, sometimes together, to identify such speciation genes (Michel et al. 2010). In bottom-up approaches, NGS is used to associate phenotypes with RI through test crosses mapping quantitative trait loci (QTLs), manipulative selection and transplant experiments on phenotypes testing for genetic responses, or scoring candidate loci for divergence (Coyne & Orr 2004, Barrett & Hoekstra 2011, Gagnaire et al. 2013). Top-down approaches use genome scans to identify gene regions displaying elevated divergence under the assumption that they contain RI loci (Lawniczak et al. 2010, Ellegren et al. 2012).

Although progress has been made in identifying speciation genes, many seminal questions remain. For example, does RI most often result from differential adaptation to the environment or from incompatible interactions between loci? What proportions of incompatibilities are associated with the external environment versus the internal genomic environment, sexual selection, meiotic drive, and genetic drift (Nosil & Schluter 2011, Schluter 2009)? What are the roles of coding versus regulatory changes in population divergence (Jones et al. 2012b)?

Table 1 Description of systems where the nature and genetic basis of traits involved in adaptation or speciation and genome-wide patterns of divergence have been described[a]

Taxa	Traits	Genomic divergence	References for traits	References for divergence
Acyrthosiphon pisum	Habitat choice, selection against migrants and hybrids, chemoreception	Localized divergence around QTLs for traits involved in speciation; some clustering of divergent chemoreception genes	Hawthorne & Via 2001, Smadja et al. 2012	Hawthorne & Via 2001, Via & West 2008, Smadja et al. 2012, Via et al. 2012
Anopheles gambiae	Adaptation to broad climate and vegetation zones	Widespread divergence; elevated divergence at regions containing genes associated with immune function, insecticide resistance, chemoreception, and inversions	Coluzzi et al. 1979	Coluzzi et al. 1979; White et al. 2007, 2009, 2010; Lawniczak et al. 2010; Neafsey et al. 2010; Reidenbach et al. 2012; Weetman et al. 2012
Arabidopsis thaliana	Flowering time, freezing tolerance, climate effects on fitness	Widespread divergence associated with climate; numerous genes have been identified influencing flowering time and freezing tolerance. Less is known about levels of differentiation at these loci	Weinig et al. 2003, Stinchcombe et al. 2004, Le Corre 2005, Hannah et al. 2006, Korves et al. 2007	Mitchell-Olds & Schmitt 2006, Van Buskirk & Thomashow 2006, Fournier-Level et al. 2011, Hancock et al. 2011
Coregonus sp.	Feeding morphology, body size, growth rate, swimming behavior	Widespread divergence; divergence tends to be accentuated at regions associated with adaptive traits; divergent regions with unknown functions also exist	Campbell & Bernatchez 2004; Rogers & Bernatchez 2006, 2007	Campbell & Bernatchez 2004; Rogers & Bernatchez 2006, 2007; Renaut et al. 2011; Gagnaire et al. 2013
Gasterosteus aculeatus	Lateral bony plates, pelvic spines, body size and shape, thyroid hormones, male nuptial color	Widespread divergence; elevated at regions of reduced recombination; elevated at regions containing genes associated with adaptive traits; divergent regions with unknown functions also exist	Peichel et al. 2001, Colosimo et al. 2004, Albert et al. 2008, Barrett et al. 2008, Chan et al. 2010, Kitano et al. 2010, Malek et al. 2012	Hohenlohe et al. 2010, 2012; Deagle et al. 2012; Jones et al. 2012a,b; Roesti et al. 2012
Heliconius erato and *H. melpomene*	Warning coloration	Divergence isolated to two genomic regions associated with color and pattern variation	Joron et al. 2006, 2011; Reed et al. 2011	Baxter et al. 2010, Counterman et al. 2010, *Heliconius* Genome Consort. 2012, Nadeau et al. 2012
Lycaeides sp.	Male genitalic morphology, oviposition preference	Numerous divergent regions; regions associated with traits tend to be more divergent than regions with unknown functions; divergent regions with unknown functions also exist	Gompert et al. 2013	Gompert et al. 2012a, 2013

(*Continued*)

Table 1 (*Continued*)

Taxa	Traits	Genomic divergence	References for traits	References for divergence
Rhagoletis pomonella	Host choice, diapause development and eclosion time	Elevated divergence at regions associated with diapause development and eclosion time	Feder et al. 2003a, Michel et al. 2010	Feder et al. 2003a, Michel et al. 2010

[a]For several of the systems, further information is required to clarify whether gene flow has been continuous and primary or secondary during the divergence process.
Abbreviation: QTL, quantitative trait locus.

whole genome perspective. The topic of the genome-wide architecture of speciation has also been reviewed (Luikart et al. 2003, Li et al. 2008, Stinchcombe & Hoekstra 2008, Nosil et al. 2009, Butlin 2010, Feder et al. 2012a). However, past reviews have not focused on the major assumption often used to interpret patterns of divergence in genome scans, namely that ongoing gene flow between populations homogenizes variation in genomic regions not affected by divergent selection or RI (Noor & Bennett 2009, Turner & Hahn 2010). We therefore focus on this issue here and organize our discussion around three major topics concerning the relationship between gene flow, genomic divergence, and speciation: (*a*) a comparison of patterns of genomic divergence expected in the presence versus absence of gene flow (i.e., the effects of geography on genomic divergence when speciation occurs in sympatry versus allopatry), (*b*) the processes allowing for divergence with gene flow and how they relate to genome architecture and stages of speciation, and (*c*) the consequences of the timing of gene flow on speciation, as for example, when populations initially diverge in allopatry, but then subsequently come into secondary contact.

We begin with an introduction of the critical relationship of linkage disequilibrium to effective gene flow, patterns of genomic divergence, and speciation. We then examine each of the three main issues of geography, process, and timing, discussing theory and existing data for each issue. Our take-home message is that although progress is being made in understanding genomic divergence during speciation, key elements of theory and data are still missing. We therefore highlight directions for future work throughout this review.

2. LINKAGE DISEQUILIBRIUM AND THE SPREAD OF REPRODUCTIVE ISOLATION ACROSS THE GENOME

Speciation occurs as genetically based barriers to gene flow evolve between populations. Thus, a key measure of how far speciation has progressed is gene flow or, more accurately, the effective gene flow rate, which distinguishes the gross migration rate, m (of individuals moving between populations), from the effective migration rate, m_e (of introgression by the alleles these migrants carry into the alternative population) (Bengtsson 1985, Zhivotovsky & Christiansen 1995, Gavrilets & Cruzan 1998, Hendry et al. 2000, Gavrilets 2003, Bierne et al. 2011, Kobayashi & Telschow 2011, Flaxman et al. 2012). Given migration, the degree of genetic differentiation for a variable nucleotide site (abbreviated SNP hereafter, for single nucleotide polymorphism) is often assumed to be a relative measure of m_e, with regions showing increased frequency differences experiencing lower m_e and thus containing loci under divergent selection or contributing to RI.

Another measure of population divergence of higher-order complexity than individual SNP frequency differences is the extent of statistical association or coupling between SNPs, i.e., linkage disequilibrium (LD). There are two general ways that LD can be promoted among SNPs

m: migration rate

m_e: effective migration rate

SNP: single nucleotide polymorphism

LD: linkage disequilibrium

(Barton 1979, 1983; Barton & Bengtsson 1986). One is when positive assortative or habitat-specific mating occurs, resulting in individuals choosing to breed with others of like genotype. The second is selection, or more specifically, when external ecological or internal genomic environments favor alternate suites of alleles.

Selection-recombination antagonism: selection acts to build up, whereas recombination breaks down, adaptive associations between loci contributing to reproductive isolation

Selection can affect LD on two different scales that are not mutually exclusive (reviewed by Smadja & Butlin 2011). The first is locally in the genome. When a SNP is under divergent selection, it creates a region of locally reduced m_e, and thus elevated LD, for nearby physically linked sites. This occurs because sites in the neighborhood of a selected SNP may not have had enough time for recombination to break up its association with the linked selected SNP and thus evolve independently from it. The second scale is globally across the genome. When several genes are under divergent selection, the combined selection exerted by these loci can also reduce m_e across the genome for all sites relative to that in the absence of selection. This occurs because sites associated with the entire disfavored genome of a migrant may not have had enough time to disassociate from all of the selected loci before being eliminated from the alternate population. A critical transition for speciation-with-gene-flow can therefore occur when the global reduction in m_e becomes great enough for genetic differentiation and for RI to begin to accumulate genome-wide and not be restricted to local sites that are physically linked to divergently selected SNPs (Feder et al. 2012a). At this stage, the genomes of taxa start to differentially congeal, with LD and divergence becoming universally higher among all selected sites throughout the genome (Turner 1967, Barton 1983, Feder & Nosil 2010, Abbott et al. 2013, Flaxman et al. 2013). In essence, RI increasingly becomes a characteristic of the entire genome rather than of individually selected loci.

3. ISSUE 1—DIVERGENCE WITH OR WITHOUT GENE FLOW: THEORY

Genome architecture is most relevant when speciation occurs with gene flow because in such cases an antagonism (i.e., selection-recombination antagonism) exists between divergent selection building up favorable combinations of locally adapted genes and migration and recombination breaking them down and homogenizing populations (Felsenstein 1976, 1981; Gavrilets 2004). Hence, genomic features that reduce recombination between populations (e.g., chromosomal inversions, translocations or centromeres) can enhance the effectiveness of divergent selection by initially creating and also maintaining LD (Noor et al. 2001, Rieseberg 2001, Feder & Nosil 2009, Nachman & Payseur 2012). By contrast, there is no antagonism between selection and interpopulation recombination among allopatric populations because geographic barriers preclude gene flow (Kirkpatrick & Ravigné 2002). As a result, physical linkage is not as critical for allopatric divergence because genome-wide LD is generated between populations by individuals mating and evolving independently in the physically separated demes. As such, allopatric populations are expected to readily differentiate in many genomic regions via selection, as well as by drift.

The above considerations generate the following predictions: (*a*) Populations undergoing speciation-with-gene-flow should be more sensitive to homogenizing gene flow and physical linkage, resulting in differentiated loci being concentrated into a smaller number of highly diverged regions (e.g., a more "L-shaped" F_{ST} distribution) compared with allopatrically speciating populations (Via 2001, Savolainen et al. 2006); (*b*) some high-F_{ST} outliers between allopatric populations will exhibit reduced gene flow if they are studied in interbreeding populations (i.e., those that contribute to RI), but others will not. This is because not all genes that diverge in allopatry will contribute to RI, and some that do, such as those generating unfit hybrids, can be eliminated by selection. We note that even these predictions are not completely straightforward, as factors, such

as recent divergence, can result in patterns for allopatric populations that do not differ markedly from those undergoing gene flow (Nosil 2012).

4. ISSUE 1—DIVERGENCE WITH OR WITHOUT GENE FLOW: DATA

Many genome scans focus on populations believed to be undergoing gene flow because such systems provide natural laboratories for examining the speciation process (Harrison 1991). In the absence of gene flow, loci causing RI may be mapped between allopatric populations, but it cannot be directly confirmed that these genes are involved in reducing gene flow. Moreover, it can be unclear when these loci diverged and, thus, whether they were integral to speciation or arose when the process was more or less complete (Nosil & Schluter 2011).

Nevertheless, empirical studies of allopatric populations can still be very valuable for helping to understand speciation, particularly if combined with data from hybridizing populations. Empirical studies contrasting genomic divergence under different geographic modes of divergence are few, but they are beginning to accumulate and provide initial support for the predictions above. For example, Nosil et al. (2012a) showed that the distribution of locus-specific F_{ST} values tended to be L-shaped, with most loci showing little or no divergence between adjacent parapatric populations experiencing high levels of gene flow (**Figure 1*a***). By contrast, the distribution was more highly skewed to the right with more loci displaying a higher F_{ST} for allopatric populations experiencing lower or no gene flow (**Figure 1*b***). The number and size of F_{ST} outlier regions have similarly been shown to vary with levels of gene flow among molecular forms of the mosquito *Anopheles gambiae* (Weetman et al. 2012), ecotypes of *Coregonus* whitefish (Gagnaire et al. 2013), and host races versus species of *Rhagoletis* flies (Powell et al. 2013).

Considering the second prediction, Gompert et al. (2012a) used data from two species of butterflies to pioneer an approach testing whether regions of exceptional divergence between allopatric parental populations undergo atypical patterns of introgression in admixed hybrid zones (Gompert & Buerkle 2009, 2011). As expected, they found some correspondence between locus-specific divergence among allopatric populations and locus-specific introgression in admixed populations. However, this correspondence was partial, and some loci departed strongly

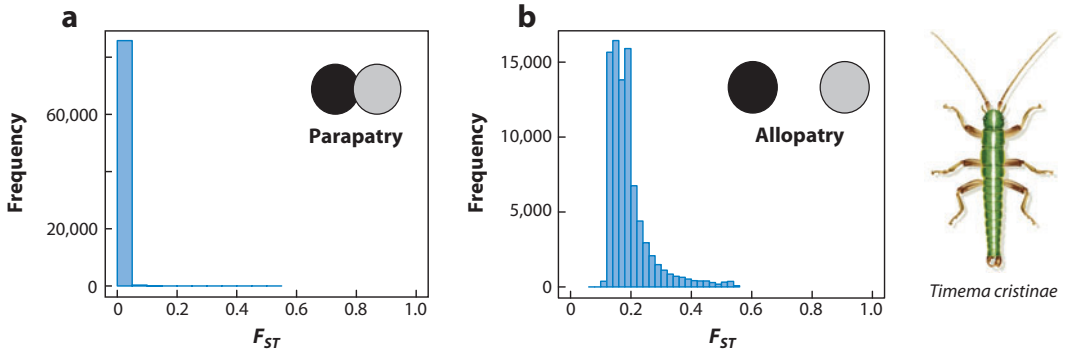


Figure 1

Genomic divergence of *Timema cristinae* stick insect populations under different geographic settings, as indicated by the distribution of F_{ST} values across loci ($n = 86{,}130$ single nucleotide polymorphisms) for comparisons between (*a*) parapatric and (*b*) allopatric pairs of populations [see Nosil et al. (2012a) for details]. The F_{ST} distributions tended to be L-shaped for geographically adjacent parapatric population pairs and skewed to the right for allopatric populations experiencing lower gene flow. *Timema cristinae* drawing by Rosa Ribas.

from the relationship. Very similar trends were reported in *Timema cristinae* stick insects (Nosil et al. 2012b). Thus, geographic variation in selective regimes and genome architecture, coupled with the potential for genetic drift between allopatric populations, can uncouple associations between locus-specific genetic divergence and locus-specific gene flow. The implication is that some strongly divergent gene regions can be "incidental" to the speciation process (Turner & Hahn 2010, Barrett & Hoekstra 2011), whereas others are not. Further studies contrasting patterns of genomic divergence in relation to patterns of gene flow are required (Lasky et al. 2012).

DS: direct effect of divergent selection

***s*:** selection coefficient

5. ISSUE 2—PROCESSES DRIVING DIVERGENCE WITH GENE FLOW: THEORY

5.1. Basic Theory of Genomic Divergence

Once gene flow is demonstrated between populations, a major issue is that of determining the processes driving and/or constraining genetic differentiation. Three general processes can aid the evolution of RI with gene flow (Feder et al. 2012a). The first involves divergent selection acting directly on a locus (DS hereafter). Here, a major consideration is the size of the selection coefficient, *s*, which describes the relative fitnesses of the alternate favored homozygotes in two populations. This is generally with respect to habitat performance, as compared with the gross migration rate, *m* (**Figure 2**) (Yeaman & Otto 2011, Yeaman & Whitlock 2011). (In this respect, *s* equates with the strength of divergent selection acting on a given homozygous variant at a locus and does not describe the relationship between genotype and phenotype, per se, except for the latter's consequences on fitness.) When $s > \sim 0.5m$, then there is at least a fair chance that a

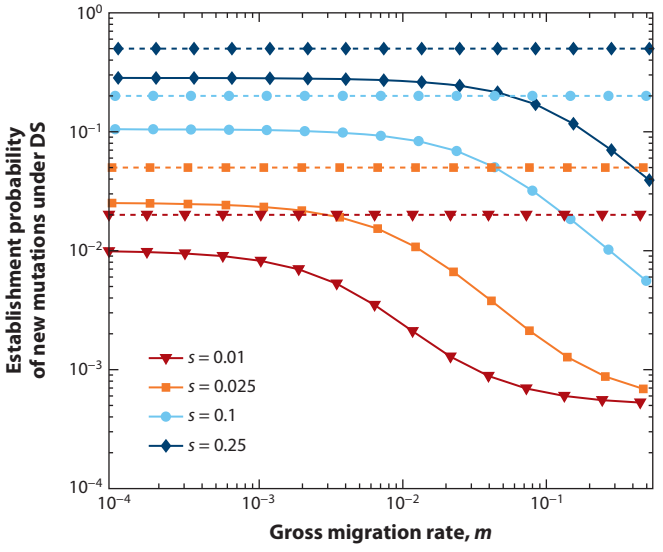


Figure 2

Establishment probabilities of divergently favored new mutations under direct selection (DS) between populations in sympatry with variable migration rates (*solid lines*) versus populations in allopatry with no migration (*dashed lines*). Four different strengths of selection (*s*) are shown. The allopatric fixation probability shown is 2*s*, whereas the establishment probabilities with gene flow were estimated using equation 9 from Yeaman & Otto (2011) on the basis of the diversification coefficient they derived. Assumptions and parameters: total population size = 1,000, codominance of alternative alleles ($h = 0.5$), and symmetrical divergent selection in two ecologically different habitats.

variant will establish and come to differentiate populations adapted to different habitats (Yeaman & Otto 2011). Thus, if new mutations with $s > 0.5m$ arise not infrequently in populations or exist as standing variation, then the effects of DS alone may often be sufficient for RI to increasingly evolve between taxa through time until gene flow between them ceases. This process may also be facilitated by the replacement of small-effect alleles by larger-effect ones through time (Holt & Barfield 2011, Yeaman & Whitlock 2011).

Divergence hitchhiking (DH): localized region of reduced effective gene flow created around selected sites enhancing the potential for differentiation

Genome hitchhiking (GH): divergence across the entire genome is facilitated by a global reduction in effective gene flow caused by all loci under selection

Genomic island of divergence: a region of the genome whose divergence exceeds neutral background expectations based on overall divergence across the genome

The next two factors affecting divergence are forms of genetic hitchhiking: divergence hitchhiking (DH) and genome hitchhiking (GH). We use the term hitchhiking here in the broad sense of "...the indirect effects of selection at one or more loci on the rest of the genome" (Barton 2000, p. 1553). DH invokes a key role for physical linkage (Via 2009, 2012; Via & West 2008; Via et al. 2012). In light of DH, direct selection on already diverged genes reduces m_e locally for nearby surrounding sites. As a result, the chance that a new variant with selection coefficient s will establish in this window of reduced m_e is greater than that for direct selection acting alone on the mutation. Instead of m, s must now be only $>\sim 0.5m_e$. GH occurs when the combined effects of divergent selection on all loci reduce m_e globally to the point that many new mutations distributed across the genome have $s >\sim 0.5m_e$ and thus manage to establish (**Figure 3**) (Feder et al. 2012b).

As discussed below in Section 6 on empirical data, patterns of genomic differentiation across the genome can be used to distinguish the processes driving genomic differentiation, where localized clusters of divergence can be indicative of DS or DH and genome-wide differentiation indicative of GH. Nonetheless, much heterogeneity in levels of differentiation across the genome is still expected under GH owing to variation across the genome in the distribution of sites under selection, their s values, and recombination rates.

5.2. Four-Phase Model

GH and DH are not mutually exclusive processes and may act simultaneously to aid speciation-with-gene-flow. The seminal question then is what is their relative importance at different points in the speciation process? This issue is conceptualized in a four-phase model of speciation-with-gene-flow (Feder et al. 2012a). Initially in phase 1, a few loci differentiate between populations due to strong divergent selection (DS dominant phase). After this, in phase 2, local reductions in m_e for sites surrounding these few, initially diverged loci can facilitate the establishment of new mutations with lower s values than would be possible by DS alone. Differentiation thus sequentially builds in magnitude and spreads in width for these genomic islands of divergence during this DH-dominated phase. Then, phase 3 is reached in which the sum total of divergent selection across loci is sufficient for new mutations to effectively establish across the genome by strong or weak selection or even by genetic drift. During this GH-dominated phase, differentiation may still be heterogeneous, but the variation between regions of high and low divergence steadily decreases as the baseline level of divergence between populations rises. It is important to note that this baseline level does not necessarily reflect neutral expectation. Consequently, outlier analyses in genome scans can statistically overlook a potential role for regions of lower divergence in speciation (Michel et al. 2010). In phase 4, the congealing of the genome begun in phase 3 comes to completion, and the two taxa reach a state of low or no gene flow across the genome. We stress that transitions between the four phases are not necessarily sharp and can be diffuse. Again, DS, DH, and GH may all contribute partly to divergence in all the phases, and it is their relative contributions that differ. Nonetheless, it has been argued that reaching a stage where GH is enabled may be important for speciation-with-gene-flow (Feder et al. 2012a, 2013), resulting in RI and LD spreading across congealing genomes (**Figure 3**), as individual selected sites become genomically "coupled" together as a barrier to gene flow (Flaxman et al. 2013).

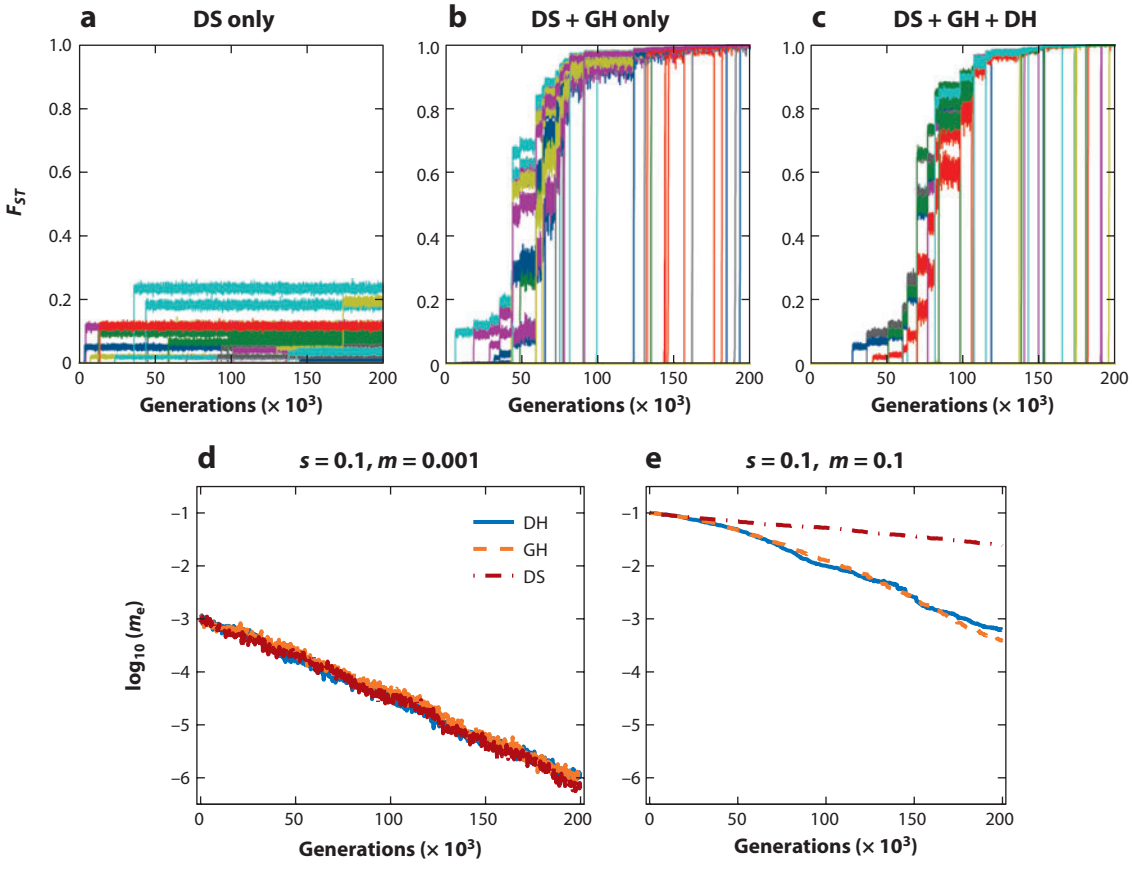


Figure 3

(*a–c*) The accumulation of locus-by-locus differentiation and (*d*,*e*) the total reduction in effective migration rate, m_e, over time for simulations of primary divergence with gene flow [see Flaxman et al. (2013) for details]. Results are shown for the effects of DS (direct selection) only, DS + GH (genome hitchhiking) together, and DS + GH + DH (divergence hitchhiking) combined. The simulations modeled DS on new mutations in a population with two demes exchanging migrants at a rate of (*a–c*,*e*) 10% or (*d*) 0.1% per generation. A new divergently selected mutation enters one of the two demes every 200 generations with a selection coefficient (s) drawn from an exponential distribution with a mean of 0.1. Panels *a–c* show results for a single simulation run in which lines in different colors represent different loci. Lines that elevate from zero depict the successful establishment of a new mutation. With the effects of DS only (*a*), successful mutations reach F_{ST} values reflecting single-locus migration-selection balance. In contrast, with GH present (*b*,*c*) and after several mutations establish, the genome starts to congeal, as the sum total of selection on all mutations causes reductions in gene flow to increase divergence for all loci. This congealing occurs similarly without (*b*) or with (*c*) DH. Panels *d* and *e* display the reduction in the backward migration rate over time, a genome-wide measure of reproductive isolation, averaged over more than 10 independent simulation runs for DS, GH, and DH scenarios. When m_e is low and selection is strong (*d*), reproductive isolation develops and m_e decreases rapidly over time with or without the effects of hitchhiking (i.e., the lines for all three scenarios overlap). However, when the average strength of selection and are both high (*e*), GH (with or without DH) aids mutation establishment and m_e is reduced ~100 fold (by the end of the simulations) compared with the action of DS alone. Reprinted with permission from Wiley.

5.3. Details of Theory and Its Predictions

Theoretical models have implied that DH can facilitate speciation-with-gene-flow, but only under certain conditions (Charlesworth et al. 1997; Feder & Nosil 2010; Feder et al. 2012a,b, 2013). We discuss details of existing theory about the buildup of divergence from new mutations below,

considering divergently selected and then neutral mutations in turn, and note that further work on divergence from standing genetic variation is required (Barrett & Schluter 2008).

cM: centi-Morgan

n_e: effective population size

For DH to have substantial effects on increasing the probability of the establishment of a new divergently selected mutation, the mutations need to be close [≤1 centi-Morgan (cM)] to an already diverged site experiencing strong selection ($s \geq \sim 0.1$) (Feder & Nosil 2010; Feder et al. 2012a,b, 2013). The distribution of fitness effects for new mutations also has an important impact for DH: If new mutations have large s values, DS largely determines their fates. Thus, for DH to play a greater role in speciation, the vast majority of new mutations must have s values much less than m. However, when this is the case, a large number of new mutations must establish to significantly reduce RI due to their low individual s values. Given that only a few regions of the genome will initially contain diverged loci under strong selection, the majority of new mutations will occur at sites outside the influence of DH. Thus, though these unlinked mutations will have very low probabilities of establishment, their sheer numbers may mean that many come to differentiate populations due to DS rather than DH. Hence, even when the mutational spectrum is favorable for DH, it does not mean that it will be the predominant process facilitating speciation. Finally, when m is high during early stages of speciation-with-gene-flow, tight physical linkage can actually have an overall negative effect on the rate that new divergently selected mutations establish, for example due to Hill-Robertson effects (Feder et al. 2012a,b; Flaxman et al. 2013).

For neutral sites, the conditions under which DH can generate divergence are very sensitive to the following parameters: effective population size (n_e), m, s, and the number of selected loci (Charlesworth et al. 1997, Feder & Nosil 2010). When a single locus is under divergent selection, n_e is not large (≤1,000), m is low (≤0.001), and selection on a linked diverged site is strong ($s = 0.5$), then linked neutral SNPs can display high levels of differentiation relatively far away (10–20 cM) from the selected locus. However, increasing n_e or m above these levels or decreasing s for the selected locus dramatically reduces the window of neutral divergence. Increasing the number of selected loci readily results in effects of GH overcoming those of DH.

Nonetheless, there are still circumstances that are expected to enhance the contribution of DH to speciation-with-gene-flow, and these are less well explored in formal models. For example, structural features of the genome, such as chromosomal inversions and translocations, decrease recombination rates and, thus, can increase the effective range of DH along a chromosome (Feder et al. 2003b, Nosil et al. 2009, Kirkpatrick 2010, Joron et al. 2011). However, elevated divergence in chromosomal inversions could be due to factors other than divergent selection (Noor & Bennett 2009, Guerrero et al. 2012); thus, measures of selection acting on inverted regions can help to clarify their role in speciation (Lowry & Willis 2010, Ayala et al. 2013). In addition, certain forms of epistatic fitness interactions among loci can result in conditions that are favorable for DH. The role of small effective population sizes is more nuanced because although this might increase neutral divergence via DH, it could actually slow differentiation from divergently selected mutations because small populations exhibit a longer waiting time for the emergence of favored mutations and there is a higher chance for their stochastic loss. These considerations all require further theoretical analysis.

An additional factor to consider when evaluating the relative importance of DH is that when a few strongly selected sites have established in the genome, GH begins to become enabled (Feder & Nosil 2010, Flaxman et al. 2013). This has been used to argue that when speciation-with-gene-flow does occur, it may often involve a rapid transition (even a jump) from phase 1 to 3 (Feder et al. 2013, Flaxman et al. 2013). It is important to caution, however, that merely observing a number of selected sites distributed across the genome is not sufficient to verify GH, as such a pattern could still be a consequence of DS alone. In this regard, estimates are needed of m and of s

for individual mutations to test whether several of these mutations have s values lower than $0.5m$, but higher than the genome-wide estimate of $0.5m_e$, implying GH.

6. ISSUE 2—PROCESSES DRIVING DIVERGENCE WITH GENE FLOW: DATA

Interpreting process from genome scans is complicated by the problem that many theoretical predictions concerning DS, DH, and GH apply to divergently selected sites (Feder et al. 2012b, Flaxman et al. 2013), whereas patterns of differentiation are discerned from SNPs that are mostly functionally neutral, even if sometimes physically linked to selected sites. Models that better establish expectations for neutral loci are therefore needed (e.g., Guerrero et al. 2012), but some general considerations are clear. Neutral SNPs can be affected mostly by genetic drift, which at equilibrium represents a balance between n_e, m_e (indirectly influenced by selection), and the neutral mutation rate (μ_n). When DH causes significant local reductions in n_e and m_e, the combined effect can influence the time it takes to reach equilibrium and elevate divergence for linked neutral SNPs compared to baseline expectations. However, as stressed above, it is not drift but the potential for hitchhiking to facilitate the establishment of new mutations causing RI that is critical for speciation. In this regard, smaller n_e might generally slow the evolution of RI. Other important parameters, such as m and the fitness spectrum of new mutations, are also not revealed from genome scans. These factors underscore the importance of resolving the natural history and demographics of study systems to more accurately interpret genome scans.

Other difficulties with interpreting data also exist. For example, finding multiple SNPs with outlier status in a region does not verify multiple clustered sites under divergent selection that contribute to RI, because this could be due (under the limited conditions noted above) to a single locus under selection with the surrounding (e.g., neutral) sites showing elevated differentiation due to a combination of reduced m_e and drift. For DH to be strongly enhancing speciation, it should be shown that divergently selected sites, or those contributing to RI, are sequentially accumulating near each other in a clustered distribution in the genome. This requires showing that more than one locus in a gene region is under divergent selection, a task again highlighting the need to couple studies of natural history, manipulative experiments, and mapping results with genome scans. In addition, it must be shown that such a clustering of genomic divergence is not just due to chance.

Taking these caveats and considerations into account, certain observations are nevertheless consistent with a role for DH in genomic divergence. Examples are the large regions of differentiation observed between host races of pea aphids and ecotypes of whitefish (Rogers & Bernatchez 2007, Via & West 2008, Via 2009, Renaut et al. 2011, Via et al. 2012, Gagnaire et al. 2013) and some genomic clustering of divergence in stickleback and butterfly genomes (*Heliconius* Genome Consort. 2012, Hohenlohe et al. 2012, Jones et al. 2012b, Roesti et al. 2012). However, other factors argue against a major role for DH. For example, many studies have reported individual regions of genomic divergence to be small (Turner et al. 2005; Turner & Hahn 2007; Strasburg & Rieseberg 2008; Strasburg et al. 2009, 2012), consistent with theoretical predictions that DH, when it occurs, will be limited to sites in very close proximity to a selected locus. A recent study in *Mimulus* monkeyflowers provides important support for this prediction by showing that a locus affecting intrinsic hybrid dysfunction diverged between populations via selection on a different locus conferring copper tolerance (i.e., hitchhiking), but one which was very tightly linked (Wright et al. 2013). Other lines of evidence argue against a key role for DH in genomic divergence. For example, even if not completely randomly distributed across the genome, outlier loci are often

scattered across the genome (Nosil et al. 2009; Strasburg et al. 2009, 2012; Lawniczak et al. 2010; Michel et al. 2010; Fournier-Level et al. 2011; Hancock et al. 2011; Ellegren et al. 2012), rather than restricted to a few regions. A final note of caution is that large regions of divergence may represent regions containing multiple QTL, some of which are undetected, rather than extended hitchhiking effects of neutral SNPs from a single selected locus. Such undetected QTL could be common for traits that are hard to see or measure, such as some aspects of physiology or behavior.

Isolation by adaptation (IBA): positive correlation between adaptive phenotypic or ecological divergence between populations and their genetic differentiation that can be created by GH

How widespread divergence is across the genome can help distinguish the processes affecting genomic divergence. In this regard, studies of isolation by adaptation (IBA) can be informative, where IBA is a correlation among population pairs between their degree of adaptive divergence (e.g., inferred from ecological or phenotypic divergence) and levels of molecular genetic differentiation between them (Nosil et al. 2008, 2009; Thibert-Plante & Hendry 2010; Funk et al. 2011). For example, if IBA is restricted to just a few gene regions, it implies that DS or DH is generating divergence of those regions and m_e is not sufficiently reduced by GH to generate genome-wide differentiation (Nosil et al. 2008, Thibert-Plante & Hendry 2010). A potential example concerns alfalfa and clover host races of pea aphids, where genetic divergence at loci unlinked to QTLs is near zero, implying homogenization by gene flow and a lack of GH (Via & West 2008). By contrast, IBA across most of the genome between willow and maple host forms of *Neochlamisus* leaf beetles is consistent with strong effects of GH (Funk et al. 2011). Other examples in a wide range of organisms have reported genome-wide IBA, a lack of IBA, or IBA restricted to a small portion of the genome (reviewed by Nosil et al. 2009). Further empirical work on IBA, especially in relation to the selective neutrality and genomic distribution of gene regions examined, is warranted.

7. ISSUE 3—TIMING OF GENE FLOW AND PRIMARY VERSUS SECONDARY CONTACT: THEORY

Even if a period of gene flow can be shown to have occurred during divergence, it often remains untested whether gene flow occurred during primary versus secondary contact. This distinction is important because if secondary contact is involved, then regions of divergence may not reflect the buildup of RI but rather its maintenance upon secondary contact or variation in breakdown upon secondary contact (Dieckmann & Doebeli 1999, Gavrilets 2003, Barluenga et al. 2006, Bolnick & Fitzpatrick 2007, Fitzpatrick et al. 2008, Mallet et al. 2009). The significance of DH for speciation has been mainly considered in terms of how it can help drive speciation de novo in the face of continuous gene flow. However, as discussed below, care must be taken in interpreting genome scans when secondary contact is known, suspected, or cannot be ruled out, which may be quite often. A major remaining challenge in this regard is therefore the further development of analytical approaches for distinguishing primary from secondary contact (Niemiller et al. 2008, 2010; Strasburg & Rieseberg 2010, 2011, 2013).

Of course, DH and GH could still aid in the evolution of additional RI following secondary contact. However, the question of the buildup of divergence following secondary contact is difficult to empirically address compared to primary contact because it requires not only resolving the genomic distribution of divergently selected loci but also how they arose and accumulated in already diverged regions after the onset of gene flow. Given these empirical difficulties, the question of the roles of DH and GH following secondary contact would benefit from additional theory. Models have been developed examining how new universally favored alleles causing genomic incompatibilities in hybrids could establish following second contact, with specific reference to chromosomal inversions (Feder et al. 2011). More work is needed to generalize these findings. Specifically, computer simulations could contribute greatly by establishing the

probability distributions of outcomes that are likely in the full spectrum of speciation scenarios ranging from continuous (primary) contact to prolonged allopatry and then a range of times since secondary contact (Barton 2001, Nosil & Flaxman 2011). Additionally, consideration of the effects of the spatial arrangement of populations, even if gene flow is ongoing, also warrants further modeling attention (Cain et al. 1999, Flaxman et al. 2012).

DM: Dobzhansky-Muller incompatibilities

Chromosomal inversion: rearranged chromosomal segment with an inverse order of genes compared with collinear regions that generally suppresses recombination locally

7.1. Differences and Similarities between Primary and Secondary Contact

In the section below, we discuss how genomic divergence might be different and similar when it comes to primary versus secondary contact. A major difference for secondary contact is that in the absence of the homogenizing effects of gene flow, differentiation can accumulate throughout the genome during the initial period of allopatry with no need for physical linkage. Indeed, when many sites in the genome are simultaneously under selection, increased recombination and looser linkage can actually be favorable, for example, in helping to avoid Hill-Robertson effects (Cutter & Choi 2010, Feder et al. 2012b, Flaxman et al. 2012). Nonetheless, there could be cases where linkage does facilitate evolution in allopatry, such as meiotic drive systems where tight linkage between driver and responder sequences is required for the spread of the system within a population (Burt & Trivers 2006, Crespi & Nosil 2012) or when epistatically acting adaptive polymorphisms are segregating at several loci.

A second aspect of secondary contact is that the probability of establishment of a new mutation in the initial period of allopatry is $\sim 2s$ in a large population, where s here refers to the fitness advantage gained by the initially rare heterozygote carrying the mutation. This probability can be reduced substantially with high gene flow. Thus, many new mutations of low selective advantage can establish in allopatry, especially if the distribution of fitness effects is skewed in this direction, than would readily establish in primary contact. However, there are some unresolved questions in this context. How much faster is the rate of accumulation of mutations in allopatry as compared to primary contact? How often do mutations accumulating in allopatry contribute to RI?

A third major difference for secondary contact is that Dobzhansky-Muller (DM) incompatibilities and competing meiotic drive systems can evolve during the initial geographic separation in allopatry; it is much more difficult for them to differentiate populations in primary contact, especially in early stages of speciation-with-gene-flow (Endler 1977, Turelli et al. 2001, Coyne & Orr 2004, Gavrilets 2004, Agrawal et al. 2011). Thus, the effect of new mutations on RI in allopatry is not limited to their direct effects on ecological fitness. With the exception of reinforcement on prezygotic isolation, there are thus potentially more opportunities for RI to evolve in allopatry (including as an accidental consequence of directional selection and drift) (Turelli et al. 2001, Coyne & Orr 2004).

A period of allopatry can also facilitate the spread of genomic features that reduce recombination, such as chromosomal inversions, which are predicted to promote speciation (Feder et al. 2011). Inverted genomic regions exhibit decreased breakup of adaptive combinations of alleles across loci and thus can be favored over collinear regions when gene flow occurs (Navarro & Barton 2003). If such features originate in allopatry, they subsequently can become favored and increase rapidly to high frequency via selection following secondary contact. New chromosomal inversions can also arise and be favored in primary contact (Kirkpatrick & Barton 2006). However, this is more difficult in primary contact because the new arrangement must usually capture all favored alleles together across loci within the chromosomal region it encompasses to be favored by selection. This is less of a problem for a new chromosomal inversion in allopatric populations because allopatric populations are expected to be well adapted across the genome. Moreover,

inversions originating in allopatry can be less likely to exhibit stochastic loss upon secondary contact because they can be present as prestanding variation in multiple copies (Feder et al. 2011).

Despite the above differences, there are also similarities between primary and secondary contact. Essentially, the initial period of allopatry for secondary contact systems can be thought of as providing a "head start" to reaching intermediate or later phases of the speciation process (Feder et al. 2012a). After secondary contact, similar considerations as for primary contact may sometimes apply to the buildup of additional divergence and RI, especially if most existing RI before contact is not neutral. The possibility for developing a more unified theory of speciation genomics can be seen in verbal models of stages of allopatric isolation and secondary contact that incorporate patterns of divergence that are analogous to those discussed for primary contact above, i.e., the spread of RI from the gene to genome level (Wu 2001). Moreover, previous work on cline theory has shown how RI-causing loci can become coupled across spatial and genome clines in hybrid zones, causing the genome to "congeal" (Barton 1983, Bierne et al. 2011, Gompert et al. 2012b, Abbott et al. 2013). These results may thus bear on the expected outcomes of primary and secondary contact once comparable levels of divergence and RI have evolved.

8. ISSUE 3—TIMING OF GENE FLOW AND PRIMARY VERSUS SECONDARY CONTACT: DATA

Studies of genome-wide differentiation are accumulating and have focused on how many genomic regions are affected by adaptive divergence (Hohenlohe et al. 2010, Lawniczak et al. 2010, Gompert et al. 2012a, Jaquiery et al. 2012), how genomic divergence varies at different stages of the speciation process (Nadeau et al. 2012), what the relationship is between selection on adaptive phenotypes and genomic divergence (Roesti et al. 2012, Gompert et al. 2013), and whether divergence tends to be accentuated at protein-coding or regulatory regions of the genome (Jones et al. 2012b). Although these provide valuable insight into adaptation and speciation, explicit studies of how the history of gene flow influences the buildup versus maintenance of divergence are few, even for well-studied systems such as those outlined in **Table 1**. Thus, we focus below on a few key examples that demonstrate that the history of gene flow has been more thoroughly considered.

8.1. Primary Divergence

One example where speciation with continuous gene flow appears likely is that of different aposematic races of *Heliconius* butterflies (Quek et al. 2010). Studies of divergence at a fine genomic scale between races of *Heliconius melpomene* have shown that divergence is localized to two regions containing genetic variation that has a large effect on adaptive aposematic color (*Heliconius* Genome Consort. 2012, Nadeau et al. 2012). Such divergence localized to two genomic regions that contain clusters of loci with large phenotypic effects is consistent with divergence in the face of ongoing gene flow. Although this example is highly suggestive of primary divergence with gene flow, alternative possibilities exist for divergence restricted to a few genomic clusters harboring color-pattern loci, such as genome-wide homogenization following secondary contact of all regions except those affecting warning coloration and chance or even evolved linkage of color pattern loci. In addition, divergence between host races versus sibling species of *Rhagoletis* fruit flies is also consistent with primary divergence. In this case,

sister sibling species attacking hawthorn versus flowering dogwood host plants generally display a similar, but elevated, pattern of divergence across different loci to the better-studied apple versus hawthorn host races of the species *Rhagoletis pomonella* (Powell et al. 2013). Sibling species therefore may represent host races "writ large". Notably, different populations of the flowering dogwood fly form a discrete genetic cluster distinct from *R. pomonella* across their range, whereas apple and hawthorn flies display divergence between local sympatric populations but not range-wide genetic clustering. These two different taxon pairs likely represent a difference between host races and species along the speciation-with-gene-flow continuum. Further studies in diverse systems where information on the historical context of gene flow is known are sorely needed.

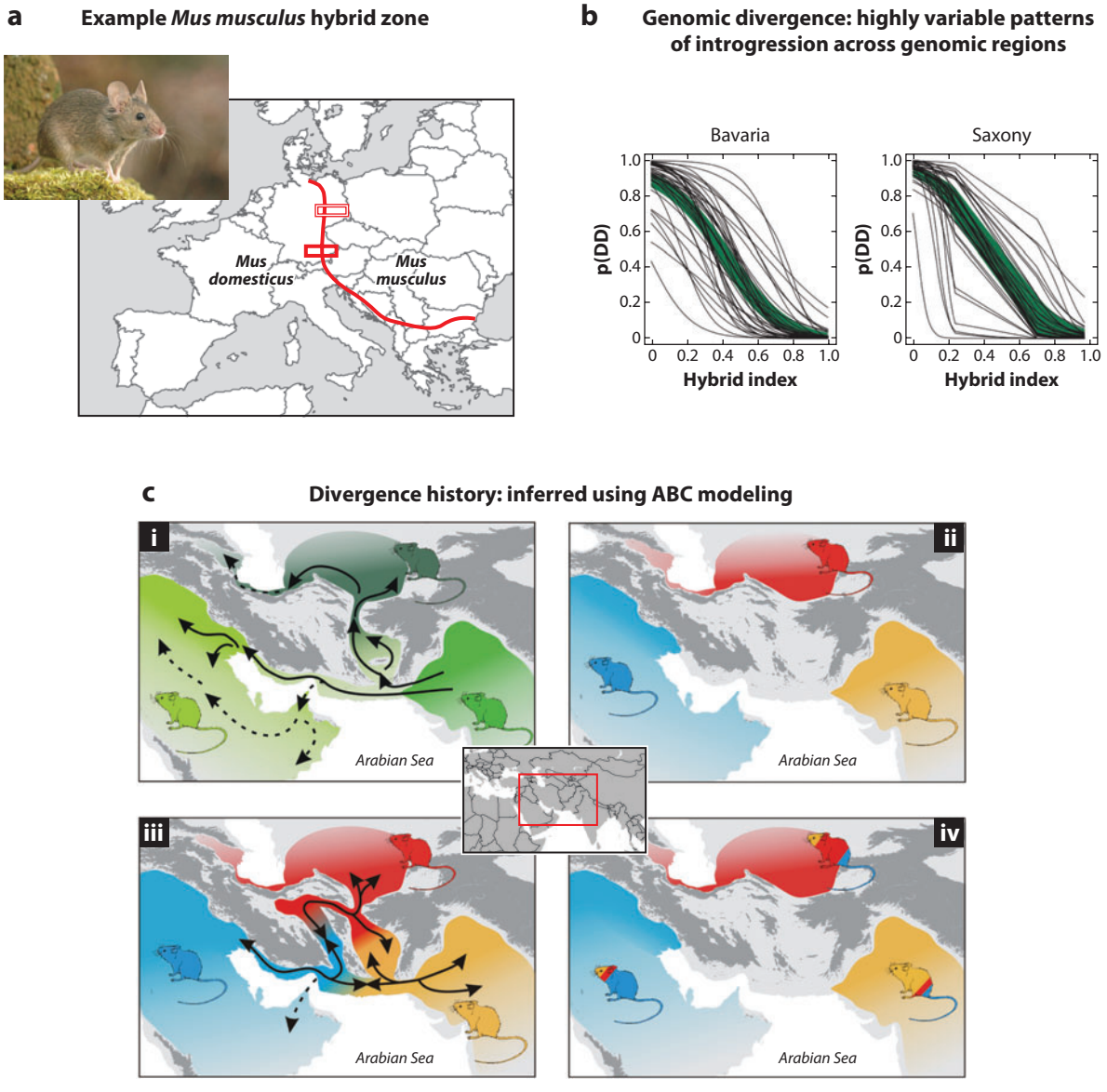
8.2. Allopatric Divergence with Secondary Contact

One system bearing on the issue of speciation-with-gene-flow following secondary contact is that of the North American lake whitefish (*Coregonus* sp.) (Rogers & Bernatchez 2007, Renaut et al. 2011, Gagnaire et al. 2013). In Canada, benthic normal and limnetic dwarf ecotypes of the fish repeatedly and independently differentiated in allopatry ~60,000 years ago across a series of glacially separated lakes. They came into secondary contact ~12,000 years ago as the ice sheets moved. The extent to which currently sympatric pairs of these ecotypes are phenotypically diverged varies among lakes, likely owing to differences in selective and gene flow regimes among lakes. Genome scans imply that the number of outlier genomic islands does not vary greatly among lakes, but the baseline level of genetic divergence and the size of regions of accentuated divergence increase with increasing phenotypic differentiation among lakes (Gagnaire et al. 2013). These patterns are consistent with widespread IBA across the genome due to GH as well as with clusters of particularly divergent regions, which are perhaps driven by DH. Most of the observed morphological divergence is believed to have evolved following secondary contact. However, intrinsic hybrid dysfunction, which evolves most readily in allopatry, also exists between ecotypes. Thus, questions remain regarding which stage of divergence the lake whitefish may have attained prior to secondary contact (i.e., one where GH was enabled upon secondary contact versus after the buildup of additional divergence in sympatry). The lake whitefish present an opportunity to study the consequences of variation in the degree of genetic divergence and RI established in allopatry on further differentiation following secondary contact.

In another recent study, Duvaux et al. (2011) explicitly tested alternate divergence scenarios between the house mouse species *Mus musculus domesticus* and *Mus musculus musculus*. Specifically, Duvaux et al. (2011) utilized sequence data from 57 loci and Approximate Bayesian Computation (ABC) to model support for each of four different divergence scenarios: (*a*) allopatric divergence, (*b*) divergence with ongoing and continuous gene flow, (*c*) an initial period of gene flow followed by allopatric divergence, and (*d*) an allopatric phase of divergence followed by secondary contact. Simulations strongly supported an allopatric phase of divergence followed by secondary contact and gene flow (Duvaux et al. 2011). In line with the predictions outlined above, it was suggested that postzygotic barriers between these subspecies of mice evolved during the relatively extended period of allopatric divergence observed in the system (**Figure 4**). Other studies that have described genome-wide patterns of divergence have frequently reported that divergence is accentuated at regions of the genome with reduced recombination, such as within inversions or near centromeres (Noor et al. 2001, Feder et al. 2003a, Machado et al. 2007, Hoffmann & Rieseberg 2008, Kirkpatrick 2010, Lowry & Willis 2010, Michel et al. 2010, Joron et al. 2011, Nachman

& Payseur 2012, Powell et al. 2013). An important question remains as to whether inversions, or more generally, regions of suppressed recombination, act as seeds to facilitate the buildup of adaptive divergence as populations experience gene flow or whether they mainly function to maintain divergence that built up during an allopatric phase of speciation upon secondary contact.

To better distinguish the roles of primary versus secondary contact in speciation-with-gene-flow we recommend studies that adopt explicit modeling approaches and, when possible, comparison of genomic patterns for populations in allopatry to those undergoing gene flow. Experiments mimicking secondary contact of lineages with known but variable patterns of



genomic divergence are also likely to be informative. In some instances, anthropogenic disturbances can create "natural experiments" of such secondary contact and speciation reversal (Seehausen et al. 1997, 2008). From such experiments, natural or otherwise, parameters like selection, gross migration, etc., might be directly measured or inferred and combined with theoretical expectations to help interpret patterns of genomic divergence among populations.

9. CONCLUSIONS

Major issues in the study of speciation-with-gene-flow, whether divergence occurs via a primary or secondary contact mode, are determining (*a*) whether and how genome architecture and different forms of hitchhiking aid in the evolution of differentiation and RI and (*b*) at which phases of divergence genome architecture and hitchhiking are most important. We have discussed how the circumstances for secondary contact are sometimes similar to sympatric divergence as well as different in other ways (e.g., more DM incompatibilities, meiotic drive, etc., when an allopatric period occurs). Generally speaking, theory regarding the genomics of primary speciation-with-gene-flow will apply with respect to loci involved in divergent ecological adaptation following secondary contact. It remains to be determined how congruent the results are for cases of DM incompatibilities, meiotic drive, and reinforcement. Analyses of hybrid zones have previously recognized that populations can reach a point in which the genome begins to congeal as barriers to gene flow couple. We have highlighted that this may be an important transition for speciation-with-gene-flow. However, much more theoretical and empirical work is needed in this area: How similar are the genomic patterns produced by these different processes? Is it possible to derive probability statements about which processes are more likely from a given pattern? Existing studies point to some intriguing similarities, including a potentially unifying view of species and speciation-with-gene-flow based on the processes generating RI as divergence proceeds. With the development of better analytical frameworks to deal with massive NGS data sets (Strasburg & Rieseberg 2013), and with increased attention to details of natural history and well-designed experimental manipulations, answers to these issues may emerge to yield a genomic synthesis of our understanding of population divergence and speciation.

Figure 4

Patterns of introgression and divergence history in *Mus* species and subspecies. (*a*) Hybrid zone (depicted by *red line*) used to study patterns of introgression between species of *Mus* in central Europe across two replicated transects. The outlined box shows the location of the transect in Saxony, and the solid box shows the location of the transect in Bavaria. (*b*) Patterns of introgression within the hybrid zone vary widely across the genome. The x-axis represents the genome-wide hybrid index of individuals and the y-axis the probability of homozygous *Mus domesticus* ancestry (the relationship between these variables is termed a "genomic cline"). The green shaded areas represent the expected probability of homozygous ancestry given the hybrid index (expected genomic clines with 95% CI; CI = confidence interval). Black lines represent genomic clines for individual loci; clines falling outside the expected 95% CI represent loci with atypical patterns of introgression. Panels *a* and *b* adapted from Teeter et al. (2010) and reprinted with permission from Blackwell Publishing. (*c*) Depiction of the divergence history of *Mus musculus* subspecies as inferred using Approximate Bayesian Computation (ABC) analyses. White areas indicate water, light gray indicates elevations up to 1,500 meters, and dark gray indicates higher elevations. Subpanels illustrate (*i*) initial colonization of geographic regions, (*ii*) subsequent genomic divergence during periods of isolation (represented by different colors), (*iii*) introgression between genomic backgrounds occurring during periods of secondary contact, and (*iv*) cycles of additional repeated allopatry during glacial maxima resulting in partially admixed genomes (illustrated by mice drawn with a mixture of colors). Adapted from Duvaux et al. (2011) and reprinted with permission from Blackwell Publishing.

SUMMARY POINTS

1. The geographic context of speciation can have important consequences for inferring process from observed patterns of genomic differentiation. Genome scans are often interpreted assuming that migration is ongoing between populations, an assumption that is usually not independently verified. Thus, it is important to confirm whether divergence indeed occurred in the face of gene flow or was purely allopatric.
2. Several processes can promote genetic divergence during speciation with gene flow, including selection acting directly on a locus and different forms of genetic hitchhiking. Divergence hitchhiking (DH) involves the physical linkage of regions affected by selection, whereas genome hitchhiking (GH) does not. A four-phase model conceptualizes changes in the relative importance of these processes as speciation proceeds.
3. Theory and data point to an important role for GH in speciation-with-gene-flow, but key evidence is still missing.
4. In addition to geography, correct interpretation of genome scans also requires information on whether gene flow was continuous during the divergence process or occurred following secondary contact after an initial period of allopatric divergence without gene flow.

FUTURE ISSUES

1. Additional studies are needed to compare patterns of genomic divergence for populations that are known to have diverged with versus without gene flow. Ideally, these empirical studies should be grounded by more predictive theoretical models distinguishing expectations for divergence with versus without gene flow.
2. Discerning the roles of direct selection and different forms of genetic hitchhiking on genomic divergence will be aided by moving beyond purely observational genome scans toward integrative studies that combine ecological, mapping, transplant, experimental, and sequencing approaches.
3. Studies of single taxon pairs at one point in the speciation process must be extended to multiple, closely related pairs spanning the speciation continuum. This will allow inferences to be made on how genomic divergence and RI build through time.
4. Better methods are needed for distinguishing between primary and secondary contact.
5. Although theory regarding the genomics of primary speciation-with-gene-flow should generally apply to loci experiencing divergent ecological selection following secondary contact, it remains to be determined how congruent the results are for cases of DM incompatibilities, meiotic drive, and reinforcement.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors would like to thank their mentors, students, and colleagues for helping us through the years to develop many of the thoughts and ideas expressed in this review. J.L.F. was funded, in part, with support from the National Science Foundation (NSF) and the US Department of Agriculture. S.P.E. was funded by the NSF and the Univ. of Notre Dame. P.N. was funded by a European Research Council (ERC) starter grant.

LITERATURE CITED

Abbott R, Alback D, Ansell S, Arntzen J, Baird S, et al. 2013. Hybridization and speciation. *J. Evol. Biol.* 26:229–46

Agrawal AA, Feder JL, Nosil P. 2011. Ecological divergence and the evolution of intrinsic postmating isolation with gene flow. *Int. J. Ecol.* 2011:1–15

Albert AYK, Sawaya S, Vines TH, Knecht AK, Miller CT, et al. 2008. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution* 62:76–85

Ayala D, Guerrero RF, Kirkpatrick M. 2013. Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito. *Evolution* 67:946–58

Barluenga M, Stolting KN, Salzburger W, Muschick M, Meyer A. 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439:719–23

Barrett RDH, Hoekstra HE. 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat. Rev. Genet.* 12:767–80

Barrett RDH, Rogers SM, Schluter D. 2008. Natural selection on a major armor gene in threespine stickleback. *Science* 322:255–57

Barrett RDH, Schluter D. 2008. Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23:38–44

Barton N, Bengtsson BO. 1986. The barrier to genetic exchange between hybridizing populations. *Heredity* 57:357–76

Barton NH. 1979. Dynamics of hybrid zones. *Heredity* 43:341–59

Barton NH. 1983. Multilocus clines. *Evolution* 37:454–71

Barton NH. 2000. Genetic hitchhiking. *Philos. Trans. R. Soc. Lond. Ser. B* 355:1553–62

Barton NH. 2001. The role of hybridization in evolution. *Mol. Ecol.* 10:551–68

Baxter SW, Nadeau N, Maroja L, Wilkinson P, Counterman BA, et al. 2010. Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in the *Heliconius melpomene* clade. *PLoS Genet.* 6:e1000794

Bengtsson BO, ed. 1985. *The Flow of Genes Through a Genetic Barrier*. Cambridge, UK: Cambridge Univ. Press

Bierne N, Welch J, Loire E, Bonhomme F, David P. 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol. Ecol.* 20:2044–72

Bolnick DI, Fitzpatrick BM. 2007. Sympatric speciation: models and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.* 38:459–87

Burt A, Trivers R. 2006. *Genes in Conflict: The Biology of Selfish Genetic Elements*. Cambridge, MA: Harvard Univ. Press

Butlin RK. 2010. Population genomics and speciation. *Genetica* 138:409–18

Cain ML, Andreasen V, Howard DJ. 1999. Reinforcing selection is effective under a relatively broad set of conditions in a mosaic hybrid zone. *Evolution* 53:1343–53

Campbell D, Bernatchez L. 2004. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Mol. Biol. Evol.* 21:945–56

Chan YF, Marks ME, Jones FC, Villarreal G, Shapiro MD, et al. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. *Science* 327:302–5

Charlesworth B, Nordborg M, Charlesworth D. 1997. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. *Genet. Res.* 70:155–74

Colosimo PF, Peichel CL, Nereng K, Blackman BK, Shapiro MD, et al. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol.* 2:635–41

Coluzzi M, Sabatini A, Petrarca V, Dideco MA. 1979. Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans. R. Soc. Trop. Med. Hyg.* 73:483–97

Counterman BA, Araujo-Perez F, Hines H, Baxter SW, Morrison CM, et al. 2010. Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in *Heliconius erato*. *PLoS Genet.* 6:e1000796

Coyne JA, Orr HA. 2004. *Speciation*. Sunderland, MA: Sinauer

Crespi BJ, Nosil P. 2012. Conflictual speciation: species formation by intragenomic conflict. *Trends Ecol. Evol.* 28:48–57

Cutter AD, Choi JY. 2010. Natural selection shapes nucleotide polymorphism across the genome of the nematode *Caenorhabditis briggsae*. *Genome Res.* 20:1103–11

Darwin C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London, UK: John Murray

Deagle BE, Jones FC, Chan YF, Absher DB, Kingsley DM, Reimchen TE. 2012. Population genomics of parallel phenotypic evolution in stickleback across stream-lake ecological transitions. *Proc. R. Soc. Lond. Ser. B* 297:1277–86

Dieckmann U, Doebeli M. 1999. On the origin of species by sympatric speciation. *Nature* 400:354–57

Duvaux L, Belkhir K, Boulesteix M, Boursot P. 2011. Isolation and gene flow: inferring the speciation history of European house mice. *Mol. Ecol.* 20:5248–64

Ellegren H, Smeds L, Burri R, Olason PI, Backström N, et al. 2012. The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* 491:756–60

Endler JA. 1977. *Geographic Variation, Speciation, and Clines*. Princeton, NJ: Princeton Univ. Press

Feder JL, Berlocher SH, Roethele JB, Dambroski H, Smith JJ, et al. 2003a. Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. *Proc. Natl. Acad. Sci. USA* 100:10314–19

Feder JL, Egan SP, Nosil P. 2012a. The genomics of speciation-with-gene-flow. *Trends Genet.* 28:342–50

Feder JL, Flaxman SM, Egan SP, Nosil P. 2013. Hybridization and the build-up of genomic divergence during speciation. *J. Evol. Biol.* 26:261–66

Feder JL, Geiji R, Powell THQ, Nosil P. 2011. Adaptive chromosomal divergence driven by mixed geographic mode of evolution. *Evolution* 65:2157–70

Feder JL, Geiji R, Yeaman S, Nosil P. 2012b. Establishment of new mutations under divergence and genome hitchhiking. *Philos. Trans. R. Soc. Lond. Ser. B* 367:461–74

Explores the idea that reproductive isolation should not just be in inverted regions of the genome, but elsewhere, as well.

Feder JL, Nosil P. 2009. Chromosomal inversions and species differences: When are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions? *Evolution* 63:3061–75

Examines how divergence hitchhiking generates differentiation, but conditions are limited. With several loci under selection, genome-wide divergence accrues.

Feder JL, Nosil P. 2010. The efficacy of divergence hitchhiking in generating genomic islands during ecological speciation. *Evolution* 64:1729–47

Feder JL, Roethele FB, Filchak K, Niedbalski J, Romero-Severson J. 2003b. Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella*. *Genetics* 163:939–53

Felsenstein J. 1976. Theoretical population genetics of variable selection and migration. *Annu. Rev. Genet.* 10:253–80

Provides classic theoretical work describing the antagonism between selection and recombination during speciation-with-gene-flow.

Felsenstein J. 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* 35:124–38

Fitzpatrick BM, Fordyce JA, Gavrilets S. 2008. What, if anything, is sympatric speciation? *J. Evol. Biol.* 21:1452–59

Flaxman SM, Feder JL, Nosil P. 2013. Genetic hitchhiking and the dynamic buildup of genomic divergence during speciation-with-gene-flow. *Evolution* 67:2577–91

Flaxman SM, Feder JL, Nosil P. 2012. Spatially explicit models of divergence and genome hitchhiking. *J. Evol. Biol.* 25:2633–50

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek A. 2011. A map of local adaptation in *Arabidopsis thaliana*. *Science* 333:86–89

Funk DJ, Egan SP, Nosil P. 2011. Isolation by adaptation in *Neochlamisus* leaf beetles: host-related selection promotes neutral genomic divergence. *Mol. Ecol.* 20:4671–82

Gagnaire P-A, Pavey SA, Normandeau E, Bernatchez L. 2013. The genomic architecture of reproductive isolation across the speciation continuum in lake whitefish species pairs assessed by RAD-sequencing. *Evolution* 67:2483–97

Gavrilets S. 2003. Perspective: Models of speciation: What have we learned in 40 years? *Evolution* 57:2197–215

Gavrilets S. 2004. *Fitness Landscapes and the Origin of Species*. Princeton, NJ: Princeton Univ. Press

Explores many of the theoretical issues about speciation discussed in this review.

Gavrilets S, Cruzan MB. 1998. Neutral gene flow across single locus clines. *Evolution* 52:1277–84

Gompert Z, Buerkle CA. 2009. A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Mol. Ecol.* 18:1207–24

Gompert Z, Buerkle CA. 2011. Bayesian estimation of genomic clines. *Mol. Ecol.* 20:2111–27

Gompert Z, Lucas LK, Nice CC, Buerkle CA. 2013. Genome divergence and the genetic architecture of barriers to gene flow between *Lycaeides idas* and *L. melissa*. *Evolution* 67:2498–514

Gompert Z, Lucas LK, Nice CC, Fordyce JA, Forister ML, Buerkle CA. 2012a. Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution* 66:2167–81

Gompert Z, Parchman TL, Buerkle CA. 2012b. Genomics of isolation in hybrids. *Philos. Trans. R. Soc. Lond. Ser. B* 367:439–50

Guerrero RF, Fousset F, Kirkpatrick M. 2012. Coalescent patterns for chromosomal inversions in divergent populations. *Philos. Trans. R. Soc. Lond. Ser. B* 367:430–38

Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, et al. 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. *Science* 333:83–86

Hannah MA, Wiese D, Freund S, Fiehn O, Heyer AG, Hincha DK. 2006. Natural genetic variation of freezing tolerance in *Arabidopsis*. *Plant Physiol.* 142:98–112

Harrison RG. 1991. Molecular changes at speciation. *Annu. Rev. Ecol. Syst.* 22:281–308

Hawthorne DJ, Via S. 2001. Genetic linkage of ecological specialization and reproductive isolation in pea aphids. *Nature* 412:904–7

Heliconius Genome Consort. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487:94–98

Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP. 2000. Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science* 290:516–18

Hoffmann AA, Rieseberg LH. 2008. Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annu. Rev. Ecol. Evol. Syst.* 39:21–42

Hohenlohe PA, Bassham S, Currey M, Cresko WA. 2012. Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philos. Trans. R. Soc. Lond. Ser. B* 367:395–408

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PloS Genet.* 6:e1000862

Discusses the genome scan of the threespine stickleback revealing genomic regions subject to divergent selection associated with speciation.

Holt RD, Barfield M. 2011. Theoretical perspectives on the statics and dynamics of species' borders in patchy environments. *Am. Nat.* 178:S6–25

Hudson ME. 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Mol. Ecol. Res.* 8:3–17

Jaquiery J, Stoeckel S, Nouhaud P, Mieuzet L, Maheo L, et al. 2012. Genome scans reveal candidate regions involved in the adaptation to host plant in the pea aphid complex. *Mol. Ecol.* 21:5251–64

Jones FC, Chan YF, Schmutz J, Grimwood J, Brady SD, et al. 2012a. A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Curr. Biol.* 22:83–90

Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, et al. 2012b. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61

Joron M, Frezal L, Jones RT, Chamberlain NL, Lee SF, et al. 2011. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* 477:203–6

Joron M, Papa R, Beltran M, Chamberlain N, Mavarez J, et al. 2006. A conserved supergene locus controls colour pattern diversity in *Heliconius* butterflies. *PloS Biol.* 4:1831–40

Kirkpatrick M. 2010. How and why chromosome inversions evolve. *PloS Biol.* 8:e1000501

Discusses theory on how inversions may establish within populations due to selection for reduced recombination.

Kirkpatrick M, Barton N. 2006. Chromosome inversions, local adaptation and speciation. *Genetics* 173:419–34

Kirkpatrick M, Ravigné V. 2002. Speciation by natural and sexual selection: models and experiments. *Am. Nat.* 159:S22–35

Kitano J, Lema SC, Luckenbach JA, Mori S, Kawagishi Y, et al. 2010. Adaptive divergence in the thyroid hormone signaling pathway in the stickleback radiation. *Curr. Biol.* 20:2124–30

Kobayashi Y, Telschow A. 2011. The concept of effective recombination rate and its application in speciation theory. *Evolution* 65:617–28

Korves TM, Schmid KJ, Caicedo AL, Mays C, Stinchcombe JR, et al. 2007. Fitness effects associated with the major flowering time gene FRIGIDA in *Arabidopsis thaliana* in the field. *Am. Nat.* 169:E141–57

Lasky JR, Des Marais DL, McKay JK, Richards JH, Juenger TE, Keitt TH. 2012. Characterizing genomic variation of *Arabidopsis thaliana*: the roles of geography and climate. *Mol. Ecol.* 21:5512–29

Lawniczak MKN, Emrich SJ, Holloway AK, Regier AP, Olson M, et al. 2010. Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. *Science* 330:512–14

Le Corre V. 2005. Variation at two flowering time genes within and among populations of *Arabidopsis thaliana*: comparison with markers and traits. *Mol. Ecol.* 14:4181–92

Li YF, Costello JC, Holloway AK, Hahn M. 2008. "Reverse ecology" and the power of population genomics. *Evolution* 62:2984–94

Shows that chromosomal inversions play an important role in adaptation and speciation in nature.

Lowry DB, Willis JH. 2010. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PloS Biol.* 8:e1000500

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. 2003. The power and promise of population genomics: From genotyping to genome typing. *Nat. Rev. Genet.* 4:981–94

Machado CA, Haselkorn TS, Noor MAF. 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* 175:1289–306

Malek TB, Boughman JW, Dworkin I, Peichel CL. 2012. Admixture mapping of male nuptial colour and body shape in a recently formed hybrid population of threespine stickleback. *Mol. Ecol.* 21:5265–79

Mallet J, Meyer A, Nosil P, Feder JL. 2009. Space, sympatry and speciation. *J. Evol. Biol.* 22:2332–41

Mayr E. 1963. *Animal Species and Evolution*. Harvard, MA: Harvard Univ. Press

Provides a counter example to island view showing widespread genomic divergence between ecological insect host races.

Michel AP, Sim S, Powell THQ, Taylor MS, Nosil P, Feder JL. 2010. Widespread genomic divergence during sympatric speciation. *Proc. Natl. Acad. Sci. USA* 107:9724–29

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* 441:947–52

Nachman MW, Payseur BA. 2012. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos. Trans. R. Soc. Lond. Ser. B* 367:409–21

Nadeau NJ, Whibley A, Jones RT, Davey JW, Dasmahapatra KK, et al. 2012. Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. *Philos. Trans. R. Soc. Lond. Ser. B* 367:343–53

Navarro A, Barton NH. 2003. Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution* 57:447–59

Neafsey DE, Lawniczak MKN, Park DJ, Redmond SN, Coulibaly MB, et al. 2010. SNP genotyping defines complex gene-flow boundaries among African malaria vector mosquitoes. *Science* 330:514–17

Niemiller ML, Fitzpatrick BM, Miller BT. 2008. Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene genealogies. *Mol. Ecol.* 17:2258–75

Niemiller ML, Nosil P, Fitzpatrick BM. 2010. Recent divergence-with-gene-flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene genealogies. *Mol. Ecol.* 19:1513–14

Noor MAF, Bennett SM. 2009. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity* 103:439–44

Provides one of the first papers to argue for an important role of inversions in speciation.

Noor MAF, Grams KL, Bertucci LA, Reiland J. 2001. Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. USA* 98:12084–88

Nosil P. 2012. *Ecological Speciation*. Oxford, UK: Oxford Univ. Press

Nosil P, Egan SP, Funk DJ. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: "isolation by adaptation" and multiple roles for divergent selection. *Evolution* 62:316–36

Nosil P, Flaxman S. 2011. Conditions for mutation-order speciation. *Proc. R. Soc. Lond. Ser. B* 278:399–407

Nosil P, Funk DJ, Ortíz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* 18:375–402

Nosil P, Gompert Z, Farkas TE, Comeault AA, Feder JL, et al. 2012a. Genomic consequences of multiple speciation processes in a stick insect. *Proc. R. Soc. London Ser. B* 279:5058–65

Nosil P, Parchman TL, Feder JL, Gompert Z. 2012b. Do highly divergent loci reside in genomic regions affecting reproductive isolation? A test using next-generation sequence data in *Timema* stick insects. *BMC Evol. Biol.* 12:164

Nosil P, Schluter D. 2011. The genes underlying the process of speciation. *Trends Ecol. Evol.* 26:160–67

Orr HA, Masly JP, Presgraves DC. 2004. Speciation genes. *Curr. Opin. Genet. Dev.* 14:675–79

Peichel CL, Nereng KS, Ohgi KA, Cole BLE, Colosimo PF, et al. 2001. The genetic architecture of divergence between threespine stickleback species. *Nature* 414:901–5

Powell THQ, Hood GR, Murphy MO, Heilveil JS, Berlocher SH, et al. 2013. Genetic divergence along the speciation continuum: the transition from host races to species in *Rhagoletis* (Diptera: Tephritidae). *Evolution* 67:2561–76

Presgraves DC 2007. Speciation genetics: epistasis, conflict and the origin of species. *Curr. Biol.* 17:R125–27

Quek SP, Counterman BA, De Moura P, Cardoso MZ, Marshall CR, et al. 2010. Dissecting comimetic radiations in *Heliconius* reveals divergent histories of convergent butterflies. *Proc. Natl. Acad. Sci. USA* 107:7365–70

Reed RD, Papa R, Martin A, Hines HM, Counterman BA, et al. 2011. *Optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* 333:1137–41

Reidenbach KR, Neafsey DE, Costantini C, Sagnon N, Simard F, et al. 2012. Patterns of genomic differentiation between ecologically differentiated M and S forms of *Anopheles gambiae* in West and Central Africa. *Genome Biol. Evol.* 4:1202–12

Renaut S, Maillet N, Normandeau E, Sauvage C, Derome N, et al. 2011. Genome-wide patterns of divergence during speciation: the lake whitefish case study. *Philos. Trans. R. Soc. Lond. Ser. B* 367:343–53

Rieseberg LH. 2001. Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* 16:351–58

Discusses how inversions are important for suppressing recombination among selected genes.

Rieseberg LH, Blackman BK. 2010. Speciation genes in plants. *Ann. Bot.* 106:439–55

Roesti M, Hendry AP, Salzburger W, Berner D. 2012. 2012. Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. *Mol. Ecol.* 21:2852–62

Rogers SM, Bernatchez L. 2006. The genetic basis of intrinsic and extrinsic postzygotic reproductive isolation jointly promoting speciation in the lake whitefish species complex (*Coregonus clupeaformis*). *J. Evol. Biol.* 19:1979–94

Rogers SM, Bernatchez L. 2007. The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonas sp* Salmonidae) species pairs. *Mol. Biol. Evol.* 24:1423–38

Savolainen V, Anstett MC, Lexer C, Hutton I, Clarkson JJ, et al. 2006. Sympatric speciation in palms on an oceanic island. *Nature* 441:210–13

Schluter D. 2009. Evidence for ecological speciation and its alternative. *Science* 323:737–41

Seehausen O, Takimoto G, Roy D, Jokela J. 2008. Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Mol. Ecol.* 17:30–44

Seehausen O, Vanalphen JJM, Witte F. 1997. Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* 277:1808–11

Smadja CM, Butlin RK. 2011. A framework for comparing processes of speciation in the presence of gene flow. *Mol. Ecol.* 20:5123–40

Smadja CM, Canbäck B, Vitalis R, Gautier M, Ferrari J, et al. 2012. Large-scale candidate gene scan reveals role of chemoreceptor genes in host plant specialization and speciation in the pea aphid. *Evolution* 66:2723–38

Stinchcombe JR, Hoekstra HE. 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100:158–70

Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, et al. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene FRIGIDA. *Proc. Natl. Acad. Sci. USA* 101:4712–17

Strasburg JL, Rieseberg LH. 2008. Molecular demographic history of the annual sunflowers *Helianthus annuus* and *H. petiolaris*—large effective population sizes and rates of long-term gene flow. *Evolution* 62:1936–50

Strasburg JL, Rieseberg LH. 2010. How robust are "isolation with migration" analyses to violations of the IM model? A simulation study. *Mol. Biol. Evol.* 27:297–310

Strasburg JL, Rieseberg LH. 2011. Interpreting the estimated timing of migration events between hybridizing species. *Mol. Ecol.* 20:2353–66

Strasburg JL, Rieseberg LH. 2013. Methodological challenges to realizing the potential of hybridization research. *J. Evol. Biol.* 26:259–60

Strasburg JL, Scotti-Saintagne C, Scotti I, Lai Z, Rieseberg LH. 2009. Genomic patterns of adaptive divergence between chromosomally differentiated sunflower species. *Mol. Biol. Evol.* 26:1341–55

Strasburg JL, Sherman NA, Wright KM, Moyle LC, Willis JH, Rieseberg LH. 2012. What can patterns of differentiation across plant genomes tell us about adaptation and speciation? *Philos. Trans. R. Soc. Lond. Ser. B* 367:364–73

Teeter KC, Thibodeau LM, Gompert Z, Buerkle SE, Nachman MW, Tucker PK. 2010. The variable genomic architecture of isolation between hybridizing species of house mice. *Evolution* 64:472–85

Thibert-Plante X, Hendry AP. 2010. When can ecological speciation be detected with neutral loci? *Mol. Ecol.* 19:2301–14

Turelli M, Barton NH, Coyne JA. 2001. Theory and speciation. *Trends Ecol. Evol.* 16:330–43

Turner JRG. 1967. Why does the genotype not congeal? *Evolution* 21:645–56

Turner TL, Hahn MW. 2007. Locus- and population-specific selection and differentiation between incipient species of *Anopheles gambiae*. *Mol. Biol. Evol.* 24:2132–38

Turner TL, Hahn MW. 2010. Genomic islands of speciation or genomic islands and speciation? *Mol. Ecol.* 19:848–50

Turner TL, Hahn MW, Nuzhdin SV. 2005. Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biol.* 3:1572–78

Van Buskirk HA, Thomashow MF. 2006. *Arabidopsis* transcription factors regulating cold acclimation. *Physiol. Plant.* 126:72–80

Via S. 2001. Sympatric speciation in animals: the ugly duckling grows up. *Trends Ecol. Evol.* 16:381–90

Via S. 2009. Natural selection in action during speciation. *Proc. Natl. Acad. Sci. USA* 106:9939–46

Via S. 2012. Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Philos. Trans. R. Soc. Lond. Ser. B* 367:451–60

Via S, Conte G, Mason-Foley C, Mills K. 2012. Localizing F-ST outliers on a QTL map reveals evidence for large genomic regions of reduced gene exchange during speciation-with-gene-flow. *Mol. Ecol.* 21:5546–60

Via S, West J. 2008. The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Mol. Ecol.* 17:4334–45

Weetman D, Wilding CS, Steen K, Pinto J, Donnelly MJ. 2012. Gene flow-dependent genomic divergence between *Anopheles gambiae* M and S forms. *Mol. Biol. Evol.* 29:279–91

Weinig C, Dorn LA, Kane NC, German ZM, Hahdorsdottir SS, et al. 2003. Heterogeneous selection at specific loci in natural environments in *Arabidopsis thaliana*. *Genetics* 165:321–29

White BJ, Cheng C, Sangaré D, Lobo NF, Collins FH, Besansky NJ. 2009. The population genomics of trans-specific inversion polymorphisms in *Anopheles gambiae*. *Genetics* 183:275–88

White BJ, Cheng C, Simard F, Costantini C, Besansky NJ. 2010. Genetic association of physically unlinked islands of genomic divergence in incipient species of *Anopheles gambiae*. *Mol. Ecol.* 19:925–39

White BJ, Hahn MW, Pombi M, Cassone BJ, Lobo NF, et al. 2007. Localization of candidate regions maintaining a common polymorphic inversion (2La) in *Anopheles gambiae*. *PLoS Genet.* 3:2404–14

Wright KM, Lloyd D, Lowry DB, Macnair MR, Willis JH. 2013. Indirect evolution of hybrid lethality due to linkage with selected locus in *Mimulus guttatus*. *PLoS Biol.* 11:e1001497

Wu C. 2001. The genic view of the process of speciation. *J. Evol. Biol.* 14:851–65
Yeaman S, Otto SP. 2011. Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution* 65:2123–29
Yeaman S, Whitlock MC. 2011. The genetic architecture of adaptation under migration-selection balance. *Evolution* 65:1897–911
Zhivotovsky LA, Christiansen FB. 1995. The selection barrier between populations subject to stabilizing selection. *Evolution* 49:490–501

Annual Review of Ecology, Evolution, and Systematics

Volume 44, 2013

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Ecology, Evolution, and Systematics* articles may be found at http://ecolsys.annualreviews.org/errata.shtml