Genetics and the Evolution of Prezygotic Isolation

Richard M. Merrill,1 Henry Arenas-Castro,2 Anna F. Feller,3,4 Julia Harenčár,5 Matteo Rossi,1 Matthew A. Streisfeld,6 and Kathleen M. Kay5

1Faculty of Biology, Division of Evolutionary Biology, LMU Munich, 82152 Planegg-Martinsried, Germany
2School of Biological Sciences, University of Queensland, St. Lucia, Queensland 4072, Australia
3Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA
4Arnold Arboretum of Harvard University, Boston, Massachusetts 02131, USA
5Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, Santa Cruz, California 95060, USA
6Institute of Ecology and Evolution, University of Oregon, Eugene, Oregon 97403-5289, USA

Correspondence: merrill@bio.lmu.de; kmkay@ucsc.edu

The significance of prezygotic isolation for speciation has been recognized at least since the Modern Synthesis. However, fundamental questions remain. For example, how are genetic associations between traits that contribute to prezygotic isolation maintained? What is the source of genetic variation underlying the evolution of these traits? And how do prezygotic barriers affect patterns of gene flow? We address these questions by reviewing genetic features shared across plants and animals that influence prezygotic isolation. Emerging technologies increasingly enable the identification and functional characterization of the genes involved, allowing us to test established theoretical expectations. Embedding these genes in their developmental context will allow further predictions about what constrains the evolution of prezygotic isolation. Ongoing improvements in statistical and computational tools will reveal how pre- and postzygotic isolation may differ in how they influence gene flow across the genome. Finally, we highlight opportunities for progress by combining theory with appropriate data.

Prezygotic isolation (Box 1) includes all barriers to gene flow between populations that occur before fertilization. By acting early in the life cycle, prezygotic barriers are expected to have a disproportionate effect on overall reproductive isolation (RI), as they have the potential to limit gene flow before other barriers can act (Ramsey et al. 2003; Coyne and Orr 2004). A key goal of evolutionary genetics is to understand the historical, developmental, and ecological mechanisms that generate adaptive divergence and reproductive isolation. However, a major obstacle limiting our understanding of prezygotic isolation is that it tends to involve diverse phenotypes, including physiology, color, morphology, and behavior. In addition, the types of traits contributing to prezygotic isolation can vary among organisms, which has led to a lack of communication among scientists who work with different study systems.

In this work (which is also available in Spanish; see Supplemental Material), we provide an overview of features shared across different organ-
isms that may constrain or facilitate the evolution of prezygotic isolation. We begin with the classic problem of selection-recombination antagonism, how different types of allelic variation and genetic architecture may overcome it, and contributions of recent research in this area. We then focus on proximate considerations, including the origins of genetic variation and how its developmental context may constrain the evolution of prezygotic isolation. Finally, we consider how prezygotic barriers affect gene flow, and ask how we can distinguish their effects from those of postzygotic isolation. We conclude with opportunities we see for significant advances.

RECOMBINATION AS THE KEY CONSTRAINT ON THE EVOLUTION OF PREZYGOTIC ISOLATION

A rich body of theoretical work now exists concerning the evolution of prezygotic isolation (see Kirkpatrick and Ravigné 2002; Kopp et al. 2018 for excellent reviews). Although prezygotic isolation can evolve in allopatry (Knowlton et al. 1993; Lang erhans et al. 2007), most models consider how speciation may proceed when populations continue to exchange alleles (whether under full sympatry, parapatry, or after secondary contact). This is not because allopatry is insignificant; given enough time, traits will diverge to the extent that populations may no longer interbreed if their ranges overlap again. However, as Felsenstein (1981) famously identified, when populations remain in contact, the evolution of prezygotic isolation faces a more fundamental genetic constraint. This is “recombination, which acts to randomize the association between the prezygotic isolating mechanism (assortative mating) and the adaptations to the two environments” (Felsenstein 1981).

Despite Felsenstein’s skepticism, it is increasingly clear that speciation can proceed despite gene flow (Pinho and Hey 2010; Abbott et al. 2013; Arnold 2015). At the same time, there has been a renewed appreciation for the role of divergent natural selection in driving population divergence (Nosil 2012). These observations are related, because theory predicts that speciation with gene flow typically requires both the evolution of assortative mating and divergence in ecological traits (Kopp et al. 2018). Although “ecology-free” models of speciation with gene flow exist (e.g., Higashi et al. 1999), the assumptions are highly restrictive, and speciation with gene flow relying solely on sexual selection may be considered unrealistic (Kopp et al. 2018; but see Yang et al. 2019 for an example of a potentially widespread mechanism in which sexual imprinting also causes divergent selection). Finally, specia-
tion with gene flow also normally requires the maintenance of genetic associations (i.e., linkage disequilibrium [LD]) between alleles that contribute to assortative mating and those under divergent selection. Together, these requirements present an enduring conceptual challenge: If populations continue to interbreed, recombination will break down LD between alleles for traits causing assortative mating and those under divergent selection ("selection-recombination antagonism") (Felsenstein 1981). In other words, gene flow will impede the evolution of prezygotic reproductive barriers that keep populations separate. The number and distribution of barrier loci (i.e., loci-causing RI) and the nature of the alleles at these loci can profoundly influence the evolution of prezygotic isolation (Smadja and Butlin 2011).

HOW MANY LOCI CONTRIBUTE TO PREZYGOTIC ISOLATION?

One fundamental question we can ask about prezygotic isolation is, how many loci contribute to its evolution? Genetic architectures involving fewer loci of large effect are expected to be more robust to the homogenizing effects of gene flow than highly polygenic architectures, in which loci have individually small effects and are distributed broadly across the genome. This is because fewer loci offer fewer targets for recombination, and (correlational) selection is concentrated on fewer targets (Gavrilets 2004; Gavrilets and Vose 2007; Yeaman and Whitlock 2011).

Quantitative trait locus (QTL) mapping is one major tool used to identify loci contributing to RI and has improved our understanding of prezygotic isolation (for review, see Arbuthnott 2009; Widmer et al. 2009). However, QTL mapping has a number of well-known limitations. Although generating genetic markers is now relatively straightforward, the large number of phenotyped offspring required to robustly detect QTLs is often difficult, limiting our ability to detect smaller effect QTLs and resulting in upward biases in estimated effect sizes (Beavis et al. 1994). Resulting QTLs may contain hundreds of genes, limiting our ability to estimate the number of mutations underlying traits or to distinguish pleiotropy from linkage (Shahandeh and Turner 2020). Moreover, studies in which no QTLS are identified may remain unpublished, leading to a biased view of effect sizes.

There seems to be variation in the number and effect sizes of loci contributing to prezygotic isolation in both plants and animals, although QTL studies are difficult to compare directly, because of differences in methodology, sample size, and the types of traits targeted. Mating cues and preferences in animals can be polygenic (Chenoweth and Blows 2006; Chenoweth and Mcguigan 2010), but there is also evidence for loci of large effect (Merrill et al. 2019; Xu and Shaw 2019). Similarly, in flowering plants, floral isolation often involves a mix of large effect loci controlling color and scent with numerous small effect loci controlling morphology (Klahre et al. 2011; Yuan et al. 2013; Wessinger and Hileman 2020; Kay and Surget-Groba 2022). Divergence in habitat affinity contributing to ecogeographic isolation or immigrant inviability is likely to be highly polygenic because of the multivariate phenotypes involved (Savolainen et al. 2013; Barghi et al. 2020), but major effect loci have also been identified in some cases (e.g., Colosimo et al. 2005; Selby and Willis 2018).

Although they may provide greater resolution, genome-wide association studies (GWASs) and admixture mapping can also suffer from a lack of power and rely on the availability of naturally occurring variation. Combining mapping with functional tests and population genomic approaches, when feasible, may provide the best opportunity for understanding the genetic architecture of prezygotic isolation (Stinchcombe and Hoekstra 2008; Bomblies and Peichel 2022). Comparative phylogenomic approaches may also provide a useful tool for understanding prezygotic barriers that have repeatedly evolved within a clade (Smith et al. 2020).

WHAT TYPES OF ALLELIC VARIATION CONTRIBUTES TO PREZYGOTIC ISOLATION?

A second key contribution by Felsenstein (1981) was the observation that, regardless of the overall number of loci or the traits or taxa involved, prezygotic isolation must evolve at individual genetic loci via either a "one-allele" or a "two-allele" mechanism (Fig. 1; Felsenstein 1981). As noted
elsewhere (Kopp et al. 2018; Butlin et al. 2021), these terms are often not well-understood, but “the critical distinction . . . is whether reproductive isolation is strengthened by substituting the same or different alleles in the two nascent species” (Felsenstein 1981). These two mechanisms need not act in isolation, and variation in different components of prezygotic isolation, or even individual traits, may involve both one- and two-allele scenarios. Nevertheless, the distinction has important implications for the evolution of prezygotic isolation with gene flow, because when the same

![Diagram](image-url)

**Figure 1.** One-allele versus two-allele mechanisms of prezygotic isolation. (A) Substitutions at multiple loci (which may influence multiple phenotypic traits) can strengthen prezygotic isolation. At each locus, the ancestral allele \(a\) can be replaced by the substitution of a derived allele \(d\). (B) At each individual locus, prezygotic isolation must evolve by either the substitution of the same allele (“one-allele mechanism”) or different alleles (“two-allele mechanism”). In the hypothetical example shown here, alleles at a locus on chromosome 2 influence flower color and cause divergence in the two daughter species, thereby strengthening assortative mating. This can be achieved through the substitution of the same derived alleles (perhaps through the evolution of habitat-induced phenotypic plasticity) or through the fixation of the derived allele in one population. One-allele mechanisms are expected to greatly facilitate the evolution of prezygotic isolation, because there is no requirement for linkage disequilibrium (LD) with other components (such as local adaptation). One- and two-allele mechanisms are not mutually exclusive, and both types of variation can contribute to prezygotic barriers or even the same phenotypes. (RI) Reproductive isolation. (C) Although allelic variation of the one-allele type is often harder to comprehend, examples are potentially widespread and could include alleles for increased choosiness, reduced migration, stronger imprinting, or decreased variance in flowering time, among others. Nevertheless, our ability to detect this type of genetic variation may be limited because of the typical focus on characterizing differences between species (including quantitative trait locus mapping, genome-wide association studies, “genome scan” analyses, etc.). As a result, although one-allele mechanisms are broadly accepted as the easiest route to strengthen prezygotic isolation in the face of gene flow, strong empirical evidence remains limited.
allele strengthens prezygotic isolation in both diverging populations (a one-allele mechanism), the requirement for LD between loci under divergent selection and those increasing assortative mating is sidestepped. Because such alleles will strengthen RI even if they are recombined into the other population, gene flow poses no obstacle to the substitution of alleles that increase isolation.

One-allele mechanisms are broadly accepted as the easiest route to strengthen prezygotic isolation in the face of gene flow (Butlin et al. 2021). Examples are potentially widespread and could include alleles for increased choosiness, reduced migration, stronger imprinting, or decreased variance in flowering time. However, current empirical evidence for an explicit one-allele mechanism of prezygotic isolation is limited to a single experiment in flies. Ortiz-Barrientos and Noor (2005) first mapped within-species variation in female mating discrimination between Drosophila pseudoobscura populations, which are either sympatric or allopatric with respect to the sister species Drosophila persimilis. They then tested for a one-allele assortative mating mechanism by introgressing either strong discrimination (sympatric) or weak discrimination (allopatric) alleles from D. pseudoobscura into D. persimilis. D. persimilis females with the strong discrimination D. pseudoobscura alleles were much less likely to mate with heterospecific males than those with weak discrimination D. pseudoobscura alleles, directly showing that the same sympatric alleles could increase prezygotic isolation in both D. pseudoobscura and D. persimilis (but see Barnwell and Noor 2008 for a failed attempt at replication). These experiments highlight the difficulty of testing for one-allele mechanisms, especially because they rely on proxy ancestral populations, in this case the allopatric D. pseudoobscura, in experimentally tractable organisms.

Because one-allele mechanisms will be missed by typical approaches investigating differences between diverging taxa, understanding these mechanisms poses a significant empirical challenge and represents a major gap in the study of speciation genetics. We suggest using appropriate outgroups to identify derived traits and alleles shared among the ingroup that consistently increase assortative mating. This approach requires a careful understanding of how traits affect mating patterns, because phenotype-naive approaches, such as genome scans, will miss these mechanisms. It may be that traits currently characterized as key innovations that increase speciation rates within clades are essentially one-allele mechanisms. For example, bilateral floral symmetry is associated with more specialized pollination and higher diversification rates (Kay et al. 2006; Yoder et al. 2020). Genetic studies across independent transitions from radial to bilateral floral symmetry have shown similar regulatory changes affecting CYCLOIDEA-like genes (for review, see Hileman 2014), which may function as one-allele mechanisms strengthening reproductive isolation among taxa in these clades. This hypothesis could be tested through manipulations similar to those described above (or indirectly through phenotypic manipulation).

**HOW ARE GENETIC ASSOCIATIONS BETWEEN COMPONENTS OF PREZYGOTIC ISOLATION MAINTAINED?**

Where prezygotic isolation evolves via the substitution of different alleles in the diverging populations (a two-allele mechanism), LD between alleles under divergent selection and those underpinning components of assortative mating must be maintained. One way this can be achieved is if the same traits under divergent selection also contribute to assortative mating. Although such scenarios were considered unlikely (and hence referred to as “magic trait models” [Gavrilets 2004]), it is now apparent that assortative mating traits are frequently under divergent selection (Servedio et al. 2011). For example, the bright wing patterns of Heliconius butterflies contribute to ecological postzygotic isolation, because hybrids with intermediate warning patterns are not recognized as distasteful, but they also act as cues during mate choice (Jiggins et al. 2001; Merrill et al. 2012). Similarly, in cichlid fish, adaptation of the visual sensory system to local environments has been hypothesized to contribute to divergent mate preferences (Seehausen et al. 2008; Maan et al. 2017). Floral isolation will also often fit a magic trait model, because divergent adaptation
to local pollinators will naturally contribute to assortative mating. For example, flower color in monkeyflowers is under divergent selection by local pollinators, which simultaneously contributes to assortative mating (Schemske and Bradshaw 1999; Streisfeld and Kohn 2007).

Associations between traits involved in prezygotic isolation and those under divergent selection may also be maintained through genetic architectures that reduce recombination, such as tight genetic linkage, inversions, or pleiotropy (Maynard Smith 1966; Felsenstein 1981; Smadja and Butlin 2011; Wellenreuther and Bernatchez 2018; Huang and Rieseberg 2020). Evidence exists for these kinds of genetic architectures, largely through QTL mapping studies. For example, Hawthorne and Via (2001) identified loci for host preference and performance in pea aphids that colocalized to the same regions of the genome. These insects mate on their host, providing a rapid path to speciation. Since then, others have reported evidence for physical linkage between loci underlying assortative mating and ecological traits, including in monkeyflowers (Lowry and Willis 2010; Ferris et al. 2017) and Heliconius butterflies (Merrill et al. 2019). An enduring question is whether physical linkage typically facilitates the substitution of coadapted alleles or whether structural rearrangements or recombination suppressors typically increase linkage after allelic substitutions (Charlesworth and Charlesworth 1979; Kirkpatrick and Barton 2006). To address this, one approach might be to examine homologous loci in an outgroup. For example, Hermann et al. (2013) found five tightly linked loci controlling differences in flower color, scent, and morphology in Petunia species adapted to hummingbird versus hawkmoth pollination. By examining the location of these loci in more distantly related relatives, they show the linkage to be unique to Petunia, suggesting that structural rearrangements may have locked in these coadapted alleles (Hermann et al. 2013), although this needs to be confirmed by synteny studies with closer outgroups.

At a phenotypic level, assortative mating can be further characterized as following “trait-preference” rules, in which coordinated divergence in both male and female traits is necessary for assortative mating, or “matching rules,” in which individuals mate with like individuals on the basis of shared traits (Kopp et al. 2018). This has genetic consequences: Whereas distinct male and female traits are likely to be controlled by different loci, phenotype matching will involve shared loci. Under the trait-preference scenario, which may be a common feature of behavioral isolation in animals, the number of genetic associations between loci required for prezygotic isolation to evolve is increased, impeding speciation (Smadja and Butlin 2011). In plants, pollen–pistil incompatibilities may be analogous to animal trait-preference systems. In these situations speciation may be facilitated by genetic architectures, such as tight linkage or pleiotropy, that reduce the dissociation of male and female traits (Pryke 2010; McNiven and Moehring 2013; Merrill et al. 2019; Xu and Shaw 2019). LD between unlinked trait and preference alleles will arise as a natural consequence of nonrandom mating (Kirkpatrick 1982), and if one of these components is subject to divergent selection (a “magic trait scenario”), it will also help overcome the selection-recombination antagonism. However, the strength of LD will depend on the effect size of preference alleles, and LD generated in these scenarios may not be robust to recombination without physical linkage or pleiotropy (Wiley et al. 2011). Nevertheless, compelling data are provided by corn borer moths, for which alleles for pheromone variation and the corresponding preference are found at loci on different chromosomes but remain in strong LD (Unbehend et al. 2021; see also Hench et al. 2019).

When mating follows a matching rule, LD is required between fewer pairs of loci. Flower color is a likely widespread example (e.g., Schemske and Bradshaw 1999), because both male and female components of a hermaphroditic flower share the same signal to attract pollinators; however, divergence in other floral traits may also contribute to matching rules. For example, Kay and Surget-Groba (2022) found QTLs for flower length divergence in two closely related spiral ginger species, which simultaneously determines whether pollen is placed on, and then subsequently retrieved from, either the bill or forehead of the shared hummingbird pollinator. Other examples of matching rules come from habitat or ecological isolation, in which individuals with similar affinities mate be-
cause of spatial proximity and/or phenological overlap. A classic example involves phytophagous insects that mate on their host (Matsubayashi et al. 2010). In sticklebacks adapted to benthic and limnetic habitats within the same lake, Conte and Schluter (2013) revealed phenotype matching by manipulating body size. In the same system, Bay et al. (2017) found that female F2 hybrids mated with males that were similar in body size and shape, and mate choice QTL map to one of the same regions as benthic versus limnetic morphology, which is best explained under a scenario of phenotype matching. Because body size is under divergent selection in these fish, this example also corresponds to a magic trait model. In addition, although divergence in body size involves different alleles (a two-allele mechanism), assortative mating might conceivably be strengthened through the substitution of the same allele in both populations (a one-allele mechanism). As such, stickleback fish nicely demonstrate how these distinct concepts, involving matching rules, magic traits, and one- and two-allele mechanisms, can simultaneously act within a single taxon pair (Fig. 2).

Despite solid theoretical expectations, distinguishing among genetic mechanisms that contribute to prezygotic isolation in natural populations remains difficult. In most cases, it will be necessary to move beyond traditional mapping studies to incorporate gene expression and population genomic and functional genomic studies. For example, gene expression studies across Heliconius species revealed candidate genes underlying a mating preference QTL (Rossi et al. 2020) that were independently implicated as barrier loci through population genomic methods (Laetsch et al. 2022). These results suggest that tight linkage between these candidates and the color pattern gene responsible for mimicry (which was not found to be differentially expressed in the brains of these butterflies) is driving this isolation, rather than pleiotropy. Similarly, in monkeyflowers, a major effect locus ("YUP") controlling pigment deposition (and the presence or absence of nectar guides that contribute to floral isolation) was mapped to a genomic region that also controls other floral traits and hybrid male sterility factors (Bradshaw et al. 1995; Bradshaw and Schemske 2003). YUP was resistant to further fine-scale genetic dissection because it occurs in a region of suppressed recombination (Fishman et al. 2013). Recently, however, Liang et al. (2023) used a combination of near-isogenic line (NIL) construction, RNA sequencing (RNA-seq), RNA interference (RNAi), transformation, complementation tests, confocal fluorescence microscopy, and comparative genomics to show that YUP produces small interfering RNAs (Liang et al. 2023). However, all these approaches depend on the prior identification of target loci, the ability to manipulate large numbers of experimental organisms, and a firm understanding of the phenotypes underlying prezygotic isolation.

HOW IS THE EVOLUTION OF PREZYGOTIC ISOLATION CONSTRAINED BY EVOLUTIONARY HISTORY AND DEVELOPMENT?

Once we identify the genetic basis of traits contributing to prezygotic isolation, an important next step is to investigate the evolutionary history of these variants, which can have significant implications for determining the tempo and mode of speciation. Although classic models generally assume de novo mutation (for review, see Orr 2005), it is now clear that preexisting, standing genetic variation can play an important role (Barrett and Schluter 2008). For example, repeated losses of lateral plates in freshwater stickleback populations were facilitated by existing variation at the Eda locus in ancestral marine populations (Colosimo et al. 2005; see Turbek et al. 2021 for a similar example in birds). Similarly, it is increasingly appreciated that hybridization and introgression can promote divergence via the reassembly of old genetic variants into novel combinations (the "combinatorial view" of speciation; Marques et al. 2019). Recent genomic data provide support for this mechanism, particularly in radiations of Heliconius butterflies, Darwin’s finches, cichlid fishes, and monkeyflowers, where interspecific gene flow seems to have led to the exchange of beneficial alleles, thereby facilitating further divergence (The Heliconius Genome Consortium et al. 2012; Lamichhaney et al. 2015; Stankowski and Streisfeld 2015; Meier et al. 2017). Ancient hybrid-
Figure 2. Three broad classes of conceptual models have been proposed that reduce the number of genetic associations (linkage disequilibrium [LD]) that must be maintained for prezygotic isolation to evolve in the face of gene flow. These models include (A) phenotype matching, in which assortative mating depends on the presence of traits that both sexes have in common (Kopp et al. 2018); (B) one-allele mechanisms, in which prezygotic isolation is strengthened by the substitution of the same allele in the two nascent species (Fig. 1; Felsenstein 1981); and (C) magic trait models, which assume that a trait under divergent selection also contributes to assortative mating (Gavrilets 2004). These scenarios are not mutually exclusive and may simultaneously contribute to the evolution of prezygotic isolation during a single speciation event. Shapes represent different phenotypes involved, in which circles are traits not directly related to mating (on which divergent selection, depicted by arrows, may act), and squares and triangles represent sex-specific mating traits (which may be one and the same in matching scenarios). Brackets represent genetic associations (LD) that must be maintained for prezygotic isolation to evolve when gene flow persists. We assume that allelic substitutions (a for ancestral or d for derived) evolving under a one- or two-allele mechanism influence one of the sex-specific phenotypes, but they might equally influence all components of prezygotic isolation (i.e., here: squares, triangles and circles). Examples are provided for illustration. (i) The same allele has experimentally been shown to strengthen female preference for conspecific males in the sister species *Drosophila subobscura* and *Drosophila persimulans* (Ortíz-Barrientos and Noor 2005) (photo from Darren J. Obbard [obbard.bio.ed.ac.uk/photo_gallery/Drosophila_subobscura.html] and reprinted under the terms of the Creative Commons CC BY-NC 4.0 License). (Legend continues on following page.)
ization may also have triggered entire adaptive radiations by generating new allelic combinations, followed by periods of sorting into distinct ecological environments, as seems to be the case for opsins gene evolution in Lake Victoria cichlids (Meier et al. 2017). Similarly, ancient hybridization is thought to have initiated the rapid evolution of host shifts among races of Rhagoletis flies (Feder et al. 2003).

Despite these examples of the origins of adaptive variation, a deeper understanding of the history of adaptive traits contributing to prezygotic isolation is possible when genetic studies are integrated with the field of evolutionary developmental biology (evo-devo). In particular, because organisms are constructed through genetic programs that unfold sequentially during development, pleiotropy can constrain the genetic changes that contribute to phenotypic evolution. Indeed, it has been argued that adaptation is more likely to proceed through changes in gene regulation, as these mutations are often less likely to incur fitness penalties due to pleiotropy compared to changes in protein-coding sequences (Prud’homme et al. 2007; Stern and Orgogozo 2008).

Although emerging evidence suggests gene regulatory elements may be more pleiotropic than previously thought (Nagy et al. 2018; Preger-Ben Noon et al. 2018; Lewis et al. 2019; Fuqua et al. 2020; Mazo-Vargas et al. 2022), diversification of numerous phenotypic traits, including those involved in prezygotic isolation, have often been linked to changes in gene regulation affecting development rather than mutations in protein-coding regions (Abzhanov et al. 2004; Reed et al. 2011; Martin et al. 2012; Unbehend et al. 2021). In addition, variation in gene regulatory network structure can greatly influence the trajectory of adaptation, potentially resulting in predictable evolutionary outcomes, including the reuse of certain types of mutations or specific genes (Martin and Orgogozo 2013; Sobel and Streisfeld 2013).

One example of how pleiotropy and gene regulatory network organization can impact the genetics of prezygotic isolation comes from flower color transitions causing pollinator isolation. Anthocyanins are common floral pigments responsible for red, pink, blue, and purple flowers (Grotewold 2006). Most plants also produce anthocyanin in vegetative tissues, where they are in-

Figure 2. (Continued) (ii) Interactions between sperm and egg that contribute to prezygotic postmating isolation might conceivably represent a trait-preference scenario, but are unlikely to be under direct divergent selection themselves (photo from Unknown via Wikimedia Commons [commons.wikimedia.org/wiki/File:Sperm-egg.jpg] and reprinted under the terms of the Creative Commons CC0 License). (iii) Divergent selection acting on beak morphology influences song in Darwin’s finches (Podos 2001), which is learned by females. Conceivably, alleles that increase learning ability could spread in both species, thereby strengthening reproductive isolation (photo from Kamster via Wikimedia Commons [commons.wikimedia.org/wiki/File:Evolution_theory_300.jpg] and reprinted under the terms of Creative Commons CC0 License). (iv) In Heliconius cydno and Heliconius melpomene, different alleles determine divergent visual mating preferences for bright warning patterns, which are under divergent selection. In this case, tight linkage between wing pattern and preference alleles is known to help maintain LD (Merrill et al. 2019) (photo from Geoff Gallice via Wikimedia Commons [commons.wikimedia.org/wiki/Category:Photographs_by_Geoff_Gallice] and reprinted under the terms of the Creative Commons CC0 License). (v) Flower color (controlled by different alleles) is under divergent selection by local pollinators, which simultaneously contributes to assortative mating (Schemske and Bradshaw 1999) (photo courtesy of Dena Louise Grossenbacher). (vi) Different species of brood parasitic Vilia finches have evolved a number of adaptations, such as gape coloration, allowing them to parasitize nests of different host species. Both male and female chicks learn the song of their foster parents, which then contributes to assortative mating (Sorenson et al. 2003). Conceivably, the same allele could spread through different species to strengthen the ability to learn, or the strength of preference for different hosts (photo from Jamie et al. 2020) and reprinted under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License). (vii) Pea aphids have adapted to different host plants, on which they mate. LD between alleles for performance and preference are maintained by physical linkage (Hawthorne and Via 2001) (photo from Andy Murray via Wikimedia Commons [commons.wikimedia.org/wiki/File:Pea_aphid__6851672066_.jpg] and reprinted under the terms of the Attribution-ShareAlike 2.0 Generic License).
volved in a variety of physiological responses to stress (Winkel-Shirley 2002). The structural and regulatory components necessary for anthocyanin production are highly conserved, and the network coordinating regulation of the anthocyanin enzymes has become a paradigm for understanding combinatorial gene regulation in plants (Koes et al. 2005). Three types of transcription factors form a multiprotein complex (known as the MBW complex) that regulates features of epidermal cell differentiation, including anthocyanin synthesis (Ramsay and Glover 2005). Among gene families that code for the proteins forming this complex, one (the R2R3-MYBs) contains multiple copies that are known to regulate anthocyanins (Stracke et al. 2001). These duplications result in redundancy of function and generate tissue-specificity in anthocyanin pigmentation. This redundancy implies that each MYB protein in the network has lower connectivity and fewer pleiotropic effects than other members of the MBW complex (Fig. 3A; Sobel and Streisfeld 2013). Indeed, de-
Genetics and the Evolution of Prezygotic Isolation

spite the potential for mutations in numerous genes to generate similar flower color phenotypes, all examples involving divergence in floral anthocyanin pigment intensity between species have been caused by mutations in MYBs (Streisfeld and Rausher 2011). Thus, the organization of the MBW complex and the reduced pleiotropy of MYB mutations appear to determine which genes are most likely to be involved in prezygotic isolation.

The need to maintain functionality at higher levels of biological organization may also determine which mutations contribute to prezygotic isolation. For example, genetic changes in the sensory periphery of animals, particularly protein-coding changes in olfactory/gustatory receptors, have repeatedly been shown to underlie the evolution of behavioral isolation (Leary et al. 2012; Fan et al. 2013; Prieto-Godino et al. 2017; Ahmed et al. 2019; Brand et al. 2020). This could be because changes in chemoreceptor genes have fewer maladaptive effects on neural functioning compared to changes in downstream/central brain circuitry (Fig. 3B). However, it remains uncertain whether these emerging patterns simply reflect bias arising from the experimental tractability of the sensory periphery (Cande et al. 2013; Zhao and McBride 2020), and whether they extend across different sensory modalities. More research is needed to confirm general patterns underlying the evolutionary history of behavioral alleles involved in prezygotic isolation. For example, changes in central neural circuitry downstream of sensory receptors are also involved in the evolution of divergent olfactory-guided mating preferences between species (Seeholzer et al. 2018; Khalaf et al. 2020). In conclusion, like other forms of phenotypic evolution, the genetic source of the traits contributing to RI will probably be the result of evolutionary forces that favor alleles with minimal pleiotropic effects, while maximizing adaptive shifts in a given environment.

**HOW DO PREZYGOTIC BARRIERS AFFECT GENE FLOW?**

Although identifying the loci contributing to prezygotic isolation is important for testing long-standing questions about the genetics of speciation, ultimately we want to know the extent to which prezygotic barriers reduce gene flow between diverging populations. Estimates of the strength of RI are intended to quantify the reduction in potential gene flow between populations (Coyne and Orr 2004; Sobel and Chen 2014). Although measuring the components of RI has proven useful for comparing the prevalence and strength of different types of barriers across systems (Coyne and Orr 1997; Christie et al. 2022), we still know little about how the strength of RI corresponds to reduced gene flow. A lack of any RI should result in genetic homogenization, whereas complete RI should impede all gene flow, allowing populations to evolve independently. However, at intermediate values, as is commonly found between taxon pairs exhibiting at least some geographic overlap, RI may not have a linear relationship with gene flow. Importantly, with incomplete RI, patterns of gene flow and/or divergence across the genome may be complex and vary among neutral loci, loci under divergent selection, and loci linked to selected loci (Nosil and Feder 2012; Cruickshank and Hahn 2014). The genetic architecture of RI and local recombination rates will also determine how and when divergently selected loci remain distinct in the face of gene flow (Schumer et al. 2018; Kautt et al. 2020). Therefore, understanding the relationship between the strength and type of prezygotic barriers and realized gene flow across the genome is critical for understanding the circumstances under which prezygotic isolation will cause speciation.

Pre- and postzygotic isolation may have different impacts on gene flow. Early-acting prezygotic barriers may be more effective at preventing gene flow because they preclude hybrid formation and act early in the life cycle before other barriers can operate (Coyne and Orr 1997; Ramsey et al. 2003). However, they may also vary in strength with the ecological context more so than postzygotic barriers (e.g., Hausmann et al. 2021; Sianta and Kay 2021). In contrast, intrinsic postzygotic barriers are expected to be consistent across environments, but hybrids, especially beyond the F1 generation, may show extensive genotypic, phenotypic and fitness variation, such that overall gene flow may be extensive even when mean hybrid fitness is low (for review, see Rieseberg et al.
In many cases, the impact of prezygotic isolation on gene flow will be confounded with that of postzygotic barriers, either through their independent effects on gene flow in the same taxon pair or by the same traits contributing to both prezygotic and extrinsic postzygotic isolation.

A further challenge to understanding the effects of prezygotic isolation on gene flow is whether patterns of gene flow are related to genetic divergence within and between taxa, but this may be confounded by shared ancestry and within-population characteristics, like population size and mating system. Few studies directly assess migration rates or, even more importantly, variation in migration rates across the genome. With whole-genome data, demographic modeling can be used to estimate a variety of population parameters, including directional migration rates and variation in migration rate across the genome. However, the field has yet to coalesce around a standard approach (for review, see Westram et al. 2022). Analytical methods for assessing gene flow also vary widely. Many studies that make conclusions about gene flow only report patterns of genetic divergence within and between taxa, but this may be confounded by shared ancestry and within-population characteristics, like population size and mating system. Few studies directly assess migration rates or, even more importantly, variation in migration rates across the genome. With whole-genome data, demographic modeling can be used to estimate a variety of population parameters, including directional migration rates and variation in migration rate across the genome (e.g., Gutenkunst et al. 2009; Excoffier et al. 2021; Laetsch et al. 2022). However, the field has yet to coalesce around a standard approach (for review, see Westram et al. 2022). Thus, deposition of raw data into public repositories is key for future comparative analyses of the complex relationships between gene flow and RI.

One way to circumvent the variability in sampling and analysis may be to assess both directional gene flow and directional estimates of prezygotic isolation within the same study (e.g., Sundqvist et al. 2016). For example, focusing on interspecific pollen transfer in plants, Moreira-Hernández and Muchhala (2019) compared the strength and asymmetry of prezygotic barriers to the predominant direction of gene flow for 10 systems. They found that only four out of 10 systems showed concordance in asymmetry direction between pollen transfer barriers and gene flow, three showed contradictory patterns, and the rest were mixed. In two of the three cases in which patterns did not match, postzygotic barriers were responsible for the observed pattern of gene flow. Another example comes from sunflowers, in which Sambatti et al. (2012) found substantial gene flow between the two focal species despite very high prezygotic isolation and nearly complete cumulative RI. The directionality of gene flow does not match symmetry in prezygotic isolation, again potentially because of opposing asymmetry in postzygotic isolation. These examples highlight the importance of considering both pre- and postzygotic barriers in the same system to disentangle their impacts on gene flow. Future investigation into the efficacy of prezygotic barriers could benefit from focusing on systems in which postzygotic isolation is known to be weak or absent.

The complementary viewpoint to asking how prezygotic isolation relates to gene flow is to ask whether patterns of gene flow or genetic divergence can help us understand the traits or loci contributing to reproductive isolation. Genome scans can help identify loci under divergent selection that may contribute to prezygotic isolation and that conform to two-allele models of RI, although they must be interpreted carefully to avoid confounding factors and false positives (Margalida et al. 2018). Heterogeneous patterns of genomic divergence have been found in a variety of systems (e.g., Martin et al. 2013; Poelstra et al. 2014; Malinsky et al. 2015; Marques et al. 2016; Vijay et al. 2016; Riesch et al. 2017; Westram et al. 2018), but rarely do we have corresponding information on the traits contributing to RI to connect with these patterns (but see Stankowski et al. 2023 for an attempt to do this).

CONCLUSIONS

Despite decades of progress in understanding the genetics of prezygotic isolation, we see opportunities for future advances. Improved genetic, genomic, and phenotyping technologies will allow finer dissection and functional characterization of prezygotic isolation, providing answers to basic questions about its genetic architecture across varied systems and helping us move beyond difficult-to-compare QTL studies (see also Delmore et al. 2023). We highlight a gap between the theoretical
expectation that one-allele mechanisms provide the easiest route for prezygotic isolation to evolve in the face of gene flow, and our ability to detect this type of genetic variation with prevailing approaches that characterize differences between species. We also see opportunities for further integration of evo-devo with speciation genetics. Understanding the developmental programs in which prezygotic isolating traits are embedded will lead to better predictions about constraints on their evolution. In addition, there remains much work to be done to understand how both pre- and postzygotic isolation shape gene flow and patterns of genetic divergence across the genome. The increasing accessibility of whole genome sequencing and development of computational approaches to explicitly estimate migration rates will make it possible to compare the strength of RI with levels of gene flow across multiple taxa and types of isolating barriers. Finally, we included examples across plants and animals, in the hope of better integrating our understanding of prezygotic isolation. We believe this highlights common principles despite divergent biologies. Whether there are substantial differences in the genetics of prezygotic isolation among taxonomic groups independent of methodological biases remains an open question.

OPEN QUESTIONS FOR FUTURE RESEARCH?

- How many loci, at the level of individual mutations, contribute to the evolution of prezygotic isolation? How are they distributed across the genome? And how does this vary with respect to taxonomic group, levels of gene flow, and the type selection driving divergence?
- How important are the different mechanisms that can overcome the fundamental constraint of recombination? In particular, how common are one-allele mechanisms in nature?
- Are physical linkage, and other recombination suppressors a cause or consequence of speciation? Can we distinguish between tight linkage and pleiotropy?
- How does development, and the need to maintain functionality at higher levels of biological organization, constrain the evolution of prezygotic isolation?
- How do prezygotic barriers shape patterns of genetic divergence across the genome, and is this different from patterns associated with postzygotic barriers?

AUTHOR CONTRIBUTIONS

All authors contributed to conceptual development, writing, and generating figures.

ACKNOWLEDGMENTS

We thank A. Holtz for citation management. The authors were funded by the following sources while working on this project: H.A.C. by Australian Research Council DP190103039; M.R. and R.M.M. by DFG GZ: ME 4845/1-1; R.M.M. by ERC Starter Grant 851040; K.M.K. and J.G.H. by National Science Foundation (NSF) DEB 1737889; M.A.S. by NSF DEB 2051242; and A.F.F. by Swiss National Science Foundation Postdoc Mobility Grant 203023.

To produce the first working draft of this manuscript in Spanish, which is available as Supplemental Material, we used the free version of DeepL to translate it from English.

REFERENCES

Arnold ML, Ballerini ES, Brothers AN. 2012. Hybrid fitness, adaptation and evolutionary diversification: lessons
Kautt AF, Hernández RD, Williamson SH, Bustamante CD. 2009. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLOS Genet* 5: e1000695. doi:10.1371/journal.pgen.1000695


R.M. Merrill et al.


Genetics and the Evolution of Prezygotic Isolation

Richard M. Merrill, Henry Arenas-Castro, Anna F. Feller, Julia Harencár, Matteo Rossi, Matthew A. Streisfeld and Kathleen M. Kay

Cold Spring Harb Perspect Biol published online October 17, 2023

Subject Collection Speciation

Genomic Approaches Are Improving Taxonomic Representation in Genetic Studies of Speciation

Mechanisms of Intrinsic Postzygotic Isolation: From Traditional Genic and Chromosomal Views to Genomic and Epigenetic Perspectives
Radka Reifová, S. Lorena Ament-Velásquez, Yann Bourgeois, et al.

Quantitative Analyses of Coupling in Hybrid Zones
Thomas J. Firneno, Jr., Georgy Semenov, Erik B. Dopman, et al.

Favored Races in the Struggle for Life: Racism and the Speciation Concept
Joseph L. Graves, Jr.

Genetics and the Evolution of Prezygotic Isolation

Anthropogenic Change and the Process of Speciation
Murielle Álund, Meredith Cenzer, Nicolas Bierne, et al.

The Impact of Chromosomal Rearrangements in Speciation: From Micro- to Macroevolution
Kay Lucek, Mabel D. Giménez, Mathieu Joron, et al.

For additional articles in this collection, see http://cshperspectives.cshlp.org/cgi/collection/