Science
AAAS

Supplementary Materials for

Sexual selection promotes reproductive isolation in barn swallows

Drew R. Schield *et al*.

Corresponding author: Drew R. Schield, drew.schield@virginia.edu

Science **386**, eadj8766 (2024)
DOI: 10.1126/science.adj8766

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Background Information

In this section, we provide an overview of previous work studying the role of divergent sexual selection in barn swallows, as well as information on geographic plumage variation, reproductive isolation in hybrid zones, and the presence of migratory divides in Eurasia. We also provide a detailed explanation of the operational definition of genetic coupling used in this study, in light of various uses in speciation literature.

Evidence for divergent sexual selection in barn swallows. Studies of sexual selection often provide indirect information about how this evolutionary mechanism works either through proxies of sexual selection (e.g., sexual dimorphism; *101*), assumptions about how traits perform in the wild based on mate-choice trials in a captive setting (e.g., *102*, *103*), or by inferring behavioral responses to simulated intrusions using playback experiments and decoys (e.g., *104*). While these approaches can provide promising evidence consistent with the role of sexual selection, they are also limited in their ability to distinguish specific processes, outcomes, or both. It is rare to have studies which example trait variation in a wild setting and its association with reproductive outcomes (including direct measures relevant to fitness like offspring production). Here, barn swallows (*Hirundo rustica*) are an outlier: there have not only been within-population studies of sexual selection, but also experimental manipulations to confirm the nature of sexual selection in different subspecies. Recent independent meta-analyses have further confirmed divergent sexual selection (*38*, *39*). Below, we summarize the extensive work on sexual selection in natural barn swallow populations across their immense breeding distribution.

Barn swallows are one of the most widespread songbirds in the world, with breeding distributions that span North America, Eurasia, and the Middle East. Across this broad distribution, populations are categorized into six subspecies represented by phenotypic variation in two plumage traits: the length of outer tail feathers (i.e., tail streamers) and the extent of ventral plumage color (see subspecies phenotype data portrayed in Scordato and Safran, *37*). These birds are commonly observed by humans, as they construct mud cup nests exclusively on human built structures. In large part due to the tractability of barn swallows as a study system, there are hundreds of publications related to various aspects of the biology of this species. A primary focus among these studies has been the role of sexual selection in shaping variation in phenotype within and among subspecies. We have been able to sample from populations of all subspecies in the present dataset, but detailed analyses of reproductive behaviors have thus far taken place in four of the six subspecies: European *rustica*, North American *erythrogaster*, Asian *gutturalis*, and Israeli *transitiva*. Reproductive behavior is understudied in two subspecies residing in Egypt along the Nile River Valley (*savignii*) and in northeastern Asia, primarily Siberia (*tytleri*).

As described above and outlined in a recent meta-analysis by Romano et al. (*38*), several plumage traits have been studied in barn swallows including: tail length, tail asymmetry, size of white spots on tail, ventral plumage color, throat plumage, color and throat patch size. Of these traits, tail length and plumage color have been the most consistently studied among populations and subspecies. In the 1980s, Møller pioneered experimental studies paired with observational analyses to examine the role of sexual selection on tail streamers in European populations of *rustica*. The experimental manipulation of tail streamers, along with many other studies

conducted in relation to tail streamers and seasonal reproductive performance, established barn swallows as a textbook example of sexual selection (*30*, *105*). Additional studies in Europe have fortified the result that longer tail streamers are associated with several reproductive benefits (*31*). Observational studies demonstrated that darker plumage color variation, rather than the length of tail streamers, was associated with greater seasonal reproductive success in a North American population of *erythrogaster* (*106*, *107*). Experimental manipulations of plumage color (*32*) and both tail streamers and plumage color (*35*) confirm reproductive benefits associated with darker plumage color in males and a lack of evidence for directional selection on streamer length. Additional observational and experimental studies in Israeli *transitiva* provide evidence for directional selection on darker plumage color and longer tail streamers (*33*, *108*). Further, correlational and experimental studies have taken place in *gutturalis* populations in China and Japan, showing that, while this subspecies is pale in color relative to others, color is quite variable and under directional selection (*109*), whereas tail streamer length is not associated with reproductive performance (*110*). Taken together, the direction and degree of sexual selection is associated with the degree to which these two phenotypes differ across subspecies.

A direct comparison of the same experimental manipulations in Israel (*transitiva*) and a North American population (*erythrogaster*) revealed divergent sexual selection (*35*, *39*). Further, the Romano et al. (*38*) meta-analysis confirms divergent patterns of sexual selection among various aspects of phenotype, with stronger directional sexual selection on males. Several additional comparative studies of selection coefficients (*36*) and correlational analyses between plumage color, tail streamer length, and measures of seasonal reproductive success confirm patterns of divergent selection, wherein darker plumage color, not tail streamer length is under sexual selection in North America, tail streamer length and not plumage color is under sexual selection in Europe and a combination of these traits are associated with reproductive success in Israel (*69*).

Plumage traits and reproductive isolation in hybrid zones. In addition to within-population and comparative studies on barn swallow plumage traits, previous investigations of hybrid zone transects in Eurasia have illustrated transitions in phenotype and genotype between populations of *rustica*, *gutturalis*, and *tytleri* that have experienced secondary contact (*42–45*). Geographic cline analyses showed associations between transitions in genome-wide ancestry and ventral plumage color across multiple hybrid zones, suggesting that color may be involved in reproductive isolation (*43*). It was also shown in *rustica-tytleri* and *rustica-gutturalis* hybrid zones that a proxy for migratory distance (carbon isotopes as a measure of overwintering habitat) is associated with genome-wide ancestry. These results dovetail with another previous study of populations across the breeding distribution, which demonstrated that wing length (another proxy migratory distance) is associated with genome-wide differentiation among populations (*68*). A detailed analysis of the *rustica-gutturalis* hybrid zone in China further revealed roles of both migratory behavior and ventral plumage color as potential barriers to gene flow (*44*). Additional details on migratory behavior and its putative role in reproductive isolation are provided in the 'Migratory divides' section below.

Plumage traits and ecology. Investigations in avian systems have inferred combined effects of both sexual and natural selection on plumage variation (*2*, *111*). Indeed, it is very rare to find a trait in nature that is related to mate choice and associated with relevant measures of

reproductive success (and thus under sexual selection) that does not also evolve in response to the environment in some way (*112*). In other words, there are few models or examples which demonstrate that a trait is governed purely by sexual selection in the absence of the environmental context in which the trait is developed, displayed, or both. Most sexually-selected traits will be constrained by natural selection or the ability to survive in a specific environmental context (*2*, *4*).

It is plausible that barn swallow plumage traits evolve in response to ecological selection pressures in addition to divergent sexual selection, yet populations are found throughout the northern hemisphere and exist in a range of climate conditions including high elevation and low elevation deserts (*37*, *113*). Further, a previous analysis found no association between broad climatic variation and ventral coloration across populations throughout the breeding range of the species (*68*). During early development, barn swallow plumage color is impacted by additive genetic variation (genetic contributions by both parents) and is also sensitive to the natal environment (*46*). A likely variable that is important in sexual selection in this system is variation in local parasite pressure. In one analysis, fine-scale plumage trait variation was analyzed in association with parasite communities, revealing that divergent male traits respond to local parasite pressure in allopatry (*69*); this provides strong evidence for a role of sexual selection on male traits which signal information about parasites. It is unclear whether populations in secondary contact zones share overlapping parasite communities.

A broad observation that would be consistent with a role of natural selection on sexual traits would be that melanin-based ventral coloration follows Gloger's Rule, which predicts that more melanistic populations are found at lower latitudes with greater humidity (*114*). However, previous results (*42*, *68*) and our current analyses (Fig. S5) do not support an association between ventral color and latitude (or longitude), indicating that geographic variation in plumage color defies simple ecological patterns of melanism predicted by Gloger's Rule. Thus, while it is possible that plumage traits also experience ecological selection pressures, previous evidence does not suggest that these are strongly divergent among populations and subspecies.

Song variation. In songbirds, acoustic signals are relevant for mate selection and reproductive isolation (*115*). Barn swallow songs are among the most complex of all songbirds. Wilkins et al. (*116*) studied the function of song variation in a population of North American *erythrogaster* and found that various aspects of this complex vocalization are associated with plumage traits and with several measures of reproductive success. Various aspects of song complexity differ among barn swallow subspecies (*117*), but it is unclear as to whether or not these traits play a role in reproductive isolation, as these traits did not show patterns of reproductive character displacement in the Russian *rustica-tytleri* hybrid zone. For this reason, we have focused on plumage traits in a hybrid zone context in the current study.

Migratory divides. Several Eurasian barn swallow hybrid zones are situated at migratory divides associated with the Tibetan Plateau (*43*, *44*), where parental populations on either side of the divide migrate to different overwintering grounds in Africa and southern Asia. To characterize migratory divides in barn swallows, Scordato et al. (*43*) first examined carbon isotopes across multiple hybrid zone transects. Swallows molt and new plumage is generated during overwinter (*118*), and mature feathers sampled in the breeding distribution bear isotopic signatures of the

overwintering habitat. Accordingly, carbon isotopes can be used as a proxy for migratory route and overwinter habitat, and migratory divides can be inferred when isotope signatures are distinct between populations on either side of a putative divide boundary. Based on these signatures, Scordato et al. (*43*) found support for migratory divides between *rustica-tytleri* in Russia and Mongolia and *rustica-gutturalis* in China, but not between *tytleri-gutturalis* in eastern Siberia, Mongolia, and China. They also found narrower geographic clines of genome-wide ancestry across the hybrid zone transects with migratory divides (i.e., *rustica-tytleri* and *rustica-gutturalis*) than those without, with evidence of extensive hybridization between *tytleri* and *gutturalis* in the absence of apparent migratory divides. They further detected strong concordance between the cline for ventral color and clines for ancestry and carbon isotopes in *rustica-tytleri* in Russia, while the ventral color cline for *rustica-gutturalis* was offset and eastward from ancestry and isotope clines in China. In a detailed analysis of the China *rustica-gutturalis* hybrid zone, Turbek et al. (*44*) confirmed evidence of a migratory divide, with geolocator tracks demonstrating that individuals on the western side of the divide migrate to Africa while individuals on the eastern side migrate to southern India. They found associations between genome-wide ancestry and migratory behavior across the hybrid zone, showing that individuals with majority *rustica* versus *gutturalis* ancestry migrate to distinct overwinter habitats along divergent migratory routes. They further showed associated transitions in plumage, with longer tail streamers and less saturated ventral coloration in the western extent of the hybrid zone dominated by *rustica* ancestry.

Together, previous studies support that migratory divides contribute to reproductive isolation in Eurasian barn swallow hybrid zones, where divergence in migratory behavior may promote both prezygotic isolation through differential timing of arrival to breeding grounds and assortative mating and postzygotic isolation through reduced fitness of hybrid offspring with intermediate migratory traits (*43*, *44*). These studies further demonstrate associations between plumage traits and migratory behaviors, and it is plausible that divergent sexual selection on plumage traits may limit the production of hybrids with lower fitness (whether due to reduced survival or reproduction) as a form of reinforcement. Butlin and Smadja (*24*) offer an extended view of reinforcement as any case where the presence of one barrier effect promotes the evolution of a second barrier effect. This is an intriguing possibility in barn swallows, and aligns with our current findings indicating more substantial barriers to gene flow at sexual trait loci in hybrid zones with migratory divides (i.e., *rustica-tytleri* and *rustica-gutturalis*). Yet, the combined impacts of divergent sexual selection and migratory divides on reproductive isolation remain tentative for two primary reasons. First, reduced hybrid fitness as a consequence of intermediate migratory phenotypes has not been demonstrated, as this would require knowledge of the genetic ancestry of individuals that have experienced at least one annual migration and extensive tracking of hybrid migration routes. This is confounded by dispersal from the natal environment, wherein migrating barn swallows in their first year rarely return to the same breeding location at which they were born (*113*), making it challenging to discern lower survival from dispersal of hybrid offspring. Second, the genetic basis of migratory behavior in barn swallows (and indeed the degree to which migratory routes have a genetic basis) is unknown and not tractable to characterize using existing data for the system. This second limitation presents a key direction for future work to characterize the genetic architecture of the timing, direction, and distance of seasonal migration in swallows.

Despite the inherent challenges to confirming combined impacts of divergent sexual selection and migratory divides on reproductive isolation, patterns in the current and previous studies are consistent with their individual impacts, and also with barriers to gene flow at sexual trait loci being strongest in the presence of migratory divides. Here, extensive hybridization between *tytleri* and *gutturalis* (no migratory divide present) presents a distinct pattern from *rustica-tytleri* and *rustica-gutturalis* hybrid zones (migratory divides present), including the observation of higher rates of F1 and early generation hybrids between *tytleri* and *gutturalis* (*42*, *43*). By contrast, few F1 and early generation hybrids have been detected in the *rustica-tytleri* and *rustica-gutturalis* hybrid zones, with much higher incidence of backcrossing to major parent ancestry. Together, these patterns are consistent with the prediction that interactions between sexual and natural selection promote stronger reproductive isolation in the presence of migratory divides as a consequence of reduced hybrid fitness.

***Definition of genetic coupling of barrier loci*.** The concept of coupling has been developed to describe how individual barriers to gene flow operate together to produce stronger reproductive isolation between populations. In principle, coupling can involve any process that leads to the coincidence of barrier effects (i.e., barrier traits or loci) and produces a greater overall barrier to gene flow (*24*, *26*, *29*, *119*). A uniting feature of coupling concepts is that coupled barrier effects produce stronger reproductive isolation because they experience both direct selection and indirect selection through association with other barriers (i.e., the total effect of selection is amplified; (*26*, *27*, *66*, *120*). However, approaches to diagnosing the presence and effects of coupled barriers differ, and multiple definitions of coupling have become established in the literature. Dopman et al. (*119*) recently characterized three main uses (or 'perspectives') of coupling in speciation research, alternatively defined as the build-up of linkage disequilibrium among loci underlying barriers to gene flow (Perspective 1; Felsenstein, *25*), the build-up of genome-wide linkage disequilibrium (Perspective 2; Barton, *26*), or the process that drives the coincidence of distinct barrier effects (Perspective 3; Butlin and Smadja, *24*). The first two perspectives explicitly consider forms of 'genetic coupling', focusing on genetic associations among barrier loci (Perspective 1) or between barrier loci and neutral loci (Perspective 2), though it is important to point out that these are not mutually-exclusive processes (i.e., a build-up of genome-wide linkage disequilibrium can occur following the initial coupling of specific barrier loci; *61*, *67*). Instead, they differ primarily in their focus on linkage disequilibrium among barrier loci or between barrier loci and neutral loci throughout the genome.

In this study, we use an operational definition of genetic coupling following Perspective 1, derived from Felsenstein's three-locus model (*25*) and Barton's (*26*) later definition: genetic coupling is the buildup and maintenance of linkage disequilibrium among barrier loci. This perspective provides a useful framework for testing for evidence of coupling when the genetic architecture of barrier effects is known, and further provides the means to investigate whether the coupling of multiple loci underlying oligogenic or polygenic traits produces stronger reproductive isolation. Accordingly, this definition can be meaningfully applied to investigate the genetic coupling of multiple barrier loci underlying individual barrier effects (*119*) (e.g., ventral color or tail streamer length, respectively, in this study). We focus on the coupling of barrier loci underlying sexual traits in barn swallow hybrid zones, and interpret evidence for genetic coupling among barrier loci when they exhibit greater linkage disequilibrium than genome-wide loci with matching distributions of allele frequency differences between parental

populations (i.e., to account for linkage disequilibrium generated merely as a consequence of admixture; see Main Text and Materials and Methods).

Materials and Methods

Genome sequencing and variant calling. We analyzed whole genomes of barn swallows sampled across the species breeding distribution (*n* = 336; Fig. 2, Fig. S2) including populations of all six subspecies, with a focus on three hybrid zone transects in Eurasia described in (*42*, *43*). The hybrid zones occur in Russia between *rustica* and *tytleri*, in China between *rustica* and *gutturalis*, and in Russia and Mongolia between *tytleri* and *gutturalis*. Our total sampling included representatives of subspecies and hybrid zone, including 31 parental *rustica*, 21 parental *tytleri*, 39 parental *gutturalis*, 58 hybrid *rustica-tytleri*, 74 hybrid *rustica-gutturalis*, and 29 hybrid *tytleri-gutturalis* (Data S1). We sequenced a female wire-tailed swallow (*H. smithii*) as an outgroup. Genomic data for a subset of barn swallows (*n* = 160) were available from previous studies (*41*, *45*) and the remaining samples were newly generated in this study (*n* = 176). For newly sequenced swallows, we sampled blood using medial brachial venipuncture and stored each sample in DNA lysis buffer. We then extracted genomic DNA using Qiagen DNeasy kits following the manufacturer's protocol and prepared Illumina Nextera XT sequencing libraries at the University of Colorado Boulder BioFrontiers Institute using sample-specific barcodes. We sequenced the multiplexed libraries using 150 bp paired-end reads on Illumina NovaSeq 6000 S4 lanes. Raw sequencing data are available on the NCBI short-read archive (accession PRJNA323498).

We quality trimmed and filtered sequencing reads using Trimmomatic v0.39 (*74*) with the settings LEADING:20 TRAILING:20 MINLEN:32 AVGQUAL:30, then mapped filtered reads to the North American barn swallow (*Hirundo rustica erythrogaster*) reference genome assembly bHirRus1 (*75*) using BWA 'mem' v0.7.17 (*76*) with default settings. Only females were mapped to the W chromosome. We sorted outputs and quantified mapping statistics using the Samtools v1.10 (*78*) 'sort' and 'stat' tools, respectively. We called genomic variants using the GATK v4.0.8.1 best-practices workflow (*77, 121*). We first used HaplotypeCaller to generate raw individual variant calls with the '--ERC GVCF ' option to output a genomic VCF per sample, then called variant sites across the cohort of samples and generated an 'all-sites' VCF using GenotypeGVCFs. We otherwise ran GATK HaplotypeCaller and GenotypeGVCFs with default settings.

We filtered raw variant calls using the GATK 'VariantFiltration' tool, with the following variant filters: variant confidence by depth (QD < 2.0), strand-bias (FS > 60.0), among sample mapping quality (MQ < 40.0), mapping quality of heterozygous sites (MQRankSum < -12.5), and distance of variant sites from ends of reads (ReadPosRankSum < -8.0). We identified spurious female heterozygous sites on the sex chromosomes and conservatively masked these sites in all individuals. We further masked genotypes within repetitive regions annotated using RepeatMasker (*122*) then recoded indels and SNPs that did not pass filter settings as missing genotypes using BCFtools v1.10.2 (*78*). We used VCFtools v0.1.17 (*79*) to process SNPs for specific analyses, including removal of singletons (--mac 2), non-biallelic SNPs (--min-alleles 2, --max-alleles 2), and SNPs with minor allele frequencies < 0.05 (--maf 0.05) among all samples. Singleton and minor allele frequency filters were applied after removing sites genotyped in fewer than 80% of individuals (--max-missing 0.2).

We aligned scaffolds from the barn swallow genome to the zebra finch (*Taeniopygia guttata*) reference assembly bTaeGut1.4 (*123*) using MashMap v2.0 (*124*, *125*) to assign scaffolds to chromosomes in the passerine karyotype (Table S13). Part of chromosome 4A in zebra finch is sex-linked in barn swallow due to the neo-sex chromosome transition in the ancestor of Sylvioidea (*126*, *127*). We removed scaffolds that could not be assigned to chromosomes from further analysis.

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 ($K = 1 – 10$) after converting SNP data using Plink (*128*). We evaluated the most likely number of genetic clusters as the *K* model with the lowest cross-validation error (Fig. S4).

Hybrid index and individual heterozygosity. We characterized the genetic composition of admixed populations by estimating individual hybrid index and interspecific heterozygosity at ancestry-informative sites using the R package 'introgress' (*92*). We calculated SNP-based Weir and Cockerham's F_{ST} (*129*) between parental populations using VCFtools (*79*) and SNPs with $F_{ST} \geq 0.6$ were considered ancestry-informative and extracted for analysis of parental and hybrid zone populations. Individual hybrid index and heterozygosity estimates were highly concordant when estimated from SNPs with higher F_{ST} thresholds (i.e., $F_{ST} \geq 0.8, 0.9$).

Demographic inference. We performed demographic inference using the diffusion approximation framework in ∂a∂i (*82*) to investigate the history of divergence and gene flow between parental populations. We fit seven two-population demographic models to the unfolded joint site frequency spectrum (JSFS) between *rustica*, *tytleri*, and *gutturalis* population pairs, using modifications to the original ∂a∂i framework implemented in (*130*, *131*). 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 second timepoint where gene flow occurred (Fig. S8). The other models (AM2*m*, IM2*m*, and SC2*m*) were equivalent to the AM, IM, and SC models, except that two effective migration rates were inferred to simulate scenarios where some regions of the genome are porous to gene flow while others present barriers to gene flow (i.e., genomic 'islands' models). Greater model support for 2*m* models over 1*m* models supports that barriers to gene flow are concentrated in a subset of loci.

Preliminary analyses using the whole genome dataset provided strong support for two-population models with gene flow over the SI model for each pair of populations. However, support among models with gene flow was ambiguous. This was unsurprising given the relatively limited sample size of parental populations and previous inference of extremely recent divergence between barn swallow subspecies (*41*). Indeed, a simulation study demonstrated that larger sample sizes are required to distinguish support for two-population histories with periods of gene flow after recent divergence (*132*). To address this, we reanalyzed reduced representation sequencing (RADseq) data for barn swallows from previous studies (*42*, *44*, *68*), including higher numbers of individuals from parental localities represented in the whole genome dataset (132 *rustica*, 54 *tytleri*, and 328 *gutturalis*; Table S3). We processed the RADseq data as

described in (*45*), including demultiplexing using the Stacks v2.5 'process_radtags' module (*133*) and read quality trimming using Trimmomatic v0.39 (*74*). We mapped reads to the *H. r. erythrogaster* reference genome using BWA 'mem' v0.7.17 and called SNPs using BCFtools v1.10.2 (*78*), recoding individual calls with < 5 mapped reads as missing genotypes. We merged variant calls with the outgroup *H. smithii* sample, then filtered SNPs with missing genotypes in greater than 20% of individuals and retained a single SNP per autosomal RAD locus to prune for linkage disequilibrium. Finally, we polarized ancestral versus derived allele states based on genotypes in *H. smithii* using a custom Python script (https://github.com/kullrich/bio-scripts/blob/master/vcf/polarizeVCFbyOutgroup.py), yielding 15,147 polarized SNPs for analysis.

We used easySFS (https://github.com/isaacovercast/easySFS) to generate the unfolded JSFS for each pair of parental populations. To achieve a balance between the number of segregating sites and sample size, we projected 96 alleles per population. We then performed analysis using ∂a∂i to fit each of the seven models to each JSFS across 20 independent runs per model with random starting parameters and grid sizes of 110, 120, and 130. Model fitting included a cold and hot annealing procedure prior to a BFGS optimization step, with a maximum of 2,000 rounds of optimization. For each model, we retained the run with the lowest log-likelihood and compared fit among models using the Akaike Information Criterion (AIC), and converted unscaled parameters to biological estimates following (*134*), assuming a per-generation mutation rate of 2.3×10^{-9} (*87*), a generation time of one year (*40*), and an effective sampled genome length (L) of 7,279,472 bp.

As in preliminary analyses, demographic inference strongly supports two-population models with gene flow over the SI model for each pair of populations (Figs. S7, S9-S11; Table S4; Data S2). SC models have the highest goodness-of-fit among the models with gene flow, supporting that parental populations diverged in allopatry for a period of time followed by the formation of hybrid zones upon secondary contact. The 1*m* SC model is the best-fit model for *tytleri* and *gutturalis*. The best-fitting model for divergence between both *rustica* and *tytleri* and *rustica* and *gutturalis* is the SC2*m* model, consistent with heterogeneous gene flow across the genome upon secondary contact. Scaled parameter estimates from these models allowed us to infer the approximate timing of initial divergence and secondary contact, effective migration rates upon secondary contact, and effective population sizes. These estimates support that barn swallow populations diverged from a common ancestor roughly 9,200 – 12,500 generations ago, with population growth following initial divergence (Fig. S7; Table S4). The inferred secondary contact times are 2,066 generations ago between *rustica* and *gutturalis*, 981 generations ago between *rustica* and *tytleri*, and 885 generations ago between *tytleri* and *gutturalis*, indicating younger hybrid zone ages between *tytleri* and the other subspecies. These results also align with previous biogeographic inferences that *tytleri* recolonized northeast Asia after the common ancestor of *tytleri* and *erythrogaster* dispersed into North America (*135*). Finally, estimates from the SC2*m* models indicate that effective migration rates in genomic islands were an order of magnitude lower than the rest of the genome between *rustica* and *tytleri* and *rustica* and *gutturalis*, respectively. We caution that these biological estimates are based on multiple assumptions and that each of the demographic models tested necessarily simplifies the true speciation history. Accordingly, we evaluate these as coarse estimates for comparative purposes, though we also note that these estimates align with previous inferences of recent population

growth in most barn swallow populations after divergence from a common ancestor during the Holocene (*41*).

Measurement of mate choice plumage traits. We examined variation in melanin-based ventral plumage coloration and tail streamer lengths, two barn swallow plumage traits shown previously to evolve under sexual selection (*32*, *39*, *136*, *137*), following previous studies (*42–44*, *68*, *106*). We quantified ventral color by sampling 5-10 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.). The plumage color dimensions brightness, hue, and chroma are highly correlated across the ventral region in barn swallows (*32*). Brightness is the most variable, however, with lower values measured from birds with more saturated, dark red feathers and higher values from birds with bright, pale feathers. Accordingly, we used average breast brightness (equivalent to percent reflectance), as a representative ventral color measure in this study. We measured left and right tail streamer lengths to the nearest millimeter, taking an average of three independent measurements per feather. Only one tail streamer was measured from several individuals used in this study. To account for this, we chose one tail streamer at random from the remaining individuals for analysis.

To understand geographic variation in the plumage measures, we tested Spearman's rank correlations between each and latitude and longitude. Ventral color shows no significant correlation with either latitude or longitude (Fig. S5; *P*-values ≥ 0.56). Tail streamer length is positively correlated with latitude (Spearman's $\rho = 0.3$, $P = 2.4 \times 10^{-7}$) and negatively correlated with longitude (Spearman's $\rho =$ -0.18, $P = 0.003$). We further tested the relationship between the plumage traits. Consistent with geographic variation in sexual preferences for each trait (*37*, *39*), we find no significant correlation between ventral color and tail streamer length ($P = 0.11$).

To further examine phenotypic variation in hybrid zones, we reanalyzed the dataset of 1,288 swallows from Scordato et al. (*43*), which includes all samples in the current study that were used for genome sequencing and genome-wide association mapping (see below). This dataset includes high-resolution sampling of hybrid zone transects and enables more detailed analyses of phenotypic differences among parental and hybrid populations than if we were to restrict these to only whole genome-sequenced individuals. We compared distributions visually (shown in Fig. 2C) and performed ANOVA to test whether ventral color and tail streamer length distributions differed among populations for each hybrid zone. ANOVA was significant for each hybrid zone and trait (Table S2), so we performed Tukey *post-hoc* tests to test pairwise differences between populations while correcting for multiple testing. The results are shown in Table S2. Ventral color is significantly different between parental populations of each hybrid zone (*P*-values $< 1 \times 10^{-4}$) and both hybrid *rustica-tytleri* and *tytleri-gutturalis* fall intermediate between, and differ significantly from, respective parental populations (*P*-values $< 1 \times 10^{-6}$). Hybrid *rustica-gutturalis* ventral color is significantly different than parental *gutturalis* (*P*-value $< 1 \times 10^{-4}$), but not parental *rustica* (*P*-value $= 0.16$). Tail streamer length differs significantly for pairs of populations across the *rustica-gutturalis* and *tytleri-gutturalis* hybrid zones (*P*-values $< 1 \times 10^{-3}$), with hybrids falling intermediate between parental distributions. Tail streamer length differs

significantly only between hybrids and parental *rustica* across the *rustica-tytleri* hybrid zone (*P*-value = 0.0038).

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. Models in GEMMA do not accommodate missing genotypes, so we first imputed missing data using BEAGLE (*84*) after filtering to remove SNPs with greater than 20% missing data among individuals, then converted the imputed data using PLINK (*128*). We first ran Bayesian sparse linear mixed models (BSLMM) to broadly summarize the genetic architecture of each trait. These models are capable of detecting genetic architectures with varying degrees of complexity, simultaneously modeling distributions including few loci of large effect (i.e., single-locus or oligogenic basis) and many loci of small effect (polygenic basis). We ran BSLMMs using ten independent chains and 25 million MCMC steps after a burn-in of 5 million steps, sampling every 1,000 steps. We incorporated a relatedness matrix in all analyses to control for population stratification. We then combined results across the independent runs and summarized the genetic architecture of each trait using posterior distributions of three main hyperparameters: the proportion of phenotypic variance explained by all SNP genotypes (PVE), the proportion of PVE explained by SNPs with detectable effect sizes or 'sparse effects' (PGE), and the number of SNPs required to explain the estimated PVE (*n* SNPs). We further broadly quantified the architecture of each trait by calculating the sum of the posterior inclusion probability (PIP) for SNPs with detectable effects (PIP $\geq$ 0.01) per chromosome compared to the chromosome length (Fig. S12).

In addition to BSLMMs, we ran linear mixed models (LMM) to detect isolated SNP associations with each trait. We again focused our primary analyses on hybrids only and ran LMMs under additional sampling schemes for comparison. The first alternative sampling scheme included all individuals with matched genotypes and phenotypes (i.e., the 'full' dataset), including all subspecies and hybrid zones. This sampling scheme provides the greatest power to detect SNP associations, but is also the most prone to potential biases introduced by population structure. To account for such biases, we reanalyzed the full dataset using randomized phenotypes among individuals. If SNP associations were driven by underlying genetic structure across genomic regions, we would expect to find these even when phenotypes were randomized. In contrast, no genomic regions are strongly associated with any trait in randomized phenotype LMMs (Fig. S13), indicating that significant associations were generally robust to genetic structure. The results from the 'full' dataset were also largely congruent with analyses of hybrids. Finally, we ran LMMs for ventral color and tail streamer length for males and females to determine if separate analyses provided additional signal due to sex differences in these traits (Fig. S13). These analyses again yield largely congruent results with the hybrid dataset for ventral coloration, however associations with tail streamer length disappear when males and females are analyzed separately. 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; Data S3).

LMMs revealed multiple significantly associated regions of chromosome 1A and the Z chromosome with ventral color, the details of which are as follows. One association on chromosome 1A contains KIT ligand (*KITLG*; Fig. 3E; $P = 1.38 \times 10^{-14}$), which controls melanin patterning through KIT receptor tyrosine kinase-expressing melanocyte proliferation, migration, and survival (*138*), and has been implicated in multiple previous studies of vertebrate pigment variation (e.g., *47*, *48*). The other chromosome 1A association houses plexin C1 (*PLXNC1*; $P = 2.94 \times 10^{-20}$), a transmembrane semaphorin receptor that regulates melanocyte adhesion (*139*), though this region also contains nearby genes with no known functions in melanogenesis (e.g., *MRPL42*; Fig. S15).

One association on the Z chromosome includes the region containing two candidate genes putatively involved in feathering phenotype, *SPEF2* and *PRLR* (*140*), nearby solute carrier family 45 member 2 (*SLC45A2*; Fig. 3E; $P = 1.45 \times 10^{-13}$), which encodes a transporter protein that supports melanin synthesis in mature melanosomes (*49*). Coding mutations in *SLC45A2* result in forms of albinism in mice (*141*), humans (*142*), and birds (*143*), and previous studies have showed that differential gene expression of *SLC45A2* also contributes to pigment variation in hybrids (*14*, *51*, *144*). Other Z-linked associations contain additional genes with known or tentative roles in melanogenesis (Fig. 3C,E). Basonuclin (*BNC2*; $P = 1.92 \times 10^{-13}$) has been linked to pigment pattern and saturation in zebrafish (*50*), anoles (*145*), and humans (*146*). G-protein subunit alpha Q (*GNAQ*; $P = 4.08 \times 10^{-13}$), APC regulator of Wnt signaling (*APC*; Fig. S20; $P = 8.5 \times 10^{-18}$), and calcium/calmodulin dependent protein kinase 4 (*CAMK4*) are also differentially expressed in melanocytes of hooded and carrion crows (*51*). The specific function of *CAMK4* in melanin production is unclear, however previous evidence indicates that *GNAQ* controls pigmentation in coordination with KIT signaling (*147*), and *APC* is an antagonist of the Wnt signaling pathway, which regulates the expression of the melanocyte inducing transcription factor (*MITF*) essential to melanin production. Finally, an association on the Z chromosome contains receptor tyrosine kinase-like orphan receptor 2 (*ROR2*; $P = 1.12 \times 10^{-13}$), which may also regulate melanin production through interaction with Wnt signaling, specifically during the differentiation of melanocyte precursor cells (*148*).

Associations with tail streamer length are in two regions on chromosome 2 (Fig. 3F; Figs. S13, S21-S22). One is upstream of little elongation complex subunit 1 (*ICE1*; $P = 5.34 \times 10^{-9}$) and downstream of a long noncoding RNA (lncRNA) and the other contains phosphodiesterase 1C (*PDE1C*; $P = 2.04 \times 10^{-20}$).

Recombination rates. We inferred a barn swallow recombination map using pyrho (*85*), which accounts for nonequilibrium 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, as avian genomic recombination landscapes are highly conserved among closely-related lineages (*149*). Pyrho incorporates population demography in recombination rate estimation, so 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 based on the sum of log-likelihoods for each distinguished lineage pair and 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. We compared recombination rates in trait loci and the genome background using Mann-Whitney U tests.

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, based on the genotype dataset including variant and invariant sites. We performed genome scans in non-overlapping sliding windows at several resolutions (1 Mb, 100 kb, and 10 kb) to examine regional variation within and among chromosomes. 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 a 5 kb step size). We compared distributions of these statistics using Mann-Whitney U tests. All statistical comparisons were performed using values calculated from non-overlapping windows.

To provide context for interpretations of population genetic differentiation across trait loci, we evaluated genome-wide relationships between parameters using Spearman's correlation coefficients. These relationships are consistent with heterogeneous differentiation being driven by linked selection across the genome, including negative relationships between F_{ST} and recombination rate, π, and d_{xy} (Fig. S25; Table S10). We also examined the overlap between association mapping peaks and F_{ST} between parental populations by sampling mean F_{ST} for the non-overlapping 10 kb window containing each significant GWA SNP and comparing these to all other windows (Table S8). We sampled each window only once in instances where multiple significant SNPs were within the same window, and compared distributions using Mann-Whitney U tests.

We compared F_{ST} values for sexual trait loci to those from putative centromere regions to test if population differentiation is elevated in trait loci beyond levels in genomic regions where the effects of recurrent background selection are expected to be strong due to low recombination rates, such as centromeres. We identified putative centromere regions for the six longest chromosome scaffolds in the barn swallow genome assembly (chromosomes 1, 1A, 2, 3, 4, and Z) based on combined signatures of high repeat content, measured as the proportion of annotated repeat element bases, and low recombination rate in 1 Mb sliding windows with a 100 kb step size (Fig. S26A). We restricted these analyses to these six chromosomes because attempts to identify putative centromeres on shorter scaffolds did not provide clear evidence of a distinct centromeric region per chromosome, however the focal set includes chromosomes with major candidate sexual trait loci (i.e., chromosomes 1A, 2, and Z). We compared mean F_{ST} values in non-overlapping 10 kb windows within putative centromeres to those within trait loci using Mann-Whitney U tests. These tests indicate higher F_{ST} values in trait loci than centromere regions between each pair of parental populations, with more pronounced differences in median F_{ST} in trait loci versus centromeres in *rustica-tytleri* and *rustica-gutturalis* than *tytleri-gutturalis* (Fig. S26B; P-values $< 2.2 \times 10^{-16}$).

Selection statistics. We calculated additional statistics to detect signatures consistent with positive selection in genomic regions associated with mate choice traits. We calculated allelic

differentiation specific to each of the parental populations that form hybrid zones (*rustica*, *tytleri*, and *gutturalis*) using the population branch statistic (PBS). PBS can detect genomic regions with long population-specific branches in the unrooted three-population tree, allowing us to disentangle which populations may have experienced positive selection in differentiation peaks across the genome. We calculated PBS using windowed pairwise F_{ST} estimates among the three parental populations as 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. For example, $PBS_{rustica} = \frac{T^{R,T}+T^{R,G}-T^{T,G}}{2}$. We calculated Tajima's *D* (*150*) to summarize the allele frequency spectrum using VCF-kit (*90*), and summarized PBS and Tajima's *D* statistics in the same sliding window resolutions used for genetic differentiation and diversity statistics. We used the R package 'rehh' (*91*) to calculate statistics based on extended haplotype homozygosity (EHH; *151*), including the integrated haplotype score within populations (iHS; *152*) and cross-population extended haplotype homozygosity (xp-EHH; *153*). These analyses also yield a per-SNP *P*-value for iHS and xp-EHH based on genome-wide ranking of scores. We phased haplotypes within the parental populations using SHAPEIT2 (*154*) after identifying phase-informative reads mapped to the reference genome with the 'extractPIRs' tool. We then performed haplotype-based tests in rehh, specifying a minor allele frequency cutoff ≥ 0.05 and 'polarized = false'. We used Mann-Whitney $\cup$ tests to detect differences in PBS, π, Tajima's *D*, and |iHS| and Komogorov-Smirnov tests to detect differences in xp-EHH between trait loci and the genome background.

Geographic and genomic clines. We used sigmoid geographic clines across hybrid zone transects to test for signal of barriers to gene flow in trait loci relative to the genomic background. These analyses required sampling of linear geographic transects, so we subsampled localities to maximize linearity and converted geographic coordinates to distance in kilometers (km) from the westernmost location of each transect. We used introgress (*92*) to estimate 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 to genome-wide patterns, we sampled at random 1,000 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 $\pm$ 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}). The six cline models differed in whether p_{min} and p_{max} were 'fixed' or 'free' parameters and how the tails of the cline were fitted (i.e., 'none', 'both', 'right', 'left', and 'mirrored'). The model specifications were: model I (fixed, none), model II (free, none), model III (free, both), model IV (free, right), model V (free, left), and model VI (free, mirrored). 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 Δ AIC ≥ 2 threshold. In cases where there was similar support for two models, clines for both models are compared in Fig. S36. We compared cline parameters from the best-fitting models for trait loci to background loci using Mann-Whitney $\cup$ tests.

Our sampling of the genome background was not restricted to neutral loci (i.e., we did not filter background loci based on allele frequency differences between populations). Nonetheless, to more explicitly compare geographic cline patterns for sexual trait loci to other loci plausibly under selection (e.g., divergent ecological selection), we extracted cline parameters from background loci with allele frequency differences between parental populations that matched those in sexual trait loci for each hybrid zone transect. Specifically, we sampled loci with maximum and mean allele frequency differences within one standard deviation of the mean among sexual trait loci. For example, matched background loci in the *rustica-tytleri* hybrid zone had maximum and mean parental allele frequency differences equal to 0.81 ± 0.16 and 0.24 ± 0.12, respectively. These are relatively high allele frequency differences in barn swallows, such that sampling matched background loci in this way should enrich for loci putatively under selection (even if this approach is agnostic to the source of selection). We compared cline width (w) estimates for trait loci and matched background loci using Mann-Whitney $\cup$ tests. These comparisons indicate significantly narrower clines for sexual trait loci than matched background loci in the *rustica-tytleri* and *rustica-gutturalis* hybrid zones, regardless of whether we compared based on maximum or mean allele frequency differences between parental populations (P-values < 0.001; Fig. S37). While these results suggest stronger barriers to gene flow at sexual trait loci in these two hybrid zones, we exercise caution in not ruling out that traits under divergent ecological selection may also contribute to reproductive isolation.

We further dissected clinal patterns between sexual trait loci and background loci plausibly under divergent selection by comparing w estimates in matching intervals of increasing allele frequency differences. These analyses necessarily focused on individual SNPs, rather than loci, as there was an insufficient number of trait and background loci in matching intervals to allow for comparison at this finer scale. Here, we sampled 50 SNPs at random from trait loci and the genome background with allele frequency differences between parental populations in intervals of 0.05 between 0.2 and 0.5 and fit sigmoid geographic cline models in HZAR to SNP genotype frequencies, as describe for hybrid indexes above. The results of these analyses are generally consistent with locus-based inferences for *rustica-tytleri*, but are rather equivocal for *rustica-gutturalis* and *tytleri-gutturalis* (Fig. S38A-C), likely driven by two main factors. First, individual SNPs may provide limited ancestry information across the hybrid zone. Indeed, even with increasing allele frequency differences between populations, genotype frequencies estimated from a single SNP provide a coarse approximation of hybrid index, thus cline models fit to SNP-based inferences may introduce more substantial noise. Second, differentiating cline parameters for matched trait and background SNPs is difficult due to the limited availability of genome-wide background SNPs in higher allele frequency difference intervals. Rather, background SNPs tend to be clustered in close linkage with the trait loci due to the concentration of genetic differentiation in these regions, and not elsewhere in the genome (Fig. S38D-E). This issue intensifies at even higher allele frequency difference intervals (i.e., 0.55 – 1.0), where it is impossible to draw a sufficient random sample of background SNPs that are not tightly physically linked to trait loci (i.e., trait locus and background SNPs are not independent due to linkage). Because of these limitations, the locus-based inferences above likely provide a more robust overall assessment of patterns at trait loci and other loci putatively under selection. Further, the difficulty in disentangling clinal patterns between trait and background SNPs with high allele frequency differences generally highlights the distinct patterns in sexual trait loci.

As a comparison to the geographically-explicit analyses above, 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 $\phi_{ij} = (h_j^{vi})/(h_j^{vi} + (1 - h_j^{vi}) * e^{ui})$, where *h* is the proportion of the genome inherited from population 2 (i.e., hybrid index), *v* is the cline slope relative to the genome average ($v = 1$), and *u* is related to the cline center. Following (*97*), we used the conversion logit(c) = *u*/*v* 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(*v*) and logit(*c*), which each have an expected mean of 0.

We fit genomic clines to trait and background loci using the same inputs as in geographic cline fitting, using a modification of the Bayesian genomic clines model implemented in bgc (*94*). Specifically, we equated the hybrid index, *h*, for each window as an estimate of the local ancestry, *z*, for that window in the relevant hybrid individual. As an example, *h* = 0.3 would indicate a 30% chance of inheriting the window segment from one of the two parental populations (and thus 70% chance of ancestry from the other population). We thus computed the log-likelihood function for the Bayesian model as $\sum_i \sum_j \log(\phi_{ij} z_{ij} + (1 - \phi_{ij}) * (1 - z_{ij})) * g_{ij}$, where g_{ij} is the ploidy (1 or 2) for individual *i* at locus *j* (this allows for hemizygous loci on the Z chromosome in females). We placed normal priors on the cline parameters (log(*v*) and logit(*c*), as noted above) with means of 0 (soft centering) and standard deviations, σ_v and σ_c, which were estimated from the data. Importantly, these parameters describe the variability in patterns of introgression across the genome and thus the degree to which selection acts more independently on individual genetic loci versus against hybrid genomes as a whole due to linkage disequilibrium and coupling among barrier loci. We placed moderately constraining priors on these parameters, both normal with a mean of 0 and standard deviation of 0.08. The model was then fit using Hamiltonian Monte Carlo (HMC) using rstan v2.21.7, an R (here v4.1.2) interface with Stan. This was done using the No-U-Turn Sampler (NUTS) with 8 chains, each comprising 3,000 HMC iterations and a 1,500 iteration warmup. Effective sample sizes and the Gelman-Rubin potential scale reduction factor were evaluated to ensure adequate HMC mixing and likely convergence to the posterior distribution.

Analysis of linkage disequilibrium. We measured inter-chromosomal linkage disequilibrium (LD) between SNPs in trait loci with allele frequency differences between parental populations ranging from 0.2 – 0.6 in intervals of 0.05. To test whether LD between trait loci was outstanding compared to background levels of admixture LD in hybrids, we also measured inter-chromosomal LD between SNPs in the genome background (i.e., outside of trait loci) in matched allele frequency difference intervals. We used VCFtools to summarize allele frequencies at biallelic SNPs in parental populations, then calculated allele frequency differences as the absolute value of the difference in reference allele frequency between each pair of populations. We then extracted SNPs in allele frequency difference intervals from trait loci and the genome background and calculated inter-chromosomal haplotype r^2 using VCFtools in each hybrid zone and parental population. Because measures of LD are sensitive to sample size, we sampled 20 individuals at random for all analyses to enable comparisons between populations. We also sampled a random number of background SNPs equal to the number of SNPs in trait loci in each interval. For additional comparison to LD patterns in genomic regions where we would expect hybridization to generate especially pronounced admixture LD, we repeated these analyses but

restricted our sampling of matched background SNPs to putative centromere regions where recombination rate is extremely low. To further investigate levels of LD between specific pairs of trait loci, we calculated both intra- and inter-chromosomal 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 to 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, and then calculated inter-chromosomal LD as a function of increasing allele frequency differences as described above. We used Mann-Whitney U tests to test for significant differences between distributions of inter-chromosomal r^2.

To compare inter-chromosomal LD to LD in linked regions, we examined the decay of LD as a function of distance between physically-linked SNPs in parental and hybrid populations. We performed separate analyses for autosomes and the Z chromosome. Autosomal analyses were based on a sample of 100 random 100 kb windows. We filtered to retain biallelic SNPs with minor allele frequency ≥ 0.05 in each population, then calculated haplotype r^2 between all SNPs within 25 kb of each other using VCFtools, and summarized LD decay by calculating mean r^2 in 100 bp distance intervals between SNPs.

***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*).

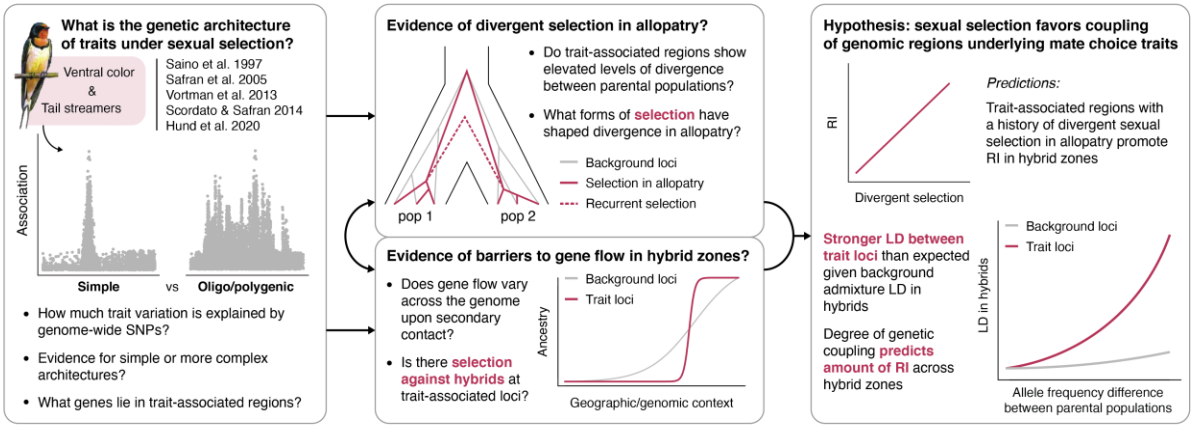

Figure S1. Study design and major research questions. Overview of our approach to testing the hypothesis that genetic coupling of mate choice traits promotes reproductive isolation (RI), including our primary research questions. We first use genome-wide association mapping to determine the genetic architecture of ventral color and tail streamer lengths, two traits previously shown to evolve under divergent sexual selection. We then test for population genetic evidence of divergent selection on trait-associated regions of the genome and test whether these are also barriers to gene flow in hybrid zones that have formed in secondary contact. Gene trees illustrating expected patterns for background loci and trait loci under alternative models of selection were redrawn from (*55*, *155*). Finally, we test core predictions of genetic coupling theory – that selection has maintained or enhanced linkage disequilibrium (LD) between trait loci in hybrid zones beyond that of background admixture LD – and that the amount of RI affected by trait loci scales with the degree of coupling across distinct hybrid zones.

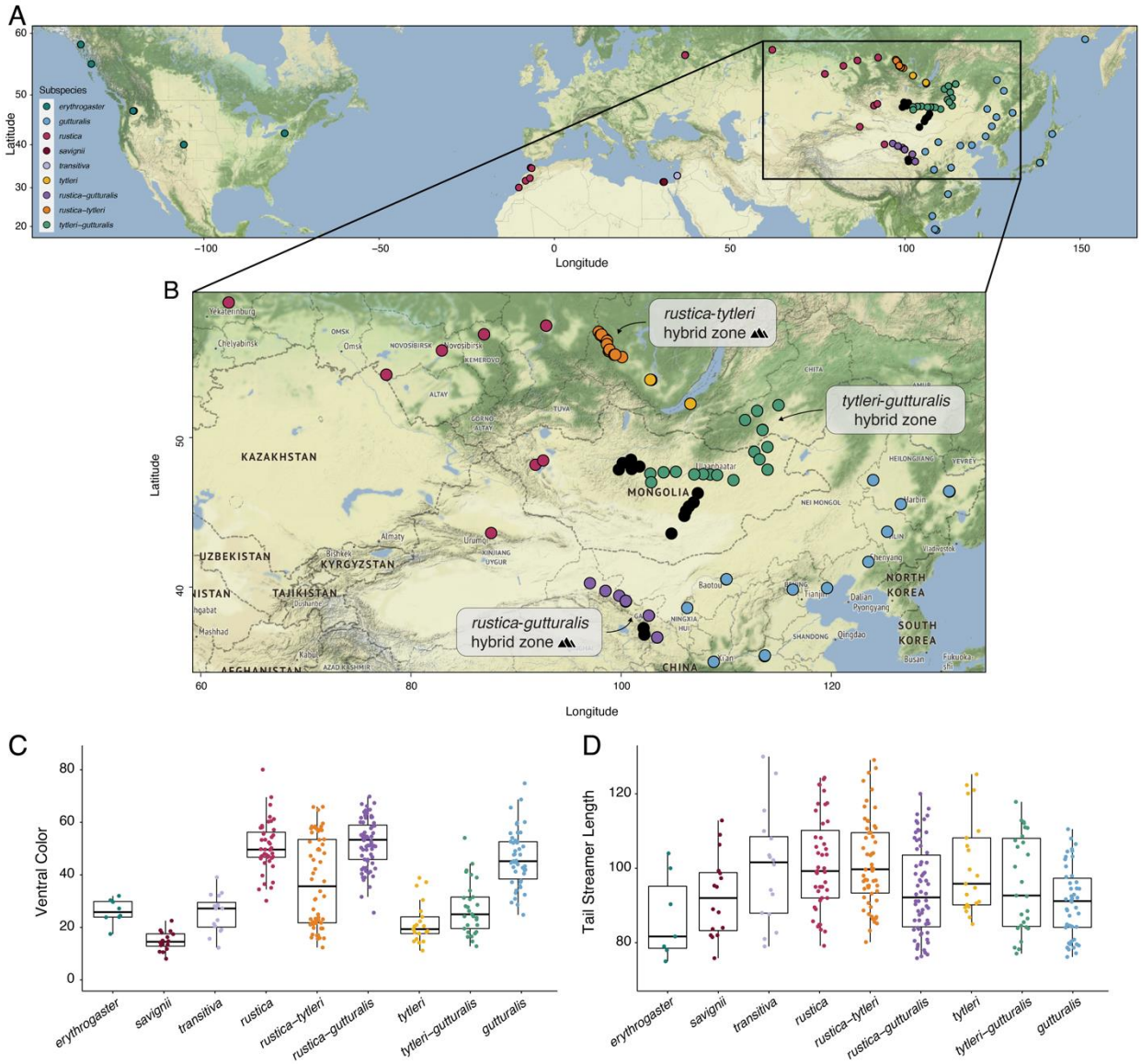

Figure S2. Geographic sampling of barn swallows in this study and phenotypic variation. (**A**) Sampling localities are shown with circles, with colors indicating the geographic zones of subspecies and hybrid zones. Black circles depict sampling localities where no breeding birds were found. The box surrounds the region shown in panel (**B**). (**B**) Inset showing hybrid zone regions in greater detail. Mountain icons indicate migratory divides at the *rustica-tytleri* and *rustica-gutturalis* hybrid zones. (**C**) Variation in ventral plumage color (measured as breast feather average brightness; % reflectance) among subspecies and hybrid zones. (**D**) Variation in tail streamer lengths in millimeters.

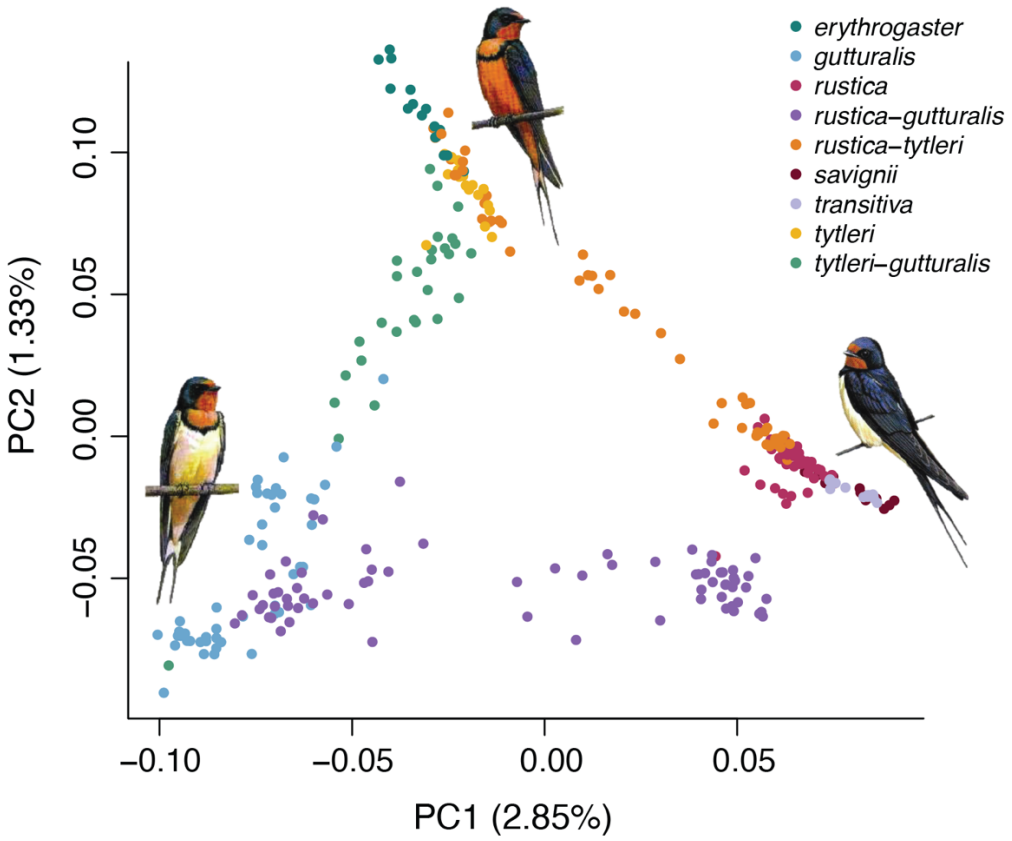

Figure S3. Summary of population genetic structure using principal components analysis. Genetic structure is summarized based on principal components one (PC1) and two (PC2). Points represent individual barn swallows.

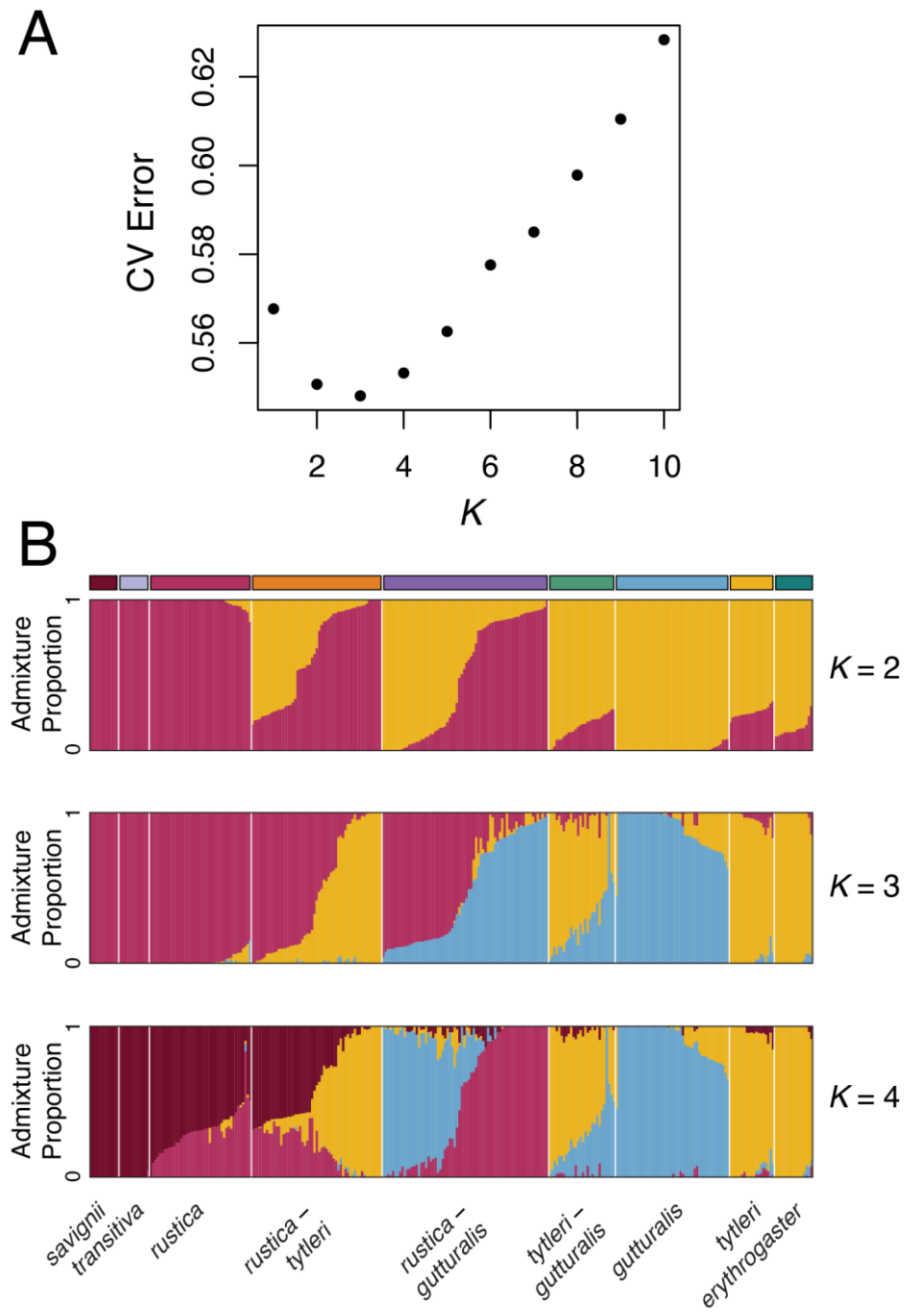

Figure S4. Cross-validation (CV) error and admixture proportions estimated from model-based ancestry inference in ADMIXTURE. A The y-axis shows the CV error associated with each of *K* genetic cluster models 1-10. The *K* = 3 model had the lowest CV error (0.548). **B** Admixture proportions under *K* = 2, *K* = 3, and *K* = 4 models.

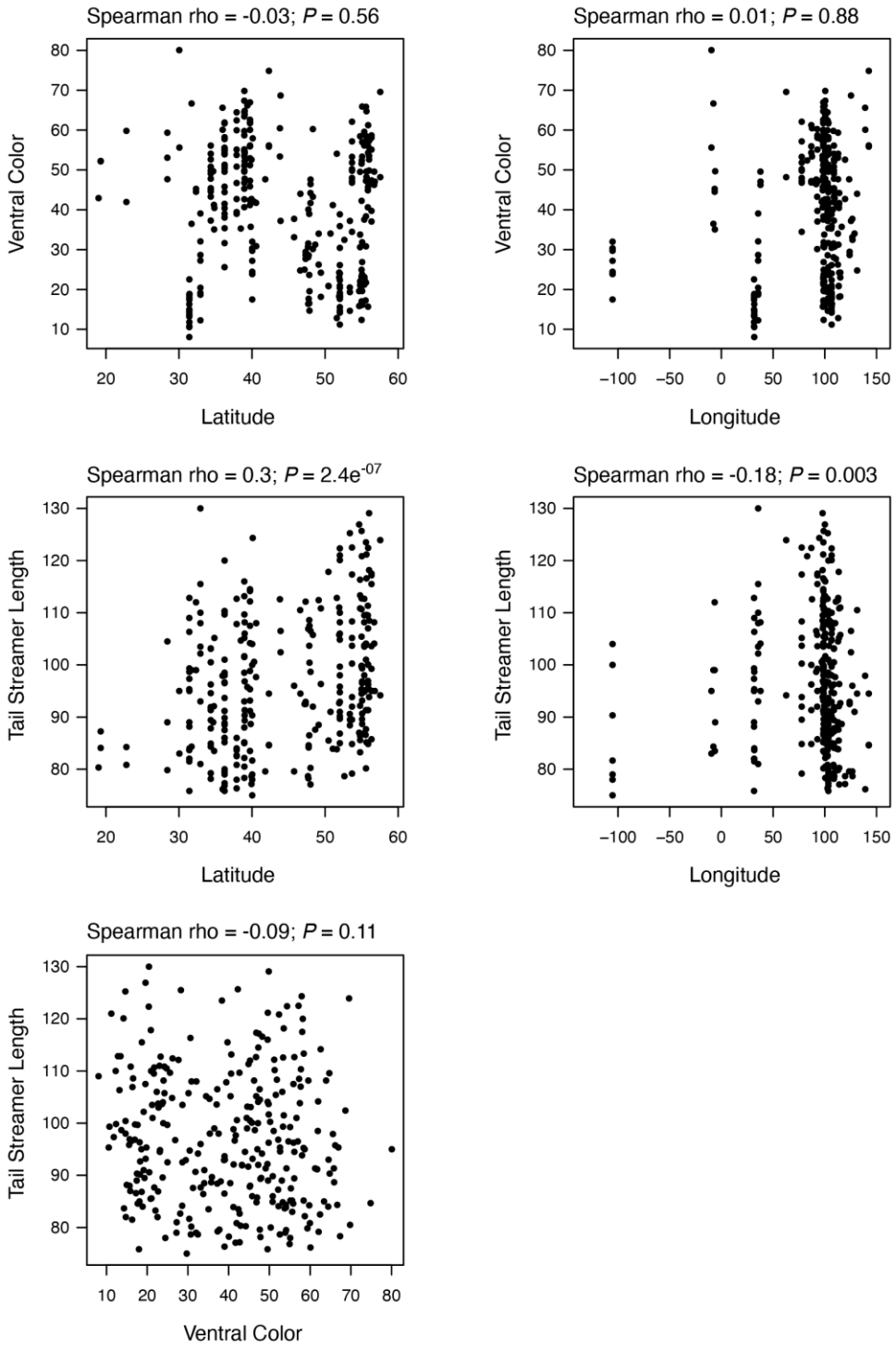

Figure S5. Geographic variation in barn swallow mate choice traits. Spearman's rank order correlations between latitude and longitude and ventral color and tail streamer length.

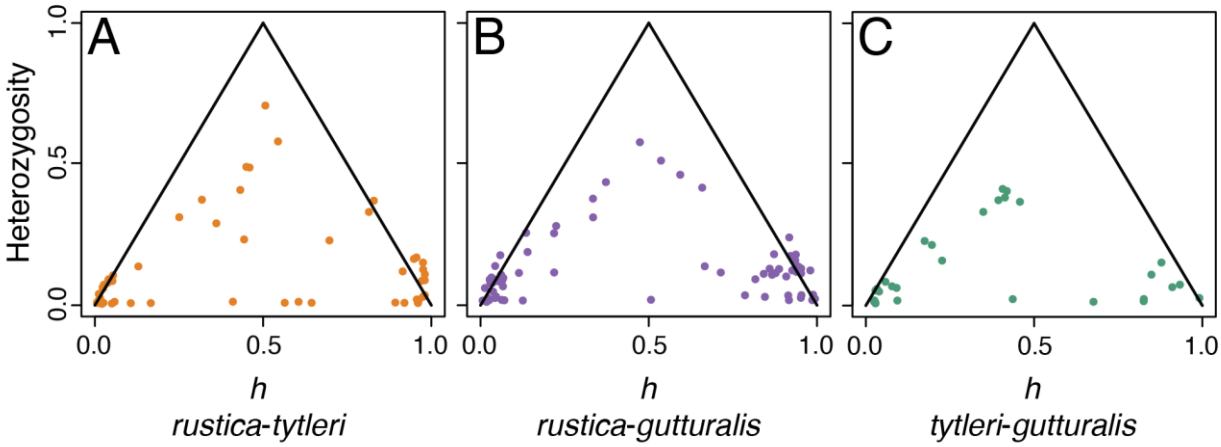

Figure S6. Genetic composition of hybrid zones. Comparisons of individual hybrid index (*h*, ranging from 0 – 1 between parental ancestries; x-axis) and interspecific heterozygosity (y-axis) in the (**A**) *rustica-tytleri* hybrid zone, (**B**) *rustica-gutturalis* hybrid zone, and (**C**) *tytleri-gutturalis* hybrid zone.

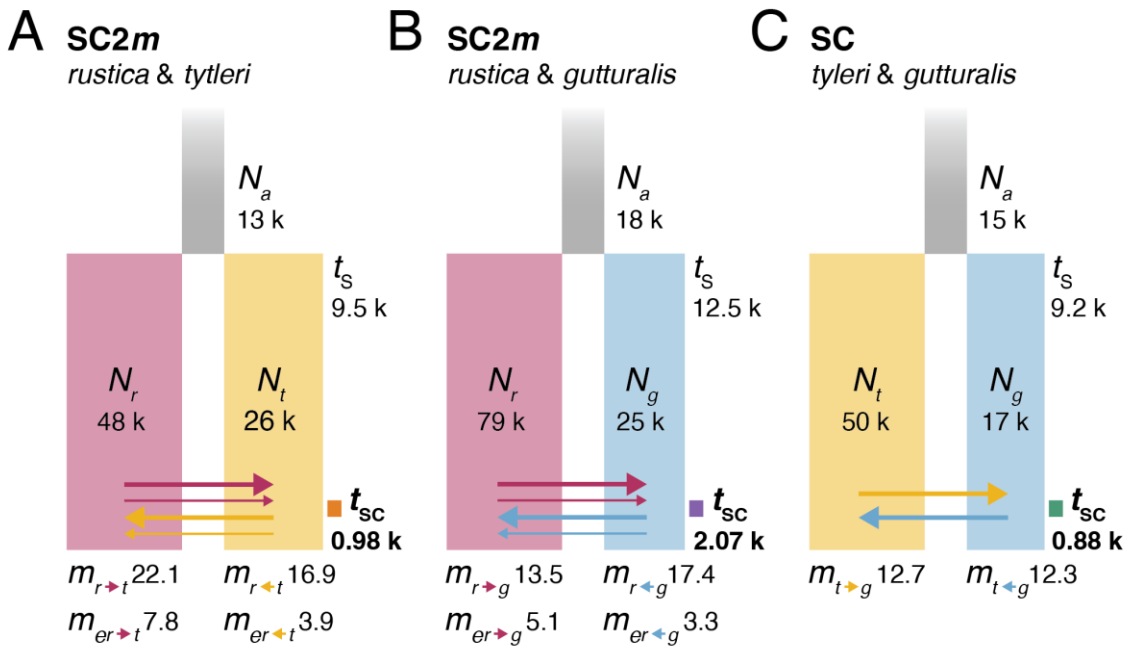

Figure S7. Two-population demographic histories estimated from the unfolded joint site frequency spectrum. Best-fitting models for (**A**) *rustica* and *tytleri*, (**B**) *rustica* and *gutturalis*, and (**C**) *tytleri* and *gutturalis*, supporting allopatric divergence followed by secondary contact with gene flow (SC = secondary contact, SC2*m* = secondary contact with heterogeneous gene flow). Abbreviations: N = effective population size, m = effective migration rate, m_e = effective migration rate in genomic islands, t_s = timing of initial divergence, t_{SC} = timing of secondary contact. Time parameters are scaled to numbers of generations (1 year generation time). Greater model support for 2*m* models over 1*m* models supports that barriers to gene flow are concentrated in a subset of the genome.

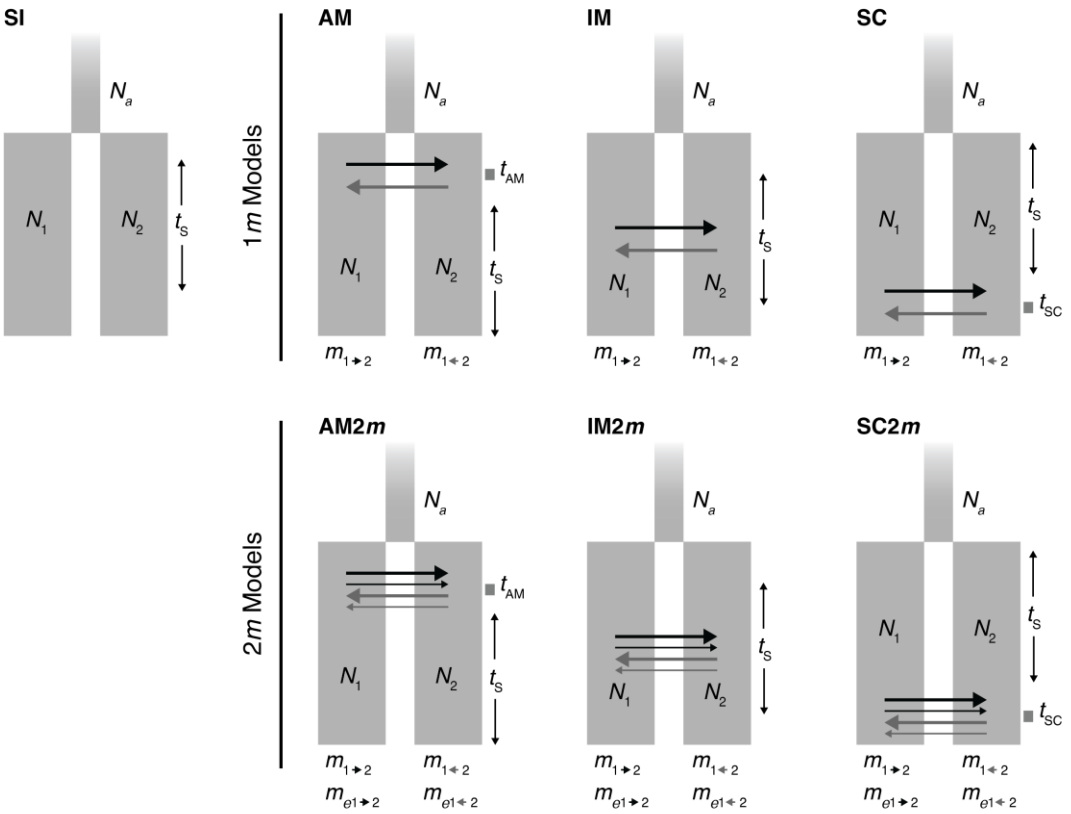

Figure S8. Two-population divergence models used in demographic inference. The models fit to the observed unfolded joint site-frequency spectra (JSFS) between parental populations were strict isolation (SI), ancient migration (AM), isolation-migration (IM), secondary contact (SC), ancient migration with heterogeneous gene flow across the genome (AM2*m*), isolation-migration with heterogeneous gene flow (IM2*m*), and secondary contact with heterogeneous gene flow (SC2*m*). Model parameter estimates included effective population size (N), effective migration rate (m; estimated in 1*m* and 2*m* models), effective migration rate in genomic islands (m_e; only estimated in 2*m* models), time since divergence (t_S), duration of ancient migration (t_{AM}), and timing of secondary contact (t_{SC}).

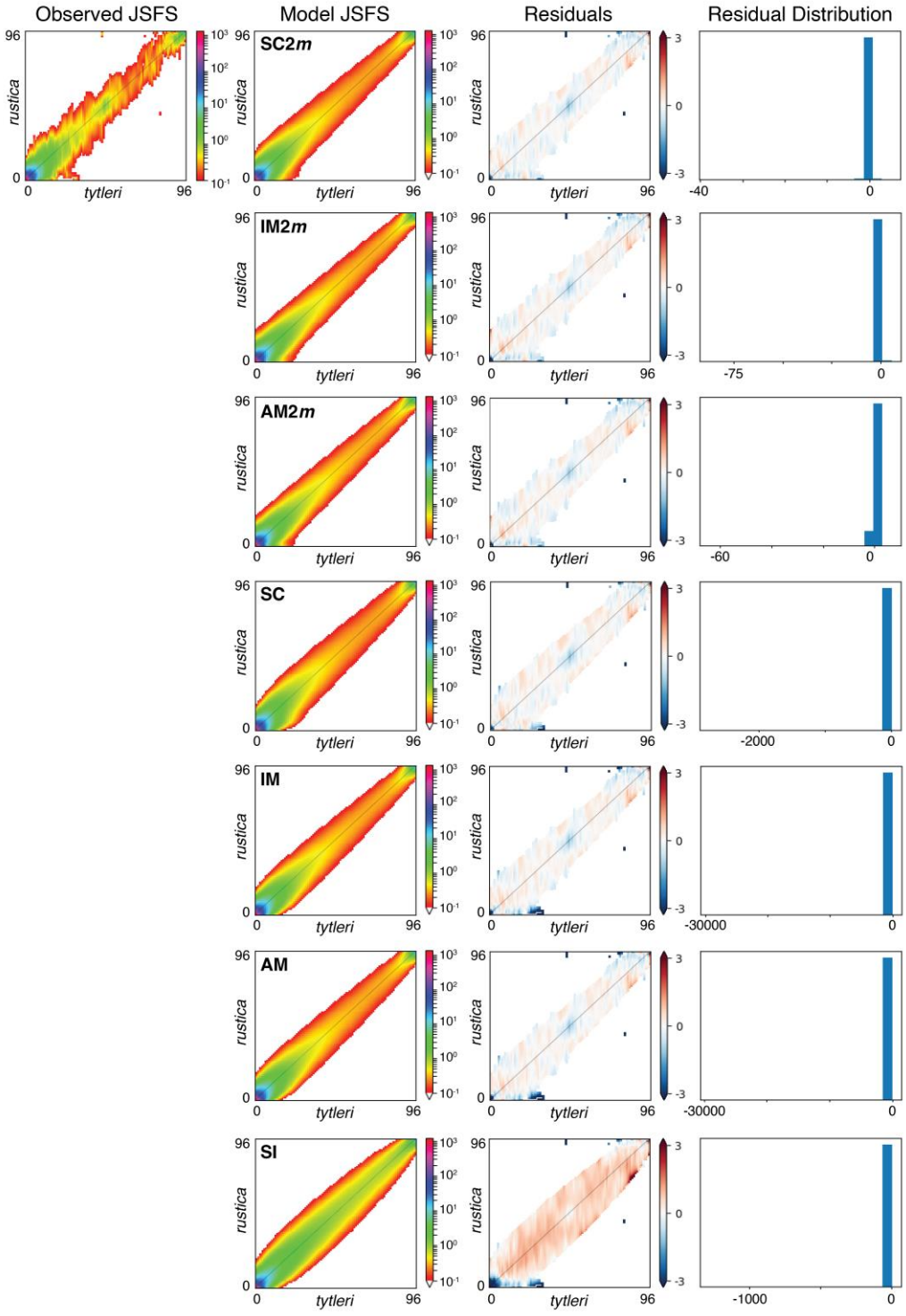

Figure S9. Comparisons of the observed unfolded joint site-frequency spectrum (JSFS) between parental *rustica* and *tytleri* and the JSFS simulated under demographic models. The JSFS from the best-fitting run for each model is shown in the left column. Residuals between the observed JSFS and the model are shown in the center column and the histogram with residual frequencies is shown in the right column. The best-fitting model is shown in the top row, with other models in descending order by -Δ AIC (see Data S2).

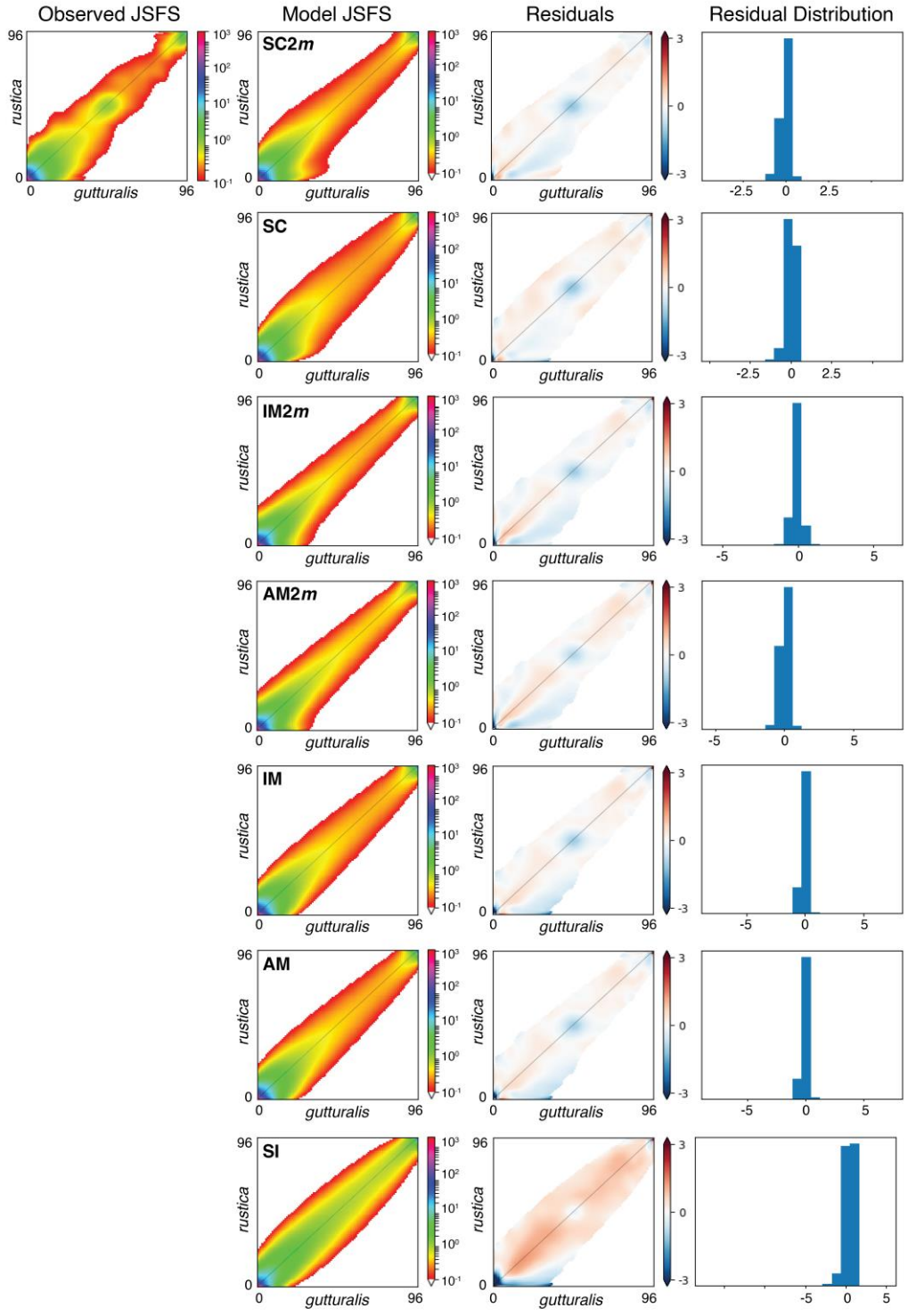

Figure S10. Comparisons of the observed unfolded joint site-frequency spectrum (JSFS) between parental *rustica* and *gutturalis* and the JSFS simulated under demographic models. The JSFS from the best-fitting run for each model is shown in the left column. Residuals between the observed JSFS and the model are shown in the center column and the histogram with residual frequencies is shown in the right column. The best-fitting model is shown in the top row, with other models in descending order by -Δ AIC (see Data S2).

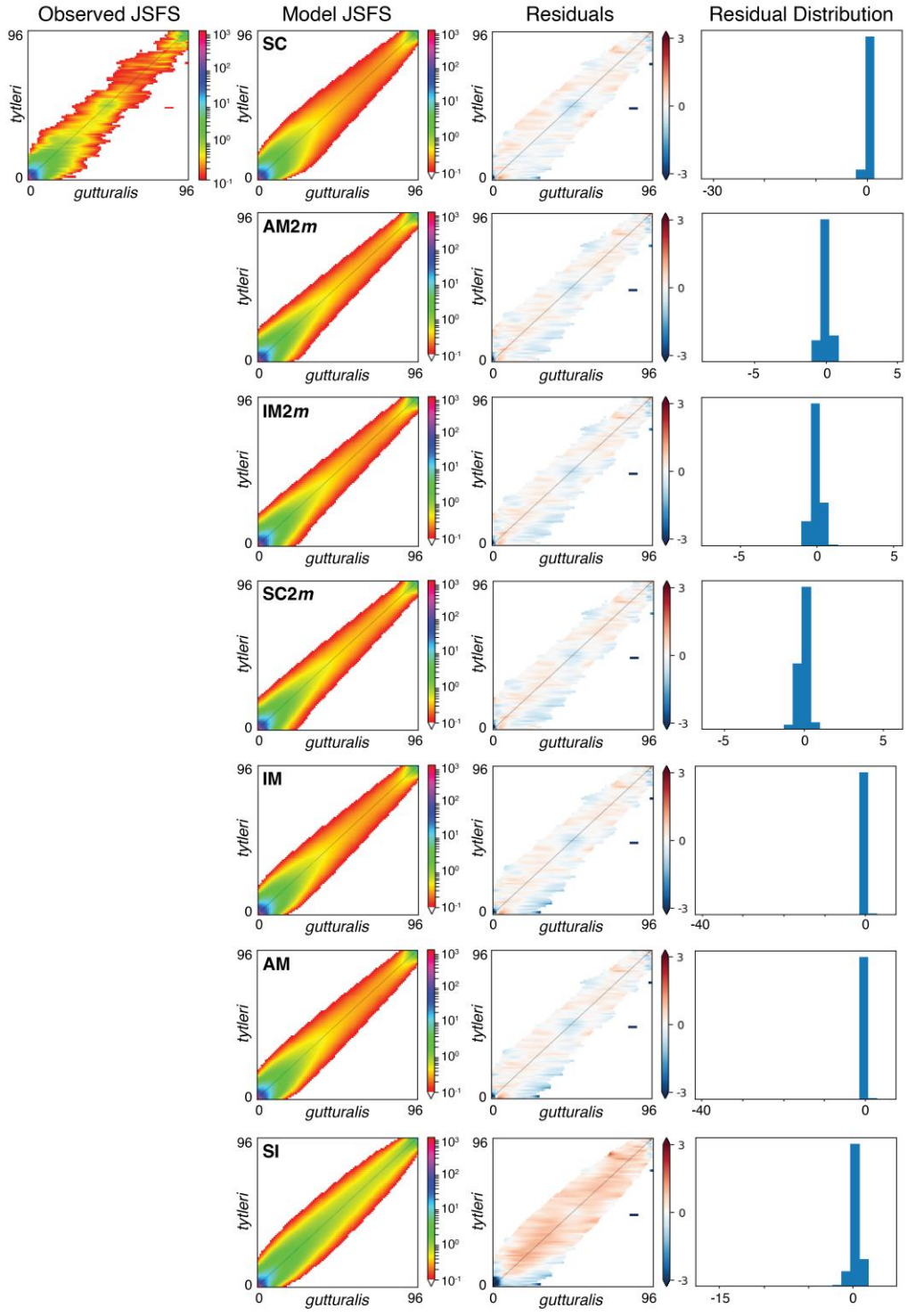

Figure S11. Comparisons of the observed unfolded joint site-frequency spectrum (JSFS) between parental *tytleri* and *gutturalis* and the JSFS simulated under demographic models. The JSFS from the best-fitting run for each model is shown in the left column. Residuals between the observed JSFS and the model are shown in the center column and the histogram with residual frequencies is shown in the right column. The best-fitting model is shown in the top row, with other models in descending order by -Δ AIC (see Data S2).

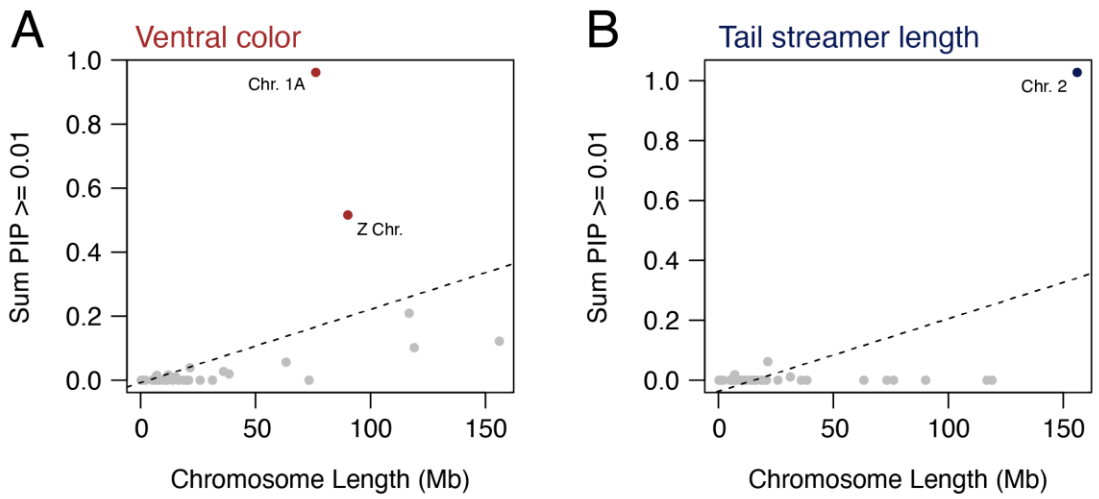

Figure S12. Concentration of mate choice trait genetic architectures on specific chromosomes. Comparison of chromosome length in megabases (Mb) and the sum of posterior inclusion probabilities (PIP) ≥ 0.01 per chromosome for (**A**) ventral color and (**B**) tail streamer length. Dashed lines represent the correlation between chromosome length and sum PIP among chromosomes. Points for chromosome 1A and the Z chromosome are shaded in dark red in (**A**) and the point for chromosome 2 is shaded in dark blue in (**B**).

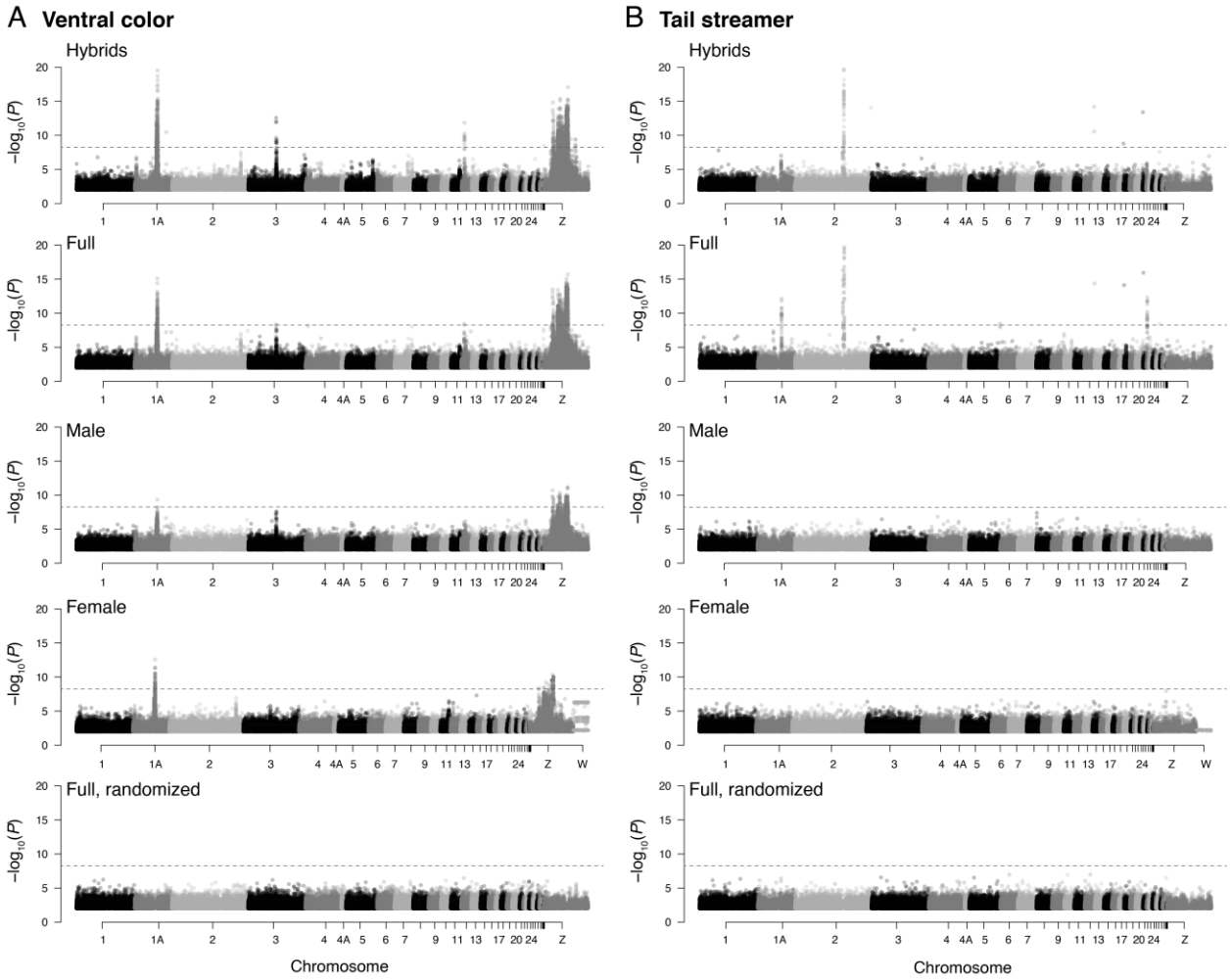

Figure S13. Genome-wide association scans with mate choice plumage traits. Association strength of SNPs from linear mixed models (LMMs) with (**A**) ventral plumage color (breast average brightness) and (**B**) tail streamer length. Results from analyses of different individuals are shown in descending panels for each trait; from top to bottom: hybrids, full (i.e., all individuals with matched genome and phenotype sampling), males, females, and the full dataset with randomized individual phenotypes. Dashed horizontal lines show the Bonferroni-corrected $-\log_{10}(P\text{-value})$ significance threshold in each analysis.

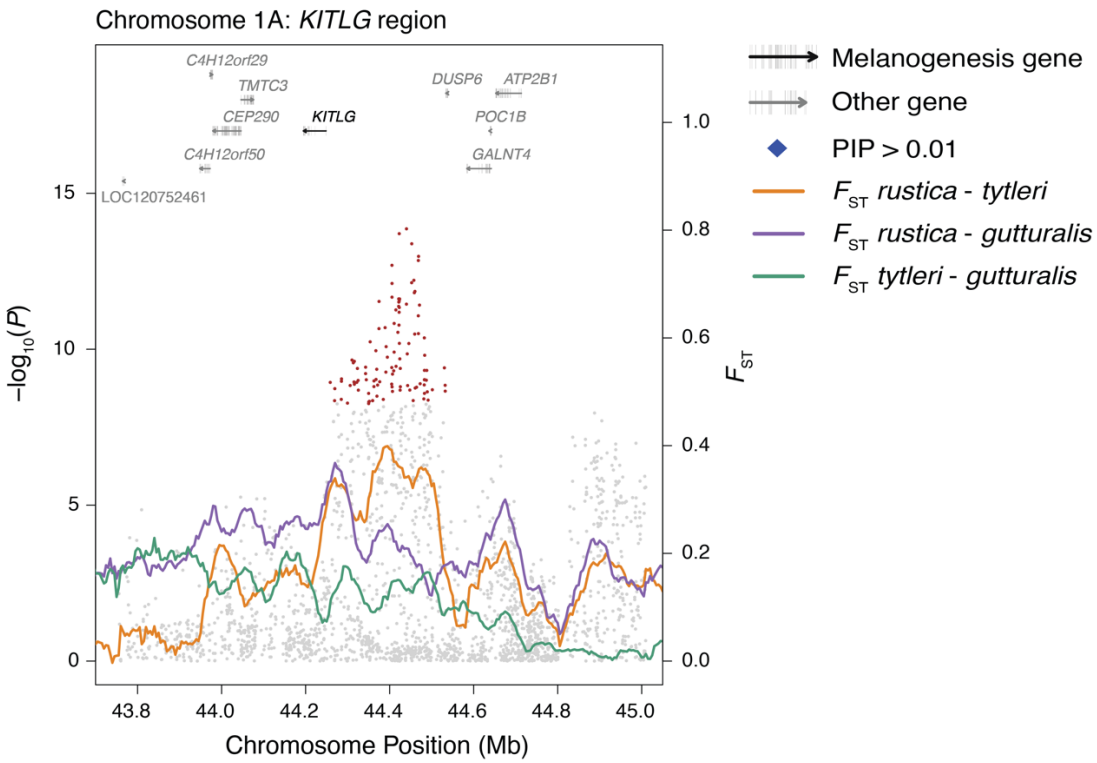

Figure S14. Detail of genome-wide associations with ventral color and genetic differentiation in the *KITLG* region of chromosome 1A. Individual SNP associations with the trait -$\log_{10}(P)$ (points; left y-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right y-axis). Significant associations after Bonferroni-correction are shaded in dark red and SNPs with PIP $\geq$ 0.01 are shown with blue diamonds. Genes are labeled, with locations and orientations shown as arrows with exon coordinates shown as vertical lines.

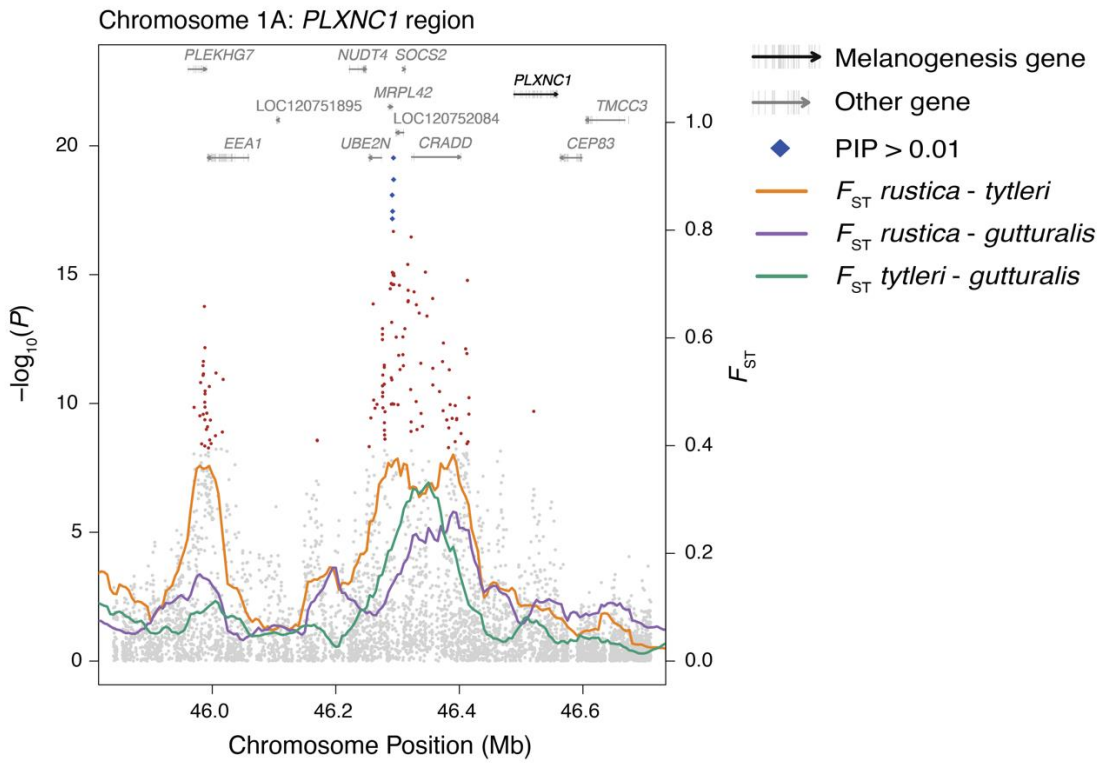

Figure S15. Detail of genome-wide associations with ventral color and genetic differentiation in the *PLXNC1* region of chromosome 1A. Individual SNP associations with the trait -$\log_{10}(P)$ (points; left y-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right y-axis). Significant associations after Bonferroni-correction are shaded in dark red and SNPs with PIP $\geq$ 0.01 are shown with blue diamonds. Genes are labeled, with locations and orientations shown as arrows with exon coordinates shown as vertical lines.

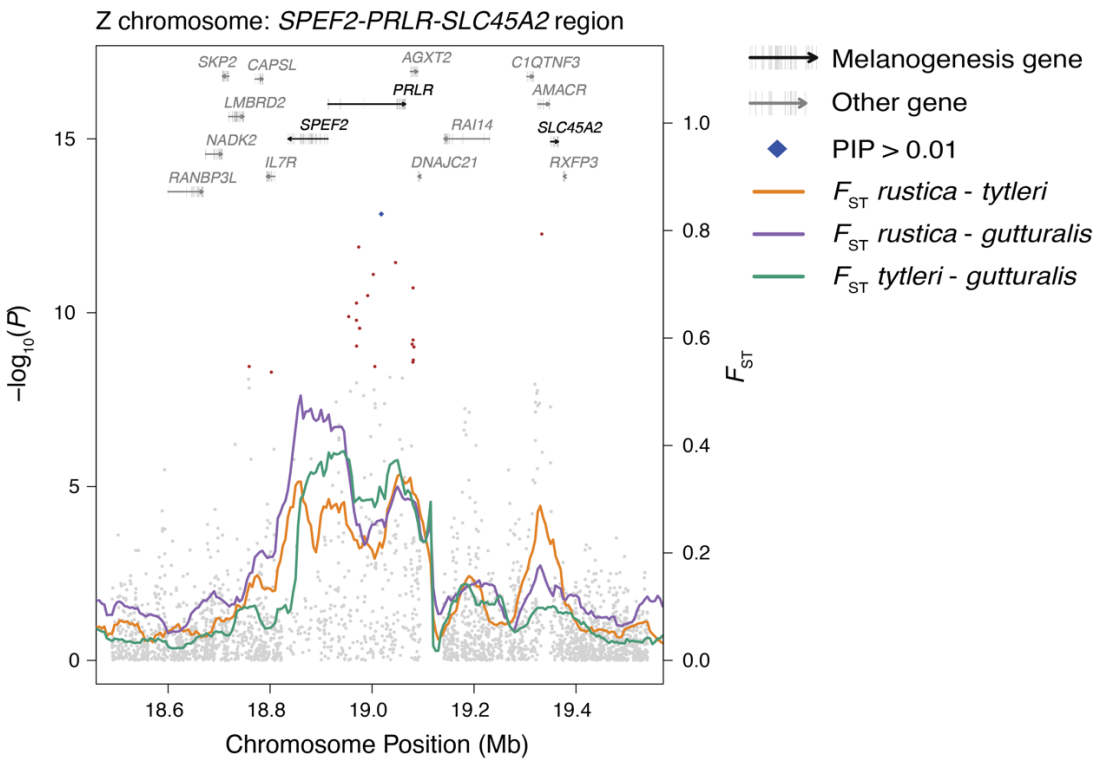

Figure S16. Detail of genome-wide associations with ventral color and genetic differentiation in the region of the Z chromosome containing *SPEF2*, *PRLR*, and *SLC45A2*. Individual SNP associations with the trait -$\log_{10}(P)$ (points; left y-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right y-axis). Significant associations after Bonferroni-correction are shaded in dark red and SNPs with PIP ≥ 0.01 are shown with blue diamonds. Genes are labeled, with locations and orientations shown as arrows with exon coordinates shown as vertical lines.

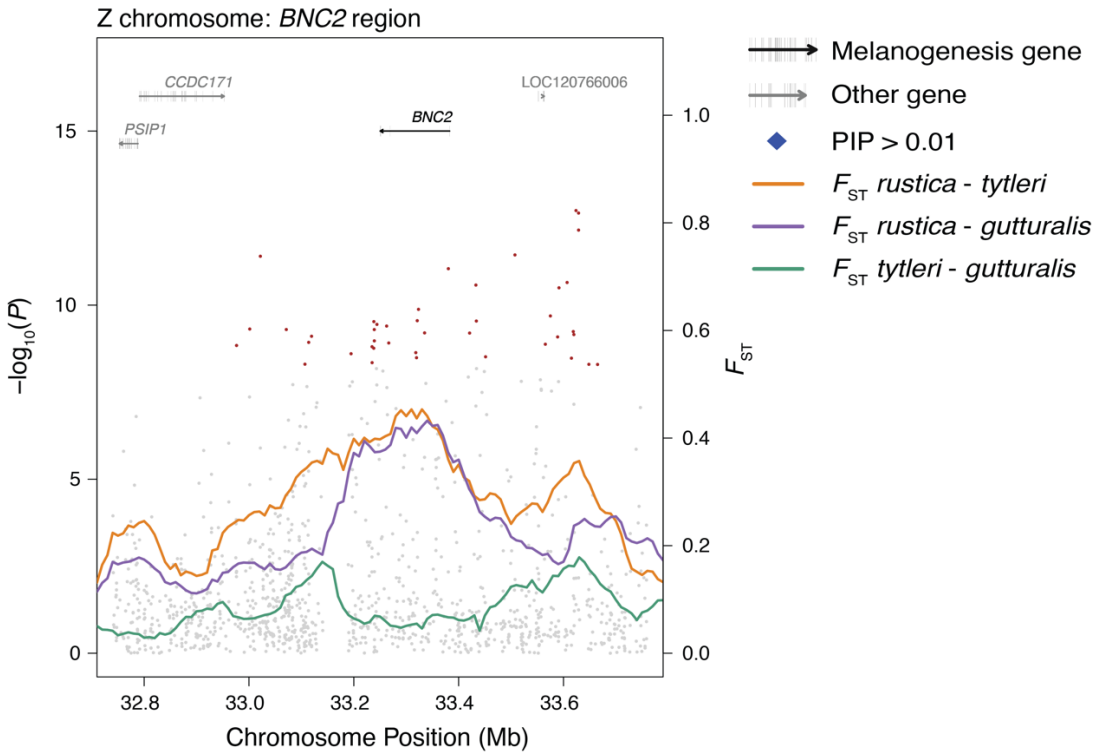

Figure S17. Detail of genome-wide associations with ventral color and genetic differentiation in the *BNC2* region of the Z chromosome. Individual SNP associations with the trait -$\log_{10}(P)$ (points; left y-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right y-axis). Significant associations after Bonferroni-correction are shaded in dark red and SNPs with PIP ≥ 0.01 are shown with blue diamonds. Genes are labeled, with locations and orientations shown as arrows with exon coordinates shown as vertical lines.

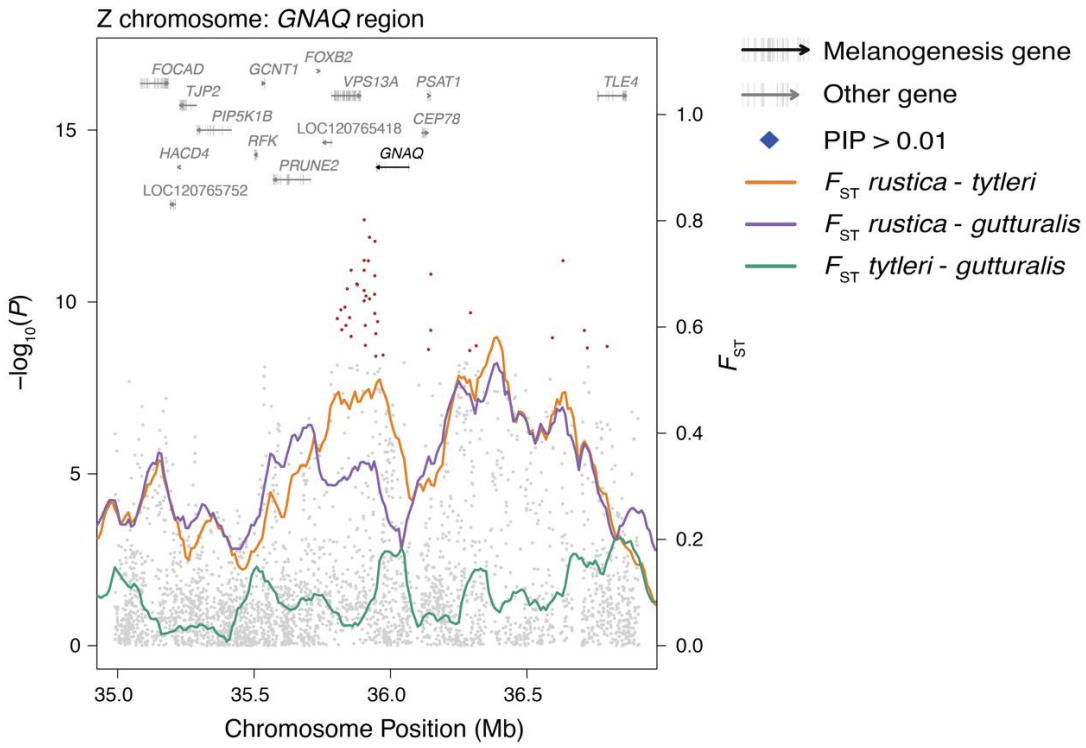

Figure S18. Detail of genome-wide associations with ventral color and genetic differentiation in the *GNAQ* region of the Z chromosome. Individual SNP associations with the trait -$\log_{10}(P)$ (points; left y-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right y-axis). Significant associations after Bonferroni-correction are shaded in dark red and SNPs with PIP $\geq$ 0.01 are shown with blue diamonds. Genes are labeled, with locations and orientations shown as arrows with exon coordinates shown as vertical lines.

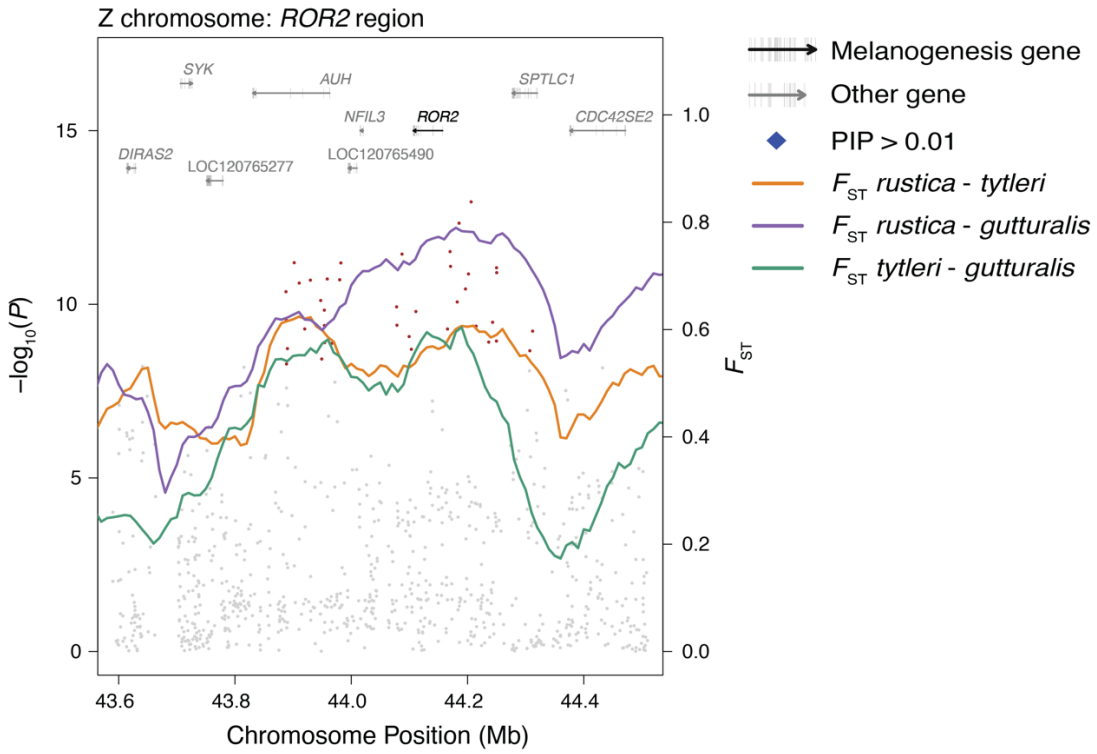

Figure S19. Detail of genome-wide associations with ventral color and genetic differentiation in the *ROR2* region of the Z chromosome. Individual SNP associations with the trait -$\log_{10}(P)$ (points; left y-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right y-axis). Significant associations after Bonferroni-correction are shaded in dark red and SNPs with PIP ≥ 0.01 are shown with blue diamonds. Genes are labeled, with locations and orientations shown as arrows with exon coordinates shown as vertical lines.

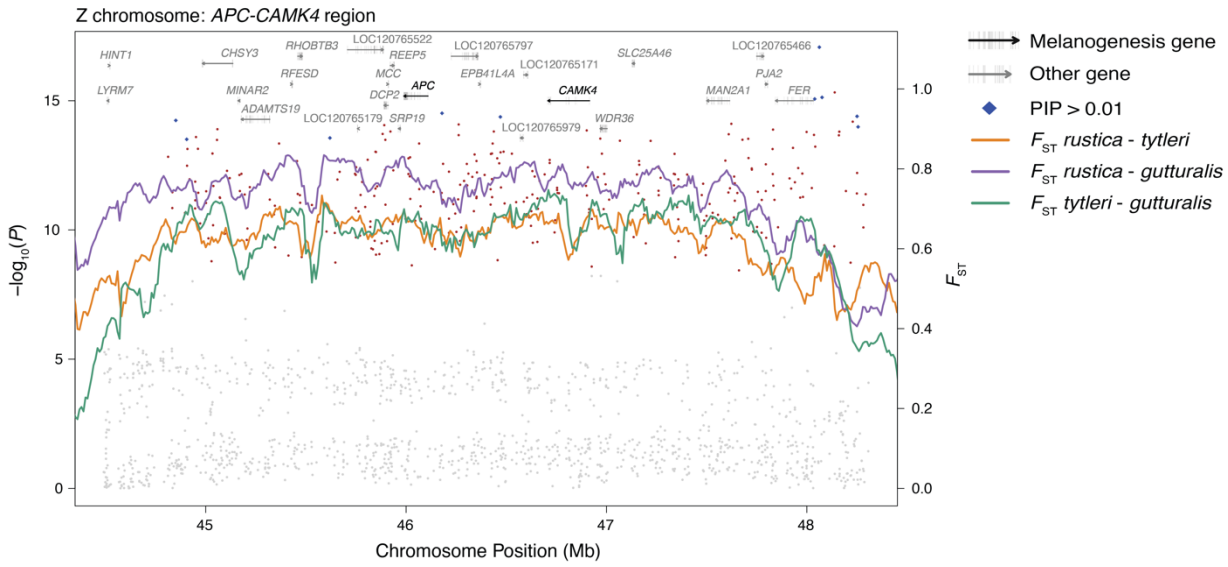

Figure S20. Detail of genome-wide associations with ventral color and genetic differentiation in the region of the Z chromosome containing *APC* and *CAMK4*. Individual SNP associations with the trait -$\log_{10}(P)$ (points; left y-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right y-axis). Significant associations after Bonferroni-correction are shaded in dark red and SNPs with PIP $\geq$ 0.01 are shown with blue diamonds. Genes are labeled, with locations and orientations shown as arrows with exon coordinates shown as vertical lines.

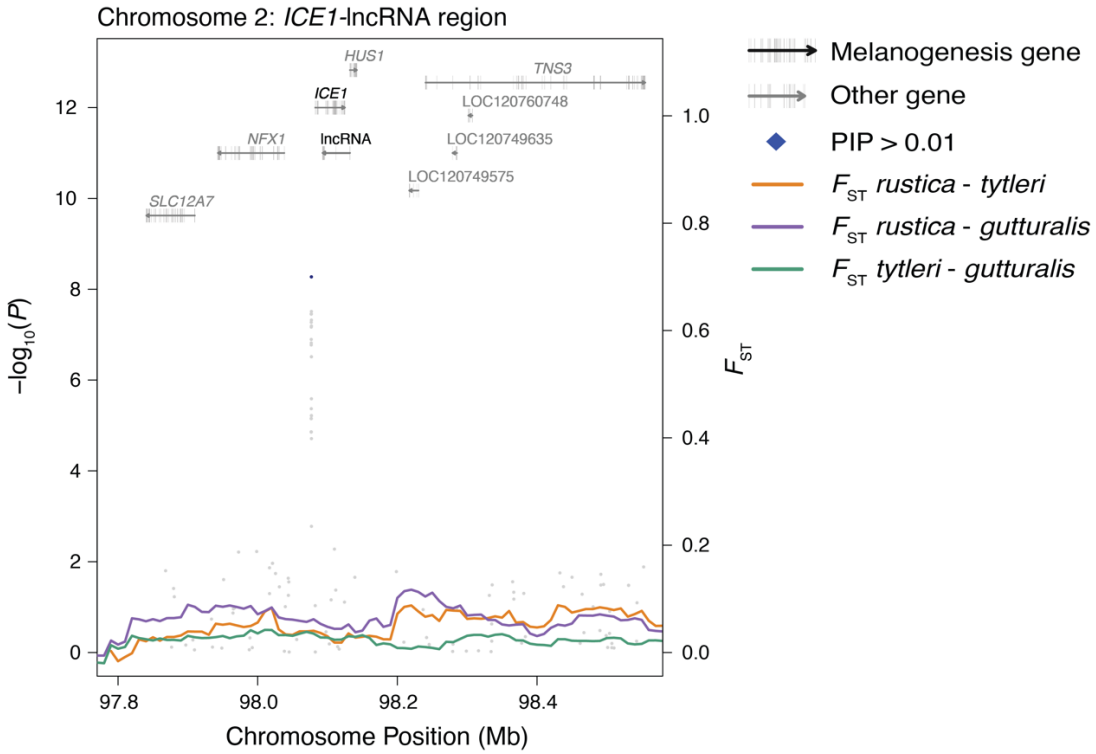

Figure S21. Detail of genome-wide associations with tail streamer length and genetic differentiation in the *ICE1*-lncRNA region of chromosome 2. Individual SNP associations with the trait -$\log_{10}(P)$ (points; left y-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right y-axis). Significant associations after Bonferroni-correction are shaded in dark blue and SNPs with PIP $\geq$ 0.01 are shown with blue diamonds. Genes are labeled, with locations and orientations shown as arrows with exon coordinates shown as vertical lines.

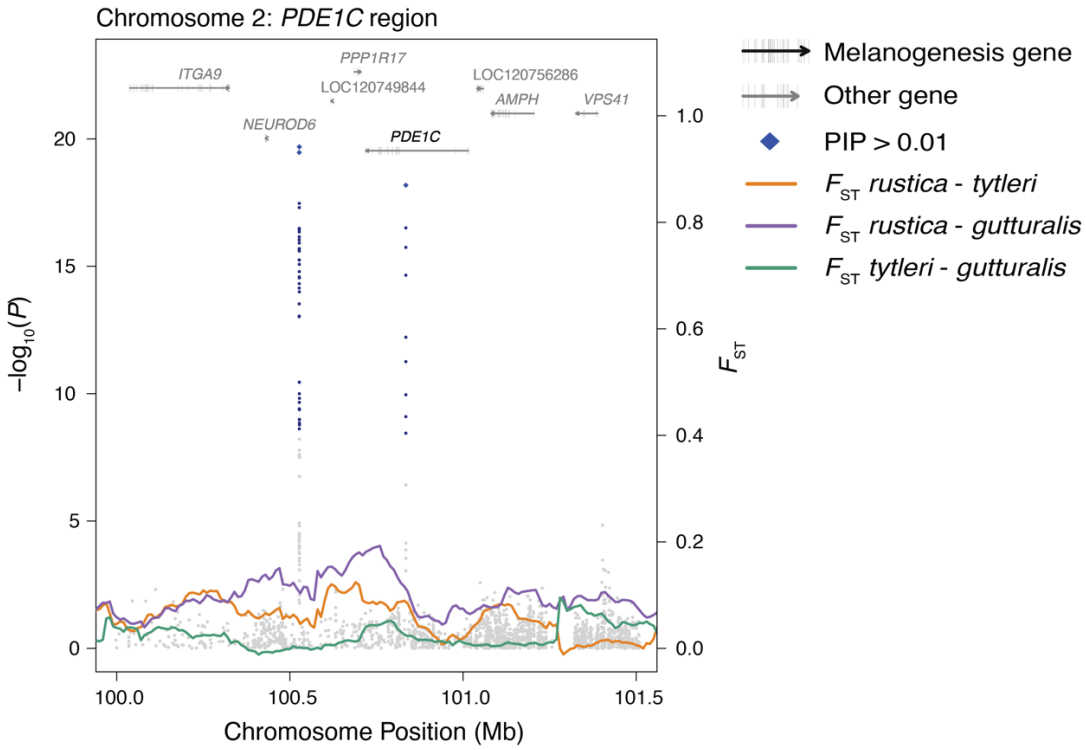

Figure S22. Detail of genome-wide associations with tail streamer length and genetic differentiation in the *PDE1C* region of chromosome 2. Individual SNP associations with the trait -$\log_{10}(P)$ (points; left y-axis) and genetic differentiation (F_{ST}) between parental populations (lines; right y-axis). Significant associations after Bonferroni-correction are shaded in dark blue and SNPs with PIP $\geq$ 0.01 are shown with blue diamonds. Genes are labeled, with locations and orientations shown as arrows with exon coordinates shown as vertical lines.

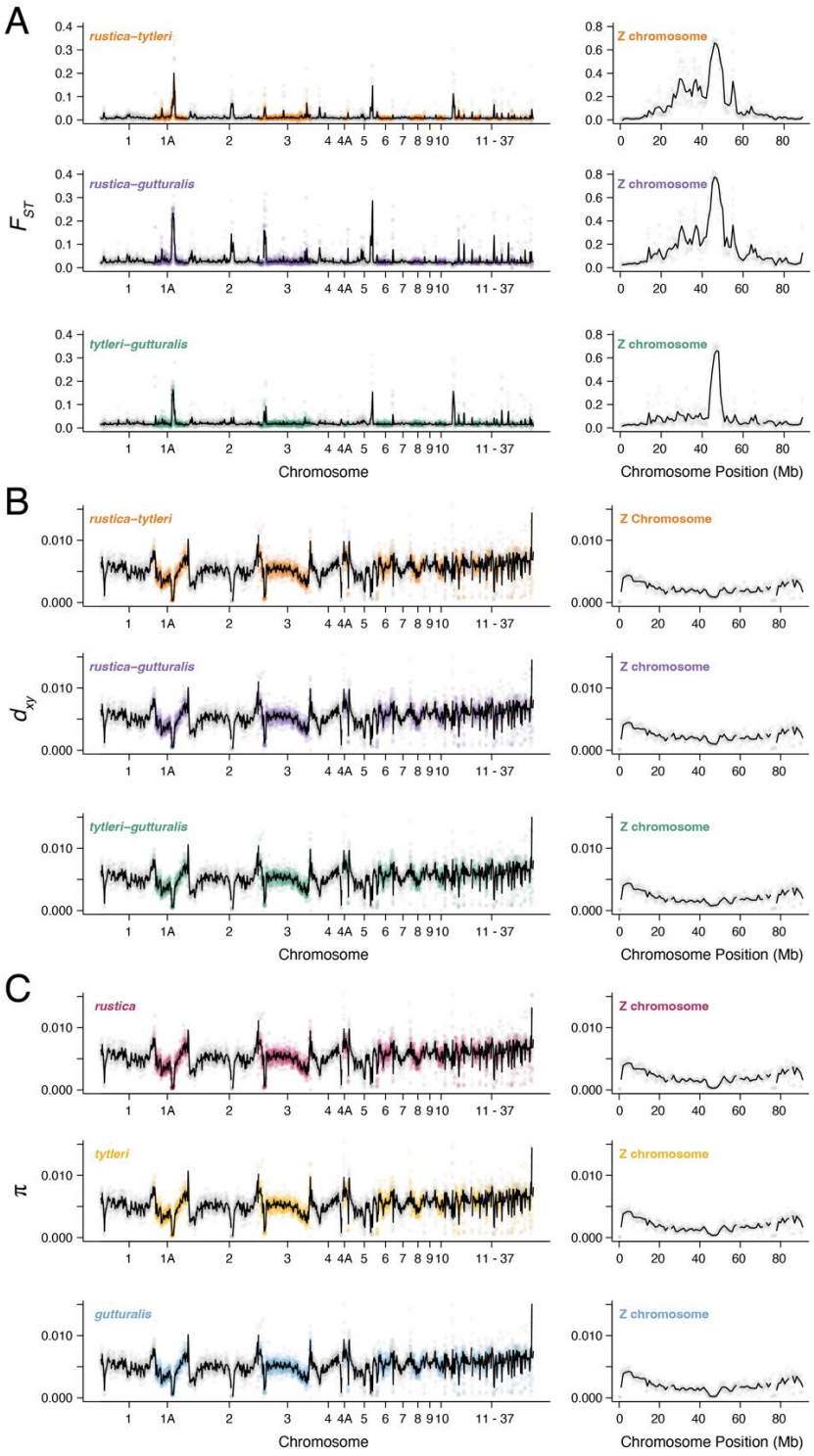

Figure S23. Genome-wide variation in population genetic divergence and diversity. Genome scans in 1 Mb (lines) and 100 kb (points) sliding windows of (**A**) F_{ST} and (**B**) d_{xy} between, and (**C**) π within, parental populations across autosomes (left panels) and the Z chromosome (right panels).

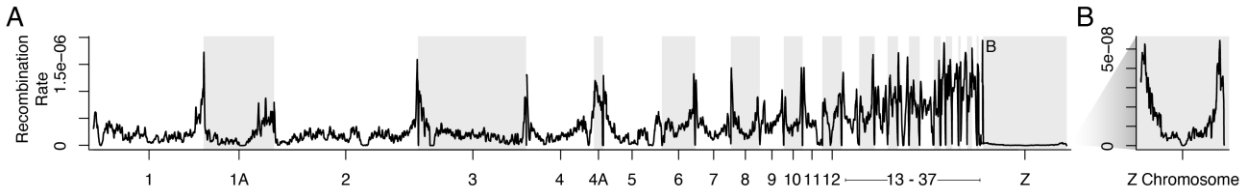

Figure S24. Recombination rate variation across the barn swallow genome. (**A**) Recombination rate was measured in 1 Mb sliding windows with a 100 kb step size across chromosomes. (**B**) Recombination rates on the Z chromosome (note the difference in y-axis scale).

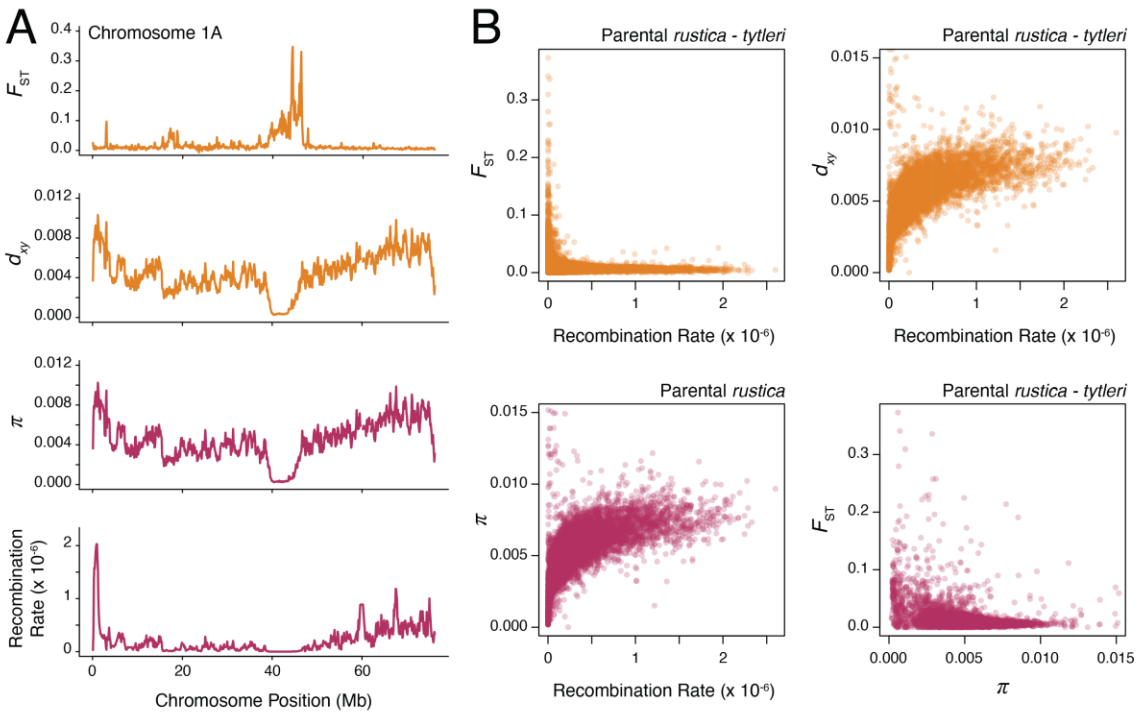

Figure S25. The genomic landscape of population differentiation and signatures of linked selection. (**A**) Variation in F_{ST}, d_{xy}, π, and recombination rate across chromosome 1A as an example of the heterogenous landscape of differentiation between parental barn swallow populations. (**B**) Genome-wide correlations between recombination rate and F_{ST}, d_{xy}, π, and between F_{ST} and π. Relationships between parameters were broadly consistent across populations (Table S10).

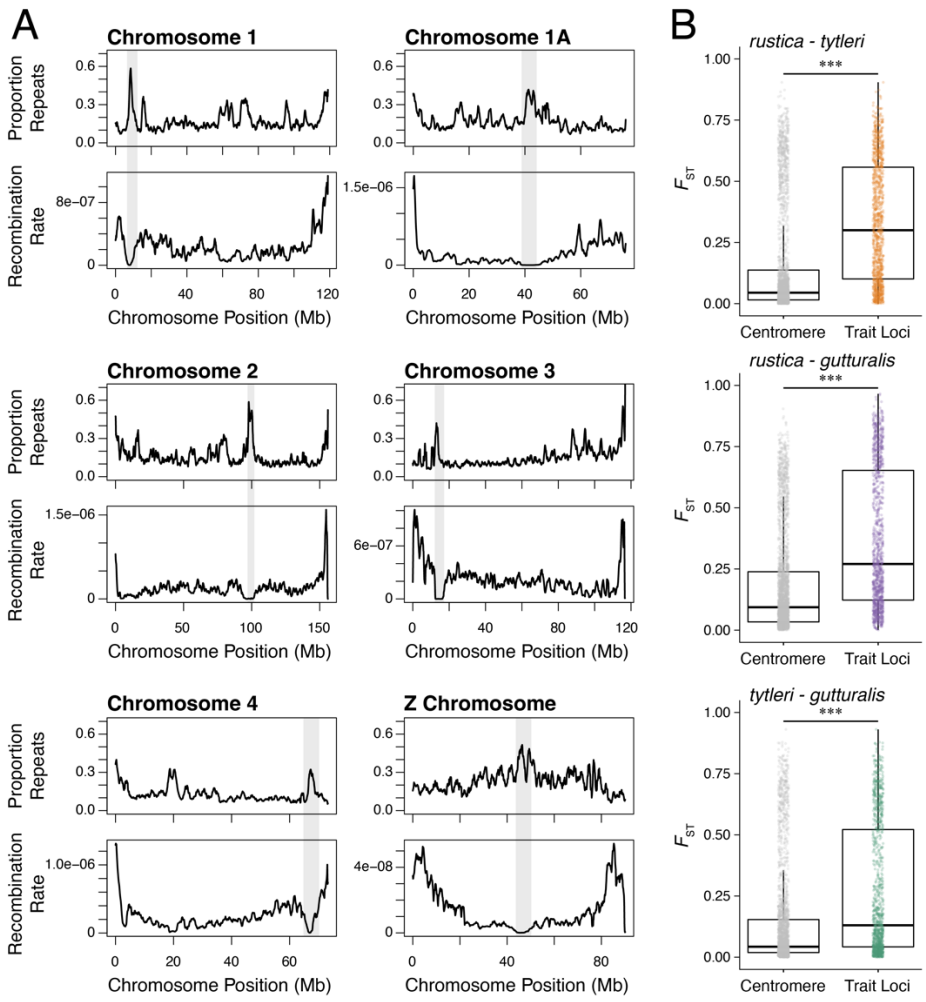

Figure S26. Putative centromere regions and comparison of population differentiation in centromeres and sexual trait loci. (**A**) Scans of repeat content, measured as the proportion of bases in sliding windows annotated as repeat elements (top panels), and recombination rate (bottoms panels) in 1 Mb sliding windows with a 100 kb step size. Putative centromere regions are shaded in grey. (**B**) Distributions of F_{ST} between parental populations in putative centromeres (grey) and sexual trait loci (colors) measured from non-overlapping 10 kb windows; asterisks summarize statistical comparisons between distributions ($^{***}P < 2.2 \times 10^{-16}$).

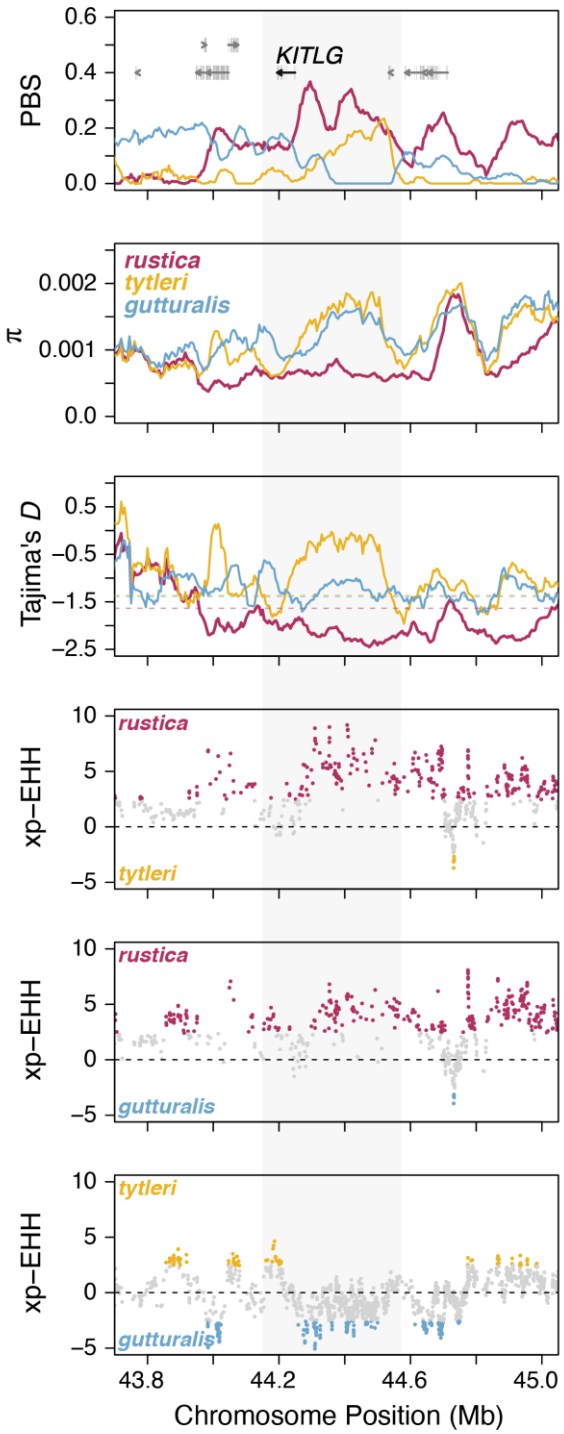

Figure S27. Detail of statistics used to detect divergent selection in the *KITLG* region of chromosome 1A. Scans of population branch statistics (PBS), π, Tajima's *D*, and cross-population effective haplotype homozygosity (xp-EHH). The locations of *KITLG* and other genes are shown, and the region associated with ventral color is shaded in grey. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Dashed horizontal lines in the Tajima's *D* panel are population means for the chromosome. In xp-EHH panels, colored points indicate SNPs that exceeded the genome-wide 99th quantile.

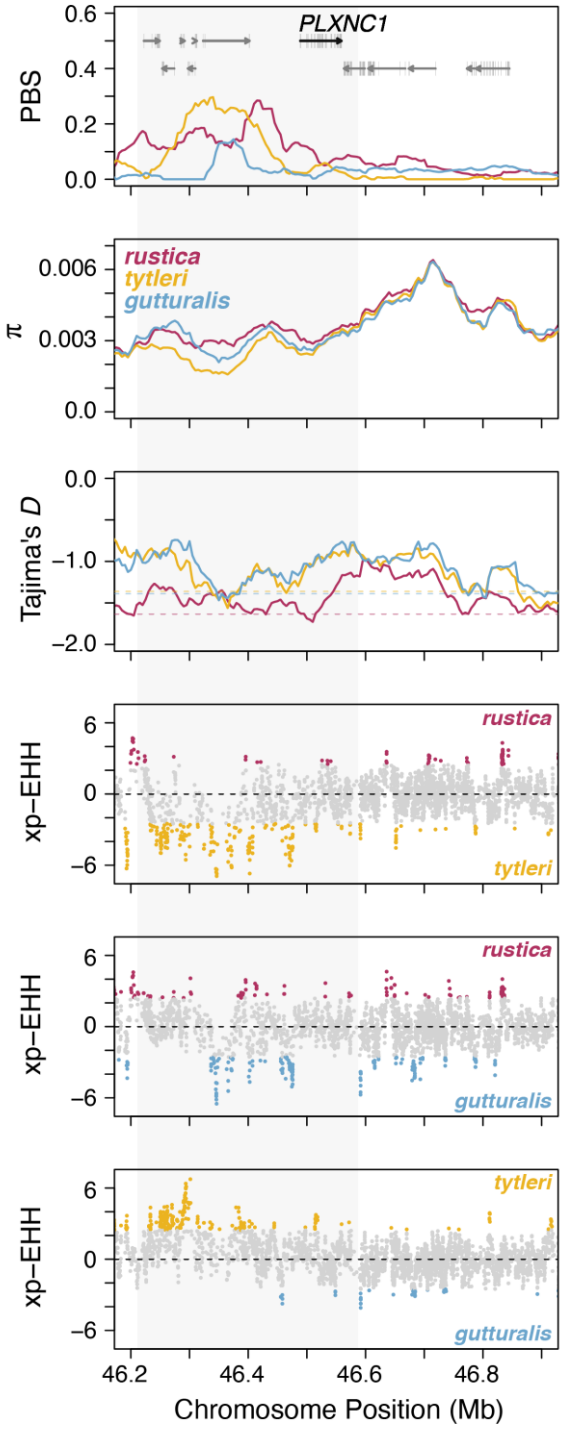

Figure S28. Detail of statistics used to detect divergent selection in the *PLXNC1* region of chromosome 1A. Scans of population branch statistics (PBS), π, Tajima's *D*, and cross-population effective haplotype homozygosity (xp-EHH). The locations of *PLXNC1* and other genes are shown, and the region associated with ventral color is shaded in grey. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Dashed horizontal lines in the Tajima's *D* panel are population means for the chromosome. In xp-EHH panels, colored points indicate SNPs that exceeded the genome-wide 99th quantile.

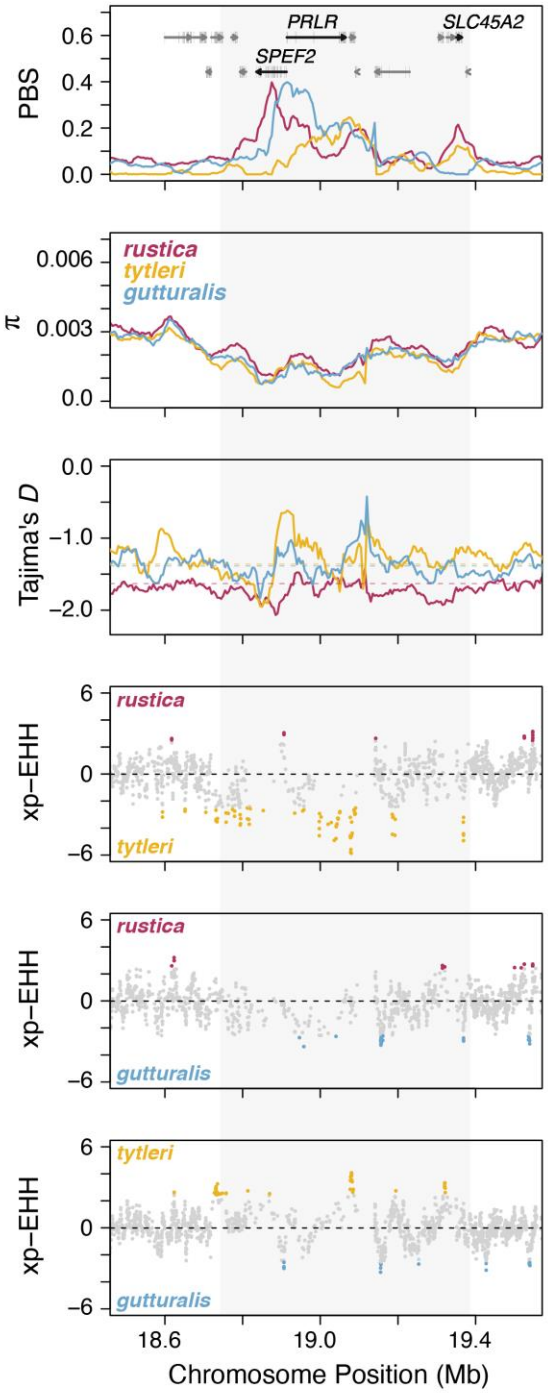

Figure S29. Detail of statistics used to detect divergent selection in the *SPEF2*, *PRLR*, and *SLC45A2* region of the Z chromosome. Scans of population branch statistics (PBS), π, Tajima's *D*, and cross-population effective haplotype homozygosity (xp-EHH). The locations of *SPEF2*, *PRLR*, *SLC45A2* and other genes are shown, and the region associated with ventral color is shaded in grey. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Dashed horizontal lines in the Tajima's *D* panel are population means for the chromosome. In xp-EHH panels, colored points indicate SNPs that exceeded the genome-wide 99th quantile.

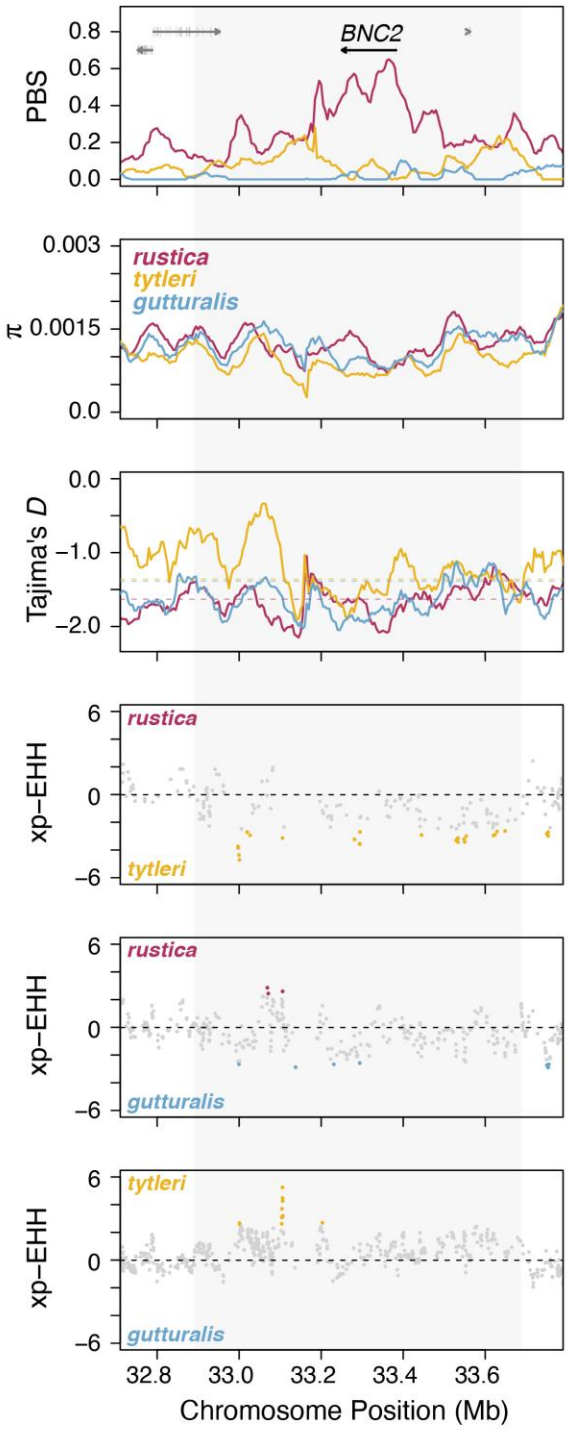

Figure S30. Detail of statistics used to detect divergent selection in the *BNC2* region of the Z chromosome. Scans of population branch statistics (PBS), π, Tajima's *D*, and cross-population effective haplotype homozygosity (xp-EHH). The locations of *BNC2* and other genes are shown, and the region associated with ventral color is shaded in grey. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Dashed horizontal lines in the Tajima's *D* panel are population means for the chromosome. In xp-EHH panels, colored points indicate SNPs that exceeded the genome-wide 99th quantile.

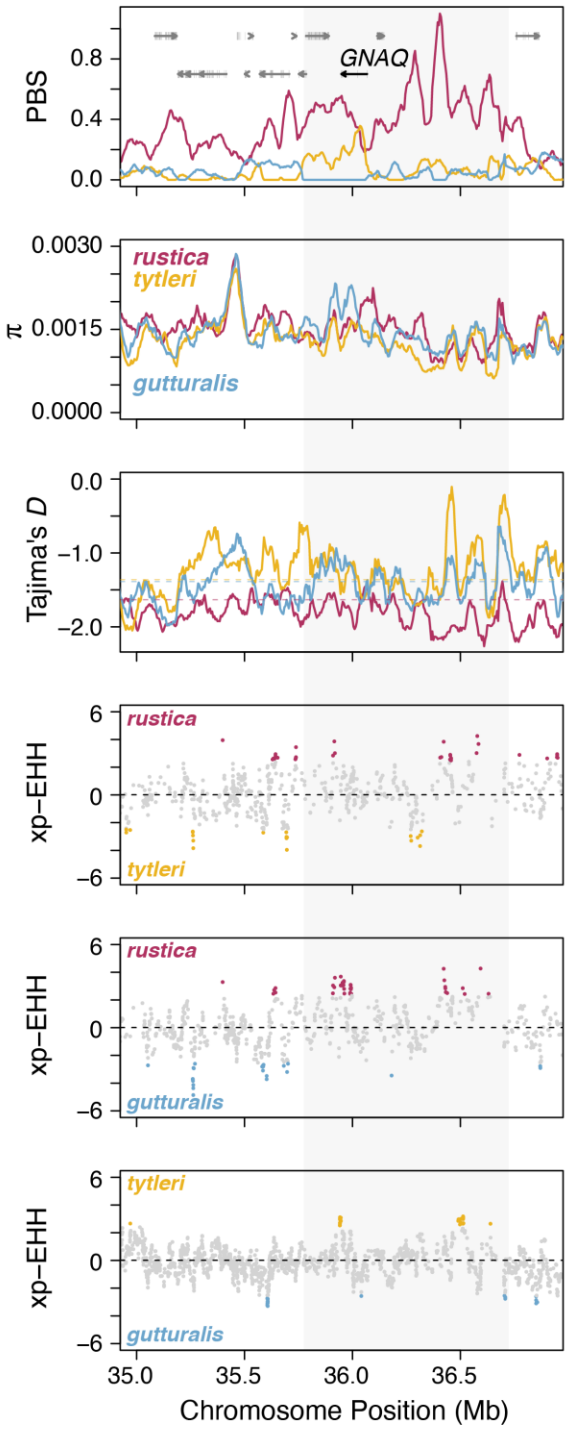

Figure S31. Detail of statistics used to detect divergent selection in the *GNAQ* region of the Z chromosome. Scans of population branch statistics (PBS), π, Tajima's *D*, and cross-population effective haplotype homozygosity (xp-EHH). The locations of *GNAQ* and other genes are shown, and the region associated with ventral color is shaded in grey. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Dashed horizontal lines in the Tajima's *D* panel are population means for the chromosome. In xp-EHH panels, colored points indicate SNPs that exceeded the genome-wide 99th quantile.

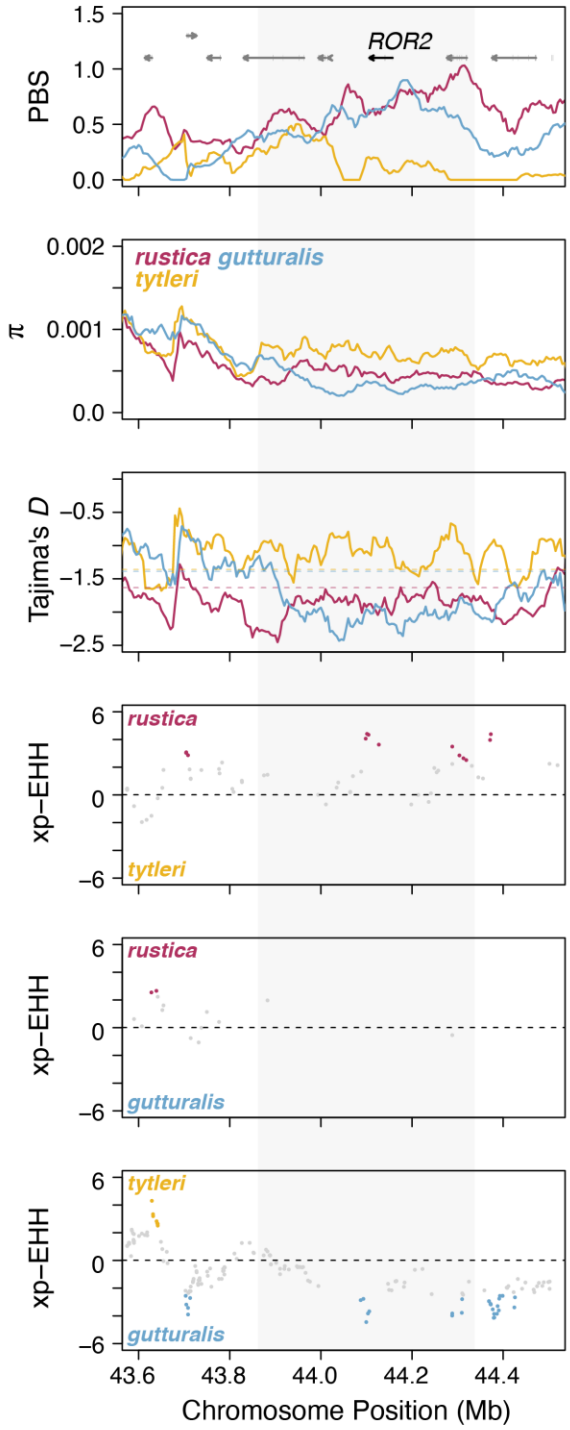

Figure S32. Detail of statistics used to detect divergent selection in the *ROR2* region of the Z chromosome. Scans of population branch statistics (PBS), π, Tajima's *D*, and cross-population effective haplotype homozygosity (xp-EHH). The locations of *ROR2* and other genes are shown, and the region associated with ventral color is shaded in grey. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Dashed horizontal lines in the Tajima's *D* panel are population means for the chromosome. In xp-EHH panels, colored points indicate SNPs that exceeded the genome-wide 99th quantile.

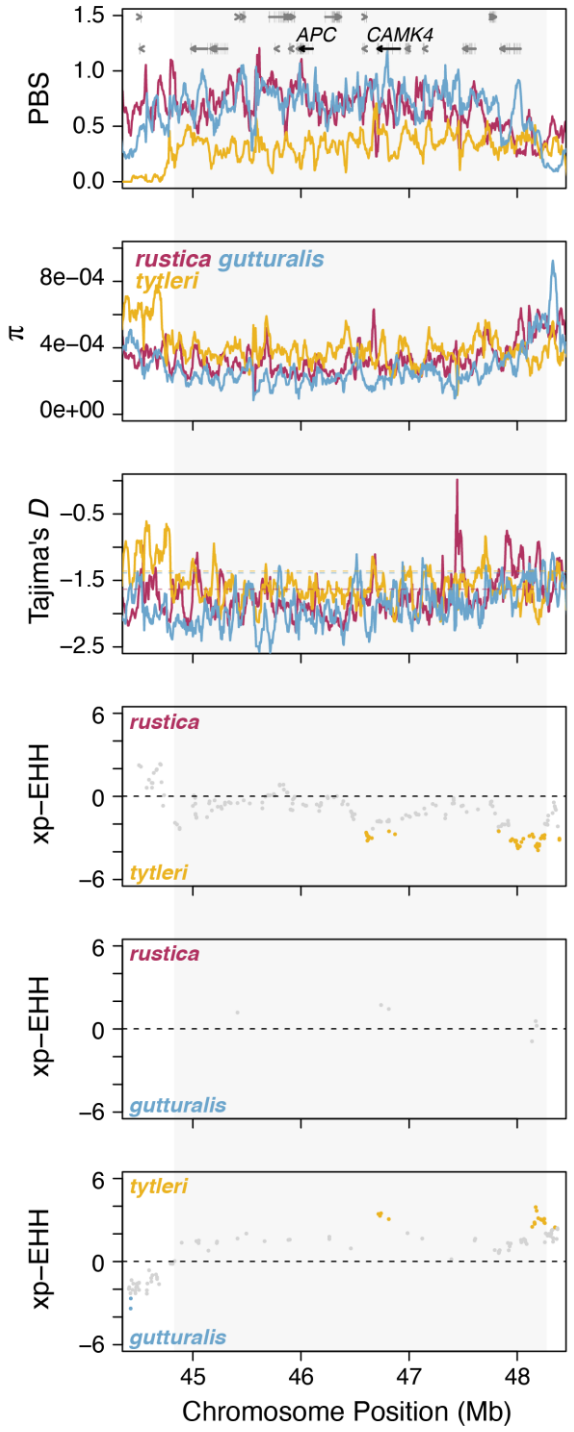

Figure S33. Detail of statistics used to detect divergent selection in the *APC* and *CAMK4* region of the Z chromosome. Scans of population branch statistics (PBS), π, Tajima's *D*, and cross-population effective haplotype homozygosity (xp-EHH). The locations of *APC*, *CAMK4*, and other genes are shown, and the region associated with ventral color is shaded in grey. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Dashed horizontal lines in the Tajima's *D* panel are population means for the chromosome. In xp-EHH panels, colored points indicate SNPs that exceeded the genome-wide 99th quantile.

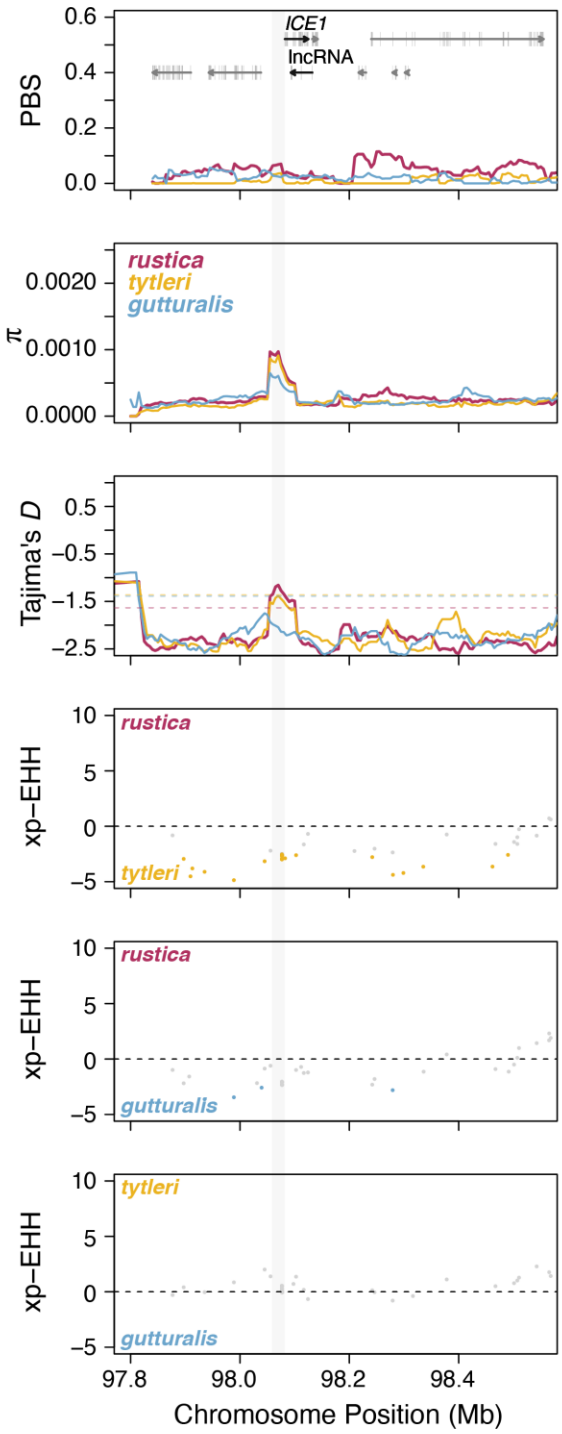

Figure S34. Detail of statistics used to detect divergent selection in the *ICE1*-lncRNA region of chromosome 2. Scans of population branch statistics (PBS), π, Tajima's *D*, and cross-population effective haplotype homozygosity (xp-EHH). The locations of *ICE1*, the lncRNA, and other genes are shown, and the region associated with tail streamer length is shaded in grey. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Dashed horizontal lines in the Tajima's *D* panel are population means for the chromosome. In xp-EHH panels, colored points indicate SNPs that exceeded the genome-wide 99th quantile.

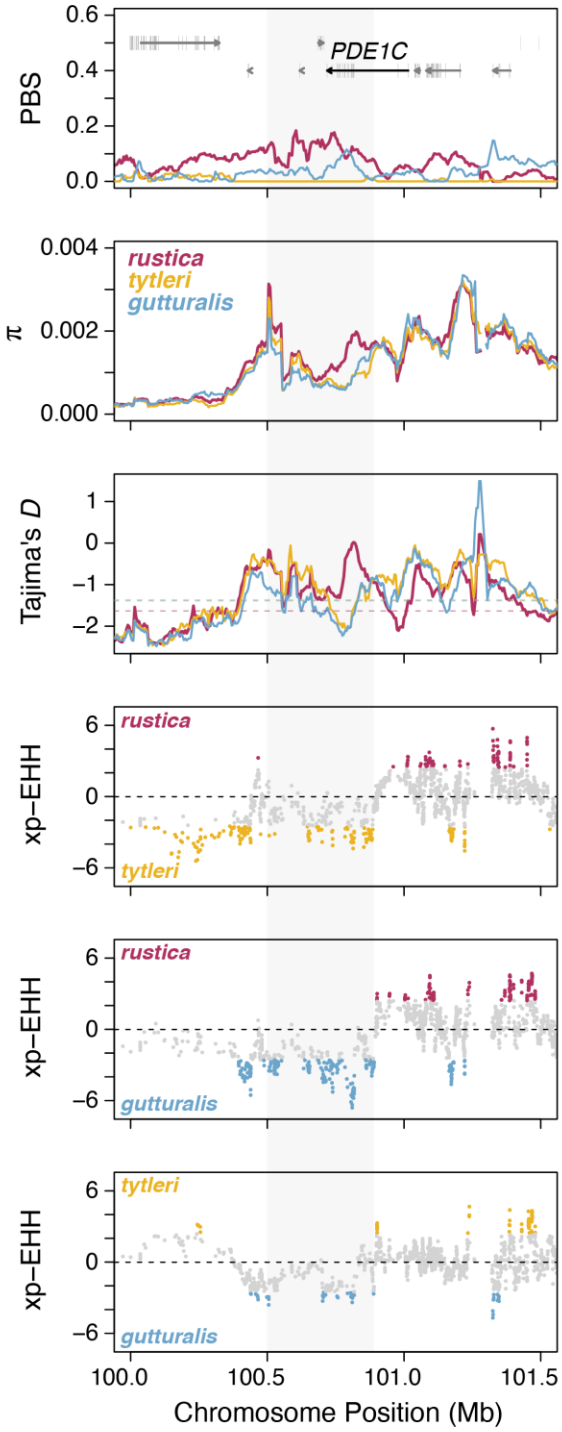

Figure S35. Detail of statistics used to detect divergent selection in the *PDE1C* region of chromosome 2. Scans of population branch statistics (PBS), π, Tajima's *D*, and cross-population effective haplotype homozygosity (xp-EHH). The locations of *PDE1C* and other genes are shown, and the region associated with tail streamer length is shaded in grey. Scans in *rustica*, *tytleri*, and *gutturalis* are shown as red, yellow, and blue lines, respectively. Dashed horizontal lines in the Tajima's *D* panel are population means for the chromosome. In xp-EHH panels, colored points indicate SNPs that exceeded the genome-wide 99th quantile.

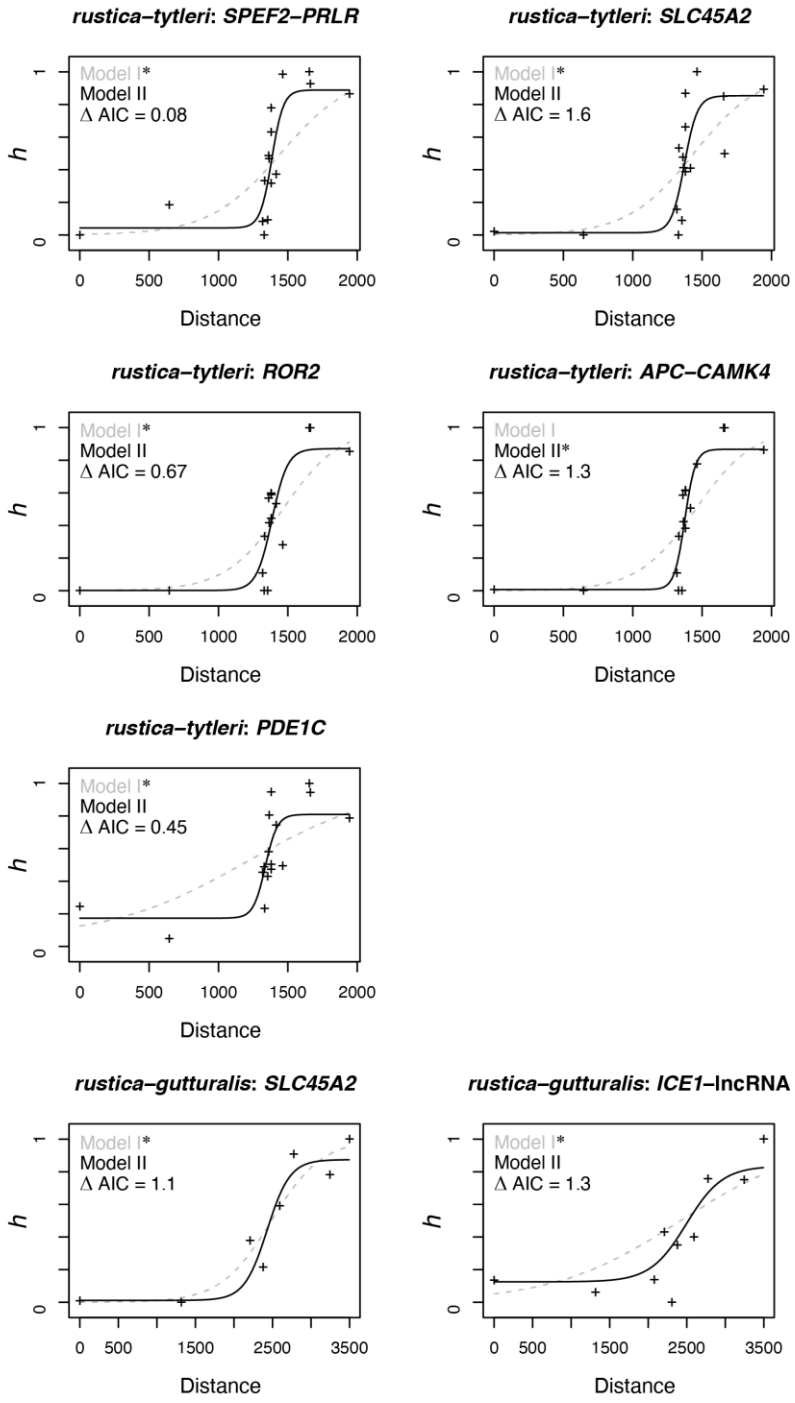

Figure S36. Geographic clines for loci with similar support for two models in the *rustica-tytleri* and *rustica-gutturalis* hybrid zones. Clines of hybrid index (*h*) with similar AIC support ($\Delta AIC \leq 2$) for models I and II, used to assess visual fit to the data. * Denotes the model with lower AIC in each panel. Cline fits under models I and II are shown as grey dashed lines and black lines, respectively. Black cross marks represent locality mean *h* values in each transect.

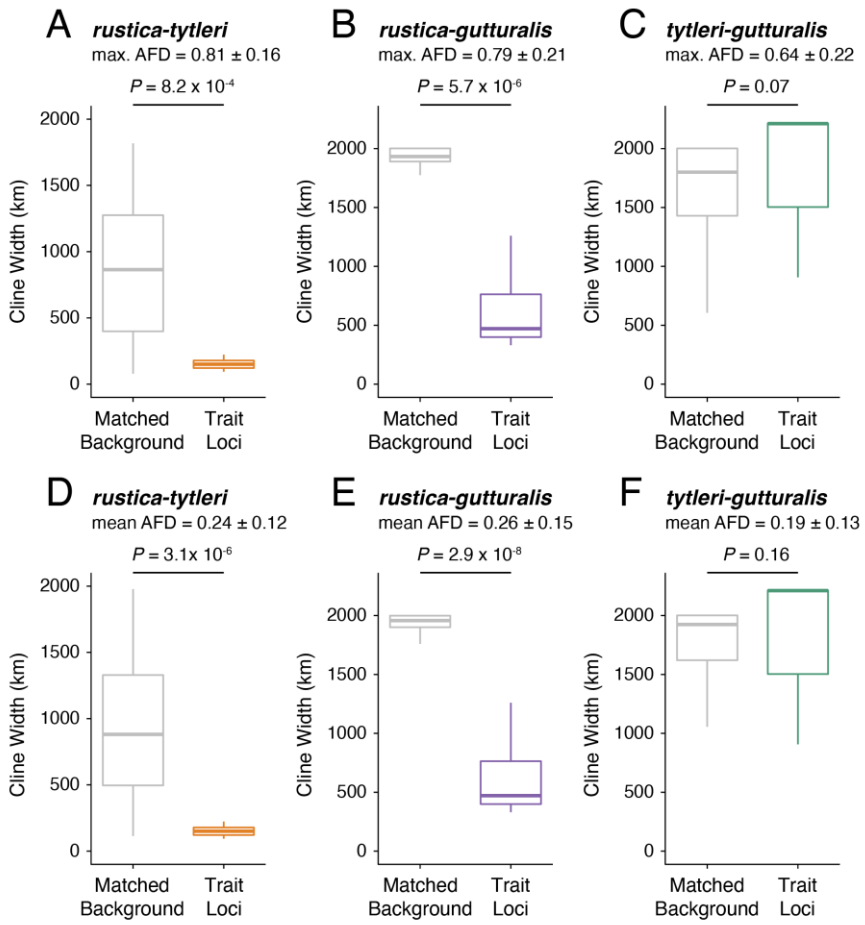

Figure S37. Comparison of geographic cline widths for trait loci and background loci with similar levels of differentiation between parental populations. Distributions of cline width (*w*) in kilometers for trait loci (colors) and background loci (grey) with matching allele frequency differences (AFD) between parental populations (i.e., within one standard deviation of mean AFD for trait loci; see text under population labels for details on AFD distributions). Results are shown for each hybrid zone based on the maximum AFD per locus (**A**-**C**) and the mean AFD per locus (**D**-**F**).

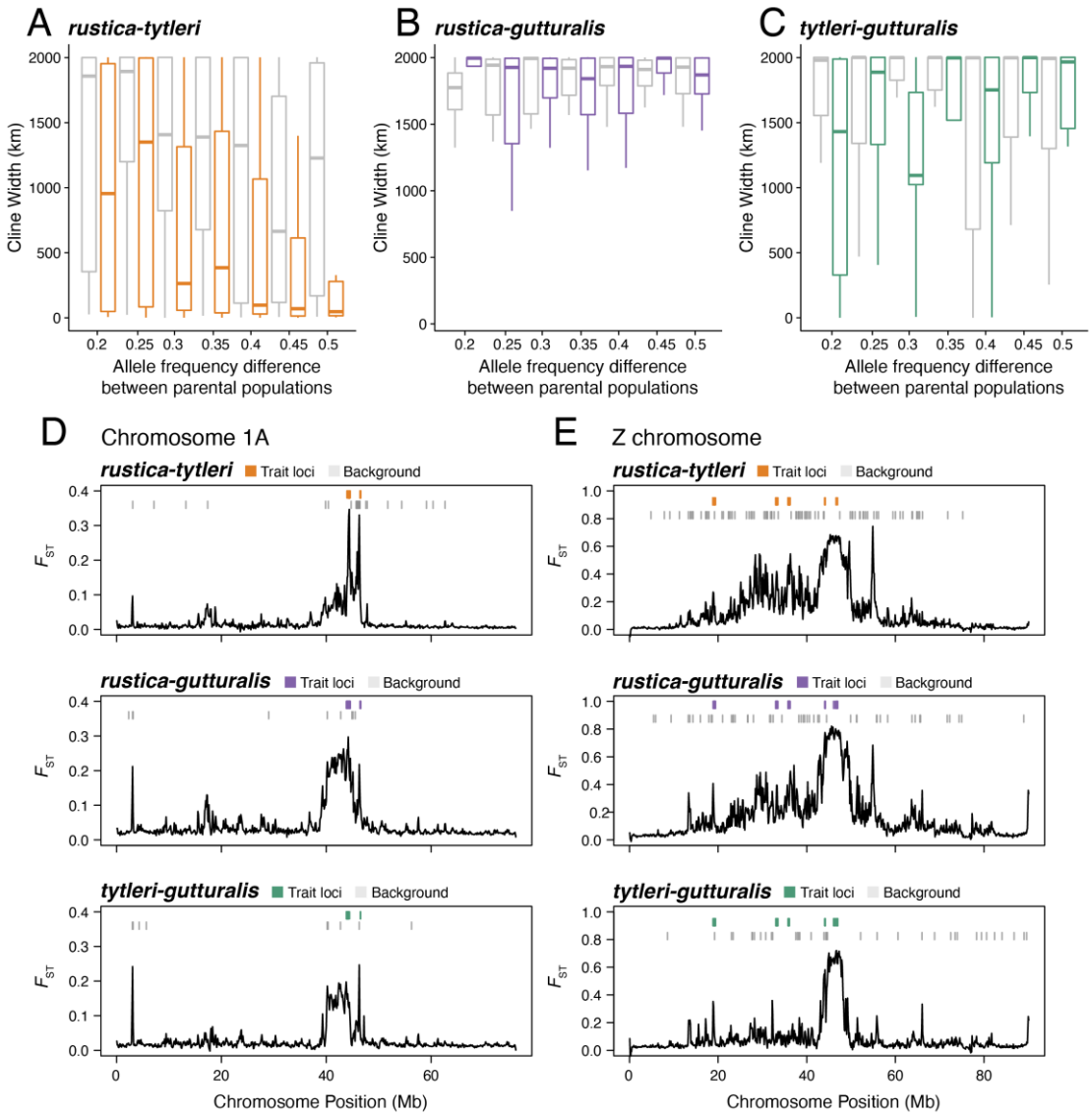

Figure S38. Comparison of geographic cline widths for SNPs with similar levels of differentiation in trait loci versus the genome background. Distributions of cline width (*w*) in kilometers for SNPs in trait loci (colors) and the genome background (grey) in bins of increasing matched allele frequency differences ($0.2 < \text{AFD} \leq 0.5$) between parental populations for the (**A**) *rustica-tytleri*, (**B**) *rustica-gutturalis*, and (**C**) *tytleri-gutturalis* hybrid zone transects. Genomic positions of trait (colors) and background SNPs (grey) are shown for (**D**) chromosome 1A and (**E**) the Z chromosome. Note that these results differ from those shown in Fig. S37; here geographic clines were inferred from individual SNPs to enable comparisons in bins of increasing matched AFD, rather than from entire loci (as in Fig. S37).

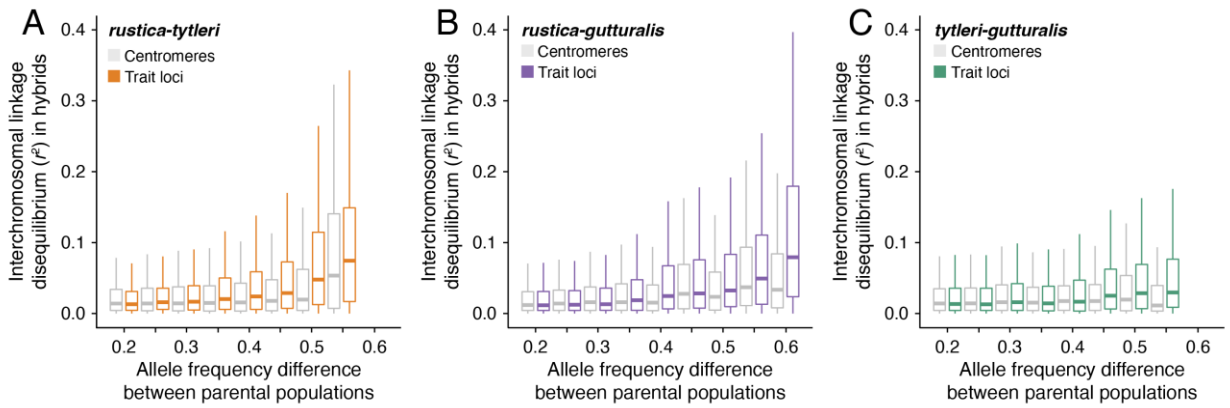

Figure S39. Inter-chromosomal linkage disequilibrium between trait loci and centromeres in hybrid zones. Inter-chromosomal linkage disequilibrium (LD; r^2) measured between SNPs in trait loci (colors) and putative centromeres (grey) with increasing allele frequency differences between parental populations (x-axis) for the (**A**) *rustica-tytleri*, (**B**) *rustica-gutturalis*, and (**C**) *tytleri-gutturalis* hybrid zones.

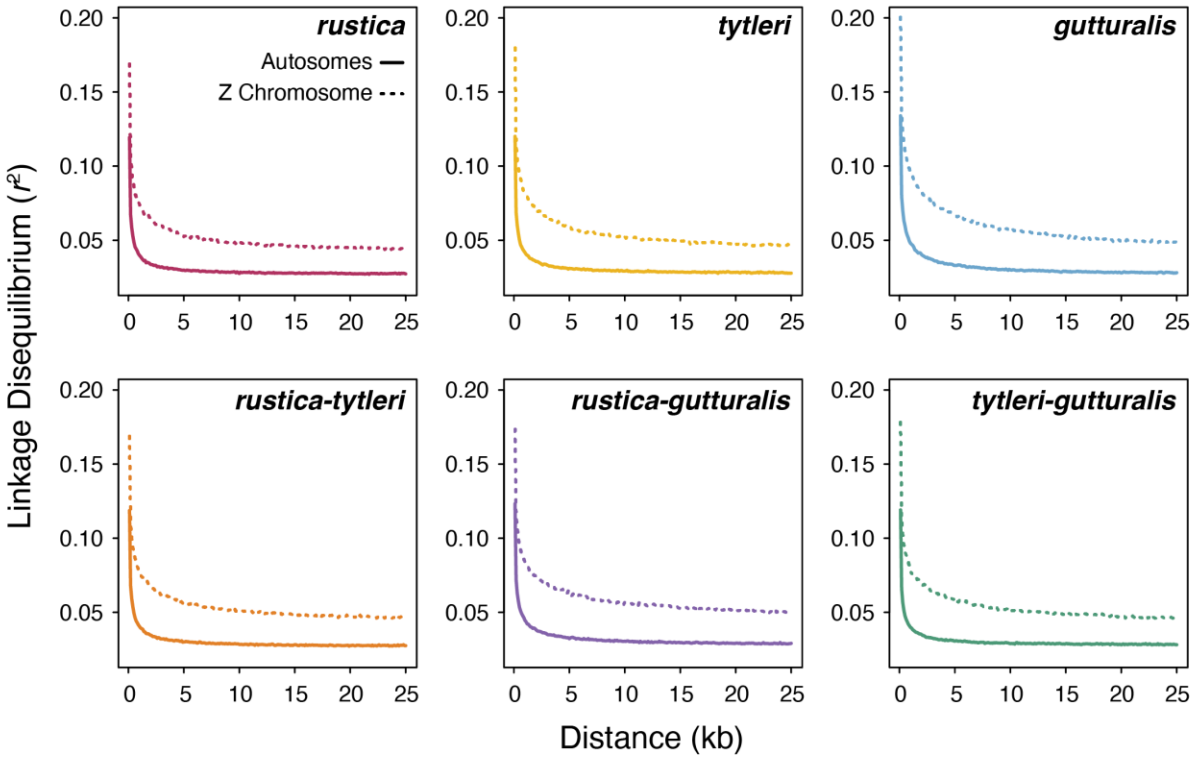

Figure S40. Decay of linkage disequilibrium with physical distance. Mean linkage disequilibrium (LD; r^2) with increasing physical distance between SNPs (x-axis) in parental (top) and hybrid zone (bottom) populations. Solid lines show LD decay on autosomes and dashed lines show decay on the Z chromosome.

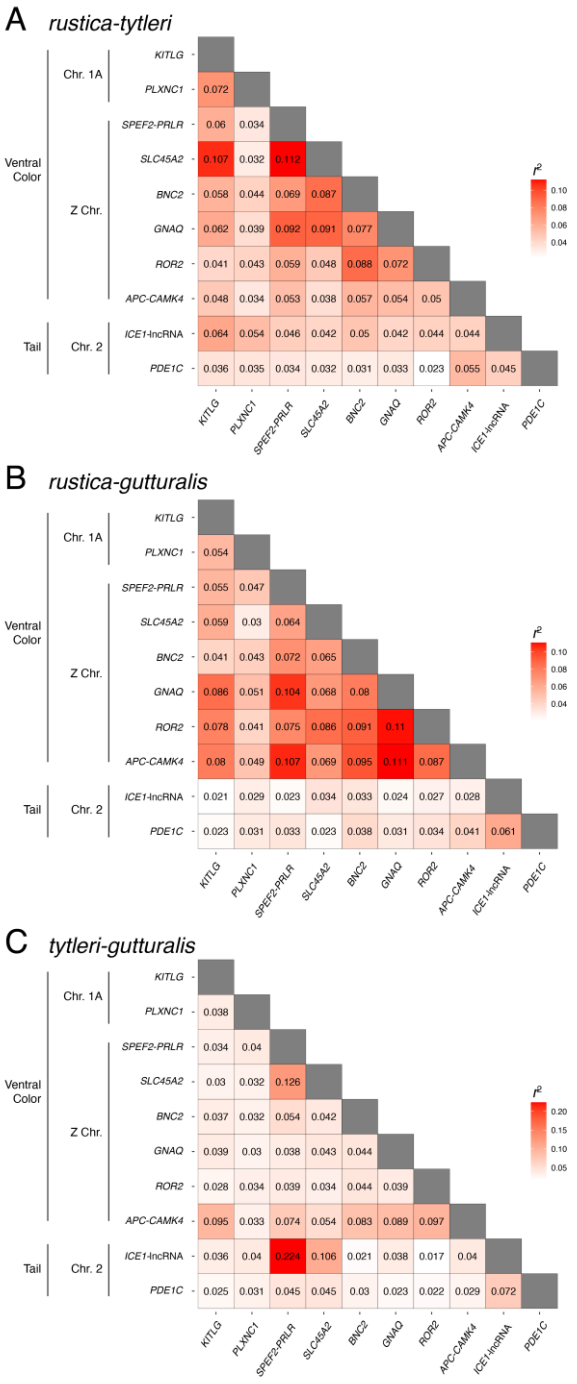

Figure S41. Pairwise linkage disequilibrium between trait loci at ancestry-informative SNPs. Intra- and inter-chromosomal linkage disequilibrium (LD; r^2) measured between SNPs in pairs of trait loci. All SNPs used in pairwise analyses had allele frequency differences ranging between 0.3 – 0.6 between parental populations. The color scales range from relatively high (red) to low (white) LD, and mean LD is labeled for each pair of loci.

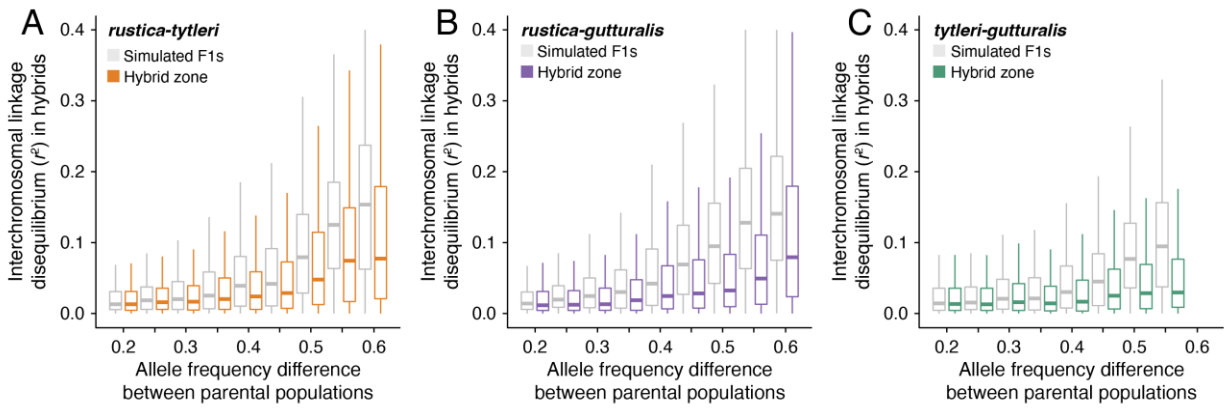

Figure S42. Inter-chromosomal linkage disequilibrium between trait loci in simulated F1s and real hybrids. Inter-chromosomal linkage disequilibrium (LD; r^2) measured between SNPs in trait loci in sampled hybrids (colors) and F1s simulated from parental haplotypes (grey) for the (**A**) *rustica-tytleri*, (**B**) *rustica-gutturalis*, and (**C**) *tytleri-gutturalis* hybrid zones. Inter-chromosomal LD is shown in intervals of increasing allele frequency differences between parental populations (x-axis).

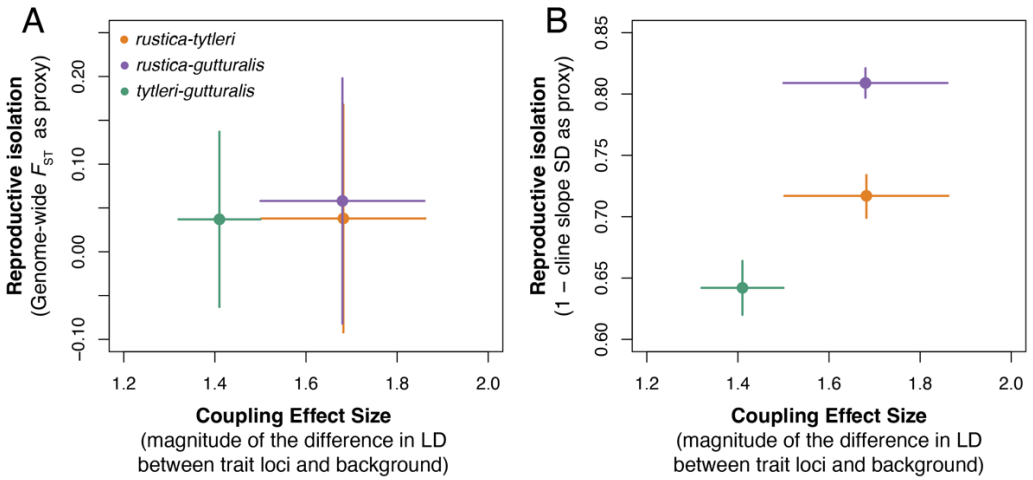

Figure S43. Comparison between degree of genetic coupling maintained among trait loci and proxies for reproductive isolation. (**A**) The x-axis shows the coupling effect size, measured as the magnitude of differences in inter-chromosomal LD among trait loci and background SNPs with matched allele frequency differences between parental populations for each hybrid zone. The y-axis shows genome-wide F_{ST} between parental populations as a proxy for reproductive isolation. (**B**) The y-axis is the same as in (**A**); The y-axis shows 1 - the genomic cline slope SD parameter as a proxy for reproductive isolation. Points show the pairwise intersection of mean values and lines show variance as either SD (coupling effect size and F_{ST}) or credible interval (cline slope) for the three hybrid zones.

Table S1. Mate choice trait variation among barn swallows. Variation in melanin-based ventral color, measured as breast average brightness (% reflectance), and tail streamer lengths in millimeters (mm) among barn swallow subspecies and hybrid zones (mean ± standard deviation).

Subspecies/zone	Ventral color (% reflectance)	Tail streamer length (mm)
erythrogaster	26.1 ± 4.7	86.9 ± 11.4
gutturalis	43.5 ± 10.9	89.4 ± 10.25
rustica	52.9 ± 9.3	103.2 ± 13
savignii	14.8 ± 3.7	92.5 ± 11
transitiva	25.4 ± 7.2	100.3 ± 15.2
tytleri	20.6 ± 6.3	100.9 ± 13.5
rustica-gutturalis	55.1 ± 8.5	97.3 ± 10.9
rustica-tytleri	30.6 ± 16.7	98.7 ± 11.9
tytleri-gutturalis	28.4 ± 10.7	94.5 ± 11.5

Table S2. Statistical tests of phenotype variation among hybrid zones. One-way ANOVA and Tukey *post-hoc* test results comparing ventral color (mean brightness; % reflectance) and tail streamer length (mm) for parental and hybrid populations for each hybrid zone.

		ANOVA		Tukey *post-hoc*		
Trait	Hybrid zone	*F*	*P*-value	Populations	*T*	*P*-value
Ventral Color	*rustica-tytleri*	222.9	$< 2.2 \times 10^{-16}$	*rustica* vs *tytleri*	-17.3	$< 1 \times 10^{-7}$
				hybrid vs *rustica*	-16.6	$< 1 \times 10^{-7}$
				hybrid vs *tytleri*	-5.1	9.9×10^{-7}
	rustica-gutturalis	97.23	$< 2.2 \times 10^{-16}$	*rustica* vs *gutturalis*	10.99	$< 1 \times 10^{-4}$
				hybrid vs *rustica*	1.8	0.16
				hybrid vs *gutturalis*	10.94	$< 1 \times 10^{-4}$
	tytleri-gutturalis	246.4	$< 2.2 \times 10^{-16}$	*tytleri* vs gutturalis	-15.1	$< 1 \times 10^{-7}$
				hybrid vs *tytleri*	4.9	2.14×10^{-6}
				hybrid vs *gutturalis*	-18.9	$< 1 \times 10^{-7}$
Tail streamer length	*rustica-tytleri*	5.22	0.0058	*rustica* vs *tytleri*	-1.19	0.45
				hybrid vs *rustica*	-3.22	0.0038
				hybrid vs *tytleri*	1.07	0.53
	rustica-gutturalis	115.9	$< 2.2 \times 10^{-16}$	*rustica* vs *gutturalis*	14.8	$< 1 \times 10^{-5}$
				hybrid vs *rustica*	6.88	1.3×10^{-5}
				hybrid vs *gutturalis*	-4.57	$< 1 \times 10^{-5}$
	tytleri-gutturalis	38.7	$< 2.2 \times 10^{-16}$	*tytleri* vs gutturalis	7.36	< 0.001
				hybrid vs *tytleri*	-3.95	< 0.001
				hybrid vs *gutturalis*	6.14	< 0.001

Table S3. Geographic sampling of parental populations included in the reduced representation genome sequencing (RADseq) dataset used for demographic inference. Samples were sampled from localities also represented in the whole genome sequencing dataset.

Population	Country	Location	Latitude	Longitude	*n*
rustica	Russia	Karasuk	53.689814	77.655419	27
		Krasnoyarsk	56.357935	92.867796	22
		Moscow	56.809699	37.825493	31
		Novosibirsk	55.031644	82.936521	2
		Yekaterinburg	57.558842	62.662777	50
tytleri	Russia	Malamolevo	53.39529	102.857567	6
		Kytyleek	53.389091	102.871307	9
		Umigan	54.67273	100.094238	9
		Zakaltoose	52.021259	106.590942	30
gutturalis	China	Boatu	40.552814	109.99833	26
		Beijing	39.820223	116.332369	12
		Changchun	43.906216	125.314668	31
		Changsha	28.400245	112.805998	18
		Harbin	45.761237	126.609063	43
		Nanning	22.791937	108.330895	17
		Qinhuangdao	39.92558	119.594752	28
		Qiqihar	47.339176	123.978505	26
		Shenyang	41.808232	123.525653	14
		Shuang	46.59295	131.248593	33
		Xian	34.345908	108.806799	22
		Yinchuan	38.461518	106.286518	30
		Zhengzhou	34.780976	113.65069	28

Table S4. Biological parameter estimates from two-population demographic models. Results of the best-fitting model to the unfolded joint site-frequency spectrum (JSFS) between each pair of parental populations. Details on the best-fit runs for all models are in Data S2. Parameters were scaled to biological estimates assuming a per-generation mutation rate of 2.3×10^{-9}, a generation time of one year, and a sampled effective genome length of 7,279,472 bp.

Populations	Model	θ	N_a	N_1	N_2	m_{12}	m_{21}	m_{e12}	m_{e21}	t_S	t_{SC}
rustica & *tytleri*	SC2*m*	441.36	13,169	48,025	26,101	16.93	22.1	3.96	7.75	9,567.82	981.21
rustica & *gutturalis*	SC2*m*	620.69	18,536	79,413	25,213	17.37	13.49	3.34	5.14	12,556.58	2,066.42
tytleri & *gutturalis*	SC	491.98	14,692	50,400	17,093	12.29	12.72	–	–	9,179.25	884.91

θ = Optimal scaling factor between the model and the empirical JSFS; N_a, N_1, N_2 = effective population size of the ancestral population, population 1, and population 2, respectively, scaled by $2N_{ref}$ units; m_{12} = effective number of migrants per generation from population 2 into population 1; m_{21} = effective number of migrants from population 1 into population 2; m_{e12} and m_{e21} = reduced effective migration rate in genomic islands between populations 1 and 2; t_S = number of generations since divergence, prior to the onset of gene flow in secondary contact; t_{SC} = number of generations since entering secondary contact.

Table S5. Results of genome-wide association mapping using Bayesian sparse linear mixed models (BSLMM) on hybrids to characterized the genetic architecture of mate choice plumage traits. Posterior hyperparameter estimates (median ± standard deviation) of the proportion of phenotypic variance explained by all SNPs (PVE), proportion of variance explained by SNPs with measurable effects (PGE), and the number of SNPs in the models.

Trait	PVE	PGE	*n* SNPs
Ventral color	0.97 ± 0.065	0.79 ± 0.132	12 ± 14.72
Tail streamer length	0.95 ± 0.09	0.49 ± 0.12	2 ± 2.77

Table S6. Results of genome-wide association mapping using linear mixed models. Numbers of SNPs with significant associations with ventral plumage color and tail steamer length across analyses of the hybrid dataset, full dataset, the full dataset with randomized phenotypes, and males versus females. Sample sizes (*n*) vary due to missing trait data for some individuals. Significant SNPs exceeded a Bonferroni-corrected $-\log_{10}(P\text{-value} < 0.05)$ threshold.

Trait	Analysis	*n*	Total SNPs	Significant SNPs	Genes containing significant SNP(s)	Genes within 50 kb of significant SNP(s)
Ventral color	Hybrid	159	9,033,285	1,051	72	149
	Full	305	9,311,628	773	63	119
	Full (randomized)	305	9,311,628	0	0	0
	Male	157	8,851,265	148	21	52
	Female	148	8,757,718	248	24	36
Tail streamer length	Hybrid	151	8,870,504	62	4	14
	Full	300	9,246,603	136	5	20
	Full (randomized)	300	9,246,603	0	0	0
	Male	151	8,743,581	0	0	0
	Female	149	8,773,176	0	0	0

Table S7. Details of candidate genes identified in genome-wide linear mixed model association tests for ventral plumage color and tail streamer length. The position and Wald test *P*-value are reported for the most significant SNP within 50 kb of the gene (*P*-values reported in the main text for associations containing candidate genes correspond to the most significant association for the broader region).

Trait	Chr.	Scaffold	SNP	*P*-value	Annotation	Gene	Start	End	Strand
Ventral color	1A	NC_053453.1	44,285,485	8.95E-10	intergenic	*KITLG*	44,194,614	44,248,428	-
	1A	NC_053453.1	46,520,417	2.03E-10	intergenic	*PLXNC1*	46,488,629	46,558,914	+
	Z	NC_053488.1	18,954,203	1.29E-10	intergenic	*SPEF2*	18,835,478	18,913,404	-
	Z	NC_053488.1	19,018,140	1.45E-13	intron	*PRLR*	18,913,803	19,065,706	+
	Z	NC_053488.1	19,332,900	5.46E-13	intergenic	*SLC45A2*	19,350,426	19,365,366	+
	Z	NC_053488.1	33,380,365	8.98E-12	intron	*BNC2*	33,250,305	33,382,837	-
	Z	NC_053488.1	35,903,569	4.08E-13	intergenic	*GNAQ*	35,948,755	36,067,437	-
	Z	NC_053488.1	44,206,228	1.12E-13	intergenic	*ROR2*	44,106,369	44,157,176	-
	Z	NC_053488.1	46,065,026	4.53E-14	intron	*APC*	45,988,157	46,109,313	-
	Z	NC_053488.1	46,901,341	1.41E-14	intron	*CAMK4*	46,706,857	46,915,499	-
Tail streamer	2	NC_053450.1	98,076,904	5.34E-09	intergenic	*ICE1*	98,082,401	98,124,818	+
	2	NC_053450.1	98,076,904	5.34E-09	intergenic	lncRNA	98,092,866	98,132,143	-
	2	NC_053450.1	100,834,905	6.55E-19	intron	*PDE1C*	100,718,916	101,015,210	-

Table S8. Relative differentiation in genomic regions associated with mate choice traits. Distributions of F_{ST} (mean $\pm$ standard deviation) in genomic regions containing significant GWA SNPs (i.e., $-\log_{10}(P) \geq$ Bonferroni-correction) and results of Mann-Whitney $\cup$ tests comparing distributions to those for non-significant SNPs. Due to the presence of association peaks with ventral color on the Z chromosome, results for autosomes and the Z chromosome are also shown separately. F_{ST} values are means from the 10 kb window containing each SNP and were sampled once in cases where multiple significant SNPs were in the same window.

Trait	Populations	All sig.	All non.	Autosome sig.	Autosome non.	Z chr. sig.	Z chr. Non.
Ventral color	*rustica* vs *tytleri*	$0.39 \pm 0.2^{**}$	0.02 ± 0.06	$0.15 \pm 0.14^{**}$	0.01 ± 0.02	$0.43 \pm 0.19^{**}$	0.13 ± 0.17
	rustica vs *gutturalis*	$0.38 \pm 0.25^{**}$	0.04 ± 0.07	$0.11 \pm 0.1^{**}$	0.03 ± 0.03	$0.43 \pm 0.24^{**}$	0.16 ± 0.18
	tytleri vs *gutturalis*	$0.23 \pm 0.23^{**}$	0.03 ± 0.05	$0.09 \pm 0.08^{**}$	0.02 ± 0.02	$0.26 \pm 0.23^{**}$	0.09 ± 0.13
Tail streamer	*rustica* vs *tytleri*	$0.01 \pm 0.04^{*}$	0.02 ± 0.06	–	–	–	–
	rustica vs *gutturalis*	0.03 ± 0.04	0.04 ± 0.07	–	–	–	–
	tytleri vs *gutturalis*	0.02 ± 0.02	0.02 ± 0.05	–	–	–	–

Abbreviations: sig. = significant SNPs; non. = non-significant SNPs; Z chr. = Z chromosome. $^{*}P < 0.05$; $^{**}P < 0.0001$.

Table S9. Genome-wide population genetic summary statistics within and between parental *rustica, tytleri*, and *gutturalis* populations. Mean ± standard deviation between-population differentiation (F_{ST}), between-population nucleotide differences (d_{xy}), and within-population nucleotide diversity (π) across the whole genome, and on autosomes, the Z chromosome, and the W chromosome, specifically. All statistics were calculated in non-overlapping 100 kb windows.

Statistic	Populations	Genome-wide	Autosome	Z chromosome	W chromosome
F_{ST}	*rustica* vs *tytleri*	0.038 ± 0.133	0.011 ± 0.018	0.134 ± 0.169	0.679 ± 0.355
	rustica vs *gutturalis*	0.058 ± 0.139	0.029 ± 0.028	0.167 ± 0.181	0.713 ± 0.355
	tytleri vs *gutturalis*	0.037 ± 0.106	0.02 ± 0.018	0.088 ± 0.131	0.536 ± 0.368
d_{xy}	*rustica* vs *tytleri*	0.0049 ± 0.0021	0.0053 ± 0.0018	0.0023 ± 0.001	0.00025 ± 0.00032
	rustica vs *gutturalis*	0.005 ± 0.0021	0.0053 ± 0.0018	0.0024 ± 0.001	0.00023 ± 0.00028
	tytleri vs *gutturalis*	0.0049 ± 0.0021	0.0053 ± 0.0018	0.0021 ± 0.001	0.00015 ± 0.00026
π	*rustica*	0.0049 ± 0.0021	0.0053 ± 0.0018	0.0021 ± 0.0011	0.00004 ± 0.00016
	tytleri	0.0048 ± 0.0021	0.0052 ± 0.0018	0.002 ± 0.0011	0.00006 ± 0.00024
	gutturalis	0.0047 ± 0.002	0.0051 ± 0.0017	0.002 ± 0.001	0.00004 ± 0.00018

Table S10. Genome-wide relationships between relative population differentiation (F_{ST}) and other population genetic parameters. Spearman's correlation coefficients (ρ) and *P*-values are reported to summarize relationships within and among pairs of parental populations. All correlations were calculated using parameter estimates in non-overlapping 100 kb windows.

Parameters	Populations	ρ	*P*
F_{ST}, recombination rate	*rustica* vs *tytleri*	-0.52	$< 2.2 \times 10^{-16}$
	rustica vs *gutturalis*	-0.57	$< 2.2 \times 10^{-16}$
	tytleri vs *gutturalis*	-0.44	$< 2.2 \times 10^{-16}$
F_{ST}, π	*rustica* vs *tytleri*, *rustica*	-0.46	$< 2.2 \times 10^{-16}$
	rustica vs *tytleri*, *tytleri*	-0.47	$< 2.2 \times 10^{-16}$
	rustica vs *gutturalis*, *rustica*	-0.53	$< 2.2 \times 10^{-16}$
	rustica vs *gutturalis*, *gutturalis*	-0.55	$< 2.2 \times 10^{-16}$
	tytleri vs *gutturalis*, *tytleri*	-0.41	$< 2.2 \times 10^{-16}$
	tytleri vs *gutturalis*, *gutturalis*	-0.42	$< 2.2 \times 10^{-16}$
F_{ST}, d_{xy}	*rustica* vs *tytleri*	-0.46	$< 2.2 \times 10^{-16}$
	rustica vs *gutturalis*	-0.52	$< 2.2 \times 10^{-16}$
	tytleri vs *gutturalis*	-0.40	$< 2.2 \times 10^{-16}$
F_{ST}, F_{ST}	*rustica* vs *tytleri*, *rustica* vs *gutturalis*	0.52	$< 2.2 \times 10^{-16}$
	rustica vs *tytleri*, *tytleri* vs *gutturalis*	0.45	$< 2.2 \times 10^{-16}$
	rustica vs *gutturalis*, *tytleri* vs *gutturalis*	0.66	$< 2.2 \times 10^{-16}$

Table S11. Model estimates from geographic cline analyses of the genome background and trait-associated candidate loci. Background estimates are from 1,000 randomly sampled 100 SNP windows with minor allele frequency ≥ 0.1. Background cline parameters are summarized using mean estimates ± standard deviation.

Hybrid zone	Chromosome	Locus	Model	Center (*c*)	Width (*w*)	p_{Min}	p_{Max}
rustica-tytleri	–	Background	–	1442.16 ± 85.34	977 ± 483.67	0.1 ± 0.07	0.94 ± 0.07
	1A	*KITLG*	I	1372.37	122.49	0.00	0.92
	1A	*PLXNC1*	I	1363.44	94.09	0.03	0.90
	Z	*SPEF2-PRLR*	II*	1384.03	165.42	0.04	0.89
	Z	*SLC45A2*	II*	1372.39	185.30	0.01	0.85
	Z	*BNC2*	I	1380.61	365.21	0.00	1.00
	Z	*GNAQ*	I	1368.55	139.15	0.00	0.95
	Z	*ROR2*	II*	1385.40	222.80	0.00	0.87
	Z	*APC-CAMK4*	II*	1375.50	150.53	0.01	0.87
	2	*ICE1*-lncRNA	I	1336.49	120.71	0.20	0.77
	2	*PDE1C*	II*	1337.63	171.11	0.17	0.81
rustica-gutturalis	–	Background	–	1961 ± 58.01	1901.02 ± 220.07	0.07 ± 0.06	0.77 ± 0.09
	1A	*KITLG*	I	2473.15	367.13	0.00	0.85
	1A	*PLXNC1*	I	2526.81	1259.48	0.02	1.00
	Z	*SPEF2-PRLR*	I	2438.50	994.72	0.00	1.00
	Z	*SLC45A2*	II†	2430.98	591.68	0.01	0.87
	Z	*BNC2*	I	2531.59	470.71	0.01	0.89
	Z	*GNAQ*	I	2537.02	591.45	0.00	0.98
	Z	*ROR2*	I	2476.39	411.95	0.00	1.00
	Z	*APC-CAMK4*	I	2490.69	386.96	0.00	1.00
	2	*ICE1*-lncRNA	II†	2495.28	934.76	0.13	0.84
	2	*PDE1C*	I	2709.07	470.07	0.23	0.84
tytleri-gutturalis	–	Background	–	1042.4 ± 219.5	1693 ± 508.26	0.03 ± 0.05	0.96 ± 0.09
	1A	*KITLG*	I	1408.30	2212.62	0.00	0.64
	1A	*PLXNC1*	I	1419.75	2213.71	0.00	0.93
	Z	*SPEF2-PRLR*	I	1020.98	1495.81	0.00	0.56
	Z	*SLC45A2*	I	1219.67	1573.82	0.11	0.94
	Z	*BNC2*	I	1312.11	1511.43	0.00	1.00
	Z	*GNAQ*	I	1274.07	2212.02	0.00	1.00
	Z	*ROR2*	I	1095.78	1275.43	0.00	1.00
	Z	*APC-CAMK4*	I	851.81	907.42	0.00	1.00
	2	*ICE1*-lncRNA	I	1228.71	2213.95	0.00	0.81
	2	*PDE1C*	I	1414.51	2214.04	0.20	1.00

* Model I had similar support in the *rustica-tytleri* zone (ΔAIC *SPEF2-PRLR* = 0.08, *SLC45A2* = 1.6, *ROR2* = 0.67, *APC-CAMK4* = 1.3, *PDE1C* = 0.45), but was a poor visual fit to the data (see Fig. S36).

† Model I had similar support in the *rustica-gutturalis* zone (ΔAIC *SLC45A2* = 1.1, *ICE1*-lncRNA = 1.3).

Table S12. Genome-wide genomic cline parameter estimates. Cline slope and center parameter standard deviation (σ; this is the actual model parameter from the hierarchical model) and related empirical standard deviations (SD; these are computed from the cline parameter estimates and are thus derived parameters and not part of the formal model) from genomic cline analyses across the three hybrid zones.

Parameter	*rustica-tytleri*	*rustica-gutturalis*	*tytleri-gutturalis*
σ_v	0.283	0.191	0.358
σ_v lower CI	0.266	0.179	0.336
σ_v upper CI	0.301	0.203	0.380
SD_v	0.273	0.187	0.329
SD_v lower CI	0.260	0.178	0.314
SD_v upper CI	0.286	0.196	0.345
σ_c	0.459	0.486	0.467
σ_c upper CI	0.431	0.460	0.439
σ_c lower CI	0.489	0.513	0.499
SD_c	0.466	0.494	0.473
SD_c lower CI	0.444	0.476	0.450
SD_c upper CI	0.488	0.512	0.497

σ_v = cline slope standard deviation parameter; σ_c = cline center standard deviation parameter; $SD_{v,c}$ = empirical standard deviation of cline slope and center parameters; CI = 95% credible interval.

Table S13. Assignment of barn swallow genome scaffolds to chromosomes. Sequence homology-based conversion of original barn swallow scaffold chromosome assignments to karyotypic chromosomes in the zebra finch (*Taeniopygia guttata*) genome assembly.

Scaffold	Original assignment	Zebra finch scaffold	Final assignment	Scaffold length (bp)
NC_053451.1	chromosome 2	NC_044998.1	Chromosome 1	119023421
NW_024403813.1	chromosome 2, unlocalized	NC_044998.1	Chromosome 1, unlocalized	43722
NC_053453.1	chromosome 4	NC_044999.1	Chromosome 1A	76187387
NW_024403819.1	chromosome 4, unlocalized	NC_044999.1	Chromosome 1A, unlocalized	89507
NC_053450.1	chromosome 1	NC_045000.1	Chromosome 2	156035725
NW_024403809.1	chromosome 1, unlocalized	NC_045000.1	Chromosome 2, unlocalized	353740
NW_024403810.1	chromosome 1, unlocalized	NC_045000.1	Chromosome 2, unlocalized	281945
NW_024403811.1	chromosome 1, unlocalized	NC_045000.1	Chromosome 2, unlocalized	109630
NW_024403812.1	chromosome 1, unlocalized	NC_045000.1	Chromosome 2, unlocalized	69765
NC_053452.1	chromosome 3	NC_045001.1	Chromosome 3	116801625
NW_024403814.1	chromosome 3, unlocalized	NC_045001.1	Chromosome 3, unlocalized	111011
NW_024403815.1	chromosome 3, unlocalized	NC_045001.1	Chromosome 3, unlocalized	92199
NW_024403816.1	chromosome 3, unlocalized	NC_045001.1	Chromosome 3, unlocalized	17442
NW_024403817.1	chromosome 3, unlocalized	NC_045001.1	Chromosome 3, unlocalized	32528
NW_024403818.1	chromosome 3, unlocalized	NC_045001.1	Chromosome 3, unlocalized	29558
NC_053454.1	chromosome 5	NC_045002.1	Chromosome 4	73257097
NW_024403820.1	chromosome 5, unlocalized	NC_045002.1	Chromosome 4, unlocalized	216157
NW_024403821.1	chromosome 5, unlocalized	NC_045002.1	Chromosome 4, unlocalized	116708
NW_024403822.1	chromosome 5, unlocalized	NC_045002.1	Chromosome 4, unlocalized	116594
NW_024403823.1	chromosome 5, unlocalized	NC_045002.1	Chromosome 4, unlocalized	41746
NC_053470.1	chromosome 21	NC_045003.1	Chromosome 4A	9617204
NW_024403828.1	chromosome 21, unlocalized	NC_045003.1	Chromosome 4A, unlocalized	35023
NC_053455.1	chromosome 6	NC_045004.1	Chromosome 5	63258489
NW_024403824.1	chromosome 6, unlocalized	NC_045004.1	Chromosome 5, unlocalized	17742
NC_053457.1	chromosome 8	NC_045005.1	Chromosome 6	36085389
NC_053456.1	chromosome 7	NC_045006.1	Chromosome 7	38459648
NC_053458.1	chromosome 9	NC_045007.1	Chromosome 8	31262510
NC_053459.1	chromosome 10	NC_045008.1	Chromosome 9	25880253
NC_053462.1	chromosome 13	NC_045009.1	Chromosome 10	20272128
NW_024403826.1	chromosome 13, unlocalized	NC_045009.1	Chromosome 10, unlocalized	72671
NC_053460.1	chromosome 11	NC_045010.1	Chromosome 11	21491857
NW_024403825.1	chromosome 11, unlocalized	NC_045010.1	Chromosome 11, unlocalized	36741
NC_053461.1	chromosome 12	NC_045011.1	Chromosome 12	20890524
NC_053463.1	chromosome 14	NC_045012.1	Chromosome 13	18810845
NC_053464.1	chromosome 15	NC_045013.1	Chromosome 14	16541138
NW_024403827.1	chromosome 15, unlocalized	NC_045013.1	Chromosome 14, unlocalized	34847
NC_053466.1	chromosome 17	NC_045014.1	Chromosome 15	13985943
NC_053469.1	chromosome 20	NC_045015.1	Chromosome 17	11194341
NC_053467.1	chromosome 18	NC_045016.1	Chromosome 18	12073725
NC_053468.1	chromosome 19	NC_045017.1	Chromosome 19	11382101
NC_053465.1	chromosome 16	NC_045018.1	Chromosome 20	15277844
NC_053471.1	chromosome 22	NC_045019.1	Chromosome 21	7507825
NW_024403829.1	chromosome 22, unlocalized	NC_045019.1	Chromosome 21, unlocalized	95381
NC_053477.1	chromosome 28	NC_045020.1	Chromosome 22	5297670
NC_053474.1	chromosome 25	NC_045021.1	Chromosome 23	6778862
NC_053472.1	chromosome 23	NC_045022.1	Chromosome 24	7098401
NC_053478.1	chromosome 29	NC_045023.1	Chromosome 25	2102120
NW_024403832.1	chromosome 29, unlocalized	NC_045023.1	Chromosome 25, unlocalized	67659
NW_024403833.1	chromosome 29, unlocalized	NC_045023.1	Chromosome 25, unlocalized	135211

NC_053473.1	chromosome 24	NC_045024.1	Chromosome 26	6843954
NC_053476.1	chromosome 27	NC_045025.1	Chromosome 27	5236451
NW_024403830.1	chromosome 27, unlocalized	NC_045025.1	Chromosome 27, unlocalized	65965
NW_024403831.1	chromosome 27, unlocalized	NC_045025.1	Chromosome 27, unlocalized	24934
NC_053475.1	chromosome 26	NC_045026.1	Chromosome 28	5553549
NC_053479.1	chromosome 30	NC_045955.1	Chromosome 29	1648998
NW_024403834.1	chromosome 30, unlocalized	NC_045955.1	Chromosome 29, unlocalized	27319
NC_053480.1	chromosome 31	–	chromosome 31	1590086
NC_053481.1	chromosome 32	–	chromosome 32	784579
NC_053482.1	chromosome 33	–	chromosome 33	437724
NW_024403835.1	chromosome 33, unlocalized	–	chromosome 33, unlocalized	276281
NC_053483.1	chromosome 34	–	chromosome 34	606149
NC_053484.1	chromosome 35	–	chromosome 35	523230
NC_053485.1	chromosome 36	–	chromosome 36	338027
NC_053486.1	chromosome 37	–	chromosome 37	276370
NC_053488.1	chromosome Z	NC_045027.1	Chromosome Z	90132487
NW_024403838.1	chromosome Z, unlocalized	NC_045027.1	Chromosome Z, unlocalized	25280
NC_053487.1	chromosome W	NC_045028.1	Chromosome W	31704074
NW_024403836.1	chromosome W, unlocalized	NC_045028.1	Chromosome W, unlocalized	626370
NW_024403837.1	chromosome W, unlocalized	NC_045028.1	Chromosome W, unlocalized	49179

Data S1. (separate file)
List of samples used in the study, with locality, sex, plumage trait, and genome sequencing details.

Data S2. (separate file)
Two-population demographic models fit to the observed unfolded joint site-frequency spectra between parental populations. Results of the best-fit run per model are shown. Model fits are ranked by -Δ AIC.

Data S3. (separate file)
Details of SNPs with significant associations with ventral color and tail streamer length in linear mixed model analyses of genome-wide associations in hybrids.

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 Medline

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 Medline

6. J. W. Boughman, Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature* **411**, 944–948 (2001). doi:10.1038/35082064 Medline

7. O. Seehausen, Y. Terai, I. S. Magalhaes, K. L. Carleton, H. D. J. Mrosso, R. Miyagi, I. van der Sluijs, M. V. Schneider, M. E. Maan, H. Tachida, H. Imai, N. Okada, Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–626 (2008). doi:10.1038/nature07285 Medline

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 Medline

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 Medline

10. M. Rossi, A. E. Hausmann, P. Alcami, M. Moest, R. Roussou, S. M. Van Belleghem, D. S. Wright, C.-Y. Kuo, D. Lozano-Urrego, A. Maulana, L. Melo-Flórez, G. Rueda-Muñoz, S. McMahon, M. Linares, C. Osman, W. O. McMillan, C. Pardo-Diaz, C. Salazar, R. M. Merrill, Adaptive introgression of a visual preference gene. *Science* **383**, 1368–1373 (2024). doi:10.1126/science.adj9201 Medline

11. T. C. Mendelson, K. L. Shaw, Sexual behaviour: Rapid speciation in an arthropod. *Nature* **433**, 375–376 (2005). doi:10.1038/433375a Medline

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 Medline

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 Medline

14. J. W. Poelstra, N. Vijay, C. M. Bossu, H. Lantz, B. Ryll, I. Müller, V. Baglione, P. Unneberg, M. Wikelski, M. G. Grabherr, J. B. W. Wolf, The genomic landscape

underlying phenotypic integrity in the face of gene flow in crows. *Science* **344**, 1410–1414 (2014). doi:10.1126/science.1253226 Medline

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 Medline

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 Medline

17. E. L. Westerman, N. W. VanKuren, D. Massardo, A. Tenger-Trolander, W. Zhang, R. I. Hill, M. Perry, E. Bayala, K. Barr, N. Chamberlain, T. E. Douglas, N. Buerkle, S. E. Palmer, M. R. Kronforst, 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 Medline

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 Medline

19. R. Riesch, M. Muschick, D. Lindtke, R. Villoutreix, A. A. Comeault, T. E. Farkas, K. Lucek, E. Hellen, V. Soria-Carrasco, S. R. Dennis, C. F. de Carvalho, R. J. Safran, C. P. Sandoval, J. Feder, R. Gries, B. J. Crespi, G. Gries, Z. Gompert, P. Nosil, Transitions between phases of genomic differentiation during stick-insect speciation. *Nat. Ecol. Evol.* **1**, 0082 (2017). doi:10.1038/s41559-017-0082 Medline

20. O. Seehausen, R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, P. A. Hohenlohe, C. L. Peichel, G. P. Saetre, C. Bank, A. Brännström, A. Brelsford, C. S. Clarkson, F. Eroukhmanoff, J. L. Feder, M. C. Fischer, A. D. Foote, P. Franchini, C. D. Jiggins, F. C. Jones, A. K. Lindholm, K. Lucek, M. E. Maan, D. A. Marques, S. H. Martin, B. Matthews, J. I. Meier, M. Möst, M. W. Nachman, E. Nonaka, D. J. Rennison, J. Schwarzer, E. T. Watson, A. M. Westram, A. Widmer, J. Schwarzer, J. Schwarzer, J. Schwarzer, E. T. Watson, A. M. Westram, A. Widmer, Genomics and the origin of species. *Nat. Rev. Genet.* **15**, 176–192 (2014). doi:10.1038/nrg3644 Medline

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 Medline

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 Medline

23. M. Ravinet, R. Faria, R. K. Butlin, J. Galindo, N. Bierne, M. Rafajlović, M. A. F. Noor, B. Mehlig, A. M. Westram, 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 Medline

24. R. K. Butlin, C. M. Smadja, Coupling, reinforcement, and speciation. *Am. Nat.* **191**, 155–172 (2018). doi:10.1086/695136 Medline

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 Medline

26. N. H. Barton, Multilocus clines. *Evolution* **37**, 454–471 (1983). doi:10.2307/2408260 Medline

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 Medline

28. T. H. Vines, A. C. Dalziel, A. Y. K. Albert, T. Veen, P. M. Schulte, D. Schluter, Cline coupling and uncoupling in a stickleback hybrid zone. *Evolution* **70**, 1023–1038 (2016). doi:10.1111/evo.12917 Medline

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 Medline

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). Medline

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 Medline

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, R. J. Safran, Y. Vortman, A. Lotem, A. McGowan, M. R. Evans, I. J. Lovette, 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 Medline

35. R. J. Safran, Y. Vortman, B. R. Jenkins, J. K. Hubbard, M. R. Wilkins, R. J. Bradley, A. Lotem, 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 Medline

36. M. R. Wilkins, H. Karaardıç, Y. Vortman, T. L. Parchman, T. Albrecht, A. Petrželková, L. Özkan, P. L. Pap, J. K. Hubbard, A. K. Hund, R. J. Safran, 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 Medline

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 Medline

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 Medline

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 Medline

41. C. C. R. Smith, S. M. Flaxman, E. S. C. Scordato, N. C. Kane, A. K. Hund, B. M. Sheta, R. J. Safran, 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 Medline

42. E. S. C. Scordato, M. R. Wilkins, G. Semenov, A. S. Rubtsov, N. C. Kane, R. J. Safran, 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 Medline

43. E. S. C. Scordato, C. C. R. Smith, G. A. Semenov, Y. Liu, M. R. Wilkins, W. Liang, A. Rubtsov, G. Sundev, K. Koyama, S. P. Turbek, M. B. Wunder, C. A. Stricker, R. J. Safran, 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 Medline

44. S. P. Turbek, D. R. Schield, E. S. C. Scordato, A. Contina, X. W. Da, Y. Liu, Y. Liu, E. Pagani-Núñez, Q. M. Ren, C. C. R. Smith, C. A. Stricker, M. Wunder, D. M. Zonana, R. J. Safran, A migratory divide spanning two continents is associated with genomic and ecological divergence. *Evolution* **76**, 722–736 (2022). doi:10.1111/evo.14448 Medline

45. D. R. Schield, E. S. C. Scordato, C. C. R. Smith, J. K. Carter, S. I. Cherkaoui, S. Gombobaatar, S. Hajib, S. Hanane, A. K. Hund, K. Koyama, W. Liang, Y. Liu, N. Magri, A. Rubtsov, B. Sheta, S. P. Turbek, M. R. Wilkins, L. Yu, R. J. Safran, Sex-linked genetic diversity and differentiation in a globally distributed avian species complex. *Mol. Ecol.* **30**, 2313–2332 (2021). doi:10.1111/mec.15885 Medline

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 Medline

47. C. T. Miller, S. Beleza, A. A. Pollen, D. Schluter, R. A. Kittles, M. D. Shriver, D. M. Kingsley, 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 Medline

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 Medline

49. L. Le, I. E. Escobar, T. Ho, A. J. Lefkovith, E. Latteri, K. D. Haltaufderhyde, M. K. Dennis, L. Plowright, E. V. Sviderskaya, D. C. Bennett, E. Oancea, M. S. Marks, 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 Medline

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 Medline

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 Medline

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:10.1098/rspb.2003.2391 Medline

53. D. C. Presgraves, Sex chromosomes and speciation in *Drosophila*. *Trends Genet.* **24**, 336–343 (2008). doi:10.1016/j.tig.2008.04.007 Medline

54. Y. Cai, D. J. Nagel, Q. Zhou, K. D. Cygnar, H. Zhao, F. Li, X. Pi, P. A. Knight, C. Yan, 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 Medline

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 Medline

56. Z. Gompert, L. K. Lucas, C. C. Nice, J. A. Fordyce, M. L. Forister, C. A. Buerkle, 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 Medline

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 Medline

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 Medline

59. D. R. Schield, R. H. Adams, D. C. Card, B. W. Perry, G. M. Pasquesi, T. Jezkova, D. M. Portik, A. L. Andrew, C. L. Spencer, E. E. Sanchez, M. K. Fujita, S. P. Mackessy, T. A. Castoe, 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 Medline

60. N. Barton, The hybrid sink effect. *Heredity* **44**, 277–278 (1980). doi:10.1038/hdy.1980.23

61. T. J. Firneno Jr., G. Semenov, E. B. Dopman, S. A. Taylor, E. L. Larson, Z. Gompert, Quantitative Analyses of Coupling in Hybrid Zones. *Cold Spring Harb. Perspect. Biol.* **15**, a041434 (2023). doi:10.1101/cshperspect.a041434 Medline

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, R. Villoutreix, C. F. de Carvalho, J. L. Feder, T. L. Parchman, Z. Gompert, 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 Medline

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 Medline

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 Medline

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 Medline

67. J. L. Feder, P. Nosil, A. C. Wacholder, S. P. Egan, S. H. Berlocher, S. M. Flaxman, 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 Medline

68. R. J. Safran, E. S. Scordato, M. R. Wilkins, J. K. Hubbard, B. R. Jenkins, T. Albrecht, S. M. Flaxman, H. Karaardıç, Y. Vortman, A. Lotem, P. Nosil, P. Pap, S. Shen, S. F. Chan, T. L. Parchman, N. C. Kane, 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 Medline

69. A. K. Hund, J. K. Hubbard, T. Albrecht, Y. Vortman, P. Munclinger, S. Krausová, O. Tomášek, R. J. Safran, Divergent sexual signals reflect costs of local parasites. *Evolution* **74**, 2404–2418 (2020). doi:10.1111/evo.13994 Medline

70. D. P. L. Toews, S. A. Taylor, R. Vallender, A. Brelsford, B. G. Butcher, P. W. Messer, I. J. Lovette, 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 Medline

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 Medline

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 Medline

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 Medline

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 Medline

75. S. Feng, J. Stiller, Y. Deng, J. Armstrong, Q. Fang, A. H. Reeve, D. Xie, G. Chen, C. Guo, B. C. Faircloth, B. Petersen, Z. Wang, Q. Zhou, M. Diekhans, W. Chen, S. Andreu-Sánchez, A. Margaryan, J. T. Howard, C. Parent, G. Pacheco, M. S. Sinding, L. Puetz, E. Cavill, Â. M. Ribeiro, L. Eckhart, J. Fjeldså, P. A. Hosner, R. T. Brumfield, L. Christidis, M. F. Bertelsen, T. Sicheritz-Ponten, D. T. Tietze, B. C. Robertson, G. Song, G. Borgia, S. Claramunt, I. J. Lovette, S. J. Cowen, P. Njoroge, J. P. Dumbacher, O. A. Ryder, J. Fuchs, M. Bunce, D. W. Burt, J. Cracraft, G. Meng, S. J. Hackett, P. G. Ryan, K. A. Jønsson, I. G. Jamieson, R. R. da Fonseca, E. L. Braun, P. Houde, S. Mirarab, A. Suh, B.

Hansson, S. Ponnikas, H. Sigeman, M. Stervander, P. B. Frandsen, H. van der Zwan, R. van der Sluis, C. Visser, C. N. Balakrishnan, A. G. Clark, J. W. Fitzpatrick, R. Bowman, N. Chen, A. Cloutier, T. B. Sackton, S. V. Edwards, D. J. Foote, S. B. Shakya, F. H. Sheldon, A. Vignal, A. E. R. Soares, B. Shapiro, J. González-Solís, J. Ferrer-Obiol, J. Rozas, M. Riutort, A. Tigano, V. Friesen, L. Dalén, A. O. Urrutia, T. Székely, Y. Liu, M. G. Campana, A. Corvelo, R. C. Fleischer, K. M. Rutherford, N. J. Gemmell, N. Dussex, H. Mouritsen, N. Thiele, K. Delmore, M. Liedvogel, A. Franke, M. P. Hoeppner, O. Krone, A. M. Fudickar, B. Milá, E. D. Ketterson, A. E. Fidler, G. Friis, Á. M. Parody-Merino, P. F. Battley, M. P. Cox, N. C. B. Lima, F. Prosdocimi, T. L. Parchman, B. A. Schlinger, B. A. Loiselle, J. G. Blake, H. C. Lim, L. B. Day, M. J. Fuxjager, M. W. Baldwin, M. J. Braun, M. Wirthlin, R. B. Dikow, T. B. Ryder, G. Camenisch, L. F. Keller, J. M. DaCosta, M. E. Hauber, M. I. M. Louder, C. C. Witt, J. A. McGuire, J. Mudge, L. C. Megna, M. D. Carling, B. Wang, S. A. Taylor, G. Del-Rio, A. Aleixo, A. T. R. Vasconcelos, C. V. Mello, J. T. Weir, D. Haussler, Q. Li, H. Yang, J. Wang, F. Lei, C. Rahbek, M. T. P. Gilbert, G. R. Graves, E. D. Jarvis, B. Paten, G. Zhang, Dense sampling of bird diversity increases power of comparative genomics. *Nature* **587**, 252–257 (2020). doi:10.1038/s41586-020-2873-9 Medline

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 Medline

77. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, 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 Medline

78. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin; 1000 Genome Project data Processing Subgroup, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009). doi:10.1093/bioinformatics/btp352 Medline

79. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011). doi:10.1093/bioinformatics/btr330 Medline

80. X. Zheng, D. Levine, J. Shen, S. M. Gogarten, C. Laurie, B. S. Weir, A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328 (2012). doi:10.1093/bioinformatics/bts606 Medline

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 Medline

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 Medline

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 Medline

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 Medline

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 Medline

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 Medline

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 Medline

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 Medline

89. X. Yi, Y. Liang, E. Huerta-Sanchez, X. Jin, Z. X. Cuo, J. E. Pool, X. Xu, H. Jiang, N. Vinckenbosch, T. S. Korneliussen, H. Zheng, T. Liu, W. He, K. Li, R. Luo, X. Nie, H. Wu, M. Zhao, H. Cao, J. Zou, Y. Shan, S. Li, Q. Yang, P. Asan, P. Ni, G. Tian, J. Xu, X. Liu, T. Jiang, R. Wu, G. Zhou, M. Tang, J. Qin, T. Wang, S. Feng, G. Li, J. Huasang, J. Luosang, W. Wang, F. Chen, Y. Wang, X. Zheng, Z. Li, Z. Bianba, G. Yang, X. Wang, S. Tang, G. Gao, Y. Chen, Z. Luo, L. Gusang, Z. Cao, Q. Zhang, W. Ouyang, X. Ren, H. Liang, H. Zheng, Y. Huang, J. Li, L. Bolund, K. Kristiansen, Y. Li, Y. Zhang, X. Zhang, R. Li, S. Li, H. Yang, R. Nielsen, J. Wang, J. Wang, Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75–78 (2010). doi:10.1126/science.1190371 Medline

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 Medline

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 Medline

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 Medline

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 Medline

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 Medline

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 Medline

97. R. Bailey, Bayesian hybrid index and genomic cline estimation with the R package *gghybrid*. *Mol. Ecol. Resour.* (2023). doi:10.1111/1755-0998.13910 Medline

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.

101. S. P. Turbek, M. Browne, A. S. Di Giacomo, C. Kopuchian, W. M. Hochachka, C. Estalles, D. A. Lijtmaer, P. L. Tubaro, L. F. Silveira, I. J. Lovette, R. J. Safran, S. A. Taylor, L. Campagna, Rapid speciation via the evolution of pre-mating isolation in the Iberá Seedeater. *Science* **371**, eabc0256 (2021). doi:10.1126/science.abc0256 Medline

102. G.-P. Stre, T. Moum, S. Bureš, M. Král, M. Adamjan, J. Moreno, A sexually selected character displacement in flycatchers reinforces premating isolation. *Nature* **387**, 589–592 (1997). doi:10.1038/42451

103. K. L. Shaw, E. Lugo, Mating asymmetry and the direction of evolution in the Hawaiian cricket genus Laupala. *Mol. Ecol.* **10**, 751–759 (2008). doi:10.1046/j.1365-294x.2001.01219.x Medline

104. J. A. C. Uy, R. G. Moyle, C. E. Filardi, Plumage and song differences mediate species recognition between incipient flycatcher species of the Solomon Islands. *Evolution* **63**, 153–164 (2009). doi:10.1111/j.1558-5646.2008.00530.x Medline

105. A. P. Møller, Male ornament size as a reliable cue to enhanced offspring viability in the barn swallow. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6929–6932 (1994). doi:10.1073/pnas.91.15.6929 Medline

106. R. J. Safran, K. J. McGraw, Plumage coloration, not length or symmetry of tail-streamers, is a sexually selected trait in North American barn swallows. *Behav. Ecol.* **15**, 455–461 (2004). doi:10.1093/beheco/arh035

107. C. R. Neuman, R. J. Safran, I. J. Lovette, Male tail streamer length does not predict apparent or genetic reproductive success in North American barn swallows *Hirundo rustica erythrogaster*. *J. Avian Biol.* **38**, 28–36 (2007). doi:10.1111/j.2007.0908-8857.03713.x

108. Y. Vortman, A. Lotem, R. Dor, I. Lovette, R. J. Safran, Multiple sexual signals and behavioral reproductive isolation in a diverging population. *Am. Nat.* **182**, 514–523 (2013). doi:10.1086/671908 Medline

109. Y. Liu, E. S. C. Scordato, Z. Zhang, M. Evans, R. J. Safran, Analysing phenotypic variation in barn swallows (*Hirundo rustica*) across China to assess subspecies status. *Biol. J. Linn. Soc.* **131**, 319–331 (2020). doi:10.1093/biolinnean/blaa112

110. M. Hasegawa, Sexual selection mechanisms for male plumage ornaments in Japanese Barn Swallows. *Ornithol. Sci.* **17**, 125–134 (2018). doi:10.2326/osj.17.125

111. A. J. Shultz, K. J. Burns, The role of sexual and natural selection in shaping patterns of sexual dichromatism in the largest family of songbirds (Aves: Thraupidae). *Evolution* **71**,

1061–1074 (2017). doi:10.1111/evo.13196 Medline

112. R. J. Safran, E. S. Scordato, L. B. Symes, R. L. Rodríguez, T. C. Mendelson, Contributions of natural and sexual selection to the evolution of premating reproductive isolation: A research agenda. *Trends Ecol. Evol.* **28**, 643–650 (2013). doi:10.1016/j.tree.2013.08.004 Medline

113. A. Turner, in *Handbook of the Birds of the World*, J. del Hoyo, A. Elliott, D. A. Christie, E. de Juana, Eds. (Lynx Edicions, 2018).

114. K. Delhey, A review of Gloger's rule, an ecogeographical rule of colour: Definitions, interpretations and evidence. *Biol. Rev.* **94**, 1294–1316 (2019). doi:10.1111/brv.12503 Medline

115. M. R. Wilkins, N. Seddon, R. J. Safran, Evolutionary divergence in acoustic signals: Causes and consequences. *Trends Ecol. Evol.* **28**, 156–166 (2013). doi:10.1016/j.tree.2012.10.002 Medline

116. M. R. Wilkins, D. Shizuka, M. B. Joseph, J. K. Hubbard, R. J. Safran, Multimodal signalling in the North American barn swallow: A phenotype network approach. *Proc. R. Soc. B.* **282**, 20151574 (2015). doi:10.1098/rspb.2015.1574 Medline

117. M. R. Wilkins, E. S. C. Scordato, G. A. Semenov, H. Karaardiç, D. Shizuka, A. Rubtsov, P. L. Pap, S.-F. Shen, R. J. Safran, Global song divergence in barn swallows (Hirundo rustica): Exploring the roles of genetic, geographical and climatic distance in sympatry and allopatry. *Biol. J. Linn. Soc* **123**, 825–849 (2018). doi:10.1093/biolinnean/bly012

118. A. Turner, C. Rose, *A Handbook to the Swallows and Martins of the World* (A&C Black, 2010).

119. E. B. Dopman, K. L. Shaw, M. R. Servedio, R. K. Butlin, C. M. Smadja, Coupling of barriers to gene exchange: Causes and consequences. *Cold Spring Harb. Perspect. Biol.* **16**, a041432 (2024). doi:10.1101/cshperspect.a041432 Medline

120. N. Bierne, J. Welch, E. Loire, F. Bonhomme, P. David, The coupling hypothesis: Why genome scans may fail to map local adaptation genes. *Mol. Ecol.* **20**, 2044–2072 (2011). doi:10.1111/j.1365-294X.2011.05080.x Medline

121. G. A. Van der Auwera, M. O. Carneiro, C. Hartl, R. Poplin, G. Del Angel, A. Levy-Moonshine, T. Jordan, K. Shakir, D. Roazen, J. Thibault, E. Banks, K. V. Garimella, D. Altshuler, S. Gabriel, M. A. DePristo, From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, bi1110s43 (2018). doi:10.1002/0471250953.bi1110s43 Medline

122. A. F. A. Smit, R. Hubley, P. Green, *RepeatMasker Open-4.0. 2013–2015* (Inst. Syst. Biol, 2015), http//repeatmasker. org.

123. W. C. Warren, D. F. Clayton, H. Ellegren, A. P. Arnold, L. W. Hillier, A. Künstner, S. Searle, S. White, A. J. Vilella, S. Fairley, A. Heger, L. Kong, C. P. Ponting, E. D. Jarvis, C. V. Mello, P. Minx, P. Lovell, T. A. Velho, M. Ferris, C. N. Balakrishnan, S. Sinha, C. Blatti, S. E. London, Y. Li, Y. C. Lin, J. George, J. Sweedler, B. Southey, P. Gunaratne, M. Watson, K. Nam, N. Backström, L. Smeds, B. Nabholz, Y. Itoh, O. Whitney, A. R. Pfenning, J. Howard, M. Völker, B. M. Skinner, D. K. Griffin, L. Ye, W. M. McLaren, P.

Flicek, V. Quesada, G. Velasco, C. Lopez-Otin, X. S. Puente, T. Olender, D. Lancet, A. F. Smit, R. Hubley, M. K. Konkel, J. A. Walker, M. A. Batzer, W. Gu, D. D. Pollock, L. Chen, Z. Cheng, E. E. Eichler, J. Stapley, J. Slate, R. Ekblom, T. Birkhead, T. Burke, D. Burt, C. Scharff, I. Adam, H. Richard, M. Sultan, A. Soldatov, H. Lehrach, S. V. Edwards, S. P. Yang, X. Li, T. Graves, L. Fulton, J. Nelson, A. Chinwalla, S. Hou, E. R. Mardis, R. K. Wilson, The genome of a songbird. *Nature* **464**, 757–762 (2010). doi:10.1038/nature08819 Medline

124. C. Jain, A. Dilthey, S. Koren, S. Aluru, A. M. Phillippy, "A fast approximate algorithm for mapping long reads to large reference databases" in *International Conference on Research in Computational Molecular Biology* (Springer, 2017), pp. 66–81.

125. C. Jain, S. Koren, A. Dilthey, A. M. Phillippy, S. Aluru, A fast adaptive algorithm for computing whole-genome homology maps. *Bioinformatics* **34**, i748–i756 (2018). doi:10.1093/bioinformatics/bty597 Medline

126. I. Pala, S. Naurin, M. Stervander, D. Hasselquist, S. Bensch, B. Hansson, Evidence of a neo-sex chromosome in birds. *Heredity* **108**, 264–272 (2012). doi:10.1038/hdy.2011.70 Medline

127. H. Sigeman, S. Ponnikas, B. Hansson, Whole-genome analysis across 10 songbird families within Sylvioidea reveals a novel autosome-sex chromosome fusion. *Biol. Lett.* **16**, 20200082 (2020). doi:10.1098/rsbl.2020.0082 Medline

128. S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, P. C. Sham, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007). doi:10.1086/519795 Medline

129. B. S. Weir, C. C. Cockerham, Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984). doi:10.2307/2408641 Medline

130. Q. Rougemont, P. A. Gagnaire, C. Perrier, C. Genthon, A. L. Besnard, S. Launey, G. Evanno, Inferring the demographic history underlying parallel genomic divergence among pairs of parasitic and nonparasitic lamprey ecotypes. *Mol. Ecol.* **26**, 142–162 (2017). doi:10.1111/mec.13664 Medline

131. S. Stankowski, M. A. Chase, H. McIntosh, M. A. Streisfeld, Integrating top-down and bottom-up approaches to understand the genetic architecture of speciation across a monkeyflower hybrid zone. *Mol. Ecol.* **32**, 2041–2054 (2023). doi:10.1111/mec.16849 Medline

132. J. D. Robinson, A. J. Coffman, M. J. Hickerson, R. N. Gutenkunst, Sampling strategies for frequency spectrum-based population genomic inference. *BMC Evol. Biol.* **14**, 254 (2014). doi:10.1186/s12862-014-0254-4 Medline

133. J. Catchen, P. A. Hohenlohe, S. Bassham, A. Amores, W. A. Cresko, Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140 (2013). doi:10.1111/mec.12354 Medline

134. C. Rougeux, L. Bernatchez, P.-A. Gagnaire, Modeling the multiple facets of speciation-with-gene-flow toward inferring the divergence history of lake whitefish species pairs (Coregonus clupeaformis). *Genome Biol. Evol.* **9**, 2057–2074 (2017).

doi:10.1093/gbe/evx150 Medline

135. R. Dor, R. J. Safran, F. H. Sheldon, D. W. Winkler, I. J. Lovette, Phylogeny of the genus Hirundo and the Barn Swallow subspecies complex. *Mol. Phylogenet. Evol.* **56**, 409–418 (2010). doi:10.1016/j.ympev.2010.02.008 Medline

136. A. Romano, N. Saino, A. P. Møller, Viability and expression of sexual ornaments in the barn swallow Hirundo rustica: A meta-analysis. *J. Evol. Biol.* **30**, 1929–1935 (2017). doi:10.1111/jeb.13151 Medline

137. A. P. Møller, H. Tegelström, Extra-pair paternity and tail ornamentation in the barn swallow Hirundo rustica. *Behav. Ecol. Sociobiol.* **41**, 353–360 (1997). doi:10.1007/s002650050395

138. B. Wehrle-Haller, The role of Kit-ligand in melanocyte development and epidermal homeostasis. *Pigment Cell Res.* **16**, 287–296 (2003). doi:10.1034/j.1600-0749.2003.00055.x Medline

139. G. A. Scott, L. A. McClelland, A. F. Fricke, Semaphorin 7a promotes spreading and dendricity in human melanocytes through β1-integrins. *J. Invest. Dermatol.* **128**, 151–161 (2008). doi:10.1038/sj.jid.5700974 Medline

140. X. Liu, Z. Wu, J. Li, H. Bao, C. Wu, Genome-Wide Association Study and Transcriptome Differential Expression Analysis of the Feather Rate in Shouguang Chickens. *Front. Genet.* **11**, 613078 (2020). doi:10.3389/fgene.2020.613078 Medline

141. J. M. Newton, O. Cohen-Barak, N. Hagiwara, J. M. Gardner, M. T. Davisson, R. A. King, M. H. Brilliant, Mutations in the human orthologue of the mouse underwhite gene (uw) underlie a new form of oculocutaneous albinism, OCA4. *Am. J. Hum. Genet.* **69**, 981–988 (2001). doi:10.1086/324340 Medline

142. T. Konno, Y. Abe, M. Kawaguchi, K. Storm, M. Biervliet, W. Courtens, M. Kono, Y. Tomita, T. Suzuki, Oculocutaneous albinism type IV: A boy of Moroccan descent with a novel mutation in SLC45A2. *Am. J. Med. Genet. A.* **149A**, 1773–1776 (2009). doi:10.1002/ajmg.a.32964 Medline

143. U. Gunnarsson, A. R. Hellström, M. Tixier-Boichard, F. Minvielle, B. Bed'hom, S. Ito, P. Jensen, A. Rattink, A. Vereijken, L. Andersson, Mutations in SLC45A2 cause plumage color variation in chicken and Japanese quail. *Genetics* **175**, 867–877 (2007). doi:10.1534/genetics.106.063107 Medline

144. S. Wu, J. Huang, Y. Li, L. Zhao, Z. Liu, Analysis of yellow mutant rainbow trout transcriptomes at different developmental stages reveals dynamic regulation of skin pigmentation genes. *Sci. Rep.* **12**, 256 (2022). doi:10.1038/s41598-021-04255-y Medline

145. P. L. H. de Mello, P. M. Hime, R. E. Glor, Transcriptomic analysis of skin color in anole lizards. *Genome Biol. Evol.* **13**, evab110 (2021). doi:10.1093/gbe/evab110 Medline

146. M. Visser, R.-J. Palstra, M. Kayser, Human skin color is influenced by an intergenic DNA polymorphism regulating transcription of the nearby BNC2 pigmentation gene. *Hum. Mol. Genet.* **23**, 5750–5762 (2014). doi:10.1093/hmg/ddu289 Medline

147. C. D. Van Raamsdonk, K. R. Fitch, H. Fuchs, M. H. de Angelis, G. S. Barsh, Effects of G-protein mutations on skin color. *Nat. Genet.* **36**, 961–968 (2004). doi:10.1038/ng1412

Medline

148. C. Hosaka, M. Kunisada, M. Koyanagi-Aoi, T. Masaki, C. Takemori, M. Taniguchi-Ikeda, T. Aoi, C. Nishigori, Induced pluripotent stem cell-derived melanocyte precursor cells undergoing differentiation into melanocytes. *Pigment Cell Melanoma Res.* **32**, 623–633 (2019). doi:10.1111/pcmr.12779 Medline

149. S. Singhal, E. M. Leffler, K. Sannareddy, I. Turner, O. Venn, D. M. Hooper, A. I. Strand, Q. Li, B. Raney, C. N. Balakrishnan, S. C. Griffith, G. McVean, M. Przeworski, Stable recombination hotspots in birds. *Science* **350**, 928–932 (2015). doi:10.1126/science.aad0843 Medline

150. F. Tajima, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989). doi:10.1093/genetics/123.3.585 Medline

151. P. C. Sabeti, D. E. Reich, J. M. Higgins, H. Z. P. Levine, D. J. Richter, S. F. Schaffner, S. B. Gabriel, J. V. Platko, N. J. Patterson, G. J. McDonald, H. C. Ackerman, S. J. Campbell, D. Altshuler, R. Cooper, D. Kwiatkowski, R. Ward, E. S. Lander, Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**, 832–837 (2002). doi:10.1038/nature01140 Medline

152. B. F. Voight, S. Kudaravalli, X. Wen, J. K. Pritchard, A map of recent positive selection in the human genome. *PLOS Biol.* **4**, e72 (2006). doi:10.1371/journal.pbio.0040072 Medline

153. P. C. Sabeti, P. Varilly, B. Fry, J. Lohmueller, E. Hostetter, C. Cotsapas, X. Xie, E. H. Byrne, S. A. McCarroll, R. Gaudet, S. F. Schaffner, E. S. Lander, K. A. Frazer, D. G. Ballinger, D. R. Cox, D. A. Hinds, L. L. Stuve, R. A. Gibbs, J. W. Belmont, A. Boudreau, P. Hardenbol, S. M. Leal, S. Pasternak, D. A. Wheeler, T. D. Willis, F. Yu, H. Yang, C. Zeng, Y. Gao, H. Hu, W. Hu, C. Li, W. Lin, S. Liu, H. Pan, X. Tang, J. Wang, W. Wang, J. Yu, B. Zhang, Q. Zhang, H. Zhao, H. Zhao, J. Zhou, S. B. Gabriel, R. Barry, B. Blumenstiel, A. Camargo, M. Defelice, M. Faggart, M. Goyette, S. Gupta, J. Moore, H. Nguyen, R. C. Onofrio, M. Parkin, J. Roy, E. Stahl, E. Winchester, L. Ziaugra, D. Altshuler, Y. Shen, Z. Yao, W. Huang, X. Chu, Y. He, L. Jin, Y. Liu, Y. Shen, W. Sun, H. Wang, Y. Wang, Y. Wang, X. Xiong, L. Xu, M. M. Waye, S. K. Tsui, H. Xue, J. T. Wong, L. M. Galver, J. B. Fan, K. Gunderson, S. S. Murray, A. R. Oliphant, M. S. Chee, A. Montpetit, F. Chagnon, V. Ferretti, M. Leboeuf, J. F. Olivier, M. S. Phillips, S. Roumy, C. Sallée, A. Verner, T. J. Hudson, P. Y. Kwok, D. Cai, D. C. Koboldt, R. D. Miller, L. Pawlikowska, P. Taillon-Miller, M. Xiao, L. C. Tsui, W. Mak, Y. Q. Song, P. K. Tam, Y. Nakamura, T. Kawaguchi, T. Kitamoto, T. Morizono, A. Nagashima, Y. Ohnishi, A. Sekine, T. Tanaka, T. Tsunoda, P. Deloukas, C. P. Bird, M. Delgado, E. T. Dermitzakis, R. Gwilliam, S. Hunt, J. Morrison, D. Powell, B. E. Stranger, P. Whittaker, D. R. Bentley, M. J. Daly, P. I. de Bakker, J. Barrett, Y. R. Chretien, J. Maller, S. McCarroll, N. Patterson, I. Pe'er, A. Price, S. Purcell, D. J. Richter, P. Sabeti, R. Saxena, S. F. Schaffner, P. C. Sham, P. Varilly, D. Altshuler, L. D. Stein, L. Krishnan, A. V. Smith, M. K. Tello-Ruiz, G. A. Thorisson, A. Chakravarti, P. E. Chen, D. J. Cutler, C. S. Kashuk, S. Lin, G. R. Abecasis, W. Guan, Y. Li, H. M. Munro, Z. S. Qin, D. J. Thomas, G. McVean, A. Auton, L. Bottolo, N. Cardin, S. Eyheramendy, C. Freeman, J. Marchini, S. Myers, C. Spencer, M. Stephens, P. Donnelly, L. R. Cardon, G. Clarke, D. M. Evans, A. P. Morris, B. S. Weir, T. Tsunoda, T. A. Johnson, J. C. Mullikin, S. T. Sherry, M.

Feolo, A. Skol, H. Zhang, C. Zeng, H. Zhao, I. Matsuda, Y. Fukushima, D. R. Macer, E. Suda, C. N. Rotimi, C. A. Adebamowo, I. Ajayi, T. Aniagwu, P. A. Marshall, C. Nkwodimmah, C. D. Royal, M. F. Leppert, M. Dixon, A. Peiffer, R. Qiu, A. Kent, K. Kato, N. Niikawa, I. F. Adewole, B. M. Knoppers, M. W. Foster, E. W. Clayton, J. Watkin, R. A. Gibbs, J. W. Belmont, D. Muzny, L. Nazareth, E. Sodergren, G. M. Weinstock, D. A. Wheeler, I. Yakub, S. B. Gabriel, R. C. Onofrio, D. J. Richter, L. Ziaugra, B. W. Birren, M. J. Daly, D. Altshuler, R. K. Wilson, L. L. Fulton, J. Rogers, J. Burton, N. P. Carter, C. M. Clee, M. Griffiths, M. C. Jones, K. McLay, R. W. Plumb, M. T. Ross, S. K. Sims, D. L. Willey, Z. Chen, H. Han, L. Kang, M. Godbout, J. C. Wallenburg, P. L'Archevêque, G. Bellemare, K. Saeki, H. Wang, D. An, H. Fu, Q. Li, Z. Wang, R. Wang, A. L. Holden, L. D. Brooks, J. E. McEwen, M. S. Guyer, V. O. Wang, J. L. Peterson, M. Shi, J. Spiegel, L. M. Sung, L. F. Zacharia, F. S. Collins, K. Kennedy, R. Jamieson, J. Stewart; International HapMap Consortium, Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913–918 (2007). doi:10.1038/nature06250 Medline

154. O. Delaneau, B. Howie, A. J. Cox, J.-F. Zagury, J. Marchini, Haplotype estimation using sequencing reads. *Am. J. Hum. Genet.* **93**, 687–696 (2013). doi:10.1016/j.ajhg.2013.09.002 Medline

155. D. E. Irwin, B. Milá, D. P. L. Toews, A. Brelsford, H. L. Kenyon, A. N. Porter, C. Grossen, K. E. Delmore, M. Alcaide, J. H. Irwin, A comparison of genomic islands of differentiation across three young avian species pairs. *Mol. Ecol.* **27**, 4839–4855 (2018). doi:10.1111/mec.14858 Medline