RESEARCH ARTICLE SUMMARY

SEXUAL SELECTION

Sexual selection promotes reproductive isolation in barn swallows

Drew R. Schield*, Javan K. Carter, Elizabeth S. C. Scordato, Iris I. Levin, Matthew R. Wilkins, Sarah A. Mueller, Zachariah Gompert, Patrik Nosil, Jochen B. W. Wolf, Rebecca J. Safran

INTRODUCTION: Assessing how sexual traits and their genetic basis contribute barriers to gene flow in secondary contact due to effects on hybrid fitness remains critical to establishing a causal role of sexual selection in in speciation. Leveraging natural systems with intraspecific variation in sexual traits at early stages of the speciation process holds promise for identifying links between patterns of phenotypic and genomic variation and the evolution of reproductive isolation.

RATIONALE: Observational and experimental studies indicate that barn swallows (*Hirundo rustica*) are a robust empirical model of divergent sexual selection and that the presence of multiple hybrid zones between populations in Eurasia enables investigation of barriers to gene flow. We investigate genotypic and phenotypic variation in barn swallows to (i) map the genetic basis of plumage traits used in sexual signaling, (ii) test whether loci underlying sexual traits have experienced divergent sexual selection in allopatry and present barriers to gene flow in secondary contact, and (iii) test the prediction that sexual selection has maintained linkage disequilibrium (LD) between barrier loci in secondary contact as a result of their effects on hybrid fitness.

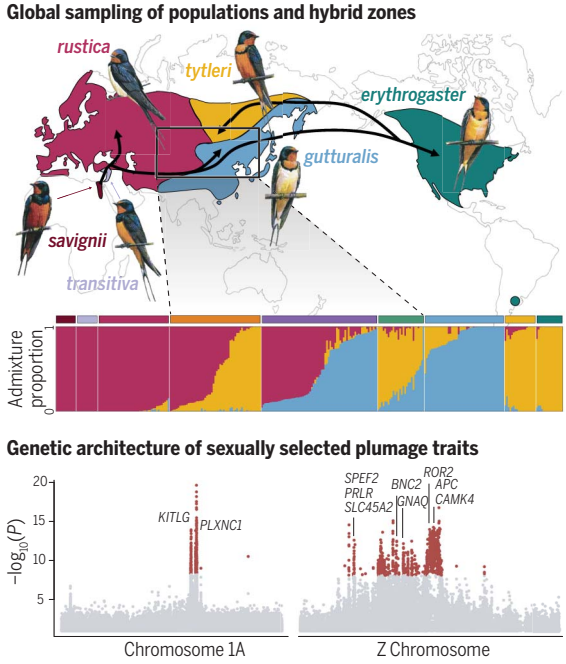
RESULTS: We sequenced the genomes of 336 barn swallows sampled across the breeding distribution of the species and quantified variation in ventral coloration and tail streamer length, two signal traits used in mate choice. Populations differ in these traits and hybrids between the subspecies *rustica*, *tytleri*, and *gutturalis* exhibit phenotypes that are intermediate between or similar to parental populations. The genetic architecture of ventral color is concentrated on chromosome 1A and the Z chromosome whereas phenotypic variation is largely explained by genotypic variation at 10 loci, including the melanogenesis genes KITLG, SLC45A2, and BNC2. Variation in tail streamer length is explained by loci on chromosome 2. Sexual trait loci—ventral color loci in particular—exhibit peaks of high differentiation between populations and signatures of divergent positive selection in allopatry. We further investigated whether loci under divergent sexual selection contribute barriers to gene flow in secondary contact using geographic and genomic cline analyses across hybrid zone transects, finding that sexual trait loci constitute barriers in the *rustica-tytleri* and *rustica-gutturalis* hybrid zones whereas gene flow is less constrained across the remainder of the genome. Clines for sexual trait loci in these hybrid zones also show a high degree of concordance, consistent with selection for specific combinations of alleles from parental populations in hybrids. Finally, we tested whether selection has generated ongoing coupling of barrier loci by investigating LD patterns in hybrid zones. These tests reveal elevated LD among sexual trait barrier loci in hybrids beyond what is expected under admixture alone, consistent with the genetic coupling of barriers being an emergent property of divergent sexual selection.

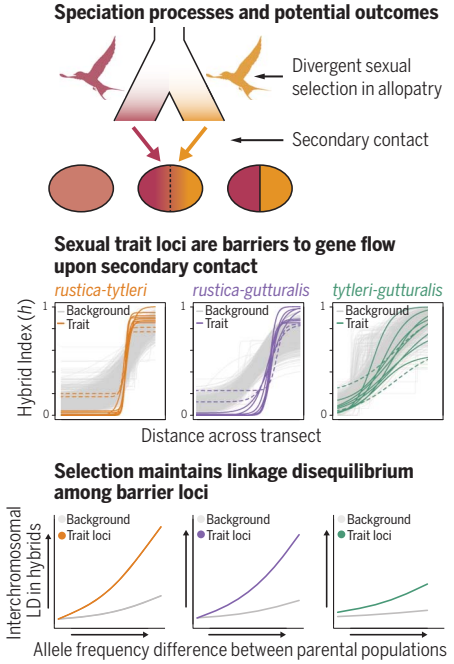
CONCLUSION: Our findings demonstrate an important role for sexual selection in speciation through the analysis of the genomic basis of sexual signal traits in barn swallows, evidence for divergent selection in geographic isolation, and evidence that loci underlying traits involved in prezygotic isolation represent barriers to gene flow. Our results further support the conclusion that the genetic coupling of sexual trait loci generated by selection promotes reproductive isolation upon secondary contact. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: drew.schield@virginia.edu
Cite this article as D. R. Schield *et al.*, *Science* **386**, eadj8766 (2024). DOI: 10.1126/science.adj8766

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adj8766

Sexual signal traits form barriers to gene flow upon secondary contact. Barn swallows breed across nearly the entire North Hemisphere and exhibit variation in sexual signal traits across populations. Hybrid zones between these populations enable the identification of the genetic basis of sexual traits and tests of the hypothesis that divergent sexual selection promotes reproductive isolation. Genetic loci underlying sexual traits show signatures of divergent selection between allopatric populations and are barriers to gene flow in secondary contact, whereas gene flow is less constrained across the rest of the genome.





RESEARCH ARTICLE

SEXUAL SELECTION

Sexual selection promotes reproductive isolation in barn swallows

Drew R. Schield[1,2]*, Javan K. Carter[1], Elizabeth S. C. Scordato[3], Iris I. Levin[4], Matthew R. Wilkins[1,5], Sarah A. Mueller[6], Zachariah Gompert[7], Patrik Nosil[8], Jochen B. W. Wolf[6], Rebecca J. Safran[1]

Despite the well-known effects of sexual selection on phenotypes, links between this evolutionary process and reproductive isolation, genomic divergence, and speciation have been difficult to establish. We unravel the genetic basis of sexually selected plumage traits to investigate their effects on reproductive isolation in barn swallows. The genetic architecture of sexual traits is characterized by 12 loci on two autosomes and the Z chromosome. Sexual trait loci exhibit signatures of divergent selection in geographic isolation and barriers to gene flow in secondary contact. Linkage disequilibrium between these genes has been maintained by selection in hybrid zones beyond what would be expected under admixture alone. Our findings reveal that selection on coupled sexual trait loci promotes reproductive isolation, providing key empirical evidence for the role of sexual selection in speciation.

Sexual selection is predicted to play a significant role in the formation of new species (*1*–*5*). Indeed, sexually selected traits have been inferred to promote reproductive isolation (*6*–*11*) and to be associated with diversification rates (*12*, *13*). Studies often examine the role of sexual selection in speciation in two ways: either by establishing links between sexual traits and reproductive isolation, or by characterizing genomic divergence and genetic associations with these traits (*6*, *7*, *11*, *14*–*19*). These approaches broadly suggest that sexual selection can promote speciation, but they do not address the fundamental assumption that divergent sexual traits form barriers to gene flow upon secondary contact (Fig. 1, processes and potential outcomes) (*3*, *20*, *21*). Tests of how sexual traits and their underlying genetic basis affect gene flow remain critical in establishing a causal role of sexual selection in speciation. Such studies are especially needed in natural systems with intraspecific variation in sexual traits at early stages of the speciation process, in which clear links between patterns of phenotypic and genomic variation and evolutionary processes can be identified (Fig. 1, A to C, predicted patterns).

Early in the speciation process, the genetic basis of reproductive isolation due to sexual selection may be concentrated in a small number of genomic regions in which selection can withstand the homogenizing effects of gene flow (barrier loci) (*22*, *23*). Upon secondary contact, genetic coupling—the buildup and maintenance of linkage disequilibrium (LD) between genomic barrier loci—may be critical for speciation to proceed (*24*–*26*). Genetic coupling exposes loci to the combined effects of both direct selection and indirect selection at other coupled barrier loci. As more barriers become coupled, the influence of indirect selection increases and the total barrier to gene flow strengthens and becomes more genomically widespread (*26*–*28*). Coupling is likely to build up as a consequence of accumulated allele frequency differences between allopatric populations when divergence occurs during periods of geographic isolation (*27*, *29*), with the potential to promote reproductive isolation when populations experience secondary contact if linkage disequilibria among barrier loci are maintained or strengthened. In the case of speciation driven by sexual selection, we expect that the strongest signal of LD maintained by selection should be related to traits important for sexual competition and mate choice. Theory postulates that the coupling of barrier loci may be a powerful mechanism in the evolution of strong reproductive isolation (*26*, *27*) and testing this prediction in the context of the genetic basis of sexual traits is critical for understanding how sexual selection contributes to the speciation process in the presence of gene flow.

We investigate genotypic and phenotypic variation in sexual signal traits in barn swallows (*Hirundo rustica*) to test predictions of the role of sexual selection in speciation (Fig. 1). Divergent sexual selection on signal traits (melanin-based ventral coloration and tail streamer length) has been demonstrated previously (*30*–*36*) and populations exhibit geographic variation in these signals and associated trait-based preferences (*37*–*39*) despite diverging from a common ancestor very recently (~11 thousand years ago) (*40*, *41*). Critically, evidence for divergent sexual selection in barn swallows comes from studies of natural populations in which variation in these sexual traits and their reproductive outcomes have been studied both observationally and experimentally through phenotype manipulations (*32*, *30*–*36*). A recent meta analysis of these studies provides additional evidence that barn swallows represent a robust empirical example of divergent sexual selection (*39*).

The presence of multiple Eurasian hybrid zones between parental populations with varying degrees of divergence in sexual traits (*42*–*44*) further enables tests to determine whether sexual selection leads to reproductive isolation upon secondary contact. Two of these hybrid zones lie at migratory divides where populations that diverged in allopatry (*40*) share overlapping breeding habitat and take divergent routes to nonbreeding habitats in Africa and southern Asia (fig. S2) (*43*, *44*). Previous studies across these hybrid zones revealed associations between transitions in genome-wide ancestry and ventral plumage color, and also between plumage traits and migratory behavior (*43*, *44*). Further, populations breeding at migratory divides exhibit less extensive hybridization than in hybrid zones without migratory divides (*43*), raising the possibility that divergent sexual selection promotes reproductive isolation between populations that have substantial differences in migratory direction and distance (i.e., due to low survival of hybrids that take nonadaptive migratory routes).

We leverage this natural system and the rich background information on sexual selection in barn swallows to map the genetic architecture of sexually selected traits and to test whether trait loci experience divergent sexual selection in allopatry (Fig. 1A) and promote reproductive isolation in secondary contact (Fig. 1B). We further test the prediction that sexual selection has maintained LD among barrier loci in secondary contact as a result of their effects on hybrid fitness (Fig. 1C and fig. S1). This series of analyses reveals that reproductive isolation is mediated by the genetic coupling of barrier loci underlying sexually selected traits, providing critical support for the hypothesis proposing a role of sexual selection in speciation.

Results

We sequenced the genomes of 336 barn swallows sampled across the Holarctic breeding distribution of the species (Fig. 2A, fig. S2A, and data S1), yielding 9,565,797 genome-wide

[1]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, USA. [2]Department of Biology, University of Virginia, Charlottesville, VA, USA. [3]Department of Biological Sciences, California State Polytechnic University, Pomona, CA, USA. [4]Department of Biology, Kenyon College, Gambier, OH, USA. [5]Galactic Polymath Education Studio, Minneapolis, MN, USA. [6]Division of Evolutionary Biology, Faculty of Biology, Ludwig Maximilian University of Munich, Munich, Germany. [7]Department of Biology, Utah State University, Logan, UT, USA. [8]CEFE, Université Montpellier, CNRS, EPHE, IRD, Montpellier, France.
*Corresponding author. Email: drew.schield@virginia.edu

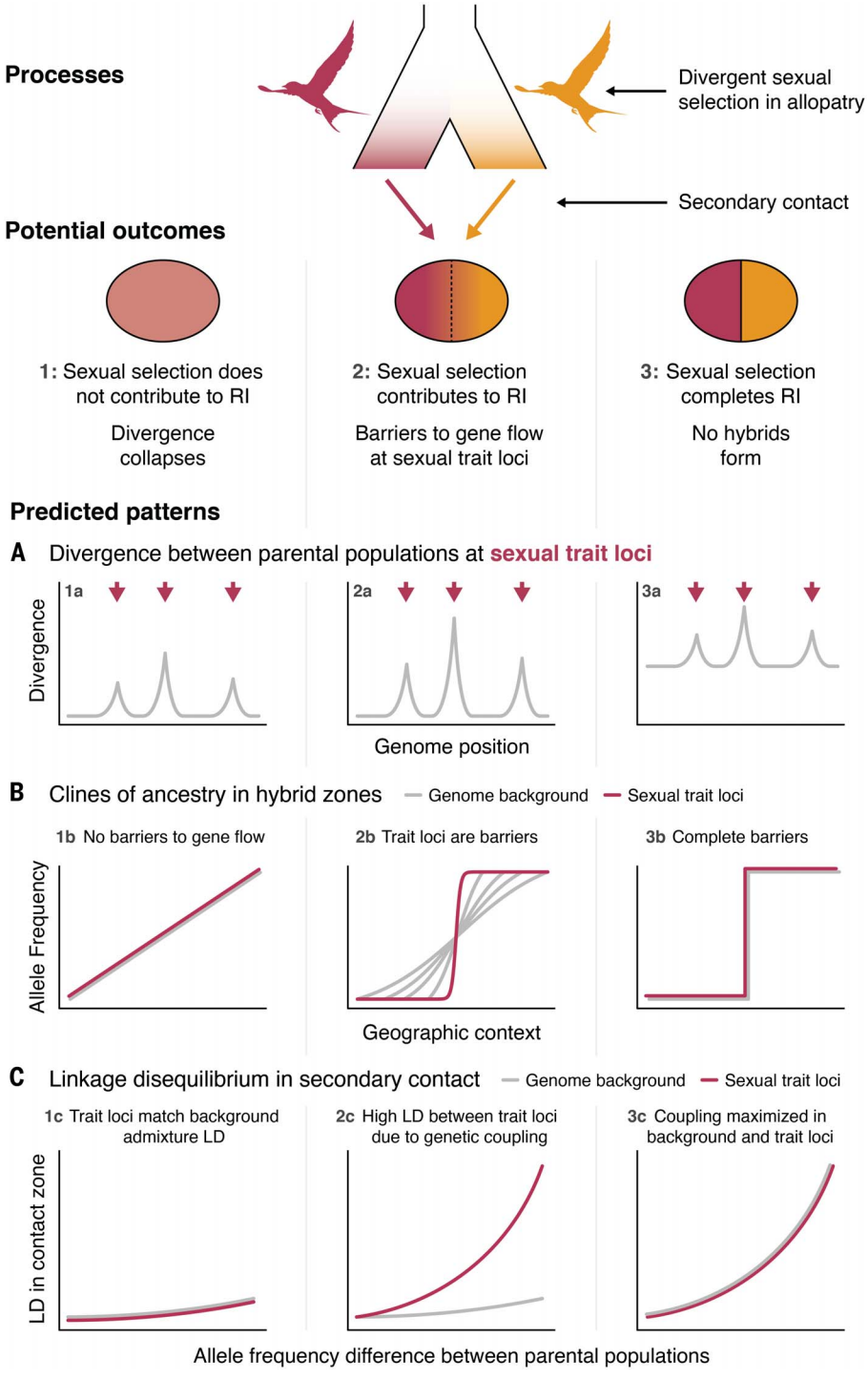


Fig. 1. Conceptual overview of the study. We investigate the roles of sexually selected traits and their genetic architectures in speciation and test predictions of genetic coupling (*24*, *26*) of sexual trait loci. We consider the process of divergent sexual selection in allopatry and potential outcomes when populations experience secondary contact, namely that a maintenance of associations between sexual trait loci through genetic coupling promotes reproductive isolation. One extreme could be that there is no reproductive isolation (RI) due to sexual selection in secondary contact, leading to complete homogenization of populations (left). The other extreme could be complete RI as a consequence of divergent sexual selection (right). A third potential outcome is that loci underlying sexual traits under divergent selection are barriers whereas the rest of the genome experiences gene flow (center). In the lower panels we show predicted patterns of (**A**) allopatric divergence, (**B**) gene flow in secondary contact, and (**C**) LD under the alternative scenarios with respect to sexual trait loci and the genome background. In the case in which sexual selection contributes to RI, sexual trait loci will present barriers to gene flow and have elevated levels of LD due to genetic coupling.

single-nucleotide polymorphisms (SNPs). Our sampling includes populations of the six subspecies and three hybrid zone transects in Eurasia between parental *H. r. rustica*, *H. r. tytleri*, and *H. r. gutturalis* populations (fig. S2B). Principal component analysis and model-based ancestry estimates support three major genetic clusters (*42*, *43*, *45*): One includes the subspecies *rustica*, *savignii*, and *transitiva*; a second includes *tytleri* and *erythrogaster*; and a third includes *gutturalis*. Individuals from hybrid zones have intermediate principal component loadings and admixture proportions between parental genetic clusters, consistent with mixed ancestry (*42*, *44*, *45*) (Fig. 2B and figs. S3 and S4). Parental populations of *rustica*, *tytleri*, and *gutturalis* differ in ventral plumage color and tail streamer lengths (Fig. 2C, fig. S2, C and D, fig. S5, and tables S1 and S2); *rustica* has light breast feathers (mean and standard deviation percent reflectance = 50.8 ± 9.6) and long tail streamers (mean = 101.3 ± 12.7 mm), *tytleri* has dark feathers (21.8 ± 7.8) and long streamers (100.5 ± 12.9 mm), and *gutturalis* has light color (though darker than *rustica*; 46.2 ± 10.8) and short streamers (90.9 ± 9.5 mm). Plumage trait distributions in hybrid zones are intermediate between (or similar to) those of parental populations (Fig. 2C and table S2) and ventral color and tail streamer lengths are not significantly correlated across populations (fig. S5; $P = 0.11$).

Comparisons of ancestry and interspecific heterozygosity at informative SNPs (i.e., $F_{ST} > 0.6$) reveal that our current sampling of each hybrid zone includes multiple hybrid classes, consisting primarily of backcrossed, late-generation hybrids (fig. S6). Demographic models support the idea that the hybrid zones formed upon secondary contact between parental populations following periods of allopatric divergence (figs. S7 to S11, tables S3 and S4, and data S2). Estimates of the timing of secondary contact indicate that the *rustica-gutturalis* hybrid zone formed roughly 2000 generations ago whereas hybrid zones with *tytleri* formed more recently (<1000 generations).

Below, we present results integrating genome-wide SNPs and phenotypic variation to characterize the genetic architecture and selective history of sexual traits and to test whether genetic coupling of trait-associated barrier loci promotes reproductive isolation. At each stage of our procedure (fig. S1), we test core predictions of the conceptual framework detailed in Fig. 1 to link patterns to processes related to the role of sexual selection in speciation, with analyses focused on populations of parental subspecies and hybrid zones in Eurasia.

The genetic architecture of sexual traits

We used genome-wide association mapping to investigate the genetic basis of ventral color

and tail streamer length, taking advantage of matched genome and phenotype data for individuals with mixed genetic ancestry. Analysis of hybrids is critical for mapping differences in phenotype between parental populations and avoids confounding factors stemming from population structure. Models using Bayesian inference provide quantitative estimates that both traits are controlled by oligogenic architectures, with high proportions of phenotypic variation explained (PVE) by SNP genotypes (mean ± Bayesian standard error ventral color PVE = 0.97 ± 0.065, tail streamer PVE = 0.95 ± 0.09; Fig. 3, A and B, and table S5) and considerable proportions of genetic effects (PGE) due to a small number of individual SNPs with measurable effects (ventral color PGE = 0.79 ± 0.13 explained by 12 SNPs, tail streamer PGE = 0.49 ± 0.12 explained by two SNPs; Fig. 3, A and B); the remaining 21 and 51% of genetic effects, respectively, are due to joint contributions of near-infinitesimal genome-wide effects. These estimates indicate that trait variation is primarily controlled by a handful of loci as opposed to having a highly polygenic basis and are consistent with estimates of heritable and environmental components of barn swallow plumage (*46*). Per-SNP posterior inclusion probabilities (PIPs) from these models indicate that SNPs with measurable effects on ventral color are concentrated on chromosome 1A and the Z chromosome whereas high-PIP SNPs for tail streamer length are on chromosome 2 (Fig. 3, C and D, and fig. S12).

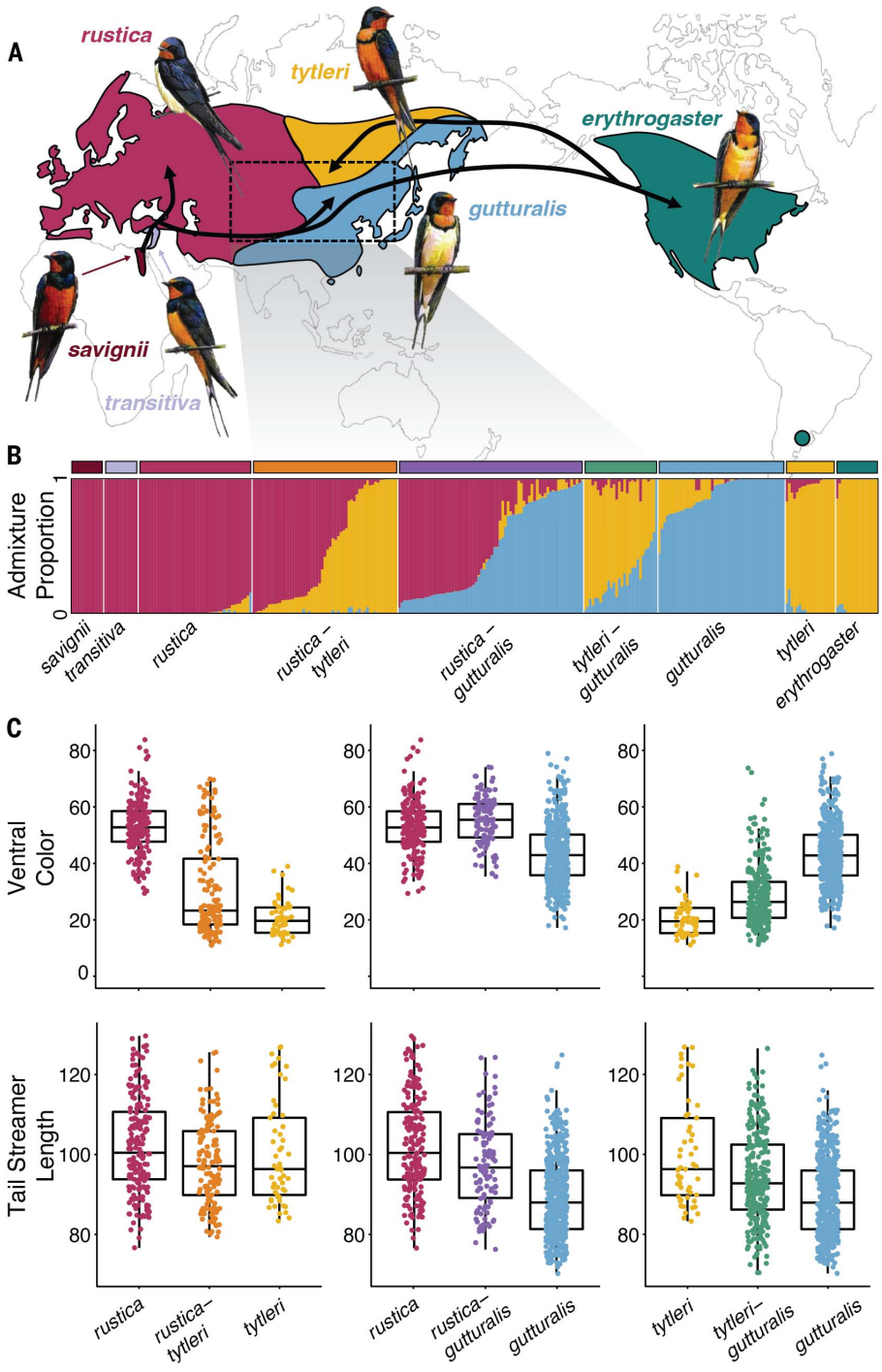


Fig. 2. The barn swallow study system, population genetic structure and composition of hybrid zones, and variation in sexual traits. (**A**) Breeding range of the barn swallow (*Hirundo rustica*). Shaded regions correspond to the breeding distributions of the six subspecies. Arrows depict the hypothesized barn swallow biogeographic history (*40*). (**B**) Population genetic structure inferred using ADMIXTURE under a $K = 3$ model, where vertical lines show individual admixture proportions between one or more genetic clusters. Horizontal bars at the top of the panel correspond to shaded regions in (A), with additional colors for hybrid zones. (**C**) Distributions of ventral color (measured as breast average brightness; percent reflectance) and tail streamer length (millimeters) in parental and hybrid zone populations. Barn swallow illustrations by Hilary Burn and used with permission.

Linear model tests of association between SNPs and ventral color further reveal multiple significantly associated regions on chromosome 1A and the Z chromosome containing genes with known or tentative roles in melanogenesis and feathering phenotype (Fig. 3, C and D, figs. S13 to S20, tables S6 and S7, and data S3). Associations on chromosome 1A contain *KITLG* and *PLXNC1*. *KITLG* is notable because it controls melanin patterning and has been implicated in previous studies of vertebrate pigmentation (*47*, *48*). Associations on the Z chromosome contain *SLC45A2*, a transporter protein that supports melanin synthesis (*49*), and *BNC2*, which has been linked to pigment pattern and saturation (*50*), among other candidate melanogenesis genes differentially expressed in melanocytes of hooded and carrion crows (*51*). The concentration of ventral color associations on the Z chromosome aligns with predictions that sex linkage is favorable for traits involved in mate choice and sexual isolation (*52*, *53*); it also aligns with previous evidence consistent with sexual selection driving pronounced Z-linked differentiation in barn swallows (*45*).

Associations with tail streamer length are in two regions on chromosome 2 (Fig. 3F, fig. S13, and figs. S21 and S22). One is upstream of *ICE1* and downstream of a long noncoding RNA (lncRNA) and the other contains *PDE1C*. Functional links between *ICE1* and *PDE1C* and feather growth have not been demonstrated; however, the enzymes encoded by both genes have regulatory functions that could influence tail streamer development through positive

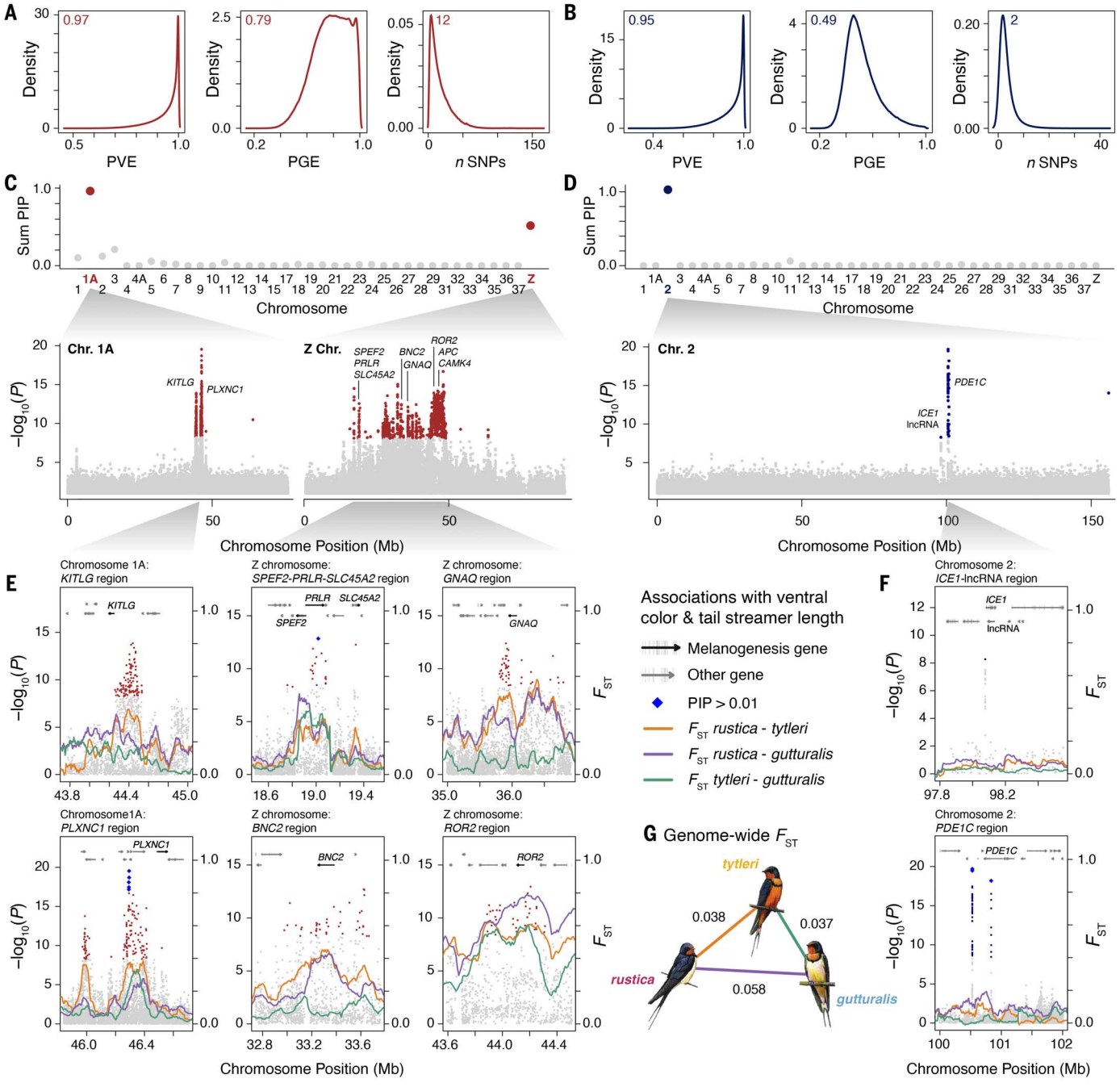


Fig. 3. Genetic architecture of sexual traits. (**A**) Posterior mean density distributions of hyperparameters from BSLMM, including the proportion of variation in ventral color explained by all SNPs in the model (PVE), the proportion of the PVE explained by SNPs with detectable effects (PGE), and the number of SNPs in the model (*n* SNPs). (**B**) Posterior distributions of BSLMM hyperparameters for tail streamer length. (**C**) Sum of SNPs with posterior inclusion probability (PIP) ≥ 0.01 in the BSLMM for ventral color across chromosomes, showing a concentration of high-PIP SNPs on chromosome 1A and the Z chromosome. (**D**) Sum of PIP ≥ 0.01 for tail streamer length concentrated on chromosome 2. Lower panels of (C) and (D) show SNP associations, $-\log_{10}(P)$, with ventral color across chromosome 1A and the Z chromosome and tail streamer length across chromosome 2. Significant associations (Bonferroni-corrected) are shaded in dark red and blue, respectively. The locations of candidate melanogenesis genes in ventral color associations and genes containing or nearby tail streamer associations are labeled. (**E**) Details of associations with ventral color, showing variation in $-\log_{10}(P)$ (points; left *y*-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right *y*-axis). Significant associations are shaded in dark red and SNPs with PIP ≥ 0.01 are indicated with blue diamonds. Genes and their orientation are shown as arrows with exon coordinates shown as vertical lines, and candidate melanogenesis genes are labeled. (**F**) Details of associations with tail streamer length, with significant associations shaded in dark blue. Additional details for panels (E) and (F) are shown in figs. S14 to S22. (**G**) Mean genome-wide F_{ST} between parental *rustica*, *tytleri*, and *gutturalis* populations.

regulation of transcription, intracellular transport, and growth factor stability (*54*). Details of all associations are provided in figs. S13 to S22, tables S6 and S7, and data S3.

Evidence for divergent selection at sexual trait loci

Next, we tested whether genomic regions associated with the mate choice traits ("trait loci" hereafter) are relevant to fitness based on signatures of divergent selection between allopatric populations, as predicted based on previous studies of sexual selection in barn swallows (*31*, *32*, *36*, *38*, *39*, *43*). Trait loci show high levels of genetic differentiation (F_{ST}) between one or more pairs of parental populations (mean F_{ST} = 0.23 to 0.39; Fig. 3, E and F), contrasting sharply with shallow genome-wide F_{ST} (mean F_{ST} = 0.037 to 0.058; Fig. 3G, figs. S23 and S24, and tables S8 and S9; $P < 0.0001$). Genome-wide patterns are consistent with the effects of linked selection in parental populations as well as in ancestral populations prior to divergence (i.e., recurrent selection; fig. S25 and table S10), indicated by negative correlations between F_{ST} and both recombination rate ($\rho = -0.44$ to 0.57, $P < 2.2\ 10^{-16}$) and within-population nucleotide diversity (π; $\rho = -0.41$ to 0.55, $P < 2.2 \times 10^{-16}$) across pairs of parental populations. F_{ST} and between-population sequence divergence (d_{xy}) are also negatively correlated ($\rho = -0.4$ to 0.52, $P < 2.2 \times 10^{-16}$), as d_{xy} is roughly equivalent to π because of extremely recent divergence (figs. S7 and S23). Given these broad patterns, elevated F_{ST} in trait loci could be driven solely by background selection in regions with reduced recombination (*55*), as is suggested by correlation of genome-wide F_{ST} among pairs of parental populations (table S10; $\rho = 0.45$ to 0.66, $P < 2.2 \times 10^{-16}$) and sharing of most discrete F_{ST} peaks by at least two population pairs (fig. S23). However, trait loci exceed F_{ST} values even in regions with extremely low recombination rates (e.g., centromeres; fig. S26; $P < 2.2 \times 10^{-16}$), consistent with the hypothesis that trait loci are highly differentiated as a result of divergent positive selection rather than attributable to the effects of background selection alone.

We measured additional parameters to specifically test for signatures of divergent selection in trait loci. We find multiple signals of positive selection at trait loci in one or more populations (Fig. 4 and figs. S27 to S35), including elevated allelic differentiation measured using population branch statistics (PBS), reductions in π and Tajima's *D*, and extended

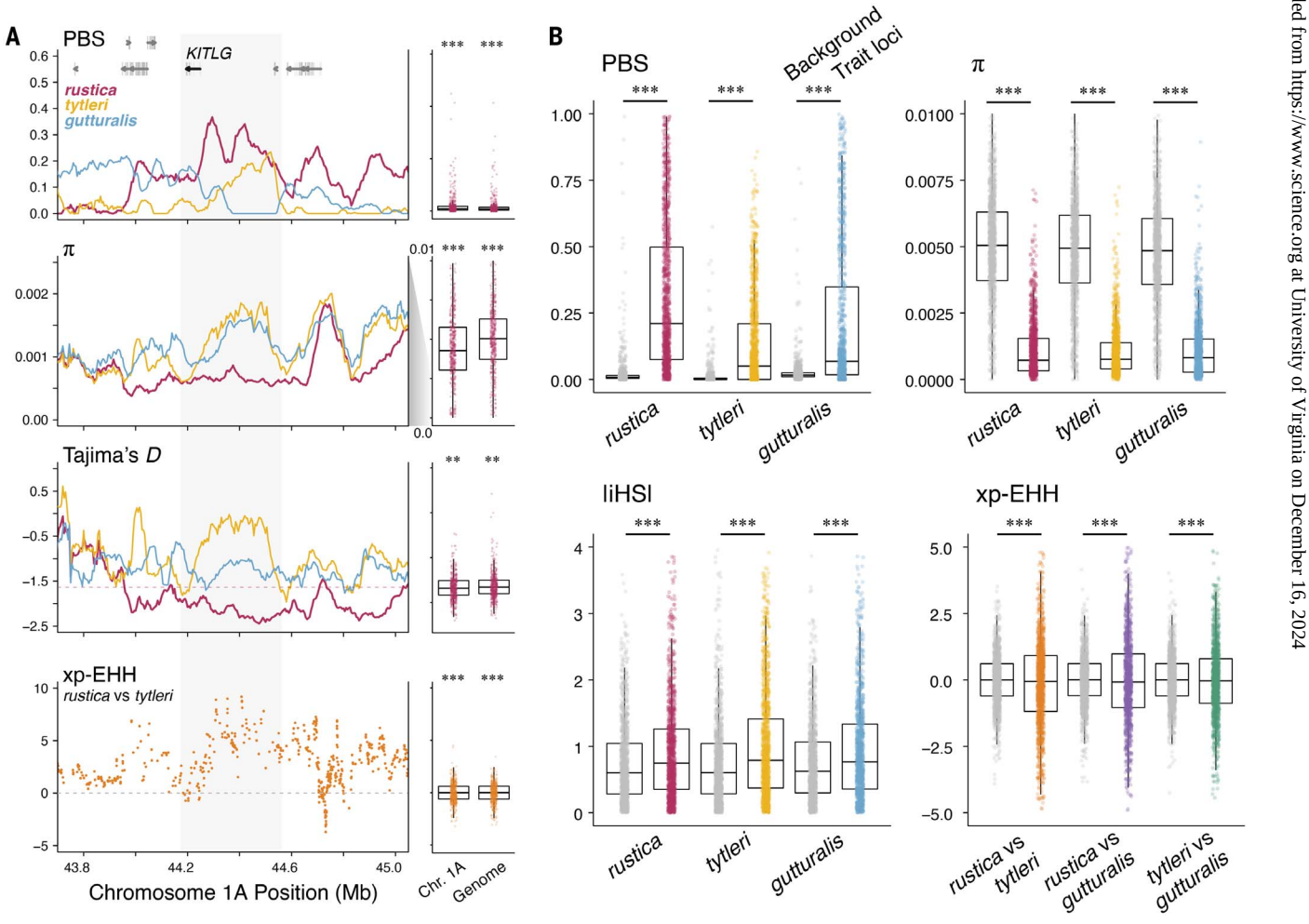


Fig. 4. Divergent selection between allopatric parental populations in genomic regions underlying sexual traits. (**A**) Scans of statistics used to detect signatures of positive selection in the *KITLG* region of chromosome 1A, including PBS, π, Tajima's *D*, and cross-population xp-EHH. The locations of *KITLG* and other genes are shown and the region associated with ventral color is shaded in gray. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Boxplots to the right of the scans show distributions of each statistic across chromosome 1A and the whole genome in *rustica* (PBS, π, and Tajima's *D*) and between *rustica* and *tytleri* (xp-EHH); asterisks summarize comparisons between values in the trait-associated region and background distributions (**$P < 0.0001$, ***$P < 2.2 \times 10^{-16}$). Scans for other trait loci are shown in figs. S26 to S34. (**B**) Comparisons of PBS, π, |iHS|, and xp-EHH between the genome background (gray) and trait loci (shaded colors).

haplotype homozygosity. For instance, the *KITLG* region shows signatures consistent with recent positive selection in *rustica*, including local deviations in PBS, π, and Tajima's *D* from chromosome 1A and genome-wide background distributions (Fig. 4A; $P < 0.0001$). This region contains 25 SNPs with significantly positive cross-population extended haplotype homozygosity (xp-EHH; $P < 0.05$ after Bonferroni-correction) between *rustica* and other populations (Fig. 4A, bottom panel, and fig. S27), consistent with a rapid increase in allele frequencies in *rustica* due to positive selection; xp-EHH in the *KITLG* region also strongly deviates from background distributions ($P < 2.2 \times 10^{-16}$). We also find signatures of positive selection across the other trait loci, including higher PBS, lower π, and higher integrated haplotype scores (|iHS|) than the genome background in parental *rustica*, *tytleri*, and *gutturalis* (Fig. 4B; $P < 2.2 \times 10^{-16}$). These regions also have more extreme xp-EHH distributions than the background ($P < 2.2 \times 10^{-16}$). Together, these findings demonstrate that most trait loci have experienced recent positive selection in one or more parental populations. Overall signatures of positive selection were more prevalent among ventral color loci (figs. S27 to S35), suggesting greater effects of ventral color on fitness than tail streamer length.

Sexual trait loci are barriers to gene flow in hybrid zones

Evidence for divergent selection on sexual trait loci implies they may affect fitness in hybrid zones. However, tests of whether these loci are also barriers to gene flow are necessary because the degree of concordance between evolutionary processes in allopatry and hybrid fitness is not clear a priori (*56*–*59*). We performed these tests using geographic and genomic cline analyses throughout the three hybrid zone transects between *rustica*, *tytleri*, and *gutturalis*, comparing trait loci to background loci sampled across the genome (Fig. 5, A to C). This combination of approaches allowed us to evaluate gene flow at loci in a geographic context and as a function of genome-wide admixture. Geographic clines for trait loci are narrower than background clines in the *rustica-tytleri* and *rustica-gutturalis* hybrid zones based on cline width (*w*; $P = 8.7 \times 10^{-5}$ and 3.8×10^{-7}, respectively; Fig. 5, A and B, fig. S36, and table S11), indicating a steeper transition between parental ancestry even when compared to background loci with parental allele frequency differences matching those found for trait loci ($P < 0.001$; Supplementary Materials; figs. S37 and S38). In principle, steep and coincident clines could be produced as a byproduct of lower population density in hybrid zones (*26*, *60*), in which case we would expect clines at trait loci to be similar to other highly differentiated loci. Our observation of steeper clines at trait loci compared with other loci plausibly under divergent selection is inconsistent with this hypothesis, supporting a role of sexual selection in promoting reproductive isolation in multiple hybrid zones. Geographic cline centers (*c*) for trait loci are also offset from the background in these hybrid zones (*rustica-tytleri* mean $c_{background}$ = 1442 km, c_{trait} = 1336 to 1385 km, $P = 0.00262$; *rustica-gutturalis* mean $c_{background}$ = 1961 km, c_{trait} = 2430 to 2709 km, $P = 2.3 \times 10^{-7}$).

Genomic clines further indicate that the majority of trait loci deviate from gradients of genome-average admixture in hybrid zones. Evidence consistent with selection against recombinants at trait loci is most pronounced in *rustica-tytleri*, in which all trait loci have steeper clines with credible deviations from the null expectation for cline slope (ν = 1) than matched background loci (Fig. 5D; only 20% of background loci have credible deviations in slope). Similarly, eight trait loci (80%) have steeper slope in *rustica-gutturalis* than null expectations (Fig. 5E). Although 70% of trait loci also have steeper slopes than expected in *tytleri-gutturalis*, evidence of selection against hybrids is generally weakest in this hybrid zone based on having the highest empirical standard deviation in ν (Fig. 5F and table S12). Specifically, genome-wide variance in ν (i.e., σ_ν) is lower in both *rustica-gutturalis* (0.19) and *rustica-tytleri* (0.28) than *tytleri-gutturalis* (0.36). This reduced variance in *rustica* hybrid zones indicates greater concordance among clines (i.e., more genetic coupling) (*61*), consistent with stronger selection and thus a more pronounced total barrier to gene flow between *rustica* and other subspecies (Fig. 5F and table S12). An alternative explanation could be the relative timing of secondary contact, although estimates of the timing for *rustica-tytleri* and *tytleri-gutturalis* are similar (fig. S7); thus, differences in the strength of selection in these hybrid zones is a plausible explanation for differences in concordance among clines. As in geographic cline analyses, a majority of trait loci also show credible shifts in genomic cline center (*c*) from the null expectation in each hybrid zone. These shifts, together with the concordance of clines among trait loci, may be explained by selection for specific combinations of alleles from the parental populations (i.e., epistasis for fitness) (*62*, *63*). Differences in cline patterns between *tytleri-gutturalis* (Fig. 5, C and F) and the other hybrid zones broadly align with the absence of a migratory divide separating parental *tytleri* and *gutturalis* populations (fig. S2) and (*59*). This, together with weaker sexual selection against hybrids, likely facilitates more extensive hybridization in this region. By contrast, analyses of hybrid zones in both geographic and genomic contexts indicate that trait loci constitute barrier effects in *rustica-tytleri* and *rustica-gutturalis* whereas gene flow is less constrained for the remainder of the genome.

Genetic coupling of barriers promotes reproductive isolation

We have shown evidence that loci underlying mate choice traits are relevant to fitness both in the context of allopatric divergence and as barriers to gene flow in hybrid zones. The concurrence of trait locus clines in multiple hybrid zones further argues that allele frequency differences among barrier loci were built up in geographic isolation and have been maintained by selection upon secondary contact (Figs. 1 and 5, and fig. S1). We thus explicitly tested the hypothesis that genetic coupling of barriers has emerged as a result of selection on combinations of alleles at trait loci by examining LD in hybrid zones.

Hybridization generates LD (Fig. 1) (*64*), which is then eroded by recombination between loci over generations. However, selection on combinations of alleles upon secondary contact can slow the erosion of (or enhance the amount of) LD between barrier loci to the degree that selection outweighs recombination (*65*, *66*), even between physically distant loci (*27*, *67*). We distinguished between genetic coupling as a consequence of selection and LD generated through admixture by comparing interchromosomal LD (r^2) between SNPs in barrier trait loci with SNPs matched by allele frequency differences between parental populations, as this is expected to scale with admixture LD. If LD were generated by admixture but not subsequently maintained by selection, we would expect to find no difference in LD between trait loci and the background in hybrids. Instead, SNPs with extreme parental allele frequency differences in trait loci have on average 40 to 70% higher interchromosomal LD than the matched genome background in hybrid zones (Fig. 6, A to C; $P \leq 2.65 \times 10^{-7}$). Similarly, trait loci have higher interchromosomal LD than matched background SNPs from centromere regions, in which we expect low recombination rates to generate pronounced admixture LD (figs. S26 and S39; $P \leq 0.02$). Interchromosomal LD between highly differentiated SNPs in trait loci is exceptional, with levels of LD similar to SNPs in tight physical linkage across the genome (e.g., SNPs < 5 kb of each other; fig. S40). The rate of increase in LD among trait loci also varies across the hybrid zones, indicating a stronger effect of coupling in *rustica-tytleri* and *rustica-gutturalis* compared with *tytleri-gutturalis* (this is also consistent with the relative degree of coupling inferred from genomic cline analyses; Fig. 5, D to F, and table S12). By contrast, we find no increase in interchromosomal LD among parental populations as a function of allele frequency differences (Fig. 6,

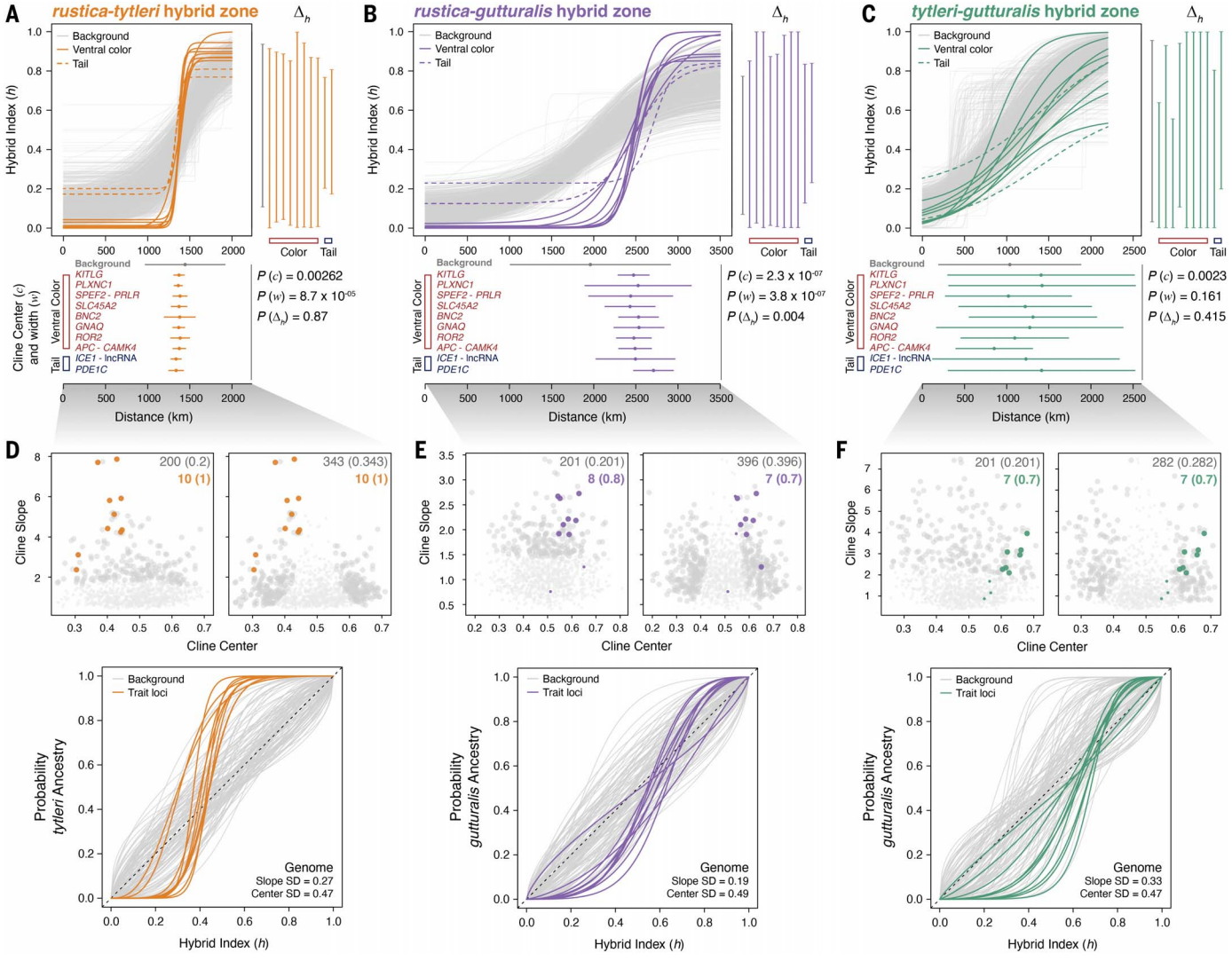


Fig. 5. Sexual trait loci are barriers to gene flow in hybrid zones. (**A**) to (**C**) Geographic clines of hybrid index (*h*) as a function of distance (kilometers) across the *rustica-tytleri*, *rustica-gutturalis*, and *tytleri-gutturalis* hybrid zone transects. In each panel, gray lines depict sigmoid clines for randomly sampled background loci across the genome, colored lines show clines for ventral color loci, and dashed colored lines show clines for tail streamer length loci. Below each panel are summaries of cline center (*c*; points) and width (*w*; whiskers) parameters for background and trait loci. To the right of each panel are *h* ranges in the cline tails (Δ_h) estimated for background and trait loci. *P*-values summarize comparisons between cline parameters estimated from background and trait loci. We note that our geographic sampling of the *tytleri-gutturalis* hybrid zone was diffuse, making geographic cline comparisons among loci less informative than in the *rustica-tytleri* and *rustica-gutturalis* transects. (**D**) to (**F**) Genomic clines for background (gray) and trait loci (colors) across the three hybrid zones. Top panels show variation in the estimated cline center and cline slope among loci, where enlarged circles were loci that deviated from null expectations for cline slope (left) and cline center (right). The number and proportion (in parentheses) of loci deviating from the null expectation are shown for each parameter. Bottom panels show genomic clines of the probability of parental ancestry as a function of hybrid index. The dashed line depicts the null expectation of no barriers to gene flow. Genome-wide summaries of the empirical standard deviation of cline slope and center are shown at the bottom right of each panel.

D to F), providing additional evidence that higher LD among trait loci was generated through admixture and has been maintained as a result of selection in hybrids specifically. A detailed examination of pairs of trait loci shows that signatures in hybrids are driven primarily by strong interchromosomal LD among ventral color loci (e.g., *KITLG* on chromosome 1A and Z-linked genes) and not between color and tail streamer length loci (fig. S41).

Finally, we sought to determine whether selection has slowed the erosion of LD among trait loci or has instead enhanced LD beyond starting levels generated by admixture. We simulated F1 hybrids and calculated interchromosomal LD among trait loci to approximate starting levels of admixture LD in each hybrid zone. Interchromosomal LD is higher among trait loci in simulated F1s (fig. S42; $P \leq 1.2 \times 10^{-13}$), clarifying that selection has maintained genetic coupling by slowing the erosion of LD (Fig. 6). Our findings together support that sexual trait loci are subject to both direct and indirect selection through spillover effects by means of coupling, producing stronger overall barriers to gene flow due to sexual selection. We also find that the apparent degree of coupling among trait loci corresponds with proxies for overall reproductive isolation (i.e., genome-wide F_{ST} and genomic cline slope, fig. S43), with more coupling and stronger barriers in the *rustica-tytleri* and *rustica-gutturalis* hybrid zones.

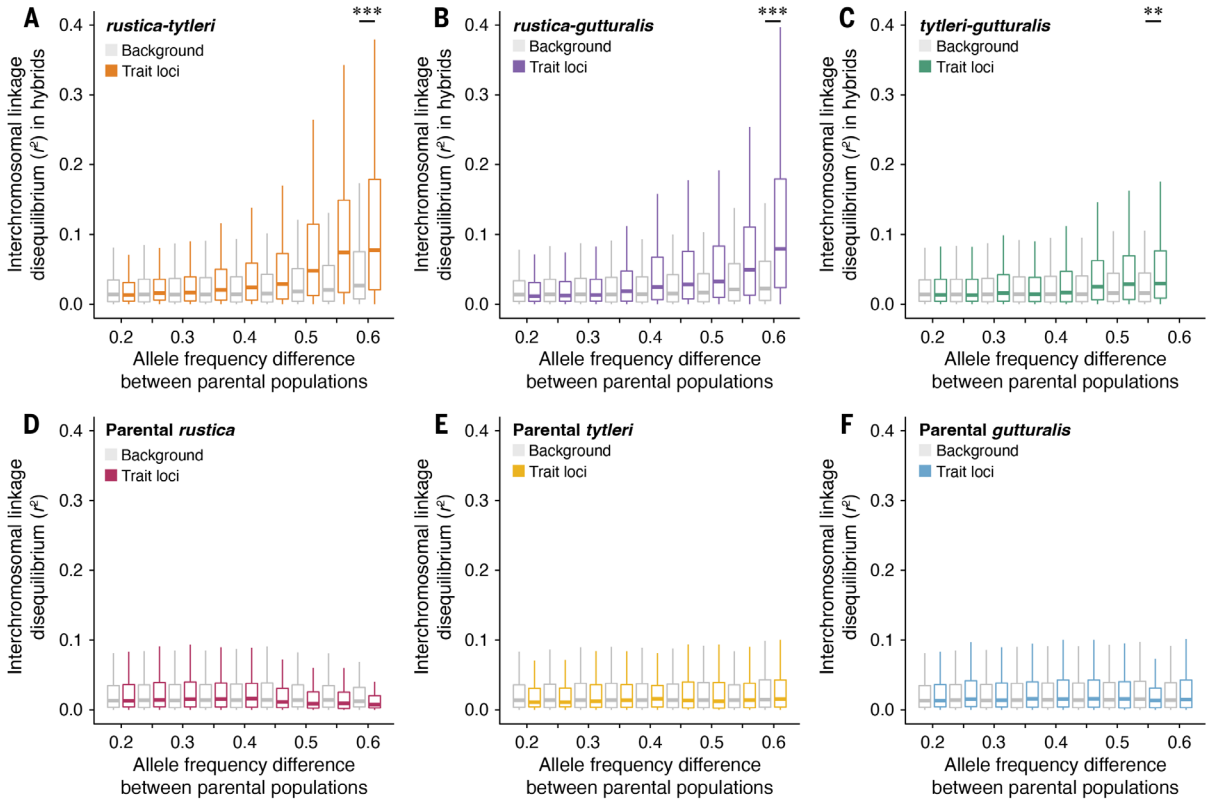


Fig. 6. Genetic coupling of barrier trait loci in hybrid zones. (**A**) to (**C**) Interchromosomal LD (r^2) measured between SNPs in trait loci (colors) and matched genome background SNPs (gray) with increasing allele frequency differences between parental populations in (A) *rustica-tytleri*, (B) *rustica-gutturalis*, and (C) *tytleri-gutturalis* hybrids; asterisks summarize comparisons between trait loci and background SNPs with extreme parental allele frequency differences (**$P < 0.0001$, ***$P < 2.2 \times 10^{-16}$). (**D**) to (**F**) Interchromosomal LD between SNPs in trait loci and the matched genome background in parental (D) *rustica*, (E) *tytleri*, and (F) *gutturalis* populations.

Discussion

Our study provides key evidence that divergent sexual traits can form barriers to gene flow upon secondary contact. We analyzed genotypic and phenotypic variation from hundreds of barn swallows sampled from across their global distribution to investigate how sexual traits and their genetic architecture contribute to reproductive isolation between populations in this young species complex. Leveraging decades of work on divergent sexual selection among closely related populations of barn swallows, our study supports an important role for sexual selection in speciation through a multifaceted analysis of the genomic basis of sexual signal traits (Fig. 3). We provide evidence of divergent selection in geographic isolation (Fig. 4) and find that loci underlying plumage traits involved in prezygotic isolation represent barriers to gene flow. By contrast, much of the genome is comparatively porous to gene flow upon secondary contact (Fig. 5). Our analyses show that sexually selected loci exceed all other signals of selection detected across the genome, consistent with a primary role of divergent sexual selection in genomic divergence and reproductive isolation. Our findings do not preclude a role for ecological selection in reproductive isolation, though a previous analysis found no associations between genomic differentiation and environmental variation across the breeding distribution of the species (*68*). We also do not rule out the possibility that sexual signal traits are also partly subject to divergent natural selection, such as selection imposed by parasites associated with plumage variation among populations (*69*). However, our results do provide empirical support for the hypothesis that divergent sexual selection can be a primary driver of reproductive isolation in hybrid zones.

Our results further support the conclusion that genetic coupling of divergent barrier loci is critical to the maintenance of reproductive isolation upon secondary contact. LD among sexual trait loci was likely established by divergent sexual selection in allopatry and has been maintained in hybrid zones, presenting a combined barrier to gene flow through genetic coupling (Fig. 6). This is emphasized by patterns in two of the three hybrid zones that we analyzed (i.e., *rustica-tytleri* and *rustica-gutturalis*; Figs. 5 and 6, and fig. S43), indicating support for the predicted association between coupling and reproductive isolation. These two hybrid zones also occur at migratory divides (*43*, *44*) where parental populations migrate along alternative routes to nonbreeding grounds in Africa or southern Asia. By contrast, a migratory divide is absent in the *tytleri-gutturalis* hybrid zone, in which there are fewer barriers to gene flow and evidence of more extensive hybridization. It is thus plausible that the coupling of sexual trait loci generates enhanced barriers when combined with ecological selection on divergent migratory traits, though it remains to be tested whether there is selection against hybrids with intermediate migratory strategies (*44*). Alternatively, postzygotic sexual selection against hybrids with intermediate signal traits could similarly reinforce prezygotic isolation, though we do not see signals of reproductive character displacement in hybrid zones.

Our study raises the question of whether genetic coupling is necessary for sexual selection to contribute to speciation. There are numerous examples of outstanding divergence and resistance to gene flow at one or few loci presumably under sexual selection (e.g., *14*, *16*, *70*–*72*), but their ultimate contribution to speciation likely depends on coupling with additional traits under divergent selection.

Periods of geographic isolation are likely important, wherein genetic associations are established with little opportunity to be broken down by gene flow and recombination (*27*, *29*). Upon secondary contact there may be genome-wide erosion of reproductive isolation outside of barrier loci where associations are maintained due to selection; we find evidence of this in barn swallows (Fig. 6 and fig. S42). Still, coupling of sexual traits with additional barriers such as traits under divergent ecological selection (e.g., migratory strategies) may be critical to the establishment of complete reproductive isolation, where the increased spread of indirect selection among barrier loci could lead to the formation of a genome-wide barrier to gene flow (*73*). Coupled barriers under sexual selection may thus represent an initial seed of reproductive isolation that "capture" other barriers through LD over time. Thoroughly addressing this question requires explicit tests of theoretical predictions for genetic coupling (*24*, *26*) in systems in which sexual selection has been demonstrated, where variation in sexual traits exists, and where it is possible to assess the influence of sexual selection on both population divergence and reproductive isolation in secondary contact. Applying the integrative analytical framework used in this study (i.e., Fig. 1) to these systems holds promise for addressing how sexual selection and genetic coupling shape the evolution of new species.

Materials and Methods summary

Genome sequencing and variant calling

We analyzed whole genomes of barn swallows sampled across the species breeding distribution (n = 336; Fig. 2 and fig. S2) including populations of all six subspecies, with a focus on three hybrid zone transects in Eurasia. Sequencing libraries were prepared using Illumina Nextera XT kits and sequenced using 150-bp paired-end reads on Illumina NovaSeq 6000 S4 lanes, yielding 10× ± 7.5× coverage per individual (data S1). We filtered raw read data using Trimmomatic v0.39 (*74*) and mapped reads to the barn swallow reference genome [assembly bHirRus1 (*75*) using BWA "mem" v0.7.17 (*76*)]. We called SNPs using the GATK v4.0.8.1 best-practices workflow (*77*). We quality-filtered variants using GATK "VariantFiltration" and used BCFtools v1.10.2 (*78*) and VCFtools v0.1.17 (*79*) to process SNPs for specific analyses.

Population structure

We estimated genetic structure after filtering to retain SNPs with minor-allele frequency ≥ 0.1 (837,275 SNPs) using PCA implemented in SNPRelate (*80*). We used ADMIXTURE (*81*) to estimate individual admixture proportions from one or more genetic clusters (K = 1 to 10). We evaluated the most likely number of genetic clusters as the K model with the lowest cross-validation error (fig. S4).

Demographic inference

We performed demographic inference using the diffusion approximation framework in ∂a∂i (*82*). We fit seven two-population demographic models to the unfolded joint site frequency spectrum (JSFS) between *rustica*, *tytleri*, and *gutturalis* population pairs. The first four models were strict isolation (SI) without gene flow, an isolation-migration (IM) model with continuous gene flow during divergence, an ancient migration (AM) model where gene flow occurred after divergence then ceased at a second timepoint, and a secondary contact (SC) model where populations diverged in isolation for a period of time followed by a second timepoint where gene flow occurred (fig. S8). The other models (AM2m, IM2m, and SC2m) were equivalent to the AM, IM, and SC models, except that two effective migration rates were inferred to simulate scenarios in which some regions of the genome are porous to gene flow whereas others present barriers to gene flow (i.e., genomic "island" models).

Measurement of mate choice plumage traits

We quantified ventral color by sampling 5 to 10 feathers from the breast region and storing them in envelopes in a dark, climate-controlled environment. We analyzed feather samples using an Ocean Optics USB4000 spectrometer to measure reflectance relative to an Ocean Optics WS-1 standard and a dark standard (no light source) and recorded all measures using Spectrasuite v2.0.125 (Ocean Optics, Inc.) We measured left and right tail streamer lengths to the nearest millimeter, taking an average of three independent measurements per feather.

Genome-wide association mapping of mate choice traits

We used genome-wide efficient mixed model association (GEMMA) (*83*) to characterize the genetic architecture of phenotypes and identify associations between genome-wide SNPs and ventral color and tail streamer length, focusing on our sampling of admixed hybrid individuals. We imputed missing data using BEAGLE (*84*) after removing SNPs with greater than 20% missing data among individuals. We first ran Bayesian sparse linear mixed models (BSLMM) to broadly summarize the genetic architecture of each trait. We ran BSLMMs using 10 independent chains and 25 million MCMC steps after a burn-in of 5 million steps, sampling every 1000 steps. We incorporated a relatedness matrix in all analyses to control for population stratification. We combined results across the independent runs and summarized the genetic architecture of each trait using posterior hyperparameter distributions. We further broadly quantified the architecture of each trait by calculating the sum of the PIP for SNPs with detectable effects (PIP ≥ 0.01) per chromosome compared with the chromosome length (fig. S12). In addition to BSLMMs, we ran linear mixed models (LMM) to detect isolated SNP associations with each trait. We assessed significance of SNP associations using Wald test P-values from LMMs and considered SNPs with $P < 0.05$ after Bonferroni correction to be significant. We annotated coding genes containing or within 20 kb of significant GWA SNPs as putative candidate loci associated with each trait (table S7 and data S3).

Recombination rates

We inferred a barn swallow recombination map using pyrho (*85*), which accounts for non-equilibrium population histories in its estimation of fine-scale recombination rate. We focused this analysis on parental *rustica* from Russia to estimate a representative recombination map. We first obtained an estimate of population size history using SMC++ (*86*), choosing diploid genotypes of five individuals at random to be assigned as "distinguished" lineages. We fit a model of population history for the last 2×10^5 generations using composite likelihood assessed using the expectation-maximization algorithm assuming a per-generation mutation rate of 2.3×10^{-9} (*87*) and a generation time of one year (*40*). We then used the pyrho "lookup" function to generate a likelihood lookup table based on the SMC++ demography and the sample size, and ran pyrho "optimize" to infer the recombination map under a block penalty of 25, window size of 50 kb, and scaled by the mutation rate.

Population genetic differentiation and diversity statistics

We used pixy v1.2.4 (*88*) to perform genome scans of relative population differentiation (F_{ST}) between-population sequence divergence (d_{xy}) and within-population nucleotide diversity (π) in nonoverlapping sliding windows at several resolutions (1 Mb, 100 kb, and 10 kb). For detailed visualization of specific genomic regions, we also performed scans using sliding windows with partial overlap between windows (e.g., 50-kb windows with 5-kb step size).

Selection statistics

We calculated allelic differentiation specific to each of the parental populations that form hybrid zones (*rustica*, *tytleri*, and *gutturalis*) using PBS. We calculated PBS using windowed pairwise F_{ST} estimates among the three parental populations as $\mathrm{PBS} = \frac{T^{1,2}+T^{1,3}-T^{2,3}}{2}$, following (*89*). Here, $T = -\log(1 - F_{ST})$ between each pair of populations. We calculated Tajima's D to summarize the allele frequency spectrum using VCF-kit (*90*). We used the R package "rehh" (*91*) to calculate statistics based on EHH, including the iHS within populations and cross-population xp-EHH.

Geographic and genomic clines

We used sigmoid geographic clines across hybrid zone transects to test for signals of barriers to gene flow in trait loci relative to the genomic background. We used introgress (*92*) to estimate the individual locus-specific hybrid index (h) from SNP allele frequencies in each trait locus after filtering to retain SNPs with minor allele frequency ≥ 0.1. To compare locus-specific clines with genome-wide patterns, we sampled, at random, 1000 windows of 100 SNPs with minor allele frequency ≥ 0.1 from each set of parental and hybrid populations. The physical distance of the background 100 SNP windows was 18,494 ± 15,419 bp. We fit six geographic cline models to ancestry estimates for candidate and background loci using the R package "HZAR" (*93*) and performed model selection between these models and the null model. Each model estimated the cline center (c) as the distance in km from the westernmost transect location, cline width, w, as 1 / the maximum slope of the cline, and the mean ancestry values in the tails of the cline ($p_{\min}$ and $p_{\max}$). We performed cline fitting using the default MCMC chain length of 100,000 steps with 10,000 burn-in steps and compared support for the null and six cline models using AIC, selecting the best model for each locus using a ΔAIC ≥ 2 threshold.

We performed hierarchical Bayesian analysis of genomic clines (*94*, *95*) to characterize locus-specific introgression relative to a genome-wide gradient of admixture using the logit-logistic cline function from (*96*) describing the probability that an allele at locus i in individual j was inherited from population 1 as $\phi_{ij} = \frac{(h_j^{vi})}{(h_j^{vi}+(1-h_j^{vi})*e^{ui})}$, where h is the proportion of the genome inherited from population 2 (i.e., hybrid index), ν is the cline slope relative to the genome average (ν = 1), and μ is related to the cline center. Following (*97*), we used the conversion logit(c) = μ/ν to define the cline center parameter, c, which specifies the value of h at which $\phi = 0.5$. We measured variance in clines as the variance in log(ν) and logit(c), which each have an expected mean of 0.

Analysis of LD

We measured interchromosomal LD between SNPs in trait loci with allele frequency differences between parental populations ranging from 0.2 to 0.6 in intervals of 0.05. We also measured interchromosomal LD between SNPs in the genome background (i.e., outside of trait loci) in matched allele frequency difference intervals. We used VCFtools v0.1.17 (*79*) to summarize allele frequencies at biallelic SNPs in parental populations, then calculated allele frequency differences between each pair of populations. We then extracted SNPs in allele frequency difference intervals from trait loci and the genome background and calculated interchromosomal haplotype r^2 using VCFtools in each hybrid zone and parental population. To further investigate levels of LD between specific pairs of trait loci, we calculated both intra- and interchromosomal haplotype r^2 between all pairwise comparisons at SNPs with allele frequency differences between 0.3 and 0.6 in each hybrid zone. As a comparison with observed LD in hybrid populations, we approximated potential starting levels of admixture LD between trait loci in hybrid zones by simulating F1 hybrids using a random sample of 10 individuals per parental population, then calculated interchromosomal LD as a function of increasing allele frequency differences as described above.

Statistical Analysis

We performed all statistical analyses in R *98*) v4.1.2 unless otherwise specified. We performed all plotting using base R graphics and ggplot2 (*99*).

REFERENCES AND NOTES

1. M. J. West-Eberhard, Sexual selection, social competition, and speciation. *Q. Rev. Biol.* **58**, 155–183 (1983). doi: 10.1086/413215
2. T. Price, Sexual selection and natural selection in bird speciation. *Philos. Trans. R. Soc. B.* **353**, 251–260 (1998). doi: 10.1098/rstb.1998.0207
3. T. M. Panhuis, R. Butlin, M. Zuk, T. Tregenza, Sexual selection and speciation. *Trends Ecol. Evol.* **16**, 364–371 (2001). doi: 10.1016/S0169-5347(01)02160-7; pmid: 11403869
4. M. R. Servedio, J. W. Boughman, The role of sexual selection in local adaptation and speciation. *Annu. Rev. Ecol. Evol. Syst.* **48**, 85–109 (2017). doi: 10.1146/annurev-ecolsys-110316-022905
5. T. C. Mendelson, R. J. Safran, Speciation by sexual selection: 20 years of progress. *Trends Ecol. Evol.* **36**, 1153–1163 (2021). doi: 10.1016/j.tree.2021.09.004; pmid: 34607719
6. J. W. Boughman, Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* **411**, 944–948 (2001). doi: 10.1038/35082064; pmid: 11418857
7. O. Seehausen *et al*., Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–626 (2008). doi: 10.1038/nature07285; pmid: 18833272
8. M. E. Maan, O. Seehausen, Ecology, sexual selection and speciation. *Ecol. Lett.* **14**, 591–602 (2011). doi: 10.1111/j.1461-0248.2011.01606.x; pmid: 21375683
9. J. A. C. Uy, R. G. Moyle, C. E. Filardi, Z. A. Cheviron, Difference in plumage color used in species recognition between incipient species is linked to a single amino acid substitution in the melanocortin-1 receptor. *Am. Nat.* **174**, 244–254 (2009). doi: 10.1086/600084; pmid: 19489704
10. M. Rossi *et al*., Adaptive introgression of a visual preference gene. *Science* **383**, 1368–1373 (2024). doi: 10.1126/science.adj9201; pmid: 38513020
11. T. C. Mendelson, K. L. Shaw, Sexual behaviour: Rapid speciation in an arthropod. *Nature* **433**, 375–376 (2005). doi: 10.1038/433375a; pmid: 15674280
12. D. F. Beltrán, A. J. Shultz, J. L. Parra, Speciation rates are positively correlated with the rate of plumage color evolution in hummingbirds. *Evolution* **75**, 1665–1680 (2021). doi: 10.1111/evo.14277; pmid: 34037257
13. K. Kraaijeveld, F. J. L. Kraaijeveld-Smit, M. E. Maan, Sexual selection and speciation: The comparative evidence revisited. *Biol. Rev.* **86**, 367–377 (2011). doi: 10.1111/j.1469-185X.2010.00150.x; pmid: 20659104
14. J. W. Poelstra *et al*., The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **344**, 1410–1414 (2014). doi: 10.1126/science.1253226; pmid: 24948738
15. K. E. Delmore, D. P. L. Toews, R. R. Germain, G. L. Owens, D. E. Irwin, The genetics of seasonal migration and plumage color. *Curr. Biol.* **26**, 2167–2173 (2016). doi: 10.1016/j.cub.2016.06.015; pmid: 27476599
16. K. Hench, M. Vargas, M. P. Höppner, W. O. McMillan, O. Puebla, Inter-chromosomal coupling between vision and pigmentation genes during genomic divergence. *Nat. Ecol. Evol.* **3**, 657–667 (2019). doi: 10.1038/s41559-019-0814-5; pmid: 30833758
17. E. L. Westerman *et al*., Aristaless controls butterfly wing color variation used in mimicry and mate choice. *Curr. Biol.* **28**, 3469–3474.e4 (2018). doi: 10.1016/j.cub.2018.08.051; pmid: 30415702
18. J. M. Gleason, R. A. James, C. Wicker-Thomas, M. G. Ritchie, Identification of quantitative trait loci function through analysis of multiple cuticular hydrocarbons differing between *Drosophila simulans* and *Drosophila sechellia* females. *Heredity* **103**, 416–424 (2009). doi: 10.1038/hdy.2009.79; pmid: 19654611
19. R. Riesch *et al*., Transitions between phases of genomic differentiation during stick-insect speciation. *Nat. Ecol. Evol.* **1**, 0082 (2017). doi: 10.1038/s41559-017-0082; pmid: 28812654
20. O. Seehausen *et al*., Genomics and the origin of species. *Nat. Rev. Genet.* **15**, 176–192 (2014). doi: 10.1038/nrg3644; pmid: 24535286
21. J. Kulmuni, R. K. Butlin, K. Lucek, V. Savolainen, A. M. Westram, Towards the completion of speciation: The evolution of reproductive isolation beyond the first barriers. *Philos. Trans. R. Soc. B.* **375**, 20190528 (2020). doi: 10.1098/rstb.2019.0528; pmid: 32654637
22. J. B. W. Wolf, H. Ellegren, Making sense of genomic islands of differentiation in light of speciation. *Nat. Rev. Genet.* **18**, 87–100 (2017). doi: 10.1038/nrg.2016.133; pmid: 27840429
23. M. Ravinet *et al*., Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *J. Evol. Biol.* **30**, 1450–1477 (2017). doi: 10.1111/jeb.13047; pmid: 28786193
24. R. K. Butlin, C. M. Smadja, Coupling, reinforcement, and speciation. *Am. Nat.* **191**, 155–172 (2018). doi: 10.1086/695136; pmid: 29351021
25. J. Felsenstein, Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* **35**, 124–138 (1981). doi: 10.2307/2407946; pmid: 28563447
26. N. H. Barton, Multilocus clines. *Evolution* **37**, 454–471 (1983). doi: 10.2307/2408260; pmid: 28563316
27. S. M. Flaxman, A. C. Wacholder, J. L. Feder, P. Nosil, Theoretical models of the influence of genomic architecture on the dynamics of speciation. *Mol. Ecol.* **23**, 4074–4088 (2014). doi: 10.1111/mec.12750; pmid: 24724861
28. T. H. Vines *et al*., Cline coupling and uncoupling in a stickleback hybrid zone. *Evolution* **70**, 1023–1038 (2016). doi: 10.1111/evo.12917; pmid: 27061719
29. P. Nosil, J. L. Feder, Z. Gompert, How many genetic changes create new species? *Science* **371**, 777–779 (2021). doi: 10.1126/science.abf6671; pmid: 33602840
30. A. P. Møller, Female choice selects for male sexual tail ornaments in the monogamous swallow. *Nature* **332**, 640–642 (1988). doi: 10.1038/332640a0
31. N. Saino, C. R. Primmer, H. Ellegren, A. P. Møller, An experimental study of paternity and tail ornamentation in the barn swallow (*Hirundo rustica*). *Evolution* **51**, 562–570 (1997). pmid: 28565341
32. R. J. Safran, C. R. Neuman, K. J. McGraw, I. J. Lovette, Dynamic paternity allocation as a function of male plumage color in barn swallows. *Science* **309**, 2210–2212 (2005). doi: 10.1126/science.1115090; pmid: 16195460
33. Y. Vortman, A. Lotem, R. Dor, I. J. Lovette, R. J. Safran, The sexual signals of the East-Mediterranean barn swallow: A different swallow tale. *Behav. Ecol.* **22**, 1344–1352 (2011). doi: 10.1093/beheco/arr139
34. R. Dor *et al*., Population genetics and morphological comparisons of migratory European (*Hirundo rustica rustica*) and sedentary East-Mediterranean (*Hirundo rustica transitiva*) barn swallows. *J. Hered.* **103**, 55–63 (2012). doi: 10.1093/jhered/esr114; pmid: 22071313
35. R. J. Safran *et al*., The maintenance of phenotypic divergence through sexual selection: An experimental study in barn swallows *Hirundo rustica*. *Evolution* **70**, 2074–2084 (2016). doi: 10.1111/evo.13014; pmid: 27436630
36. M. R. Wilkins *et al*., Phenotypic differentiation is associated with divergent sexual selection among closely related barn swallow populations. *J. Evol. Biol.* **29**, 2410–2421 (2016). doi: 10.1111/jeb.12965; pmid: 27538265
37. E. S. C. Scordato, R. J. Safran, Geographic variation in sexual selection and implications for speciation in the Barn Swallow. *Avian Res.* **5**, 8 (2014). doi: 10.1186/s40657-014-0008-4

38. A. Romano, A. Costanzo, D. Rubolini, N. Saino, A. P. Møller, Geographical and seasonal variation in the intensity of sexual selection in the barn swallow *Hirundo rustica*: A meta-analysis. *Biol. Rev.* **92**, 1582–1600 (2017). doi: 10.1111/brv.12297; pmid: 27615554
39. A. Lotem, Y. Vortman, R. J. Safran, The evidence for divergent sexual selection among closely related barn swallow populations is strong. *Evolution* **76**, 2204–2211 (2022). doi: 10.1111/evo.14506; pmid: 35561267
40. R. M. Zink, A. Pavlova, S. Rohwer, S. V. Drovetski, Barn swallows before barns: Population histories and intercontinental colonization. *Proc. R. Soc. B.* **273**, 1245–1251 (2006). doi: 10.1098/rspb.2005.3414; pmid: 16720398
41. C. C. R. Smith *et al*., Demographic inference in barn swallows using whole-genome data shows signal for bottleneck and subspecies differentiation during the Holocene. *Mol. Ecol.* **27**, 4200–4212 (2018). doi: 10.1111/mec.14854; pmid: 30176075
42. E. S. C. Scordato *et al*., Genomic variation across two barn swallow hybrid zones reveals traits associated with divergence in sympatry and allopatry. *Mol. Ecol.* **26**, 5676–5691 (2017). doi: 10.1111/mec.14276; pmid: 28777875
43. E. S. C. Scordato *et al*., Migratory divides coincide with reproductive barriers across replicated avian hybrid zones above the Tibetan Plateau. *Ecol. Lett.* **23**, 231–241 (2020). doi: 10.1111/ele.13420; pmid: 31746098
44. S. P. Turbek *et al*., A migratory divide spanning two continents is associated with genomic and ecological divergence. *Evolution* **76**, 722–736 (2022). doi: 10.1111/evo.14448; pmid: 35166383
45. D. R. Schield *et al*., Sex-linked genetic diversity and differentiation in a globally distributed avian species complex. *Mol. Ecol.* **30**, 2313–2332 (2021). doi: 10.1111/mec.15885; pmid: 33720472
46. J. K. Hubbard, B. R. Jenkins, R. J. Safran, Quantitative genetics of plumage color: Lifetime effects of early nest environment on a colorful sexual signal. *Ecol. Evol.* **5**, 3436–3449 (2015). doi: 10.1002/ece3.1602; pmid: 26380676
47. C. T. Miller *et al*., cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* **131**, 1179–1189 (2007). doi: 10.1016/j.cell.2007.10.055; pmid: 18083106
48. J. K. Hubbard, J. A. Uy, M. E. Hauber, H. E. Hoekstra, R. J. Safran, Vertebrate pigmentation: From underlying genes to adaptive function. *Trends Genet.* **26**, 231–239 (2010). doi: 10.1016/j.tig.2010.02.002; pmid: 20381892
49. L. Le *et al*., SLC45A2 protein stability and regulation of melanosome pH determine melanocyte pigmentation. *Mol. Biol. Cell* **31**, 2687–2702 (2020). doi: 10.1091/mbc.E20-03-0200; pmid: 32966160
50. M. R. Lang, L. B. Patterson, T. N. Gordon, S. L. Johnson, D. M. Parichy, Basonuclin-2 requirements for zebrafish adult pigment pattern development and female fertility. *PLOS Genet.* **5**, e1000744 (2009). doi: 10.1371/journal.pgen.1000744; pmid: 19956727
51. J. W. Poelstra, N. Vijay, M. P. Hoeppner, J. B. W. Wolf, Transcriptomics of colour patterning and coloration shifts in crows. *Mol. Ecol.* **24**, 4617–4628 (2015). doi: 10.1111/mec.13353; pmid: 26302355
52. M. R. Servedio, G.-P. Sætre, Speciation as a positive feedback loop between postzygotic and prezygotic barriers to gene flow. *Proc. R. Soc. B* **270**, 1473–1479 (2003). doi: 10.1111/mec.13353; pmid: 26302355
53. D. C. Presgraves, Sex chromosomes and speciation in *Drosophila*. *Trends Genet.* **24**, 336–343 (2008). doi: 10.1016/j.tig.2008.04.007; pmid: 18514967
54. Y. Cai *et al*., Role of cAMP-phosphodiesterase 1C signaling in regulating growth factor receptor stability, vascular smooth muscle cell growth, migration, and neointimal hyperplasia. *Circ. Res.* **116**, 1120–1132 (2015). doi: 10.1161/CIRCRESAHA.116.304408; pmid: 25608528
55. T. E. Cruickshank, M. W. Hahn, Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* **23**, 3133–3157 (2014). doi: 10.1111/mec.12796; pmid: 24845075
56. Z. Gompert *et al*., Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. *Evolution* **66**, 2167–2181 (2012). doi: 10.1111/j.1558-5646.2012.01587.x; pmid: 22759293
57. Z. Gompert, L. K. Lucas, C. C. Nice, C. A. Buerkle, Genome divergence and the genetic architecture of barriers to gene flow between Lycaeides idas and L. Melissa. *Evolution* **67**, 2498–2514 (2013). doi: 10.1111/evo.12021; pmid: 24033163
58. P. Nosil, T. L. Parchman, J. L. Feder, Z. Gompert, Do highly divergent loci reside in genomic regions affecting reproductive isolation? A test using next-generation sequence data in Timema stick insects. *BMC Evol. Biol.* **12**, 164 (2012). doi: 10.1186/1471-2148-12-164; pmid: 22938057
59. D. R. Schield *et al*., Insight into the roles of selection in speciation from genomic patterns of divergence and introgression in secondary contact in venomous rattlesnakes. *Ecol. Evol.* **7**, 3951–3966 (2017). doi: 10.1002/ece3.2996; pmid: 28616190
60. N. Barton, The hybrid sink effect. *Heredity* **44**, 277–278 (1980). doi: 10.1038/hdy.1980.23
61. T. J. FirnenoJr *et al*., Quantitative Analyses of Coupling in Hybrid Zones. *Cold Spring Harb. Perspect. Biol.* **15**, a041434 (2023). doi: 10.1101/cshperspect.a041434; pmid: 37739809
62. M. C. Whitlock, P. C. Phillips, F. B.-G. Moore, S. J. Tonsor, Multiple fitness peaks and epistasis. *Annu. Rev. Ecol. Syst.* **26**, 601–629 (1995). doi: 10.1146/annurev.es.26.110195.003125
63. P. Nosil *et al*., Ecology shapes epistasis in a genotype-phenotype-fitness map for stick insect colour. *Nat. Ecol. Evol.* **4**, 1673–1684 (2020). doi: 10.1038/s41559-020-01305-y; pmid: 32929238
64. Z. Gompert, C. A. Buerkle, Analyses of genetic ancestry enable key insights for molecular ecology. *Mol. Ecol.* **22**, 5278–5294 (2013). doi: 10.1111/mec.12488; pmid: 24103088
65. N. Barton, B. O. Bengtsson, The barrier to genetic exchange between hybridising populations. *Heredity* **57**, 357–376 (1986). doi: 10.1038/hdy.1986.135; pmid: 3804765
66. N. H. Barton, M. A. R. de Cara, The evolution of strong reproductive isolation. *Evolution* **63**, 1171–1190 (2009). doi: 10.1111/j.1558-5646.2009.00622.x; pmid: 19154394
67. J. L. Feder *et al*., Genome-wide congealing and rapid transitions across the speciation continuum during speciation with gene flow. *J. Hered.* **105**, 810–820 (2014). doi: 10.1093/jhered/esu038; pmid: 25149256
68. R. J. Safran *et al*., Genome-wide differentiation in closely related populations: The roles of selection and geographic isolation. *Mol. Ecol.* **25**, 3865–3883 (2016). doi: 10.1111/mec.13740; pmid: 27357267
69. A. K. Hund *et al*., Divergent sexual signals reflect costs of local parasites. *Evolution* **74**, 2404–2418 (2020). doi: 10.1111/evo.13994; pmid: 32385910
70. D. P. L. Toews *et al*., Plumage genes and little else djstinguish the genomes of hybridizing warblers. *Curr. Biol.* **26**, 2313–2318 (2016). doi: 10.1016/j.cub.2016.06.034; pmid: 27546575
71. A. F. Feller, M. P. Haesler, C. L. Peichel, O. Seehausen, Genetic architecture of a key reproductive isolation trait differs between sympatric and non-sympatric sister species of Lake Victoria cichlids. *Proc. R. Soc. B.* **287**, 20200270 (2020). doi: 10.1098/rspb.2020.0270; pmid: 32259470
72. A. Brelsford, D. E. Irwin, Incipient speciation despite little assortative mating: The yellow-rumped warbler hybrid zone. *Evolution* **63**, 3050–3060 (2009). doi: 10.1111/j.1558-5646.2009.00777.x; pmid: 19624726
73. P. Nosil, J. L. Feder, S. M. Flaxman, Z. Gompert, Tipping points in the dynamics of speciation. *Nat. Ecol. Evol.* **1**, 0001 (2017). doi: 10.1038/s41559-016-0001; pmid: 28812620
74. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014). doi: 10.1093/bioinformatics/btu170; pmid: 24695404
75. S. Feng *et al*., Dense sampling of bird diversity increases power of comparative genomics. *Nature* **587**, 252–257 (2020). doi: 10.1038/s41586-020-2873-9; pmid: 33177665
76. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). doi: 10.1093/bioinformatics/btp324; pmid: 19451168
77. A. McKenna *et al*., The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010). doi: 10.1101/gr.107524.110; pmid: 20644199
78. H. Li *et al*., The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009). doi: 10.1093/bioinformatics/btp352; pmid: 19505943
79. P. Danecek *et al*., The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011). doi: 10.1093/bioinformatics/btr330; pmid: 21653522
80. X. Zheng *et al*., A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328 (2012). doi: 10.1093/bioinformatics/bts606; pmid: 23060615
81. D. H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009). doi: 10.1101/gr.094052.109; pmid: 19648217
82. R. N. Gutenkunst, R. D. Hernandez, S. H. Williamson, C. D. Bustamante, Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLOS Genet.* **5**, e1000695 (2009). doi: 10.1371/journal.pgen.1000695; pmid: 19851460
83. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012). doi: 10.1038/ng.2310; pmid: 22706312
84. S. R. Browning, B. L. Browning, Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007). doi: 10.1086/521987; pmid: 17924348
85. J. P. Spence, Y. S. Song, Inference and analysis of population-specific fine-scale recombination maps across 26 diverse human populations. *Sci. Adv.* **5**, eaaw9206 (2019). doi: 10.1126/sciadv.aaw9206; pmid: 31681842
86. J. Terhorst, J. A. Kamm, Y. S. Song, Robust and scalable inference of population history from hundreds of unphased whole genomes. *Nat. Genet.* **49**, 303–309 (2017). doi: 10.1038/ng.3748; pmid: 28024154
87. L. Smeds, A. Qvarnström, H. Ellegren, Direct estimate of the rate of germline mutation in a bird. *Genome Res.* **26**, 1211–1218 (2016). doi: 10.1101/gr.204669.116; pmid: 27412854
88. K. L. Korunes, K. Samuk, pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Mol. Ecol. Resour.* **21**, 1359–1368 (2021). doi: 10.1111/1755-0998.13326; pmid: 33453139
89. X. Yi *et al*., Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75–78 (2010). doi: 10.1126/science.1190371; pmid: 20595611
90. D. E. Cook, E. C. Andersen, VCF-kit: Assorted utilities for the variant call format. *Bioinformatics* **33**, 1581–1582 (2017). doi: 10.1093/bioinformatics/btx011; pmid: 28093408
91. M. Gautier, A. Klassmann, R. Vitalis, rehh 2.0: A reimplementation of the R package rehh to detect positive selection from haplotype structure. *Mol. Ecol. Resour.* **17**, 78–90 (2017). doi: 10.1111/1755-0998.12634; pmid: 27863062
92. Z. Gompert, C. Alex Buerkle, introgress: A software package for mapping components of isolation in hybrids. *Mol. Ecol. Resour.* **10**, 378–384 (2010). doi: 10.1111/j.1755-0998.2009.02733.x; pmid: 21565033
93. E. P. Derryberry, G. E. Derryberry, J. M. Maley, R. T. Brumfield, HZAR: Hybrid zone analysis using an R software package. *Mol. Ecol. Resour.* **14**, 652–663 (2014). doi: 10.1111/1755-0998.12209; pmid: 24373504
94. Z. Gompert, C. A. Buerkle, Bayesian estimation of genomic clines. *Mol. Ecol.* **20**, 2111–2127 (2011). doi: 10.1111/j.1365-294X.2011.05074.x; pmid: 21453352
95. Z. Gompert, D. A. DeRaad, C. A. Buerkle, A next generation of hierarchical Bayesian analyses of hybrid zones enables direct quantification of variation in introgression in R. bioRxiv 2024.03.29.587395 [Preprint] (2024); doi: 10.1101/2024.03.29.587395.
96. B. M. Fitzpatrick, Alternative forms for genomic clines. *Ecol. Evol.* **3**, 1951–1966 (2013). doi: 10.1002/ece3.609; pmid: 23919142
97. R. Bailey, Bayesian hybrid index and genomic cline estimation with the R package *gghybrid*. *Mol. Ecol. Resour.* (2023). doi: 10.1002/ece3.609; pmid: 23919142
98. R Core Team, R: The R Project for Statistical Computing (R Foundation for Statistical Computing, 2020); www.R-project.org/
99. H. Wickham, *ggplot2: elegant graphics for data analysis* (Springer, 2016).
100. D. R. Schield, drewschield/hirundo_speciation_genomics: v1, Version 2024-09-22, Zenodo (2024); https://zenodo.org/records/13825420.

ACKNOWLEDGMENTS

We thank S. Stankowski for helpful discussion and suggestions. We thank E. Shogren for kindly providing an R script adapted

for our analyses. We are grateful to G. Bachurin, S. I. Cherkaoui, B. Dashnyam, U. Enkhbat, D. Enkhbayar, C. Glidden, S. Gombobaatar, S. Hajib, S. Hanane, J. Hubbard, A. Hund, Y. Inaguma, B. Jenkins, H. Karaardiç, W. Kitamura, K. Koyama, W. Liang, Y. Liu, R. Lock, A. Lotem, N. Magri, N. Markov, E. Pagani-Núñez, A. Rubtsov, G. Semenov, B. Sheta, E. Shnayder, C. Smith, T. Tanioka, S. Turbek, L. Yu, Y. Vortman, and O. Zayatseva for assistance in the field and laboratory. We are grateful to the Field Museum, Burke Museum, Bell Museum, and University of Alaska Museum for tissue loans. The support and resources from the Center for High Performance Computing at the University of Utah are gratefully acknowledged. All animal procedures were approved by protocols from the University of Colorado Institutional Animal Care and Use Committee (IACUC). **Funding:** This work was funded by the following: National Science Foundation grant DBI-1906188 (to D.R.S.), National Science Foundation grant DEB-1149942 (to R.J.S.), National Science Foundation grant IOS-DEB-1856266 (to R.J.S.), National Science Foundation grant IOS-1856254 (to I.I.L.), and National Science Foundation grant DEB-1844941 (to Z.G.). **Author contributions:** Conceptualization: D.R.S., S.A.M., Z.G., P.N., J.B.W., and R.J.S. Methodology: D.R.S., J.K.C., E.S.C.S., M.R.W., Z.G., and R.J.S. Investigation: D.R.S. and Z.G. Visualization: D.R.S. and Z.G. Funding acquisition: D.R.S., I.I.L., Z.G., and R.J.S. Project administration: D.R.S. and R.J.S. Supervision: D.R.S. and R.J.S. Writing – original draft: D.R.S. and R.J.S. Writing – review and editing: D.R.S., J.K.C., E.S.C.S., I.I.L., M.R.W., S.A.M., Z.G., P.N., J.B.W., and R.J.S. **Competing interests:** Authors declare that they have no competing interests. **Data and materials availability:** Data supporting the conclusions of this study have been deposited to the NCBI short-read archive (accession PRJNA323498). The computational workflow and associated analysis scripts used in this work are available at https://github.com/drewschield/hirundo_speciation_genomics (*100*). Additional details related to the data and analyses are presented in the Supplementary Material. Correspondence and requests for materials should be sent to D.R.S.

SUPPLEMENTARY MATERIALS

Submitted 20 July 2023; resubmitted 25 June 2024
Accepted 11 October 2024
10.1126/science.adj8766