ANNUAL REVIEWS

Annual Review of Ecology, Evolution, and Systematics

Analysis of Population Genomic Data from Hybrid Zones

Zachariah Gompert,[1] Elizabeth G. Mandeville,[2] and C. Alex Buerkle[3]

[1]Department of Biology and Ecology Center, Utah State University, Logan, Utah 84322; email: zach.gompert@usu.edu

[2]Department of Botany and Wyoming Cooperative Fish and Wildlife Research Unit, University of Wyoming, Laramie, Wyoming 82071

[3]Department of Botany and Program in Ecology, University of Wyoming, Laramie, Wyoming 82071

Annu. Rev. Ecol. Evol. Syst. 2017. 48:207–29

First published online as a Review in Advance on August 16, 2017

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

https://doi.org/10.1146/annurev-ecolsys-110316-022652

Keywords

admixture, ancestry, clines, contingency, hybridization, introgression, linkage disequilibrium, selection, reproductive isolation, speciation

Abstract

Hybrid zones provide a powerful opportunity to analyze ecological and evolutionary interactions between divergent lineages. As such, research on hybrid zones has played a prominent role in the fields of evolutionary biology and systematics. Herein, we clarify what hybrid zones are, what is (and is not) known about them, and how different types of genomic data contribute to our understanding of hybrid zones. We then review two key topics, namely, what genomic analyses of hybrid zones have revealed about the basis and dynamics of speciation and how hybrid zones directly affect evolutionary processes. In the latter case, we emphasize the importance of contingency and ecological and genomic context in outcomes of hybridization. Throughout, we highlight limitations and key unknowns, and suggest approaches most likely to advance our understanding of hybrid zones and evolutionary processes in general.

1. INTRODUCTION

The dispersal of organisms and their gametes leads to the spread of genetic variants. This gene flow between populations is a fundamental evolutionary process that interacts with other processes (deterministic natural selection, stochastic genetic drift, etc.) to shape the evolutionary trajectory of populations. The exchange of genetic variants between diverged lineages is referred to as hybridization, and the geographic areas where hybridization is localized are hybrid zones (Barton & Hewitt 1985, Harrison 1993, Gompert & Buerkle 2016). Hybrid zones provide a powerful means to observe interactions between divergent gene pools and thus are particularly relevant for understanding the basis and process of speciation (Barton & Hewitt 1985, Hewitt 2001, Harrison & Larson 2016). This is not the same as the ample evidence for gene tree incongruence that might have arisen through hybridization and introgression (e.g., Mallet et al. 2016). Although they are useful, studies of gene tree incongruence lack the direct observation of hybridization and its context that is a key strength of studying hybrid zones.

Beyond providing a means to study evolutionary processes, hybrid zones are important evolutionary phenomena in their own right. Hybrid zones show that distinct lineages can persist despite ongoing gene flow, which is difficult to reconcile with strict interpretations of the biological species concept (Harrison 1990, Harrison & Larson 2014). Hybrid zones also have evolutionary consequences (Arnold 1997, Abbott et al. 2013). They can provide a conduit for adaptive alleles to spread between species, and they can drive reinforcement or the coupling of incompatibilities, thereby completing the speciation process (Servedio 2004, Barton & De Cara 2009, Fitzpatrick et al. 2009, Pardo-Diaz et al. 2012). Hybrid zones can collapse, expand, or become cut off from parental species, resulting in wide admixture zones, genetic swamping of one species by another, or the creation of hybrid species (Allendorf et al. 2001, Mallet 2007, Abbott et al. 2013). Thus, the long history of interest in hybrid zones is justified by both what one can learn from them and how they affect evolutionary processes (Arnold 1997, Harrison 1993, Abbott et al. 2013). With the growing accessibility of population genomic data, hybrid zone research has the potential to make rapid advances in both these areas.

Herein, we review the use of genomic data in the analysis of hybrid zones. We discuss specific analytical approaches in the context of key questions and topics in evolutionary biology. We highlight ways in which results and techniques developed over the past few decades have altered the questions that we can or should ask about hybrid zones. We begin by clarifying what is meant by hybrid zones and genomic data, and by describing the utility of genomic data for studying hybrid zones. We then discuss what has been (and can be) learned about adaptation and speciation from genomic analyses of hybrid zones (i.e., the use of hybrid zones as a tool) before turning to the evolutionary consequences of hybrid zones (i.e., hybrid zones as important evolutionary phenomena), where we focus our attention on recent results concerning variation and contingency in hybrid zone dynamics. We conclude by discussing key limitations of hybrid zone research and suggesting future research directions that we think could prove particularly insightful.

2. DEFINING HYBRID ZONES AND GENOMIC DATA

2.1. What are Hybrid Zones?

The seemingly simple question of what constitutes a hybrid zone has been answered several ways, ranging from broad definitions that include a variety of population histories and selective regimes to more restrictive, narrow definitions (Harrison 1993). The usage of the term hybrid zone necessarily implies hybridization, the genetic mixing of divergent groups (Barton & Hewitt 1985,

Harrison 1993), but the specification of hybridization as occurring within a zone also implicates a geographic region. Thus, from a spatial perspective, hybrid zones comprise sets of clines or gradients in trait values or allele frequencies. The geography of hybrid zones can involve wide or narrow regions of contact and hybridization. Additionally, the zone could be a region with a more idiosyncratic distribution of populations with hybrids, sometimes referred to as a mosaic hybrid zone. The principal distinction that is drawn is relative to comparatively rare, perhaps longer-distance dispersal that can lead to hybrids within the range of only a single parental taxon. Hybrid zones can also vary in the extent to which they are maintained as a result of fitness variation that is tied to environmental gradients, due to intrinsic reductions in hybrid fitness (tension zone model), or both (reviewed by Moore 1977, Barton & Hewitt 1985). Furthermore, hybrid zones vary in their spatial extent relative to dispersal, such that some are geographically wide enough that they mostly persist because of a lack of gene flow (e.g., *Lycaeides melissa* × *L. idas* butterflies; Gompert et al. 2014).

A second question related to the definition of hybrid zones is the extent of divergence that is necessary to refer to hybrids and, by extension, to hybrid zones. Not all sexually produced zygotes are hybrids, but instead hybrids are the result of wider crosses between divergent parents. The consequences of wider crosses are that hybrid zones will contain more heterozygous individuals and more loci with intermediate allele frequencies and with atypical linkage disequilibrium (LD) between loci (**Figure 1**). Consequently, LD, including between physically unlinked loci, is a distinguishing feature of hybrid zones (Szymura & Barton 1986, Baldassarre et al. 2014, Saarman & Pogson 2015). Whereas hybrids are also recognized between much less diverged genotypes, such as commercially and agronomically important crops (e.g., heterozygous F_1 maize hybrids between inbred parental lines) or nominal taxa that represent distinct phenotypic forms but exhibit minimal genome divergence (e.g., Poelstra et al. 2014, Toews et al. 2016), such crosses do not lead to the same population genetic consequences as the wider crosses that are responsible for hybrid zones.

2.2. How Do Hybrid Zones Arise, and How Do They Change?

Hybrid zones are mostly assumed to arise from secondary contact of diverged lineages, but they can also form in situ (primary divergence; Endler 1977, Barton & Hewitt 1985). Clear evidence for whether primary or secondary divergence has predominated in the history of a particular hybrid zone is rare in empirical studies (Barton & Hewitt 1985, Gompert & Buerkle 2016). Beyond the challenge of inferring evolutionary history from genomic data, it is very possible that primary divergence and secondary contact form a false dichotomy. Instead, the dynamics of individual hybrid zones may be shaped by both primary divergence and secondary contact, potentially with multiple episodes of each.

Hybrid zones could be relatively old and persist long enough for some or all of the genome composition in hybrids to approach an equilibrium between migration of alleles and selection among them (Barton & Hewitt 1985). Alternatively, hybrid zones can be recent, and all or some fraction of the genome could depart from migration-selection equilibrium expectations. This potentially includes many hybrid zones that have arisen as a result of translocation of organisms by human activity (e.g., Riley et al. 2003, Mandeville et al. 2015). Given the changes that species' ranges have undergone repeatedly during climate fluctuations and the dynamic nature of many habitats, it is likely that the composition of contemporary hybrid zones can be a mixture of loci and individuals that reflects recent and more ancient hybridization. Additionally, if the composition of a population lies away from a particular equilibrium between migration and selection, by the time a population has evolved close to the initial equilibrium, some aspect of the system might have shifted to affect migration, selection, or both.

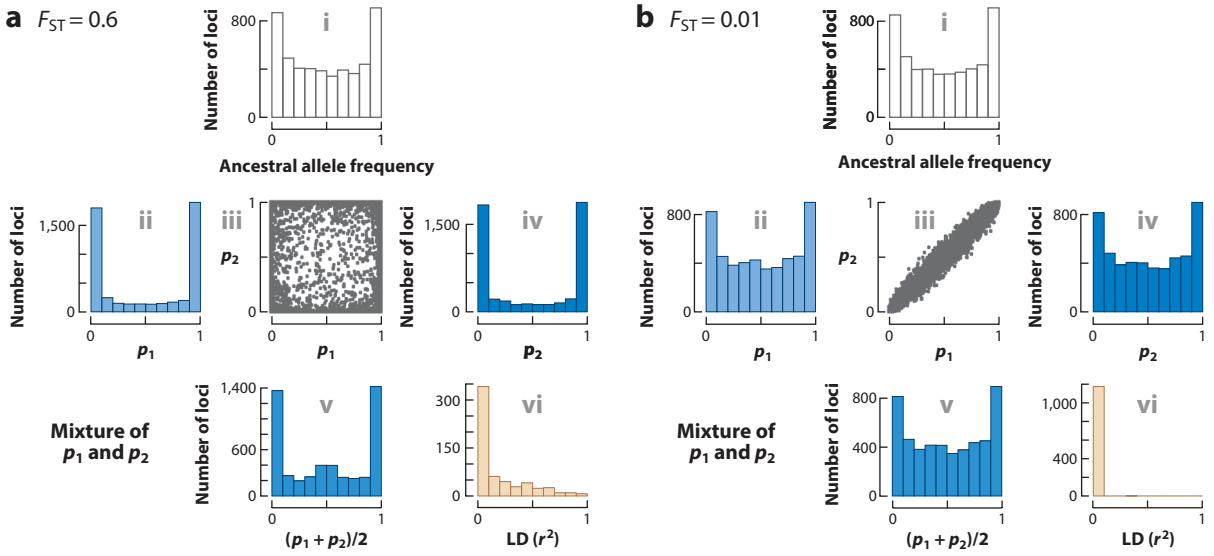

Figure 1

Plots of model expectations for the genetic composition of populations that are first generation mixtures of (*a*) strongly ($F_{ST} = 0.6$) and (*b*) weakly ($F_{ST} = 0.01$) diverged source populations, which can be taken as representative of mixtures that correspond to hybridization (panel *a*) and gene flow in a meta-population (panel *b*). Allele frequencies in the source populations p_1 and p_2 (*blue*, subpanels *ii* and *iv*) are a function of drift ($F_{ST} = 0.6$ or 0.01) from starting allele frequencies in an ancestral population (panels *i*) with $\pi \sim \text{Beta}(\alpha = 0.6, \beta = 0.6)$, with random draws for each locus in the source populations according to an *F*-model with $p \sim \text{Beta}[\alpha = \pi(-1 + 1/F_{ST}), \beta = (1 - \pi)(-1 + 1/F_{ST})]$. Allele frequencies for 5,000 loci in each pair of source populations are plotted (*blue*, subpanels *ii* and *iv*). Bivariate plots (subpanels *iii*) show the similarity of derived allele frequencies (p_1 and p_2) in the pair of populations following drift from the ancestral allele frequency for each locus. The population mixtures (subpanels *v* and *vi*) are assumed to have equal representation of both source populations, and *r* is the Pearson coefficient. Averaging between highly differentiated loci in the $F_{ST} = 0.6$ model leads to a small peak of intermediate allele frequencies that is absent in a mixture of weakly differentiated ($F_{ST} = 0.01$) sources (*blue*, subpanels *v*; this peak would be even more evident if loci with small minor allele frequencies were removed, e.g., <0.05, and only SNPs were analyzed). Likewise, pairwise LD is large between many loci in the mixture of source populations with $F_{ST} = 0.6$ and is absent from the mixture of populations with $F_{ST} = 0.01$ (subpanels *vi*; based on simulated genotypes at 50 loci for a mixture of 50 individuals from each source population). Abbreviations: LD, linkage disequilibrium; SNP, single nucleotide polymorphism.

Theory and data suggest that hybrid zones are dynamic and can traverse considerable space, even on the scale of decades (Barton & Hewitt 1985, Buggs 2007). Hybrid zones can shift in response to changes in the environment, or because of fitness differences among individuals or variation in population density or dispersal rates. Hybrid zones maintained by endogenous selection (i.e., tension zones) are expected to move toward areas of low population density, and they can be trapped by physical barriers to dispersal. Hybrid zone movement has been inferred from patterns of directionally biased introgression of neutral alleles (e.g., Wielstra et al. 2017) and from direct observation where clines have been measured in the same geographic region at two or more times (e.g., Dasmahapatra et al. 2002, Taylor et al. 2014). Evidence for hybrid zone stability also exists (e.g., Yanchukov et al. 2006). In general, hybrid zone movement provides a compelling example of evolution in action, but it also means that making inferences from hybrid zones assuming that they are stable can lead to erroneous conclusions.

Given this extensive variation in the processes operating in hybrid zones, methods for analysis of hybrid zones will vary in their applicability to particular cases in nature. Knowledge from outside of the hybrid zone, of biogeographic history, and of the biology of the organisms will shape the

HYBRID INDEXES AND ADMIXTURE PROPORTIONS

Hybrid indexes and admixture proportions measure the proportion of an individual's genome inherited from each of two (hybrid index), or two or more (admixture proportion) parental source populations. When genetic loci with fixed differences between hybridizing species are analyzed, hybrid indexes (or admixture proportions) can be estimated by simply counting the number of alleles from each taxon. Otherwise, statistical methods, such as maximum likelihood or Bayesian inference, must be used to account for uncertainty in the source of each allele. As an example, the Bayesian admixture model in the `structure` software estimates admixture proportions based on the product of a series of conditional probability statements: $\Pr(\mathbf{g}|\mathbf{z},\mathbf{p})\Pr(\mathbf{z}|\mathbf{q})\Pr(\mathbf{q},\mathbf{p}|\boldsymbol{\phi})\Pr(\boldsymbol{\phi})$, where $\Pr(\mathbf{g}|\mathbf{z},\mathbf{p})$ is the probability of observing a set of genotypes ($\mathbf{g}$) given the source of the alleles ($\mathbf{z}$) for each locus and the source population allele frequencies ($\mathbf{p}$), $\Pr(\mathbf{z}|\mathbf{q})$ denotes the probability of inheriting an allele from a given source population given admixture proportions ($\mathbf{q}$), and $\boldsymbol{\phi}$ is a set of parameters that the admixture proportions and allele frequencies depend on. This and related models are reviewed by Gompert & Buerkle (2013). Importantly, criticisms that models like `structure` and the resulting admixture proportions do not necessarily provide evidence of hybridization are not relevant in the analysis of known hybrid zones, where admixture proportions can be used to quantify patterns of admixture and introgression.

questions of interest and, along with other aspects of study design, affect the suitability of methods for analysis.

2.3. What Are Genomic Data, and What Can We Measure with Them?

A consensus on what constitutes genomic, rather than genetic, data is lacking, and the term is variously applied. When considering the analysis of hybrid zones, we think it is useful to delineate three contemporary categories of genomic data that differ in the information they provide. What we call type *i* genomic data include any large (i.e., thousands to hundreds of thousands) set of single nucleotide polymorphisms (SNPs) that is analyzed either without a reference genome or with a poor-quality, fragmented genome. This is typical for genotyping-by-sequencing or similar data sets (e.g., restriction site associated DNA sequencing, or RADseq) generated for organisms without prior genomic resources. Targeted enrichment sequence data (e.g., exomes), where only a reference transcriptome is available, fit in this category. Such data provide sufficient information to precisely estimate key measures of hybridity, such as hybrid indexes, admixture proportions, and measures of interspecific ancestry (see the sidebar Hybrid Indexes and Admixture Proportions), and to measure the average extent and variability of introgression (e.g., Parchman et al. 2013, Gompert et al. 2014, Mandeville et al. 2015).

Large sets of SNPs that have been aligned to a reference genome comprising large scaffolds or whole chromosomes (or where a high-resolution linkage map is used) constitute type *ii* genomic data (e.g., Nadeau et al. 2014, Maroja et al. 2015). Type *ii* genomic data provide information on ancestry blocks in hybrids and on how various population genetic signals, such as restricted introgression, vary along the genome and among chromosomes. Type *iii* genomic data are whole-genome resequence data aligned to a high-quality reference genome (e.g., Dasmahapatra et al. 2012, Burri et al. 2015). Type *iii* genomic data provide finer-scale resolution of patterns along chromosomes, including ancestry block boundaries, and reduce the probability that important population genetic signals are missed (elevated LD in hybrid zones might make the high-resolution information provided by whole-genome sequences less important than in other areas of population

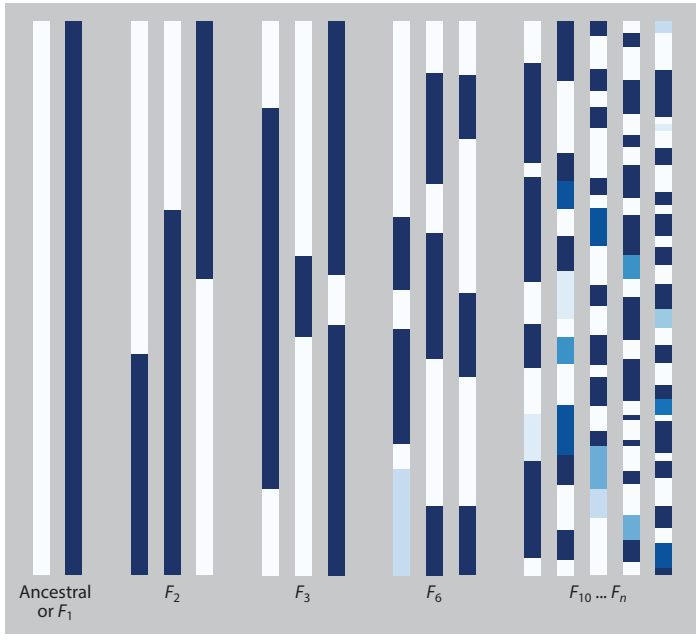

Figure 2

Plot of ancestry (*dark blue versus white*) along chromosomes with different numbers of generations of admixture ($F_1 \ldots F_n$) and a genetic map length of one Morgan. Recombination decreases block sizes from generation to generation. Intermediate shades of blue indicate ancestry blocks with different levels of certainty regarding ancestry, as would be encountered in a typical inference with some level of uncertainty about ancestry due to shared allelic states and intermediate allele frequencies.

genomics). Genomes can also be annotated with additional information, such as recombination rates or gene models, providing meaningful context for interpreting patterns in hybrid zones (e.g., Maroja et al. 2015).

Aside from an increase in the number of loci and genomic context, another key shift with genomic data is that, unlike molecular markers used in many early hybrid zone studies (Szymura & Barton 1986) or even contemporary studies using ascertained sets of SNPs (e.g., Janoušek et al. 2012, Larson et al. 2013), most polymorphisms identified from direct sequencing are not fixed for different alleles between hybridizing species. This means that sequenced variants provide a less biased view of introgression, but also that genotype is not fully indicative of the ancestry or source of an allele (that is, which of the parental species an allele was inherited from). Ancestry can be inferred from genotypes using probabilistic models, and ancestry (rather than state) can then be used to describe patterns of introgression (**Figure 2**; Gompert & Buerkle 2013). Reconstruction and analysis of ancestry blocks benefit from the fact that ancestry is autocorrelated along chromosomes, meaning that accurate inference of ancestry is possible even when some markers contain very little information about their origin (e.g., Tang et al. 2006, Paşaniuc et al. 2009, Wegmann et al. 2011). In the future, long-read sequencing methods (e.g., Huddleston et al. 2014), which provide direct observation of phase, could be used to further improve ancestry inference.

Genomic data are available because of advances in sequencing and genotyping that make it feasible and economical to sample substantial fractions of the genome and larger numbers of individuals. So, along with large numbers of SNPs, we can now sample many individuals across many locations. As a result, we have learned about multiple dimensions of hybridization (e.g.,

Gompert et al. 2014, Mandeville et al. 2015) that would have been difficult or impossible to resolve with smaller numbers of individuals, fewer sampling locations, and fewer ancestry-informative loci. For example, sparse spatial sampling along a cline can give the false impression of demic (rather than clinal) population genetic structure (Gompert & Buerkle 2016). Because of inherent trade-offs between sampling more individuals versus more of the genome, type *i* and *ii* genomic data will continue to be important for hybrid zone studies; conversely, whole-genome sequence data (type *iii* genomic data) from fewer individuals might be preferable for diagnosing historical introgression (Dasmahapatra et al. 2012, Sankararaman et al. 2014). With these considerations in mind, we turn to the use of hybrid zones as a tool to study evolutionary processes, particularly speciation.

3. WHAT CAN GENOMIC ANALYSES OF HYBRID ZONES TELL US ABOUT THE PATTERN AND PROCESS OF SPECIATION?

Natural hybridization tests the efficacy of inherent barriers to gene flow between differentiated populations or species. As such, hybrid zones can serve as natural laboratories to study speciation and to probe the nature of species and species boundaries (Barton & Hewitt 1989, Harrison 1990). The development of cline theory (see the sidebar Cline Theory) in the 1970s and 1980s provided an analytical framework that connected spatial clines to evolutionary processes, thereby generating great interest in hybrid zones (Bazykin 1969; Slatkin 1973; Endler 1977; Barton & Hewitt 1985, 1989). In a series of classic studies, cline theory was used to estimate the strength of selection on individual loci, the overall barrier to gene flow between species, and the number of loci contributing to reproductive isolation (Barton & Hewitt 1981; Szymura & Barton 1986, 1991; Mallet et al. 1990). As we have moved into the genomic era, there has been an increased emphasis on analyzing patterns of differential introgression among loci to identify genomic regions that contribute to speciation (i.e., barrier loci) or that introgress adaptively (Payseur 2010, Harrison & Larson 2016). The availability of genomic data also led to a series of new analytical developments (**Table 1**), including genomic cline methods that quantify locus-specific introgression relative to a genome-average admixture gradient (Lexer et al. 2007; Gompert & Buerkle 2009, 2011b; Fitzpatrick 2013) and methods that analyze ancestry tracts to reconstruct hybrid zone dynamics and identify putative barrier loci (e.g., Gravel 2012, Sedghifar et al. 2016). In this section we review recent empirical and theoretical results to address two sets of questions concerning genomic analyses of hybrid zones and the speciation process: (*a*) What do patterns of differential introgression tell us about evolutionary processes in general, and specifically about the genes and traits that contribute to reproductive isolation in hybrid zones, and (*b*) what do patterns of hybridization and introgression suggest about the dynamics of speciation and nature of species boundaries? We highlight ways in which population genomic data have been useful or necessary for addressing these questions. We focus on inferences from hybrid zones; other recent reviews have considered genomic studies of hybridization more generally (e.g., Payseur & Rieseberg 2016).

3.1. Patterns of Differential Introgression

Genomic studies of hybrid zones have frequently documented differential or variable patterns of introgression across the genome (**Figure 3**) (e.g., Hamilton et al. 2013, Parchman et al. 2013, Baldassarre et al. 2014, Janoušek et al. 2015, Kovach et al. 2016). This is consistent with the view of hybrid zones as semipermeable barriers to gene flow (Harrison & Larson 2014), though counterexamples with more concordant patterns of introgression exist (e.g., Zbawicka et al. 2014). As an example, genomic analyses of a mosaic hybrid zone between the field crickets *Gryllus pennsylvanicus*

CLINE THEORY

Cline models can be used to quantify patterns of gene flow in or across hybrid zones and to infer the strength of selection on a locus or sets of loci from these patterns (Endler 1977, Szymura & Barton 1986, Barton & Gale 1993). Various models where selection maintains a spatial cline in allele (or ancestry) frequencies predict a sigmoidal cline shape, such that $p_x = \frac{1}{2}(1 + \tanh(\frac{2(x-c)}{w}))$, where c is the cline center, w is the cline width, x is a spatial coordinate, and p_x is the allele (or ancestry) frequency at position x (Barton & Gale 1993, Payseur 2010). Cline width is determined by history and the balance between dispersal and selection. When clines are maintained by endogenous selection against hybrids (that is, selection arising from intrinsic incompatibilities), equilibrium cline width (w) is $\sqrt{8}l$, where $l \equiv \frac{\sigma}{\sqrt{s}}$ is referred to as the characteristic scale of selection, σ is the average dispersal distance (along a one-dimensional landscape), and s is the strength of selection (Barton & Gale 1993). Quantitatively similar results hold for other forms of selection. Thus, it is possible to infer the strength of selection from the width of a cline without knowing the form of selection. However, this approach necessitates an estimate of dispersal and assumes that an equilibrium between selection and dispersal (gene flow) holds, which can take many (e.g., thousands of) generations (Baird 1995, Gompert & Buerkle 2016) or might never occur if the environment is in flux. Neutral loci approach equilibrium very slowly, with cline width increasing indefinitely as a function of time (t) ($w = 2.51\sigma\sqrt{t}$; Barton & Gale 1993). When selection occurs on multiple loci, stepped clines are predicted, as LD among loci in the center of the cline creates a stronger barrier to gene flow and increases the selection experienced by individual loci (Barton 1983, Barton & Bengtsson 1986). In such cases, clinal patterns at the center (where LD is important) and edges (where LD is less important) of the hybrid zones can be analyzed separately to estimate direct selection on causal variants and the overall barrier to gene flow (Szymura & Barton 1986, Barton & Gale 1993).

Genomic cline models, such as Barton's concordance analysis (Szymura & Barton 1986, Macholán et al. 2011) and Bayesian genomic cline methods (Gompert & Buerkle 2011a, Gompert et al. 2012b), consider introgression along an admixture gradient and thus can be applied in hybrid zones without a clear spatial axis, such as mosaic hybrid zones (Harrison 1986, Rand & Harrison 1989). With these methods, allele or ancestry frequencies are modeled as a function of the average allele frequency or hybrid index. Two parameters, α and β, describe the departure of introgression for individual loci from the genome average. When a clear spatial axis does exist, α is twice the shift in cline center and β is the increase or decrease in cline width relative to the average (Szymura & Barton 1986). Alternative genomic cline models have been proposed with different functional forms (e.g., a logit-logistic function; Fitzpatrick 2013).

and *G. firmus* detected considerable variation in introgression among genetic markers, with nucleotide polymorphisms on the X chromosome exhibiting particularly low levels of introgression (Larson et al. 2013, Maroja et al. 2015). Likewise, differential introgression has been documented in genetic and genomic analyses of a central European hybrid zone between *Mus musculus* and *M. domesticus* (Teeter et al. 2010, Macholán et al. 2011, Janoušek et al. 2012). Introgression in this hybrid zone tends to be most restricted on the X chromosome, in genomic regions with low rates of recombination or high levels of divergence between species, and in those that harbor genes involved with certain functions, such as DNA binding (Janoušek et al. 2015).

3.2. Causes of Differential Introgression

Differential introgression among loci can result from stochastic processes, and thus evidence of restricted (or enhanced) introgression does not necessarily imply selection or an association with reproductive isolation (Long 1991, Fitzpatrick et al. 2009, Gompert et al. 2012b). Multiple

Table 1 Overview of methods (software) for genomic analyses of hybrid zones

Metric	Software	Description
Admixture proportions	`admixture`[1], `entropy`[2], `structure`[3]	Admixture proportions describe the proportion of an individual's genome derived from each of two or more source populations (i.e., genome-average ancestry) and are akin to a hybrid index when considering two parental populations. Software differs in whether Bayesian (`structure` and `entropy`) or likelihood (`admixture`) methods are used. `structure` can account for autocorrelations in ancestry along chromosomes (i.e., admixture linkage disequilibrium), and `entropy` includes a model to estimate the proportion of the genome where both allele copies were inherited from the same versus different source populations. Ancestry models of this type and in the following categories are reviewed and compared in detail elsewhere[4] (also see the sidebar Hybrid Indexes and Admixture Proportions).
Geographic clines	`analyse`[5], `hzar`[6]	`analyse` fits one- and two-dimensional equilibrium cline models (see the sidebar Hybrid Indexes and Admixture Proportions) using maximum likelihood [estimates are obtained using Markov-chain Monte Carlo (MCMC)]. `hzar` is an `R` package that fits one-dimensional clines with MCMC and uses AIC (the Akaike information criterion) to contrast clines and test their coincidence and concordance.
Genomic clines	`analyse`[5], `bgc`[7,8], `hiest`[9], `introgress`[10]	Genomic clines describe introgression along an admixture gradient. `analyse` conducts Barton's concordance analysis but only considers diagnostic loci. `bgc` fits a hierarchical Bayesian model for clines in ancestry as a function of hybrid index using the Barton cline function (i.e., the same functional form as `analyse`). `bgc` accounts for uncertainty in ancestry due to polymorphism within species and can test a null hypothesis of no differential introgression (i.e., excess ancestry) or for deviations from a null distribution based on all markers (global outliers) or nearby markers (local outliers). Autocorrelations in clines for linked loci can be accounted for through an intrinsic conditional autoregressive prior. `hiest` uses maximum likelihood to fit a variety of cline models for genotypic data (rather than ancestry) and applies multivariate methods to test for global outliers, similar to `bgc`. `introgress` uses multinomial regression to fit genomic clines for genotypes and allows distinct clines for homozygous and heterozygous genotypes. It tests the null hypothesis of no differential introgression.
Ancestry blocks and frequencies	`hapmix`[11], `popanc`[12]	A variety of methods exist to infer local ancestry using hidden Markov models (HMM); `hapmix` is one of the more popular ones. Results from such methods can be used to calculate differences between genome-average and local ancestry frequencies in admixed populations (delta-ancestry) and to identify loci with exceptional ancestry frequencies. `popanc` fits a correlated Beta process model for ancestry frequencies in admixed populations and is unique in accounting for both ancestry blocks within individuals and autocorrelations in ancestry frequencies along chromosomes. Analyses based on ancestry frequencies are most appropriate when variance in hybrid index is small, and HMM approaches are probably best when admixture is recent, whereas `popanc` may perform better in older admixed populations.

[1]Alexander et al. (2009), [2]Gompert et al. (2014), [3]Falush et al. (2003), [4]Gompert & Buerkle (2013), [5]Barton & Baird (1995), [6]Derryberry et al. (2014), [7]Gompert & Buerkle (2011b), [8]Gompert et al. (2012b), [9]Fitzpatrick (2013), [10]Gompert & Buerkle (2009) [11]Price et al. (2009), [12]Gompert (2016).

lines of evidence do support selection in some cases. For example, a role for selection in explaining restricted introgression for the X chromosome between *M. musculus* and *M. domesticus* is supported by experimental results showing that this chromosome is disproportionately responsible for hybrid sterility (Good et al. 2008, Turner & Harr 2014). In other cases, selection's role is less clear, and several factors suggest that caution is needed when making conclusions about evolutionary processes from patterns of introgression in hybrid zones.

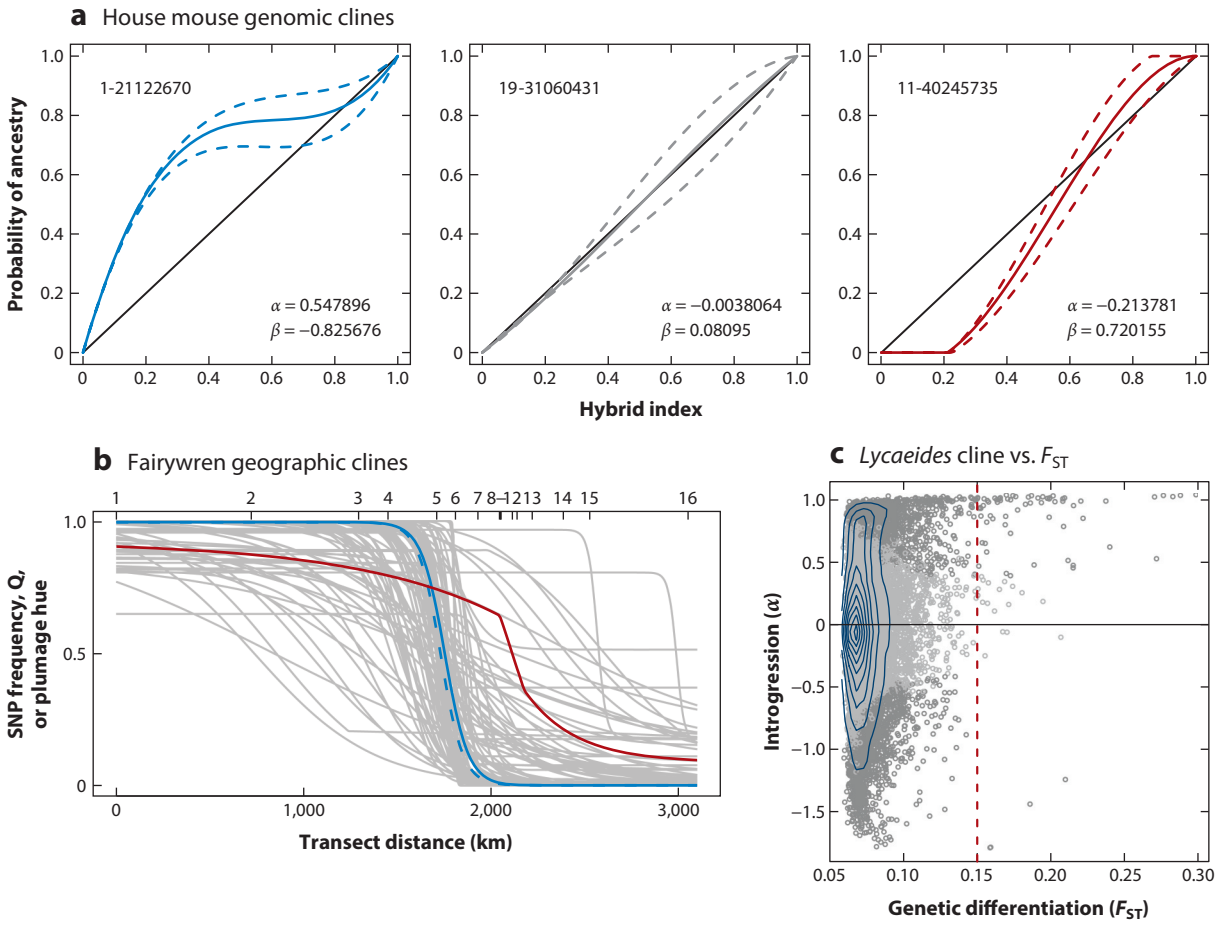

Figure 3

Patterns of differential introgression in three hybrid zones. (*a*) Genomic clines for three single nucleotide polymorphisms (SNPs) (1-21122670, 19-31060431, and 11-40245735) in the European *Mus domesticus* × *M. musculus* hybrid zone (*solid* and *dashed colored lines* show cline point estimates and 95% credible intervals; figure adapted with permission from Janoušek et al. 2015). SNP 1-21122670 exhibits a wide cline with an excess of *M. domesticus* ancestry, whereas 11-40245735 shows evidence of restricted introgression. α and β are the genomic cline parameters. (*b*) Geographic clines for ancestry-informative SNPs (SNP frequency, *gray lines*), hybrid index Q (*solid blue line*, 2,702 SNPs; *dashed blue line*, 102 diagnostic SNPs) and plumage color (*red line*) in an Australian fairywren hybrid zone (figure adapted with permission from Baldassarre et al. 2014). Clines for several SNPs and plumage color are offset from or steeper (or shallower) than the hybrid index cline. (*c*) Patterns of differentiation (F_{ST}) and introgression (α) in a hybrid zone between the butterfly species *Lycaeides idas* and *L. melissa* (Gompert et al. 2012a). Dark gray circles denote loci with 95% credible intervals for genomic cline parameter α that do not include 0. Positive and negative values of α correspond with elevated *L. idas* and *L. melissa* ancestry, respectively. Highly differentiated SNPs between allopatric parental populations (i.e., F_{ST} beyond the vertical dashed line) were more likely to exhibit excess *L. idas* ancestry (above the horizontal line) than excess *L. melissa* ancestry.

First, many common analytical approaches do not distinguish between drift and selection as causes of restricted (or enhanced) introgression (**Table 1**). For example, tests for coincidence and concordance of spatial clines and some genomic cline methods treat no variation in introgression (rather than neutral introgression) as the null hypothesis (i.e., these tests only account for sampling error, not drift; Gompert et al. 2012b). In any hybrid zone where the population size is finite,

drift will cause some variability in introgression, so the null hypothesis of no variation among loci is false a priori. A subset of genomic cline methods attempts to parse drift and selection based on variability in patterns among loci. Outlier loci can be identified relative to a genome-wide distribution of introgression (global outliers) or relative to patterns of introgression for nearby loci (local outliers relative to patterns of introgression for nearby loci); the latter requires type *ii* or *iii* genomic data (Gompert & Buerkle 2011b, Gompert et al. 2012b, Fitzpatrick 2013). A related method was proposed by Long (1991) to measure the contribution of drift to variation in ancestry among loci and to test the null hypothesis of pure drift for all loci; residuals from this null model have been examined to assess evidence for selection on individual loci (e.g., Fitzpatrick et al. 2010, Kovach et al. 2016). All of these methods implicitly assume that the majority of loci are unaffected by selection and thus define a null distribution under pure drift; this might not be reasonable (see below). More generally, successfully parsing the effects of selection versus neutral processes on patterns of differential introgression depends on demography and genetic architecture (Gompert et al. 2012b). Unfortunately, the bulk of studies suggest that barriers to gene flow are often caused by many loci and that simple genetic architectures that would make inferences easier may be rare and not representative of species barriers in general (e.g., Szymura & Barton 1986, Janoušek et al. 2012, Payseur & Rieseberg 2016).

Second, direct selection on causal variants that are responsible for the low fitness of hybrids or immigrants (i.e., barrier loci) and indirect selection from LD with causal variants can affect introgression in hybrid zones. The effects of direct and indirect selection can be difficult to parse (e.g., Gompert et al. 2017). When hybrid fitness is polygenic and LD is high (as is expected when the total barrier to gene flow is moderate or strong, or when hybridization occurs between well-differentiated populations), many or most loci could experience indirect selection via LD, making tests of a pure drift null hypothesis mostly meaningless. A more reasonable null model would be no direct selection, but this is also harder to test. Similarly, elevated LD makes causal variants easier to detect via association with marker loci, but difficult to localize. Methods such as local outlier analyses in genomic clines (Gompert et al. 2012b) that test for restricted introgression relative to neighboring linked loci can improve resolution, but only to a point. And of course, unless whole-genome data (type *iii* genomic data) are available for many individuals, even local outlier scans are unlikely to identify true causal variants (i.e., sparser marker sets will likely tag causal variants via LD, but they are less likely to include the actual causal mutations).

3.3. Differential Introgression and the Genetics of Speciation

Despite these limitations, a few patterns have emerged from genomic analyses of hybrid zones that suggest certain genetic regions are commonly involved in reproductive isolation. There is a clear tendency for restricted introgression on X (or Z) chromosomes relative to autosomes (e.g., Garrigan et al. 2012, Carneiro et al. 2013, Maroja et al. 2015). This is consistent with results from experimental crosses that often associate hybrid sterility and inviability with sex chromosomes (reviewed by Coyne & Orr 2004). However, restricted X (or Z) introgression does not necessarily imply a greater number (or average effect) of barrier loci on sex chromosomes, as the pattern could also arise from differences in dispersal, effective population size, or recombination (and thus LD and indirect selection) between sex chromosomes and autosomes (Petit & Excoffier 2009). A second general finding is that introgression is often restricted in rearranged portions of the genome, or regions that otherwise exhibit reduced recombination (e.g., Rieseberg et al. 1999, Barb et al. 2014, Lohse et al. 2015). However, this too could be explained by a greater propensity for such regions to harbor loci that contribute to reproductive isolation and speciation, or by increased effects of selection on linked loci when recombination is restricted (Noor & Bennett 2009).

There is also growing evidence that highly differentiated regions between allopatric populations tend to be more resistant to gene flow in hybrid zones, as would be predicted if differentiated genetic regions harbor barrier loci and selection acts similarly to create and maintain species differences. Studies in *Lycaeides* butterflies (Gompert et al. 2012a), manakins (Parchman et al. 2013), and field crickets (Larson et al. 2013) found that locus-specific differentiation and introgression were associated overall (this is partially but not wholly a statistical artifact, as ancestry information depends on allele frequency differences; Gompert et al. 2012a), but that introgression was not restricted for all differentiated regions. This suggests that differentiation at some of these loci was not adaptive or that selection was context dependent. In *Lycaeides*, highly differentiated loci had an overrepresentation of excess ancestry from one of the parental species (*L. idas*; see **Figure 3**) (Gompert et al. 2012a). This is hard to explain from neutral processes, but instead suggests that highly differentiated loci affected hybrid and parental fitness in ways that depended on habitat or genomic background (hybrids occur in *L. idas*-like habitat and a greater proportion of their genome is from this species).

Given the research that has been done thus far, we think further progress in understanding the causes of differential introgression and genetics of species barriers can be best made not by using cline-based genome scans to identify individual barrier loci, but rather by testing specific hypotheses about patterns of introgression. For example, differences in patterns of introgression between randomly chosen loci and quantitative trait loci (QTL) can help identify the trait differences most associated with restricted introgression and reproductive isolation in hybrid zones (e.g., Rieseberg et al. 1999, Janoušek et al. 2012, Gompert et al. 2013a). Admixture mapping could be particularly useful in this context (see the sidebar Admixture Mapping). Similarly, tests for environment-dependent patterns of introgression can be used to identify possible agents of selection in hybrid zones (e.g., Kovach et al. 2016). Further progress could also be made by extracting additional information from ancestry block lengths (e.g., Wegmann et al. 2011, Gravel 2012, Sedghifar et al. 2016) and by shifting to process-based models that attempt to infer selection and demographic history while accounting for LD (and recombination; see Payseur & Rieseberg 2016). This could done with an approximate Bayesian computation framework, either using the coalescent or forward-time simulations, but it is unclear whether joint inference of these parameters would be practical and sufficiently precise (time-series data could be useful in this context). In the end, we may need to confront the fact that precisely describing patterns of differential introgression is much easier than confidently assigning causes to the patterns, and thus to decide when a description of pattern alone is and is not useful or sufficient.

3.4. The Dynamics of Speciation and the Reality of Species

Patterns of natural hybridization suggest the possibility of a series of stages (or phases) of speciation (Jiggins et al. 2001, Wu 2001, Roux et al. 2016). In early stages, intermediates are common and hybridization results in gene flow across most of the genome. If speciation progresses, barrier loci restrict introgression across more of the genome. Late (or genomic) stages of speciation are characterized by bimodal hybrid zones with mostly parental individuals and genome-level reproductive isolation that is often due in part to strong assortative mating. An understanding of the evolutionary dynamics driving the transition from early to late stages of speciation is central to understanding the speciation process in some contexts (Nosil et al. 2017). The traditional hypothesis of reinforcement, in which selection directly favors an increase in prezygotic isolation, has received some support (e.g., Servedio & Noor 2003, Bímová et al. 2011), but recent theory suggests more general mechanisms by which genotypes at different barrier (and nonbarrier) loci can become associated or coupled, enhancing the overall barrier to gene flow (Barton 1983, Baird

ADMIXTURE MAPPING

Admixture mapping is a special case of association mapping in which phenotypes are mapped to the genome based on statistical associations with genotypes or ancestry blocks in an admixed population or hybrid zone (Buerkle & Lexer 2008). Admixture mapping is facilitated by LD between loci that affect phenotype and marker loci that are assayed, where LD is expected to extend over greater genomic distances as a result of recent admixture between populations (**Figure 1**). This has some advantages relative to controlled crosses or typical genome-wide association mapping (GWAM) in unstructured populations. In contrast to GWAM, admixture mapping can be used to study the genetics of trait differences between species, including those differences that underlie reproductive isolation (e.g., Malek et al. 2012, Gompert et al. 2013a, Pallares et al. 2014). And, by using natural mixing and recombination, admixture mapping is amenable to organisms that cannot be readily crossed and reared in the lab. Moreover, by using the same genetic data and hybrid zone for trait mapping and analyses of introgression, QTL are inherently identified in the ecological and genomic context and at the genomic resolution relevant for evolutionary dynamics (i.e., admixture LD affects the resolution of mapping and the genomic scale at which selection affects patterns of introgression similarly).

Various approaches for admixture mapping have been proposed. This includes applying standard GWAM methods to admixed populations (e.g., Gompert et al. 2013b, Lindtke et al. 2013); testing for associations between excess ancestry and trait values, disease risk, or specific gene functions (e.g., Patterson et al. 2004, vonHoldt et al. 2016); or jointly modeling the association of ancestry and genotype with phenotype (Shriner et al. 2011). Ancestry-based approaches are expected to perform particularly well for traits that differ substantially between species, but will fail to map within-species trait variation. More generally and similar to other mapping approaches, the efficacy of admixture mapping depends on trait genetic architectures. When trait differences are highly polygenic with mostly small effect variants, it will be difficult to confidently identify and localize QTL. However, combined analyses of introgression for a set of QTL could still be informative, and polygenic models can be used to infer higher-level parameters, such as the contribution of genetic variation to trait variation in a hybrid zone (Pallares et al. 2014).

1995, Barton & De Cara 2009). Selection favors coupling (associations of genotypes between loci) when fitness is multiplicative or when there is positive epistasis, and it can occur between different types of barrier loci, including those responsible for intrinsic (endogenous) incompatibilities and those with environment-dependent effects on fitness (exogenous barrier loci) (**Figure 4**, Bierne et al. 2011). Increased coupling corresponds to LD among barrier loci, which results in an increase in the total (indirect plus direct) selection experienced by each locus. This can result in a gradual or sudden shift from a weak, semipermeable hybrid zone to a late (or genomic) phase of speciation, in which most of the genome is protected from gene flow (Flaxman et al. 2014, Nosil et al. 2017). Sudden shifts can be driven by positive feedback between LD and indirect selection (Barton & De Cara 2009, Flaxman et al. 2014, Nosil et al. 2017). Feedback can also occur in the opposite direction, causing sudden collapse of species when the barrier is too weak.

Whether coupling of barrier loci is often a critical component of speciation is an empirical question that depends on opportunities for gene flow and the genetic basis of reproductive isolation. Strict allopatric speciation is unaffected by coupling. Whereas geographic isolation likely facilitates speciation in most cases (Coyne & Orr 2004), periods of potential gene flow are probably reasonably common during the speciation process, suggesting coupling could be important (e.g., Davison et al. 2005, Martin et al. 2013). Similarly, coupling may be more critical when gene flow is high relative to the strength of divergent selection (or the barrier strength more generally) on individual loci (Flaxman et al. 2014). For example, coupling allows many weakly selected loci that

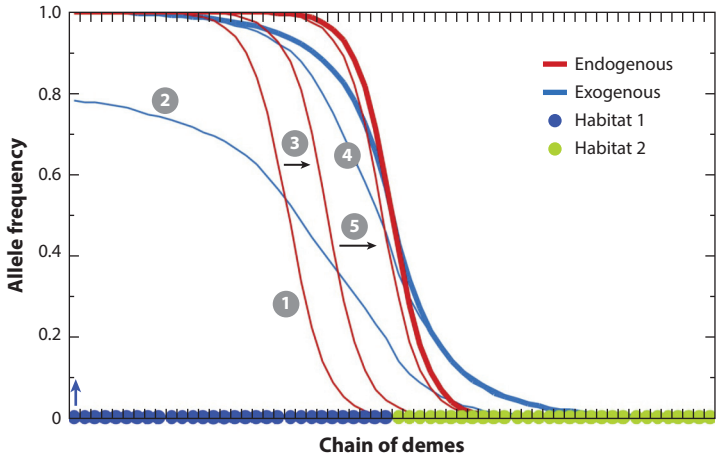

Figure 4

Coupling between an endogenous and exogenous cline in a 60-deme stepping stone model (Bierne et al. 2011). The endogenous cline (*red line*) was positioned at deme 20 (1). A new exogenous allele arose (*blue arrow*) and spread across habitat 1 (2), which caused the endogenous cline to shift and both clines to couple, narrow, and move toward the environmental boundary (3–5). Figure adapted with permission from Bierne et al. (2011). A video of coupling can be found online (see **Supplemental Video 1**).

Supplemental Material

are spread across the genome to form a strong barrier to gene flow (Barton & Bengtsson 1986). In cases where coupling is important, feedback can drive rapid transitions between early and late stages of speciation. If such rapid transitions are common, it would bolster the argument of species as real entities, as little time would be spent in intermediate stages of speciation. A prediction of this feedback hypothesis is that most hybrid zones would contain either a mixture of hybrid types with highly variable ancestry and introgression (early stage) or mostly parental types with little variability in introgression (late stage). This is consistent with recent work showing a narrow gray zone of speciation based on historical inference of gene flow (Roux et al. 2016), but the quantitative measures of cline variability needed to address this question are mostly lacking in the hybrid zone literature.

4. EVOLUTIONARY CONSEQUENCES OF GENE FLOW ACROSS HYBRID ZONES

Much historical, theoretical, and contemporary work has focused on tension zones, which are hybrid zones maintained by a balance between dispersal of individuals and selection against hybrids. Analyses in the context of tension zones incorporate the assumption that hybrid zones are stable through time and that outcomes are relatively independent of ecological conditions (Barton & Hewitt 1985). However, a stable tension zone is only one potential outcome of hybridization, and recent work shows that outcomes of hybridization can vary substantially. Hybridization outcomes can also include adaptive introgression (Borge et al. 2005, Whitney et al. 2010), extinction via hybridization (Rhymer & Simberloff 1996, Wolf et al. 2001, Buerkle et al. 2003), or formation of hybrid species (Mallet 2007). The ecological and genomic contexts for hybridization have the potential to influence which outcome occurs in a given hybrid zone, but our understanding of how contingency, ecological context, and genomic context influence evolutionary processes in

hybrid zones remains largely incomplete. High-throughput sequencing now makes it feasible to compare genomic variation across multiple instances of hybridization, and we are poised to better characterize the influence of contingency on hybrid zones.

Outcomes of hybridization can be influenced by traits of hybrid zones and the specific locations where hybrid zones occur. In general, the dynamics and consequences of gene flow across hybrid zones depend on the spatial scale of hybridization relative to the ranges of hybridizing species. In widespread species with limited geographic overlap, introgression at the local scale of the hybrid zone might have different evolutionary consequences than introgression at the scale of a species's range, or when overlap between hybridizing species constitutes a greater proportion of one or both species' ranges. The specific geographic location where alleles that contribute to isolation arise matters for outcomes of hybridization, and there is potential for polymorphism in isolation in different locations where a pair of species interacts (Cutter 2012).

Likewise, the population size and status of hybridizing populations matter to outcomes of hybridization. Introgression is likely to be asymmetric in cases where alleles from one parental species confer higher fitness in hybrid individuals (adaptive introgression; see Borge et al. 2005, Whitney et al. 2010), or when effective population size is much larger in one parental species than the other (e.g., Lepais et al. 2009). Even when barriers to interspecific reproduction are fairly strong and genomic, adaptive introgression is possible. Recent research on introgression in humans has suggested that some alleles might be negatively selected because of a higher mutation load in the originating deme (deleterious alleles harbored by Neanderthals but exposed to selection through introgresssion into humans, described by Harris & Nielsen 2016).

4.1. Variation in Hybridization Within a Species Pair

A current focus of research in hybrid zones is how consistent outcomes of hybridization are across repeated zones of contact between the same pair of species. Identifying how much variation exists in outcomes of hybridization is crucial for understanding the roles of contingency and context-dependence in hybrid zones. Variability in hybridization and introgression can be quantified using both overall estimates of ancestry in individuals and ancestry of specific loci or segments of the genome (**Table 1**).

In many systems where multiple instances of hybridization between the same pair of species have been observed, outcomes of hybridization and introgression vary across locations where the species come into contact. Hybridization outcomes vary across multiple sampled hybrid zones in many taxa, including *Helianthus* sunflowers (Buerkle & Rieseberg 2001), *Bombina* toads (Vines et al. 2003), oaks (Lepais et al. 2009), house mice (Teeter et al. 2010), *Populus* trees (Lexer et al. 2010), spruce (Haselhorst & Buerkle 2013), *Lycaeides* butterflies (Gompert et al. 2014), and *Catostomus* fish (Mandeville et al. 2015). In contrast, only a few systems where multiple instances of hybridization have been studied show relatively consistent patterns of hybridization and locus-specific introgression across locations, including consistency between hybrid zones in Connecticut and Pennsylvania in *Gryllus* crickets (Larson et al. 2014) and some consistency in locus-specific introgression between hybrid zones between *Helianthus annuus* and *H. petiolaris* sunflowers in California and Nebraska (Buerkle & Rieseberg 2001).

One common feature between these two examples of consistent hybridization across locations is that these studies did not use genomic data, but rather used a smaller number of loci. In the case of *Gryllus*, 110 loci (SNPs) with large allele frequency differences between species (putatively fixed loci) were selected from a larger, genomic data set (Larson et al. 2014). These loci are ancestry-informative and therefore are useful for understanding processes of hybridization and

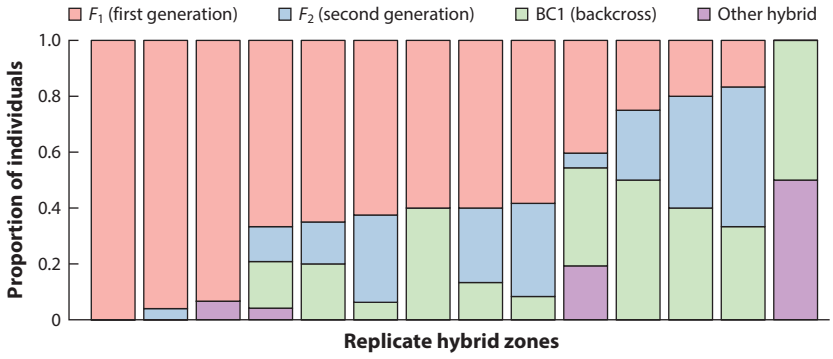

Figure 5

Catostomus latipinnis × *C. commersoni* hybridization outcomes can vary across hybrid zones where a pair of species comes into contact. *Catostomus* fishes *C. latipinnis* and *C. commersoni* hybridize at many locations in the Upper Colorado River basin (Wyoming, Colorado, and Utah), but the distribution of hybrid individuals across different classes of hybrids varies by site. It is likely that ecological context and historical contingency influence hybridization outcomes, but specific relationships between ecology, contingency, and hybrid zone dynamics are difficult to detect. Figure based on data from Mandeville et al. (2017).

identifying loci potentially associated with reproductive isolation (Larson et al. 2013, Maroja et al. 2015). However, these strongly differentiated loci might be unusually consistent in hybrid zones relative to the rest of the genome because of selection. In *Helianthus* sunflowers, locus-specific introgression due to hybridization between *H. annuus* and *H. petiolaris* is consistent in hybrid zones in Nebraska and California at 61 RAPD (rapid amplification of polymorphic DNA) loci (Buerkle & Rieseberg 2001). However, when analyzed with genomic clines (Gompert & Buerkle 2009), locus-specific introgression relative to the rest of the genome is more variable.

Genomic data have also changed our understanding of the magnitude of variation in hybrid zones simply by making more comparisons possible. Reductions in cost and an increase in the feasibility of sequencing many individuals make it easier to obtain population genomic data for multiple hybrid zones between a pair of species, which has altered our understanding of hybridization in some species pairs. For example, in *Catostomus* fish, sampling multiple geographic instances of hybridization revealed that hybridization is variable across locations (**Figure 5**; Mandeville et al. 2017). Initial genetic analysis in this system suggested an imminent breakdown of reproductive isolation among three parental species and the potential for extinction via hybridization (McDonald et al. 2008). More recent research makes it clear that hybridization outcomes vary among rivers and are much more constrained in many other locations than in the original population studied (Mandeville et al. 2015). The previous studies lacked both the genomic resolution and the geographic sampling to detect the extensive variation among hybrid zones.

4.2. Contingency in Reproductive Isolation: Ecological and Genomic Context

The extent of variation in outcomes of hybridization suggests that mechanisms of reproductive isolation are likely to be polymorphic or context-dependent within a species pair. In some study systems, a clear association exists between some feature of the environment and outcomes of hybridization. For example, in *Bombina* toads, hybridization dynamics are believed to vary according to habitat, with differing habitat preferences of parental species affecting the hybridization outcomes (Vines et al. 2003). In oaks, outcomes of hybridization are driven by relative abundance

of parental species (Lepais et al. 2009). Westslope cutthroat trout hybridization with rainbow trout is believed to be promoted by warming water temperatures that increase overlap in spawning (Muhlfeld et al. 2014). However, these clearly understood examples of ecological context influencing hybridization dynamics might be the exception rather than the rule. The mechanism driving variation in hybridization outcomes in the first example is likely differences in fitness of hybrids; in the second and third cases, prezygotic mechanisms probably drive hybridization dynamics. In many hybrid zones, identifying specific sources of context dependence may be very difficult, especially when large, heterogeneous geographic ranges of species and multiple pairs of interacting species are involved (e.g., Mandeville et al. 2015). It is also probable that reproductive isolation in many systems is not controlled by a single ecological factor, but rather by a suite of interacting ecological characteristics that might interact differently in different locations.

5. CONCLUSIONS AND FUTURE DIRECTIONS

Hybrid zones have long been of interest both as tools for studying speciation and because of potential evolutionary consequences of hybridization (Barton & Hewitt 1985, Harrison 1993, Hewitt 2001, Abbott et al. 2013, Gompert & Buerkle 2016). The increased use of population genomic data in studies of hybrid zones has brought clarity to some longstanding questions, but it has also revealed complexity that will prove challenging in future efforts to understand speciation and the evolutionary consequences of hybridization. Methods for identifying loci with unusual patterns of introgression have been widely employed (Barton & Baird 1995, Gompert et al. 2012b, Fitzpatrick 2013), but the results of these analyses have made it clear that future progress in understanding the genetic architecture of reproductive isolation will require explicitly linking patterns of introgression to the function of outlier loci. We are likely to make the best progress toward understanding the genetic architecture of reproductive isolation by using explicit process-based models and implementing tests of specific hypotheses about how patterns of introgression relate to reproductive isolation.

Genomic analyses have also advanced our understanding of hybrid zones as evolutionary phenomena that permit gene flow between species (Abbott et al. 2013, Gompert & Buerkle 2016). As studies of hybrid zones have grown to encompass multiple instances of hybridization between a pair of species, it has become increasingly clear that many, if not most, species pairs have variable outcomes across locations where hybridization occurs (Vines et al. 2003, Lexer et al. 2010, Haselhorst & Buerkle 2013, Gompert et al. 2014). In some cases, our inferences will therefore depend on both the genomic and geographic scale of studies (e.g., McDonald et al. 2008, Mandeville et al. 2015). Ecological and genomic contexts, along with historical contingency, are likely to drive much of this variation, but these processes are still poorly characterized in most systems. Several related challenges exist for future studies of the evolutionary consequences of hybrid zones. It is currently unknown to what extent genomic variation in hybridization and introgression affects speciation and maintenance of reproductive isolation, and which variation is a consequence of drift and unlikely to be functionally important. Another major unresolved question is how tractable it is to identify specific ecological contexts or sources of contingency that drive variation in hybridization outcomes. It is possible that systems with clear context dependence are unusual (e.g., Vines et al. 2003, Lepais et al. 2009, Muhlfeld et al. 2014), and that more typical hybrid zones have too many sources of ecological dependence and contingency to identify why hybridization varies. Progress in understanding context dependence and contingency in hybridization outcomes requires characterizing the extent of variability in genomic and phenotypic outcomes of hybridization, and might necessitate the development of explicit statistical frameworks for comparing hybridization across multiple, replicate hybrid zones.

Recent incorporation of population genomic data into analyses of hybrid zones has enriched our understanding of speciation and evolutionary consequences of hybridization. Relative to theoretical expectations developed prior to the availability of genomic data (Barton & Hewitt 1985, Hewitt 1988, Harrison 1993), we now know that there is significant diversity in hybrid zones, both across distinct hybridizing species pairs and across instances of contact between a single pair of species. Geography of hybrid zones can be extremely variable, including wide and narrow tension zones and also regions better described as mosaic hybrid zones. A rich collection of recent methods now allows more precise description of outcomes of hybridization and enables better comparisons across hybrid zones (**Table 1**), both at the whole-genome scale and at specific loci or genomic regions. As availability of whole-genome data and high-quality reference genomes increases, we expect additional insights into mechanisms of speciation and evolutionary consequences of hybridization, which will be most powerful if combined with analysis of genome function and phenotype, and with an awareness of potential for variability and context dependence in evolutionary processes.

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

We thank Bradley Shafer for inviting us to write this review, and Daniel Baldassarre, Nicolas Bierne, Vaclav Janoušek, and Priscilla Tucker for providing figures and media for use in this article.

LITERATURE CITED

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

Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19:1655–64. **https://dx.doi.org/10.1101/gr.094052.109**

Allendorf F, Leary R, Spruell P, Wenburg J. 2001. The problems with hybrids: setting conservation guidelines. *Trends Ecol. Evol.* 16:613–22

Arnold ML. 1997. *Natural Hybridization and Evolution*. New York: Oxford Univ. Press

Baird SJE. 1995. A simulation study of multilocus clines. *Evolution* 49:1038–45

Baldassarre DT, White TA, Karubian J, Webster MS. 2014. Genomic and morphological analysis of a semipermeable avian hybrid zone suggests asymmetrical introgression of a sexual signal. *Evolution* 68:2644–57

Barb JG, Bowers JE, Renaut S, Rey JI, Knapp SJ, et al. 2014. Chromosomal evolution and patterns of introgression in *Helianthus*. *Genetics* 197:969–79

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

Barton NH, Baird S. 1995. Analyse: an application for analysing hybrid zones. *Software for genomic analyses of hybrid zones.* **http://archive.bio.ed.ac.uk/software/Mac/Analyse/index.html**

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

Barton NH, De Cara MAR. 2009. The evolution of strong reproductive isolation. *Evolution* 63:1171–90

Barton NH, Gale KS. 1993. Genetic analysis of hybrid zones. In *Hybrid Zones and the Evolutionary Process*, ed. RG Harrison, pp. 13–45. New York: Oxford Univ. Press

Barton NH, Hewitt GM. 1981. A chromosomal cline in the grasshopper *Podisma pedestris*. *Evolution* 35:1008–18

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* 16:113–48

Barton NH, Hewitt GM. 1989. Adaptation, speciation and hybrid zones. *Nature* 341:497–503
Bazykin AD. 1969. Hypothetical mechanism of speciation. *Evolution* 23:685–87
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
Bímová BV, MacHolán M, Baird SJ, Munclinger P, Dufková P, et al. 2011. Reinforcement selection acting on the European house mouse hybrid zone. *Mol. Ecol.* 20:2403–24
Borge T, Lindroos K, Nádvorník P, Syvänen AC, Sætre GP. 2005. Amount of introgression in flycatcher hybrid zones reflects regional differences in pre and post-zygotic barriers to gene exchange. *J. Evol. Biol.* 18:1416–24
Buerkle CA, Lexer C. 2008. Admixture as the basis for genetic mapping. *Trends Ecol. Evol.* 23:686–94
Buerkle CA, Rieseberg LH. 2001. Low intraspecific variation for genomic isolation between hybridizing sunflower species. *Evolution* 55:684–91
Buerkle CA, Wolf DE, Rieseberg LH. 2003. The origin and extinction of species through hybridization. In *Population Viability in Plants: Conservation, Management, and Modeling of Rare Plants*, ed. CA Brigham, MW Schwartz, pp. 117–41. New York: Springer-Verlag
Buggs RJA. 2007. Empirical study of hybrid zone movement. *Heredity* 99:301–12
Burri R, Nater A, Kawakami T, Mugal CF, Olason PI, et al. 2015. Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers. *Genome Res.* 25:1656–65
Carneiro M, Baird SJ, Afonso S, Ramirez E, Tarroso P, et al. 2013. Steep clines within a highly permeable genome across a hybrid zone between two subspecies of the European rabbit. *Mol. Ecol.* 22:2511–25
Coyne JA, Orr HA. 2004. *Speciation*. Sunderland, MA: Sinauer
Cutter AD. 2012. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. *Trends Ecol. Evol.* 27:209–18
Dasmahapatra KK, Blum MJ, Aiello A, Hackwell S, Davies N, et al. 2002. Inferences from a rapidly moving hybrid zone. *Evolution* 56:741–53
Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, et al. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487:94–98
Davison A, Chiba S, Barton NH, Clarke B. 2005. Speciation and gene flow between snails of opposite chirality. *PLOS Biol.* 3:e282
Derryberry EP, Derryberry GE, Maley JM, Brumfield RT. 2014. HZAR: hybrid zone analysis using an R software package. *Mol. Ecol. Resourc.* 14:652–63
Endler JA. 1977. *Geographic Variation, Speciation, and Clines*. Princeton, NJ: Princeton Univ. Press
Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–87
Fitzpatrick BM. 2013. Alternative forms for genomic clines. *Ecol. Evol.* 3:1951–66
Fitzpatrick BM, Johnson JR, Kump DK, Shaffer HB, Smith JJ, Voss SR. 2009. Rapid fixation of non-native alleles revealed by genome-wide SNP analysis of hybrid tiger salamanders. *BMC Evol. Biol.* 9:176
Fitzpatrick BM, Johnson JR, Kump DK, Smith JJ, Voss SR, Shaffer HB. 2010. Rapid spread of invasive genes into a threatened native species. *PNAS* 107:3606–10
Flaxman SM, Wacholder AC, Feder JL, Nosil P. 2014. Theoretical models of the influence of genomic architecture on the dynamics of speciation. *Mol. Ecol.* 23:4074–88
Garrigan D, Kingan SB, Geneva AJ, Andolfatto P, Clark AG, et al. 2012. Genome sequencing reveals complex speciation in the *Drosophila simulans* clade. *Genome Res.* 22:1499–511
Gompert Z. 2016. A continuous correlated beta process model for genetic ancestry in admixed populations. *PLOS ONE* 11:e0151047
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. 2011a. A hierarchical Bayesian model for next-generation population genomics. *Genetics* 187:903–17
Gompert Z, Buerkle CA. 2011b. Bayesian estimation of genomic clines. *Mol. Ecol.* 20:2111–27
Gompert Z, Buerkle CA. 2013. Analyses of genetic ancestry enable key insights for molecular ecology. *Mol. Ecol.* 22:5278–94

Gompert Z, Buerkle CA. 2016. What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. *Evol. Appl.* 9:909–23

Gompert Z, Egan SP, Barrett RDH, Feder JL, Nosil P. 2017. Multilocus approaches for the measurement of selection on correlated genetic loci. *Mol. Ecol.* 26:365–82

Gompert Z, Lucas LK, Buerkle CA, Forister ML, Fordyce JA, Nice CC. 2014. Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Mol. Ecol.* 23:4555–73

Gompert Z, Lucas LK, Nice CC, Buerkle CA. 2013a. 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, Buerkle CA, Forister ML. 2013b. Geographically multifarious phenotypic divergence during speciation. *Ecol. Evol.* 3:595–613

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. B* 367:439–50

Good JM, Handel MA, Nachman MW. 2008. Asymmetry and polymorphism of hybrid male sterility during the early stages of speciation in house mice. *Evolution* 62:50–65

Gravel S. 2012. Population genetics models of local ancestry. *Genetics* 191:607–19

Hamilton JA, Lexer C, Aitken SN. 2013. Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* x *P. glauca*). *Mol. Ecol.* 22:827–41

Harris K, Nielsen R. 2016. The genetic cost of Neanderthal introgression. *Genetics* 203:881–91

Harrison RG. 1986. Pattern and process in a narrow hybrid zone. *Heredity* 56:337–49

Harrison RG. 1990. Hybrid zones: windows on evolutionary process. *Oxf. Surv. Evol. Biol.* 7:69–128

Harrison RG. 1993. Hybrids and hybrid zones: historical perspective. In *Hybrid Zones and the Evolutionary Process*, ed. RG Harrison, pp. 3–12. New York: Oxford Univ. Press

Harrison RG, Larson EL. 2014. Hybridization, introgression, and the nature of species boundaries. *J. Heredity* 105:795–809

Harrison RG, Larson EL. 2016. Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Mol. Ecol.* 25:2454–66

Haselhorst MSH, Buerkle CA. 2013. Population genetic structure of *Picea engelmannii*, *P. glauca* and their previously unrecognized hybrids in the central Rocky Mountains. *Tree Genet. Genomes* 9:669–81

Hewitt GM. 1988. Hybrid zones—natural laboratories for evolution studies. *Trends Ecol. Evol.* 3:158–66

Hewitt GM. 2001. Speciation, hybrid zones and phylogeography—or seeing genes in space and time. *Mol. Ecol.* 10:537–49

Huddleston J, Ranade S, Malig M, Antonacci F, Chaisson M, et al. 2014. Reconstructing complex regions of genomes using long-read sequencing technology. *Genome Res.* 24:688–96

Janoušek V, Munclinger P, Wang L, Teeter KC, Tucker PK. 2015. Functional organization of the genome may shape the species boundary in the house mouse. *Mol. Biol. Evol.*

Janoušek V, Wang L, Luzynski K, Dufková P, Vyskočilová MM, et al. 2012. Genome-wide architecture of reproductive isolation in a naturally occurring hybrid zone between *Mus musculus musculus* and *M. m. domesticus*. *Mol. Ecol.* 21:3032–47

Jiggins C, Naisbit R, Coe R, Mallet J. 2001. Reproductive isolation caused by colour pattern mimicry. *Nature* 411:302–5

Kovach RP, Hand BK, Hohenlohe PA, Cosart TF, Boyer MC, et al. 2016. Vive la résistance: genome-wide selection against introduced alleles in invasive hybrid zones. *Proc. R. Soc. Lond. B* 283:20161380

Larson EL, Andrés JA, Bogdanowicz SM, Harrison RG. 2013. Differential introgression in a mosaic hybrid zone reveals candidate barrier genes. *Evolution* 67:3653–61

Larson EL, White TA, Ross CL, Harrison RG. 2014. Gene flow and the maintenance of species boundaries. *Mol. Ecol.* 23:1668–78

Lepais O, Petit R, Guichoux E, Lavabre J, Alberto F, et al. 2009. Species relative abundance and direction of introgression in oaks. *Mol. Ecol.* 18:2228–42

Lexer C, Buerkle CA, Joseph JA, Heinze B, Fay MF. 2007. Admixture in European *Populus* hybrid zones makes feasible the mapping of loci that contribute to reproductive isolation and trait differences. *Heredity* 98:74–84

Lexer C, Joseph JA, van Loo M, Barbará T, Heinze B, et al. 2010. Genomic admixture analysis in European *Populus spp*. reveals unexpected patterns of reproductive isolation and mating. *Genetics* 186:699–712

Lindtke D, González-Martínez SC, Macaya-Sanz D, Lexer C. 2013. Admixture mapping of quantitative traits in *Populus* hybrid zones: power and limitations. *Heredity* 111:474–85

Lohse K, Clarke M, Ritchie MG, Etges WJ. 2015. Genome-wide tests for introgression between cactophilic *Drosophila* implicate a role of inversions during speciation. *Evolution* 69:1178–90

Long JC. 1991. The genetic structure of admixed populations. *Genetics* 127:417–28

Macholán M, Baird SJE, Dufková P, Munclinger P, Bimová BV, Piálek J. 2011. Assessing multilocus introgression patterns: a case study on the mouse X chromosome in Central Europe. *Evolution* 65:1428–46

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. 2007. Hybrid speciation. *Nature* 445:279–83

Mallet J, Barton N, Lamas G, Santisteban J, Muedas M, Eeley H. 1990. Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. *Genetics* 124:921–36

Mallet J, Besansky N, Hahn MW. 2016. How reticulated are species? *BioEssays* 38:140–49

Mandeville EG, Parchman TL, McDonald DB, Buerkle CA. 2015. Highly variable reproductive isolation among pairs of *Catostomus* species. *Mol. Ecol.* 24:1856–72

Mandeville EG, Parchman TL, Song SJ, Thompson KG, Compton RI, et al. 2017. Inconsistent reproductive isolation revealed by interactions between *Catostomus* fish species. *Evol. Lett.* In press

Maroja LS, Larson EL, Bogdanowicz SM, Harrison RG. 2015. Genes with restricted introgression in a field cricket (*Gryllus firmus/Gryllus pennsylvanicus*) hybrid zone are concentrated on the X chromosome and a single autosome. *G3* 5:2219–27

Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, et al. 2013. Genome-wide evidence for speciation with gene flow in heliconius butterflies. *Genome Res.* 23:1817–28

McDonald DB, Parchman TL, Bower MR, Hubert WA, Rahel FJ. 2008. An introduced and a native vertebrate hybridize to form a genetic bridge to a second native species. *PNAS* 105:10837–42

Moore WS. 1977. An evaluation of narrow hybrid zones in vertebrates. *Q. Rev. Biol.* 52:263–67

Muhlfeld CC, Kovach RP, Jones LA, Al-Chokhachy R, Boyer MC, et al. 2014. Invasive hybridization in a threatened species is accelerated by climate change. *Nat. Clim. Change* 4:620–24

Nadeau NJ, Ruiz M, Salazar P, Counterman B, Medina JA, et al. 2014. Population genomics of parallel hybrid zones in the mimetic butterflies, *H. melpomene* and *H. erato*. *Genome Res.* 24:1316–33

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

Nosil P, Feder JL, Flaxman SM, Gompert Z. 2017. Tipping points in the dynamics of speciation. *Nat. Ecol. Evol.* 1:0001

Pallares LF, Harr B, Turner LM, Tautz D. 2014. Use of a natural hybrid zone for genomewide association mapping of craniofacial traits in the house mouse. *Mol. Ecol.* 23:5756–70

Parchman TL, Gompert Z, Braun MJ, Brumfield RT, McDonald DB, et al. 2013. The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. *Mol. Ecol.* 22:3304–17

Pardo-Diaz C, Salazar C, Baxter SW, Merot C, Figueiredo-Ready W, et al. 2012. Adaptive introgression across species boundaries in heliconius butterflies. *PLOS Genet.* 8:1–13

Paşaniuc B, Sankararaman S, Kimmel G, Halperin E. 2009. Inference of locus-specific ancestry in closely related populations. *Bioinformatics* 25:i213–21

Patterson N, Hattangadi N, Lane B, Lohmueller KE, Hafler DA, et al. 2004. Methods for high-density admixture mapping of disease genes. *Am. J. Hum. Genet.* 74:979–1000

Payseur BA. 2010. Using differential introgression in hybrid zones to identify genomic regions involved in speciation. *Mol. Ecol. Resourc.* 10:806–20

Payseur BA, Rieseberg LH. 2016. A genomic perspective on hybridization and speciation. *Mol. Ecol.* 25:2337–60

Petit RJ, Excoffier L. 2009. Gene flow and species delimitation. *Trends Ecol. Evol.* 24:386–93
Poelstra JW, Vijay N, Bossu CM, Lantz H, Ryll B, et al. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* 344:1410–14
Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, et al. 2009. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLOS Genet.* 5:e1000519
Rand DM, Harrison RG. 1989. Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution* 43:432–49
Rhymer JM, Simberloff D. 1996. Extinction by hybridization and introgression. *Annu. Rev. Ecol. Syst.* 27:83–109
Rieseberg LH, Whitton J, Gardner K. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* 152:713–27
Riley SPD, Shaffer HB, Voss SR, Fitzpatrick BM. 2003. Hybridization between a rare, native tiger salamander (*Ambystoma californiense*) and its introduced congener. *Ecol. Appl.* 13:1263–75
Roux C, Fraisse C, Romiguier J, Anciaux Y, Galtier N, Bierne N. 2016. Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLOS Biol.* 14:e2000234
Saarman NP, Pogson GH. 2015. Introgression between invasive and native blue mussels (genus *Mytilus*) in the central California hybrid zone. *Mol. Ecol.* 24:4723–38
Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, et al. 2014. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* 507:354–57
Sedghifar A, Brandvain Y, Ralph P. 2016. Beyond clines: lineages and haplotype blocks in hybrid zones. *Mol. Ecol.* 25:2559–76
Servedio MR. 2004. The what and why of research on reinforcement. *PLOS Biol.* 2:e420
Servedio MR, Noor MAF. 2003. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. Syst.* 34:339–64
Shriner D, Adeyemo A, Rotimi CN. 2011. Joint ancestry and association testing in admixed individuals. *PLOS Comput. Biol.* 7:e1002325
Slatkin M. 1973. Gene flow and selection in a cline. *Genetics* 75:733–56
Szymura JM, Barton NH. 1986. Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. variegata*, near Cracow in southern Poland. *Evolution* 40:1141–59
Szymura JM, Barton NH. 1991. The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. *Evolution* 45:237–61
Tang H, Coram M, Wang P, Zhu XF, Risch N. 2006. Reconstructing genetic ancestry blocks in admixed individuals. *Am. J. Hum. Genet.* 79:1–12
Taylor SA, White TA, Hochachka WM, Ferretti V, Curry RL, Lovette I. 2014. Climate-mediated movement of an avian hybrid zone. *Curr. Biol.* 24:671–76
Teeter KC, Thibodeau LM, Gompert Z, Buerkle CA, Nachman MW, Tucker PK. 2010. The variable genomic architecture of isolation between hybridizing species of house mouse. *Evolution* 64:472–85
Toews DP, Taylor SA, Vallender R, Brelsford A, Butcher BG, et al. 2016. Plumage genes and little else distinguish the genomes of hybridizing warblers. *Curr. Biol.* 26:2313–18
Turner LM, Harr B. 2014. Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. *eLife* 3:e02504
Vines TH, Kohler SC, Thiel A, Ghira I, Sands TR, et al. 2003. The maintenance of reproductive isolation in a mosaic hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*. *Evolution* 57:1876–88
vonHoldt BM, Kays R, Pollinger JP, Wayne RK. 2016. Admixture mapping identifies introgressed genomic regions in North American canids. *Mol. Ecol.* 25:2443–53
Wegmann D, Kessner DE, Veeramah KR, Mathias RA, Nicolae DL, et al. 2011. Recombination rates in admixed individuals identified by ancestry-based inference. *Nat. Genet.* 43:847–53
Whitney KD, Randell RA, Rieseberg LH. 2010. Adaptive introgression of abiotic tolerance traits in the sunflower *Helianthus annuus*. *New Phytologist* 187:230–39
Wielstra B, Burke T, Butlin RK, Avcı A, Üzüm N, et al. 2017. A genomic footprint of hybrid zone movement in crested newts. *Evol. Lett.* 1:93–101
Wolf DE, Takebayashi N, Rieseberg LH. 2001. Predicting the risk of extinction through hybridization. *Conserv. Biol.* 15:1039–53

Wu CI. 2001. The genic view of the process of speciation. *J. Evol. Biol.* 14:851–65
Yanchukov A, Hofman S, Szymura JM, Mezhzherin SV. 2006. Hybridization of *Bombina bombina* and *B. variegata* (Anura, Discoglossidae) at a sharp ecotone in western Ukraine: comparisons across transects and over time. *Evolution* 60:583–600
Zbawicka M, Sańko T, Strand J, Wenne R. 2014. New SNP markers reveal largely concordant clinal variation across the hybrid zone between *Mytilus* spp. in the Baltic Sea. *Aquat. Biol.* 21:25–36