Molecular Biology and Evolution, 2025, **42**, 1–15
https://doi.org/10.1093/molbev/msaf044
Advance access publication 24 February 2025
Discoveries

MBE OXFORD

Repeatable Selection on Large Ancestry Blocks in an Avian Hybrid Zone

Stephanie A. Blain [1,2,*] **Hannah C. Justen** [1] **Quinn K. Langdon** [3,4,5] **Kira E. Delmore** [1,2]

[1]Biology Department, Texas A&M University, College Station, TX, USA
[2]Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY, USA
[3]Department of Biology, Stanford University, Stanford, CA, USA
[4]Centro de Investigaciones Científicas de las Huastecas "Aguazarca", A.C., Calnali, Hidalgo, Mexico
[5]Present address: Gladstone Institute of Virology, Gladstone Institutes, San Francisco, CA, USA

***Corresponding author:** E-mail: sab2392@columbia.edu.
Associate editor: John True

Abstract

Hybrid zones create natural tests of genetic incompatibilities by combining loci from 2 species in the same genetic background in the wild, making them useful for identifying loci involved in both intrinsic and ecological (extrinsic) isolation. Two Swainson's thrush subspecies form a hybrid zone in western North America. These coastal and inland subspecies exhibit dramatic differences in migration routes; their hybrids exhibit poor migratory survival, suggesting that ecological incompatibilities maintain this zone. We used a panel of ancestry informative markers to identify repeated patterns of selection and introgression across 4 hybrid populations that span the entire length of the Swainson's thrush hybrid zone. Two repeatable patterns consistent with selection against incompatibilities—steep genomic clines and few transitions between ancestry states—were found in large genetic blocks on chromosomes 1 and 5. The block on chromosome 1 showed evidence for inland subspecies introgression while the block on chromosome 5 exhibited coastal subspecies introgression. Some regions previously associated with migratory phenotypes, including migratory orientation, or exhibiting misexpression between the subspecies exhibited signatures of selection in the hybrid zone. Both selection and introgression across the genome were shaped by genomic structural features and evolutionary history, with stronger selection and reduced introgression in regions of low recombination, high subspecies differentiation, positive selection within the subspecies, and on macrochromosomes. Cumulatively, these results suggest that linkage among loci interacts with divergent selection and past divergent evolution between species to strengthen barriers to gene flow within hybrid zones.

Keywords: speciation, hybridization, migration, Swainson's thrush

Introduction

Understanding the genetic basis of speciation, and how individual loci contribute to or hinder this process, has been a long-standing goal in evolutionary biology (Orr et al. 2004; Noor and Feder 2006). Hybrid zones, geographic areas where species with incomplete reproductive isolation have overlapping ranges, create a natural test for potential incompatibilities between diverging species by combining loci inherited from both species in the same genome. Some loci will introgress freely while others will be restricted by selection against hybrids and/or in the other species range (Barton and Hewitt 1989). These natural experiments could be particularly valuable for identifying loci involved in ecology-based reproductive isolation (Schluter 2009), which is poorly understood relative to intrinsic mechanisms of isolation (Schluter and Rieseberg 2022; Thompson et al. 2024). Loci that underlie ecology-based postzygotic isolation evolve via divergent ecological adaptation between incipient species and are subject to selection in hybrid zones alongside selection acting on non-ecological incompatibilities (Gavrilets 1997; Barton 2001).

Within hybrid genomes, selected and introgressed loci exhibit distinct patterns of variation that are detectable in natural populations with genome-scale data (Moran et al. 2021). First, loci involved in incompatibilities are expected to show steep genomic clines, meaning a reduction in heterozygosity and a sharp transition from 1 species ancestry to the other relative to genome-wide ancestry (Gompert and Buerkle 2011, 2012). If a locus is introgressed within the hybrid zone then the genomic cline will be asymmetric, with ancestry disproportionately likely to be from 1 species. Second, incompatibilities are expected to be concentrated within regions of the genome with longer contiguous tracts of ancestry, as the reduction in heterozygosity near incompatibilities will lead to fewer opportunities for recombination between different ancestries (Hvala et al. 2018). These patterns of selection and introgression are expected to interact differently with genomic structure and the evolutionary history of genomic regions. For example, features such as low recombination rates and high gene density increase the likelihood that a locus may be genetically linked to another that is under selection, potentially leading to stronger patterns of selection and reduced introgression in these regions (Schumer et al. 2018; Martin et al. 2019). A history of positive selection within one of the hybridizing species may lead to elevated introgression in those regions if the beneficial allele conveys an advantage in the hybrid zone

Received: September 9, 2024. **Revised:** December 29, 2024. **Accepted:** January 27, 2025

(Nouhaud et al. 2022) or reduced introgression if the selected allele contributes to reproductive isolation.

Patterns of selection and introgression can arise due to mechanisms that do contribute to the speciation process, including epistatic incompatibilities, selection against heterozygosity, or selection favoring alleles of 1 species that are maladaptive in hybrids or the hybrid zone environment. However, one challenge in linking loci to reproductive isolation is that patterns consistent with selection and introgression can emerge due to unrelated processes, such as genetic drift, demography, or population structure. Replicate hybrid populations can help overcome this problem, as loci exhibiting parallel signatures of selection across replicates provide stronger evidence that they contribute to speciation (Chaturvedi et al. 2020; Westram et al. 2021). Some factors affecting the strength of selection against alleles involved in incompatibilities will likely vary among populations, including the landscape of ecological selection and the relative contribution of each species to hybrid genomes (Bolnick et al. 2018; Langdon et al. 2022). Nevertheless, in many cases, if alleles of diverging species perform poorly together in a genome, this will be the case across the full extent of the hybrid zone.

We tested for repeatability of genomic signatures of selection and introgression in a narrow hybrid zone between 2 subspecies (coastal "*Catharus ustulatus ustulatus*" and inland "*C. u. swainsoni*") of Swainson's thrush in western North America (Ruegg 2008; Delmore and Irwin 2014). This hybrid zone is centered over the Coast Mountains, with the coastal subspecies breeding in temperate rainforest while the inland subspecies breeding sites are primarily in boreal forest. Comparing the 2 breeding niches, the coastal subspecies niche has higher precipitation, lower seasonality in temperature and precipitation, and denser vegetation (Ruegg et al. 2006). The ecological divergence between these subspecies, which form a migratory divide and well-developed genomic resources make this hybrid zone ideal for investigating speciation genomics in nature. Specifically, these subspecies differ in a series of color, morphological, and behavioral traits (Ruegg et al. 2012; Delmore and Irwin 2014; Delmore et al. 2016). Differences in their migratory routes are the most prominent, as the coastal subspecies migrates along the Pacific coast to Central America while the inland subspecies follows a much longer route to Colombia via inland North America (Ruegg and Smith 2002). This leads to migratory niche differences between the subspecies, with the coastal niche characterized by dense vegetation and environmental wetness while the inland route is limited to a restricted temperature range (Justen et al. 2021). Earlier work showed that hybrids take intermediate and ecologically inferior routes, reducing their survival (Blain et al. 2024). Genomic loci underlying variation in migratory traits that distinguish coastal and inland thrushes have been identified (Delmore et al. 2016; Justen et al. 2024) and analyses of gene expression recently uncovered genes exhibiting transgressive (misexpressed) patterns in hybrids (Louder et al. 2024). Misexpressed genes may affect hybrid fitness (Turner et al. 2014; Satokangas et al. 2020). Combined, these phenotype-associated loci represent candidate ecological incompatibilities in Swainson's thrushes, but tests for selection at these loci have yet to be conducted.

We characterized patterns of selection and introgression across 4 hybrid populations that span the length of the hybrid zone between Swainson's thrushes, from Alaska to Washington (Fig. 1). We used genomic clines and ancestry tracts to identify loci under selection (i.e. those exhibiting steep clines and/or located within long ancestry tracts). Genomic clines and tests for gene flow between parapatric coastal and inland populations were used to identify introgressed loci (i.e. those exhibiting asymmetrical cline centers and/or excess gene trees grouping parapatric coastal and inland populations). Once these loci were identified in each population, we looked for those exhibiting parallel patterns across hybrid populations and tested for enrichment of outlier loci previously linked to markers associated with divergent phenotypes or misexpressed genes. We then tested for associations between genomic structure and both selection and introgression, with the expectation that stronger selection and reduced introgression would be observed in regions of the genome where linkage to a locus contributing to reproductive isolation is more likely. We also tested for a role of coastal and inland evolutionary history in shaping hybrid genomic patterns, by evaluating whether regions under positive selection in the coastal and inland subspecies or showing high divergence between the subspecies were under associated with greater selection or changes in introgression. Finally, if traits are under similar forms of selection across populations, they might also exhibit repeatability in relationships to genome-wide ancestry. Therefore, we quantified the strength of relationships between ancestry and morphological, behavioral, and color traits across populations, then compared repeatability in these relationships to repeatability in genomic cline estimates.

Results

Ancestry Informative Markers

Using a reference panel of allopatric inland and coastal Swainson's thrushes, we identified a set of loci that were fixed or nearly fixed for alternate alleles in coastal and inland thrushes (ancestry informative markers, AIMs). We then called ancestry at these AIMs in 299 hybrids from 4 populations that span the hybrid zone between Swainson's thrushes: Alaska (n = 29), Pemberton (n = 122), Hope (n = 106), and Washington (n = 42) (Fig. 1; supplementary table S1, Supplementary Material online). Following linkage pruning, we retained a set of 1,488 autosomal and 699 Z chromosome AIMs. Each of the 4 populations contained a mix of early and late generation hybrids, indicated by a mix of individuals with long tracts of heterozygous ancestry and individuals with frequent transitions between ancestry states (Fig. 1). This pattern is consistent with ongoing hybridization that did not start recently.

Steep Genomic Clines on Chromosomes 1 and 5

We fitted genomic clines, which model relationships between genome-wide ancestry and ancestry at a specific locus, to AIMs separately in each hybrid population. We were interested in loci with (i) cline gradients that were high relative to 1 (the expected slope relative to the admixture gradient), which indicate a fast transition from 1 subspecies ancestry to the other, and therefore selection against heterozygotes, and (ii) asymmetric cline centers, which indicate introgression from 1 subspecies. On average, cline gradients were higher than the null expectation of 1 (mean = 1.22, 95% CI = [1.13, 1.30]) and were shifted slightly toward inland ancestry (0.502 [0.500, 0.504]). Within each population, variability in cline gradients estimated as the standard deviation (σ_v), was low in all populations (Alaska: 0.17, Pemberton: 0.18, Hope: 0.10, and Washington: 0.13), possibly reflecting coupling among genomic clines (Firneno et al. 2023), as was

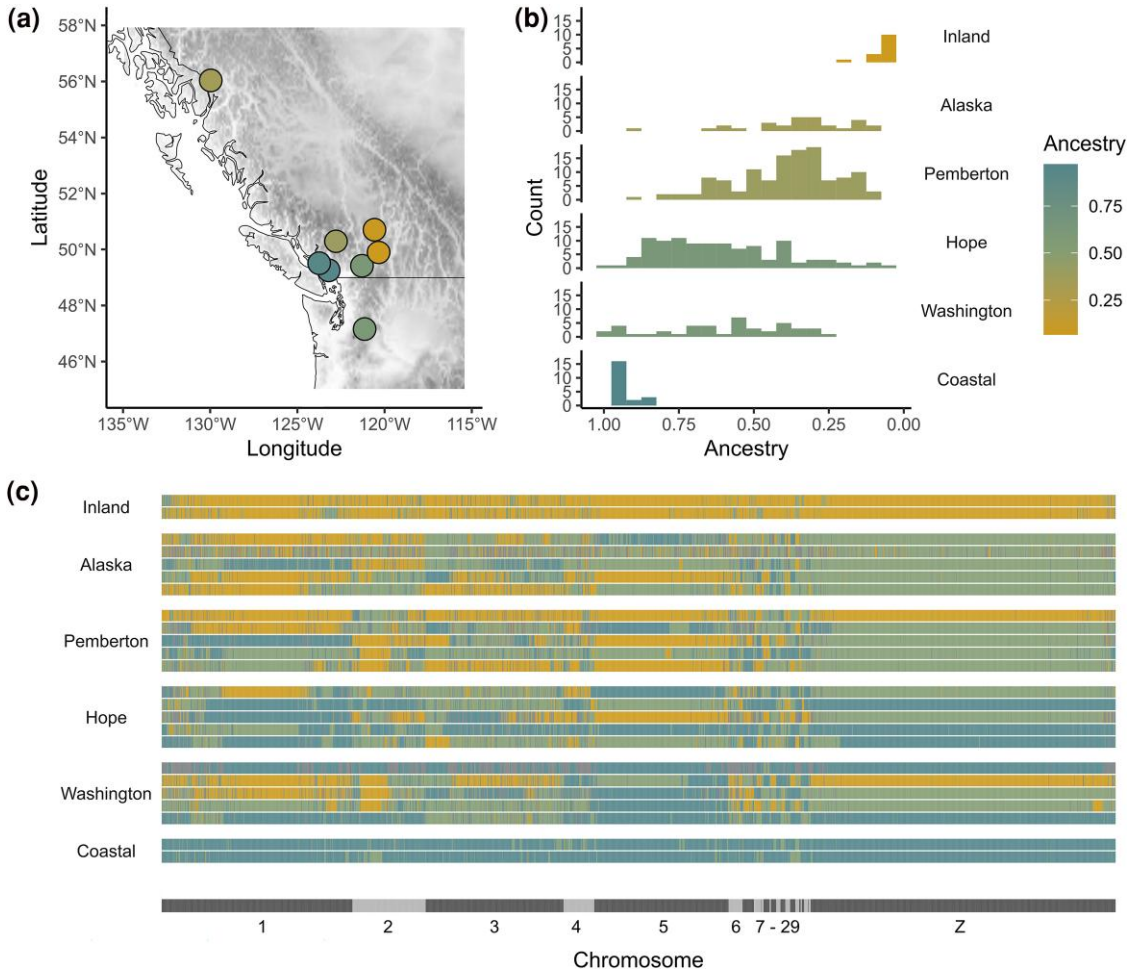

Fig. 1. Genomic variation across sampling sites. a) Map showing sampling locations of 4 hybrid populations, 2 coastal populations, and 2 inland populations. From north to south, the 4 populations are from southeastern Alaska; Pemberton, BC; Hope, BC; and northwestern Washington. Color indicates mean genome-wide ancestry in each population. Gray shading on the map indicates elevation. b) Distributions of genome-wide ancestry across inland, coastal, and hybrid populations. Both inland populations and both coastal populations were pooled. c) Ancestry along a set of 2,187 AIMs in random individuals from inland, coastal, and hybrid populations. Color indicates ancestry: inland homozygotes are yellow, heterozygotes are green, and coastal homozygotes are blue. Each bar represents 1 individual. Alternating shades of gray along the *x* axis indicate the chromosome on which AIMs are located.

variability of cline centers within populations (σ_w; Alaska: 0.03, Pemberton: 0.12, Hope: 0.15, and Washington: 0.09). At first glance, 2 large regions of the genome showed repeatable and distinct clinal patterns. Specifically, 2 contiguous sets of AIMs—one on chromosome 1 and the other on chromosome 5—showed elevated cline gradients across all 4 populations (Fig. 2). Similar consistency was documented for cline centers along the contiguous sets of AIMs; on average, the chromosome 1 clines were shifted toward inland ancestry while the chromosome 5 clines were shifted toward coastal ancestry (Fig. 2). There were 283 genes within the LD blocks around AIMs that were cline gradient outliers in 2 or more populations. A gene ontology (GO) analysis of these genes showed enrichment for involvement in xenobiotic metabolic process, with 3 genes associated with this ontology (AHR, LOC116996430, LOC116996432).

Few Ancestry Transitions in Regions on Chromosomes 1 and 5

Ancestry tracts, or contiguous regions with the same ancestry state, can be indicative of selection within hybrid populations. Incompatible loci are expected to occur in long ancestry tracts, near relatively few ancestry transitions. As in genomic cline analyses, 2 regions of the genome—one on chromosome 1 and one on chromosome 5—showed qualitatively distinctive patterns, with very few ancestry transitions in all populations (Fig. 3a). This was reflected in estimates of mean ancestry and heterozygosity of AIMs on macrochromosomes, as chromosomes 1 and 5 showed a deficit of individuals with intermediate ancestry and heterozygosity relative to other chromosomes (Fig. 3c).

To quantify the expected number of transitions within 100 kb windows, we fit asymptotic models to estimate the relationship between the number of AIMs and number of transitions within windows in each population. We identified outliers as windows with fewer transitions than the lower bound of simulated 95% confidence intervals. 2,000 out of 4,118 windows included in the analysis were identified as outliers in one or more populations, with most located on chromosomes 1, 2, and 5. 2,820 genes fell within windows that were identified as outliers in 2 or more populations. A GO analysis showed enrichment for UDP-glycosyltransferase activity (22 genes), the nucleosome (43 genes), protein heterodimerization activity (37 genes), sensory perception of taste (12 genes), extracellular space (28 genes), regulation of cytokine production (6 genes), and chemokine receptor activity (9 genes).

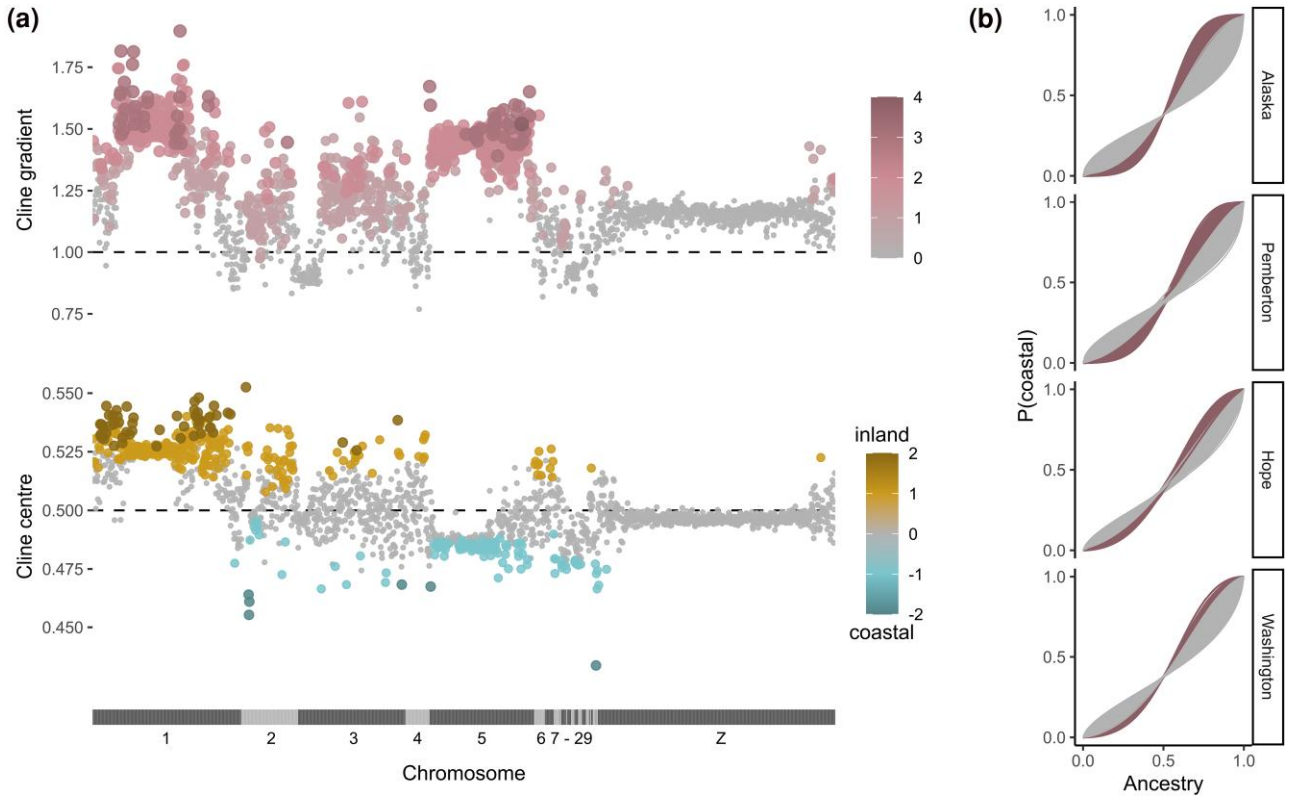

Fig. 2. Genomic cline gradients and centers. a) Mean cline parameters estimated for AIMs along the genome. Each point represents the mean parameter value for an AIM across 4 hybrid populations. Shading indicates the number of populations in which a particular AIM was an outlier as a fast cline gradient (pink), inland-shifted center (yellow), or coastal-shifted center (blue), with darker shading indicating than an AIM was an outlier in more populations. Alternating shades of gray along the *x* axis indicate the chromosome on which AIMs are located. b) Genomic cline shapes across populations. Each line represents the cline for 1 AIM, with color indicating outliers with elevated cline gradients (pink). The *x* axis represents genome-wide ancestry while the *y* axis represents the probability that a locus has a coastal genotype. Estimates of mean cline gradient were: 1.31 (Alaska), 1.24 (Pemberton), 1.12 (Hope), and 1.19 (Washington), while estimates of gradient variability (σ_v) were: 0.17 (Alaska), 0.18 (Pemberton), 0.10 (Hope), 0.13 (Washington). Estimates of mean cline center were 0.50 in all populations, while estimates of cline center variability (σ_c) were 0.03 (Alaska), 0.12 (Pemberton), 0.15 (Hope), and 0.09 (Washington).

Loci involved in speciation might have evolved through divergent evolution as part of the speciation process, or they may be a result of variation that was segregating prior to species divergence. We identified regions with very few ancestry transitions in the reference panel on chromosome 1 (46 to 59.5 Mb) and chromosome 5 (21 to 39 Mb). To determine when the coastal and inland haplotypes in these regions evolved, we built maximum likelihood phylogenies for each of the 2 regions using sequences from both subspecies as well as a set of other species in the *Catharus* genus. The coastal and inland subspecies were sister groups in both phylogenies, indicating that the 2 haplotypes evolved relatively recently (Fig. 3b).

Introgression From Coastal to Inland Subspecies Near the Hybrid Zone

We looked for introgression between parapatric coastal and inland populations using D-statistics, which compare the number of gene trees that group sympatric populations of 2 species to gene trees that group the sympatric population of 1 species with an allopatric population of the second species. An excess number of gene trees grouping the sympatric populations indicate that hybridization occurred (Fig. 4). Looking for introgression between coastal populations and parapatric inland populations, we found evidence for substantial introgression ($D = 0.022$, $Z = 11.1$, $P < 0.001$). We only looked at introgression using an allopatric inland population for comparison because the coastal breeding range is quite narrow and entirely adjacent to the hybrid zone, so there is no allopatric population available to anchor the analysis. We then evaluated patterns of introgression in sliding windows across the genome with f_d statistics, finding several regions of the genome with f_d values well above zero, indicating widespread introgression (Fig. 4). Some regions, including on chromosomes 1, 3, and 5, tended to show reduced introgression with f_d values consistently closer to zero.

Signatures of Selection Repeated Across Populations

We tested for repeatability across populations for outliers identified in analyses of cline gradients, cline centers, and ancestry transitions. We compared the observed number of outlier loci overlapping between populations to the number expected by chance, which we estimated as: # *population 1 outliers* × # *population 2 outliers* / # *AIMs*. The number of overlapping loci among trios was estimated as: *expected pairwise overlap* × # *population 3 outliers* / # *AIMs*, while the number of overlapping outliers among all 4 populations was estimated as: *expected trio overlap* × # *population 4 outliers* / # *AIMs*.

We tested for repeatability in genomic clines across hybrid populations, starting by looking at cline gradients. Of 2,187 AIMs, we found 504 that were outliers with elevated cline gradients in 2 or more populations. Starting with pairs of populations, we compared observed overlap in outlier loci to the number expected by chance and found greater overlap in

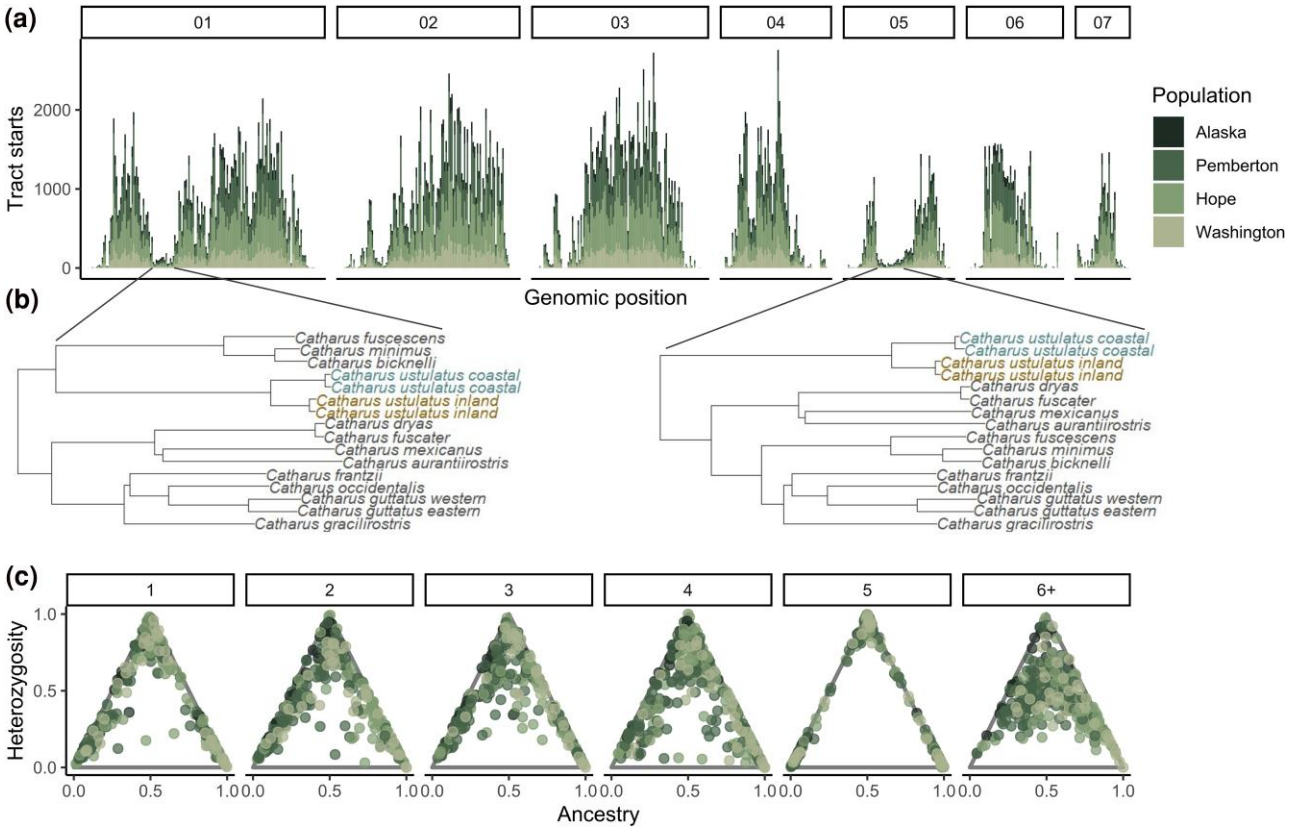

Fig. 3. Ancestry transitions across autosomal macrochromosomes. a) Histograms show the counts of ancestry tract start sites in 1 Mb bins across the 7 largest autosomes, with color indicating population. b) Phylogenies were constructed with sequences from regions with few ancestry start sites on chromosome 1 (left) and chromosome 5 (right). c) Mean ancestry and heterozygosity, measured as the proportion of heterozygous loci, of AIMs within macrochromosomes.

outlier loci than expected by chance by 2.2- to 3.0-fold in all population pairs except Alaska and Hope, which had fewer overlapping outliers than expected (Table 1). Results for overlap among trios of populations were similar, with overlap among trios that did not include both Alaska and Hope 8-fold higher than expected by chance, while overlap among trios that did include both Alaska and Hope was similar to random expectations (Table 1). Of the 66 AIMs that were shared outliers among 3 or more populations, 29 were located on chromosome 1, 36 were located on chromosome 5, and 1 was located on chromosome 2 (Fig. 2). One AIM, located on chromosome 5, was an outlier in all 4 populations. Outliers in cline centers were only observed in 2 populations—Pemberton and Hope. Of the cline center outliers, 61 overlapped between the 2 populations, compared to 21.8 expected by chance ($P < 0.001$). These were primarily clustered on chromosome 1, with other repeated cline center outliers on chromosomes 2, 3, and 5 (Fig. 2).

We additionally tested for repeatability across populations in windows identified as having few ancestry transitions. Pairwise overlap was between 2.1- and 2.6-fold higher than expected in all comparisons, with between 988 and 1,090 outliers overlapping between pairs (Table 2). Among trios of populations, between 856 and 932 outliers were shared between populations, between 5.5- and 6.8-fold the number expected by chance. 808 windows were outliers in all 4 populations, relative to 48 expected by chance. Most of these were located on chromosomes 1, 2, and 5, but repeated outliers were observed across all 7 macrochromosomes (Fig. 3).

Phenotype-associated Loci Show Patterns of Selection

Recent admixture mapping in the Swainson's thrush identified loci connected to variation in their migratory behavior and morphology (Justen et al. 2024). Twenty-nine of these loci fell within linkage blocks around our AIMs. Of these, 7 loci were cline gradient outliers in 2 or more populations; this is not more than we would expect by chance (29 AIMs associated with phenotypes × 504 outlier AIMs/2187 AIMs = 6.7 loci; $P = 0.20$) but the loci themselves may still be of interest. Two of these loci (located on chromosomes 1 and 3), were associated with the longitude of birds" nonbreeding sites, 3 (located on chromosome 1) were associated with wing pointedness, and 2 (located on chromosome 5) was associated with orientation on fall migration (supplementary fig. S2, Supplementary Material online). A transcriptomic analysis in the Swainson's thrush also identified genes misexpressed in hybrids (Louder et al. 2024). AIMs from the present study occurred within 61 of these genes and 10 were cline gradient outliers in 2 or more populations, which is 1.6-fold higher than we would expect by chance (27 AIMs linked to misexpressed genes × 504 outlier AIMs/2187 AIMs = 6.2 loci, $P < 0.001$). One was located on each of chromosomes 1 and 3, and 8 on chromosome 5 in genes LOC116996633, LOC116996682, KLHL2, SH3RF1, SH3RF1, GPM6A, CASP3, CLOCK, USP46, UBE2E1, and CCNC (supplementary fig. S2, Supplementary Material online).

For ancestry transitions, phenotype-associated SNPs fell within 35 outlier windows shared across all 4 populations,

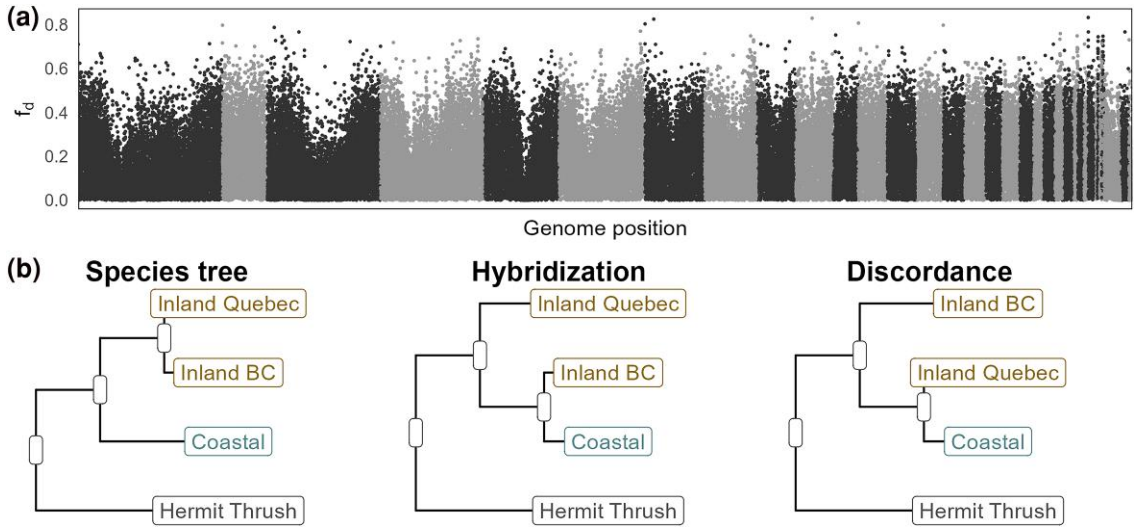

Fig. 4. Coastal to inland introgression and genomic features. a) Distribution of f_d statistics in 100 SNP windows across the genome. Each point represents 1 window, with color indicating chromosome. b) Cladogram of potential gene trees among populations included in the analyses. Windows with elevated f_d contain a greater proportion of gene trees with the "hybridization" topology relative to the "discordance" topology.

Table 1 Observed relative to expected overlap between pairs of hybrid populations in loci that are cline gradient outliers

Populations	Observed	Expected	Ratio	*P*-value
Alaska × Hope	29	58.42	0.5	1
Alaska × Pemberton	245	109.42	2.24	<0.001
Alaska × Washington	22	7.26	3.03	<0.001
Hope × Pemberton	281	122.37	2.3	<0.001
Hope × Washington	22	8.12	2.71	<0.001
Pemberton × Washington	40	15.22	2.63	<0.001
Alaska × Hope × Pemberton	27	18.91	1.43	<0.001
Alaska × Hope × Washington	1	1.26	0.8	0.5062
Alaska × Pemberton × Washington	19	2.35	8.08	<0.001
Hope × Pemberton × Washington	22	2.63	8.37	<0.001
Alaska × Pemberton × Hope × Washington	1	0.41	2.46	<0.001

Table 2 Observed relative to expected overlap between pairs of hybrid populations in 1 Mb windows that are ancestry transition outliers

Populations	Observed	Expected	Ratio	*P*-value
Alaska × Hope	1,011	474.25	2.13	<0.001
Alaska × Pemberton	988	428.22	2.31	<0.001
Alaska × Washington	1,004	485.06	2.07	<0.001
Hope × Pemberton	1,054	405.56	2.60	<0.001
Hope × Washington	1,090	459.39	2.37	<0.001
Pemberton × Washington	1,015	414.80	2.45	<0.001
Alaska × Hope × Pemberton	883	141.42	6.24	<0.001
Alaska × Hope × Washington	884	160.19	5.52	<0.001
Alaska × Pemberton × Washington	856	144.65	5.92	<0.001
Hope × Pemberton × Washington	932	136.99	6.80	<0.001
Alaska × Pemberton × Hope × Washington	808	47.77	16.91	<0.001

which is not more than expected by chance (244 windows with phenotype-associated SNPs × 808 outliers/4118 windows = 48; $P = 0.98$). These windows included SNPs on chromosomes 1, 2, and 3 that were identified in a multitrait GWAS as well as SNPs associated with orientation midway through migration (located on chromosome 1 and 5), orientation at the start of migration (located on chromosome 5), fall migration departure date (located on chromosome 3), fall migration distance (located on chromosomes 2, 3, and 5), wing pointedness (located on chromosome 1), longitude midway through spring migration (located on chromosomes 2, 3, and 7), spring migration distance (located on chromosomes 3, 4, and 6), wing length (located on chromosome 1), and nonbreeding site longitude (located on chromosomes 1, 2, 3, and 4; supplementary fig. S3, Supplementary Material online). 118 genes found to be misexpressed in hybrids, relative to parental subspecies, were located within 108 outlier windows shared across all 4 populations, 1.2-fold the number expected by chance (458 windows with misexpressed genes × 808 outliers/4118 windows = 90 windows; $P < 0.001$). These outlier windows were

distributed across all 7 macrochromosomes, with most located on chromosomes 1, 2, 3, and 5 (supplementary fig. S3, Supplementary Material online).

Patterns of Selection and Introgression are Driven by Genome Structure

Structural genomic features that increase the likelihood of linkage to loci under selection should lead to stronger patterns of selection. Therefore, we expected to observe higher cline gradients and fewer ancestry transitions in regions with low recombination and high gene density. Avian genomes are primarily comprised of a few macrochromosomes, defined as those above 40 Mb in length, several microchromosomes, which are below 20 Mb in length, and Z/W sex chromosomes (Ellegren 2013), so we also tested if selection and introgression varied among chromosome types. Cline gradients declined with recombination rate (slope = −0.012, 95% CI = [−0.014, −0.009]; Fig. 5), as expected, but also declined with gene density (−0.015 [−0.018, −0.011]; supplementary fig. S4, Supplementary Material online), contrary to expectations. Cline gradients were higher on macrochromosomes (mean = 1.27, 95% CI = [1.19, 1.35]) and on the Z chromosome (1.16 [1.07, 1.24]) than on microchromosomes (0.98 [0.90, 1.07]; supplementary fig. S4, Supplementary Material online) and showed a positive relationship to chromosome length (0.0015 [0.001, 0.002]).

In line with our expectations, we found that the number of ancestry transitions in 100 kb windows increased with recombination rate (0.092 [0.090, 0.094]; Fig. 5), and declined with gene density (−0.010 [−0.012, −0.007; supplementary fig. S5, Supplementary Material online). A higher density of AIMs increases the frequency of ancestry transitions (supplementary fig. S1, Supplementary Material online), which could provide an alternate interpretation for this pattern. However, windows with high recombination rates tended to contain fewer AIMs (−37.2 [−37.9, −36.5]), indicating that the relationship between recombination rate and the number of ancestry

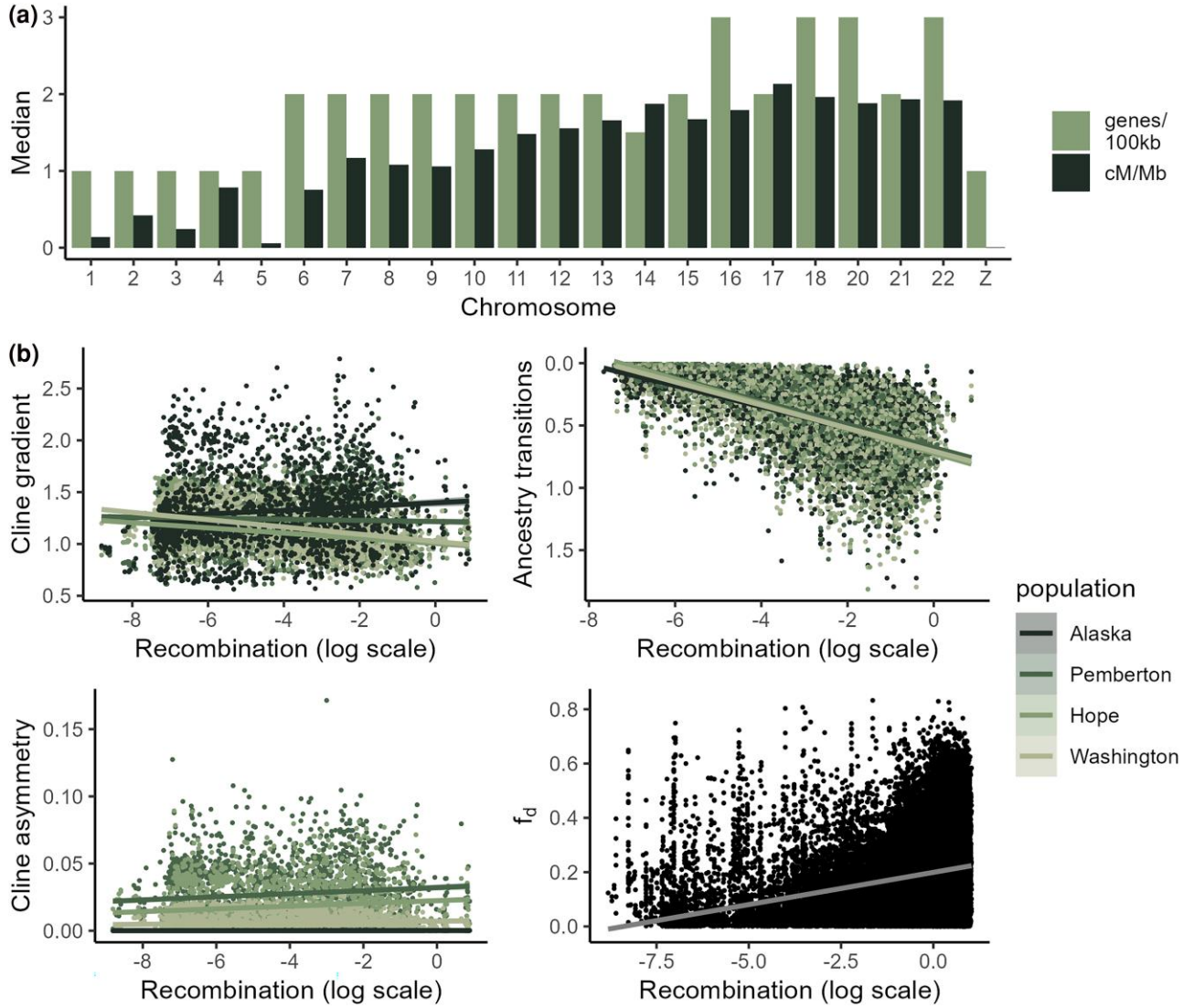

Fig. 5. Genome structural features, selection, and introgression. a) Chromosomal variation in median recombination and gene density. Bars represent the median value of gene density and recombination rate in 100 kb windows on each chromosome longer than 10 Mb. b) Relationship between recombination rate and cline gradients (top left), ancestry transitions, estimated as transitions per individual per 100 kb window (top right), cline asymmetry, estimated as the absolute value of the difference between 0.5 and the cline center (bottom left), and coastal to inland introgression, estimated as f_d in 100 SNP windows (bottom right).

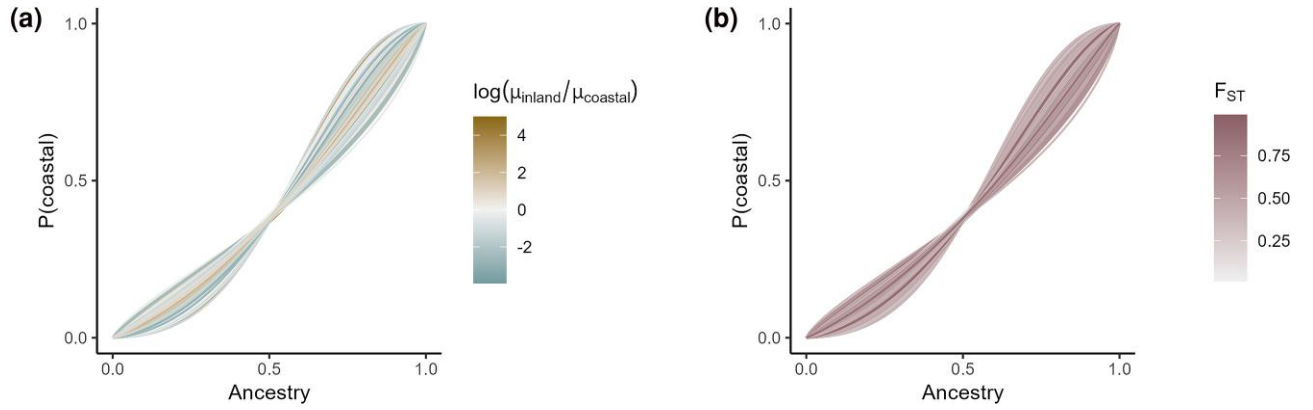

Fig. 6. Genomic clines and their relationships to coastal and inland evolutionary history. Each line represents the cline for 1 AIM, fit using the mean center and gradient parameters across the 4 populations. Color indicates estimates for the 100 kb windows containing that AIM for a) positive selection, estimated as $\log \mu_{inland}/\mu_{coastal}$, and b) genetic divergence, estimated as F_{ST}.

transitions were not due to a higher frequency of AIMs. Because windows with higher gene densities contained fewer AIMs (−0.76 [−1.8, 0.35]), it is possible that the negative relationship between ancestry transitions and gene density is driven to some extent by the density of AIMs.

If introgression is generally under purifying selection, there should be reduced introgression in regions of the genome where loci are more likely to be in linkage with alleles under selection. Therefore, we expected both introgression within the hybrid zone, indicated by asymmetrical cline centers, and introgression between parapatric subspecies populations, indicated by elevated f_d, to exhibit a positive relationship to recombination rate and a negative relationship to gene density. As predicted, asymmetry in cline centers (absolute value of difference between observed cline center and null cline center) increased with recombination rate (6.1×10^{-4} [4.8×10^{-4}, 7.5×10^{-4}]; Fig. 5). Cline center asymmetry was not related to gene density (6.6×10^{-5} [-1.1×10^{-4}, 2.4×10^{-4}]; supplementary fig. S6, Supplementary Material online) and was similar among macrochromosomes (0.016 [0.0045, 0.028]), microchromosomes (0.015 [0.0037, 0.027]), and the Z chromosome (0.0049 [−0.0068, 0.017]; supplementary fig. S6, Supplementary Material online). However, cline asymmetry did increase with chromosome length (1.1×10^{-4} [1.1×10^{-4}, 1.2×10^{-4}]).

Looking at patterns in f_d, we found some support for our predictions, as rates of coastal to inland introgression increased with recombination rate (0.024 [0.023, 0.024]; Fig. 5), but increased with gene density (slope = 0.0037 [0.0033, 0.0041]; supplementary fig. S7, Supplementary Material online). Coastal to inland introgression was highest on microchromosomes (0.23 [0.22, 0.23]), lowest on the Z chromosome (0.17 [0.16, 0.17]), comparatively intermediate on macrochromosomes (0.19 [0.18, 0.19]; supplementary fig. S7, Supplementary Material online), and showed a negative relationship to scaffold length (-3.3×10^{-4} [-3.5×10^{-4}, -3.1×10^{-4}]).

Coastal and Inland Evolutionary History Determines Hybrid Genomic Patterns

If alleles under positive selection in the coastal or inland subspecies also convey an advantage within the hybrid zone, we would expect linked regions to show elevated signatures of selection and introgression. Alternatively, there might be positive selection for alleles that reinforce species boundaries, which should lead to elevated selection and reduced introgression within the hybrid zone. Both metrics of selection (cline gradients and ancestry transitions) and 1 measure of introgression (f_d) do not have a direction associated with coastal versus inland differences, so for these variables we estimated positive selection as $|\log \mu_{inland}/\mu_{coastal}|$, with μ representing estimates of positive selection in each subspecies (Alachiotis and Pavlidis 2018). In contrast, higher cline centers indicate more inland introgression while lower cline centers indicate more coastal introgression, so we used $\log \mu_{inland}/\mu_{coastal}$ as the positive selection metric for cline centers.

Windows with positive selection in the coastal or inland subspecies had higher cline gradients (0.025 [0.019, 0.031]; Fig. 6; supplementary fig. S8, Supplementary Material online) and fewer ancestry transitions (−0.090 [−0.096, −0.083]; supplementary fig. S9, Supplementary Material online). Cline center estimates decreased with $\log \mu_{inland}/\mu_{coastal}$, indicating reduced inland introgression in windows with positive selection in the inland subspecies and reduced coastal introgression in windows with positive selection in the coastal subspecies (−0.0018 [−0.0021, −0.0014]; Fig. 6; supplementary fig. S10, Supplementary Material online). Additionally, f_d decreased with positive selection in the coastal or inland subspecies (−0.047 [−0.049, −0.044]; supplementary fig. S11, Supplementary Material online).

As a last way to examine the relationship between genomic patterns and the evolutionary history of Swainson's thrush hybrids, we looked for relationships of F_{ST} to selection and introgression. F_{ST} is a measure of genomic divergence and regions with high F_{ST} are expected to show patterns of elevated selection and reduced introgression. The relationship between cline gradients and F_{ST} was nonsignificant but in the predicted direction (0.019 [−0.003, 0.042]; Fig. 6; supplementary fig. S8, Supplementary Material online) while the number of ancestry transitions declined with F_{ST} (−0.73 [−0.75, −0.70]; supplementary fig. S9, Supplementary Material online). Windows with high F_{ST} tended to contain more AIMs (484.5 [478.6, 490.4]), indicating that the negative relationship to F_{ST} was not due to the number of AIMs. Cline asymmetry decreased with F_{ST} (−0.014 [−0.015, −0.013]; Fig. 6; supplementary fig. S10, Supplementary Material online), as did f_d (−0.37 [−0.38, −0.36]; supplementary fig. S11, Supplementary Material online).

High Variability in Migratory Behaviors

So far, we have tested for repeatability at the genomic level; in this last analysis, we extend our test to the phenotypic level. Traits under consistent selection across populations, if controlled by many loci of small effect, might be expected to show relationships to genome-wide ancestry and high repeatability in those relationships across populations. Therefore, if phenotypic divergence between the coastal and inland subspecies contributes to speciation between them, we might expect to see repeatable relationships of genome-wide ancestry to the phenotypes that distinguish the coastal and inland subspecies. First, we estimated relationships between genome-wide ancestry and normalized values from a set of traits in 3 categories: migratory, morphology, and color. Migratory traits, such as the distance of fall migration and wintering site longitude, were estimated using archival tracking tags attached to migrating adult birds (Justen et al. 2024). Morphological traits, such as wing measurements and tarsus length, were measured when birds were first tagged. Color traits were estimated from back feathers using spectrophotometry. Absolute values of trait by ancestry slopes varied among trait categories ($\chi^2_2 = 8.9$, $P = 0.01$), with migratory traits showing stronger relationships to ancestry than morphological or color traits (Fig. 7).

We then estimated phenotypic repeatability among populations by calculating the correlation between pairs of populations in ancestry by trait relationships for migratory, morphology, and color traits, and for genomic cline gradients. Repeatability varied among trait categories ($\chi^2_3 = 49.1$, $P < 0.001$; Fig. 7). We observed generally high correlations for color traits, with values ranging from 0.85 to 0.99. Correlations for morphological and behavioral traits were more variable, ranging from −0.06 to 0.74 for behavior and −0.60 to 0.51 for morphology. To compare phenotypic and genomic repeatability across populations, we estimated correlations for genomic cline gradients, which represent slopes of the relationships between locus-specific and genome-wide ancestry. Correlations for genomic cline gradients were moderate in comparison to the trait estimates, ranging from 0.13 to 0.49.

Discussion

We tested for 2 genomic patterns consistent with selection on incompatibilities—elevated cline gradients and a reduction in ancestry transitions—across a set of 4 hybrid Swainson's thrush populations ranging from Alaska to Washington. We found 2 contiguous blocks of AIMs that exhibited distinctive, repeatable patterns consistent with selection on hybrids—one on chromosome 1 and the other on chromosome 5. These blocks represent candidate genomic regions important for speciation in the Swainson's thrush. Cline centers in these blocks were shifted in opposite directions, with the block on chromosome 1 indicating inland introgression and the block on chromosome 5 indicating coastal introgression. Phylogenetic relationships in these regions indicated that these blocks

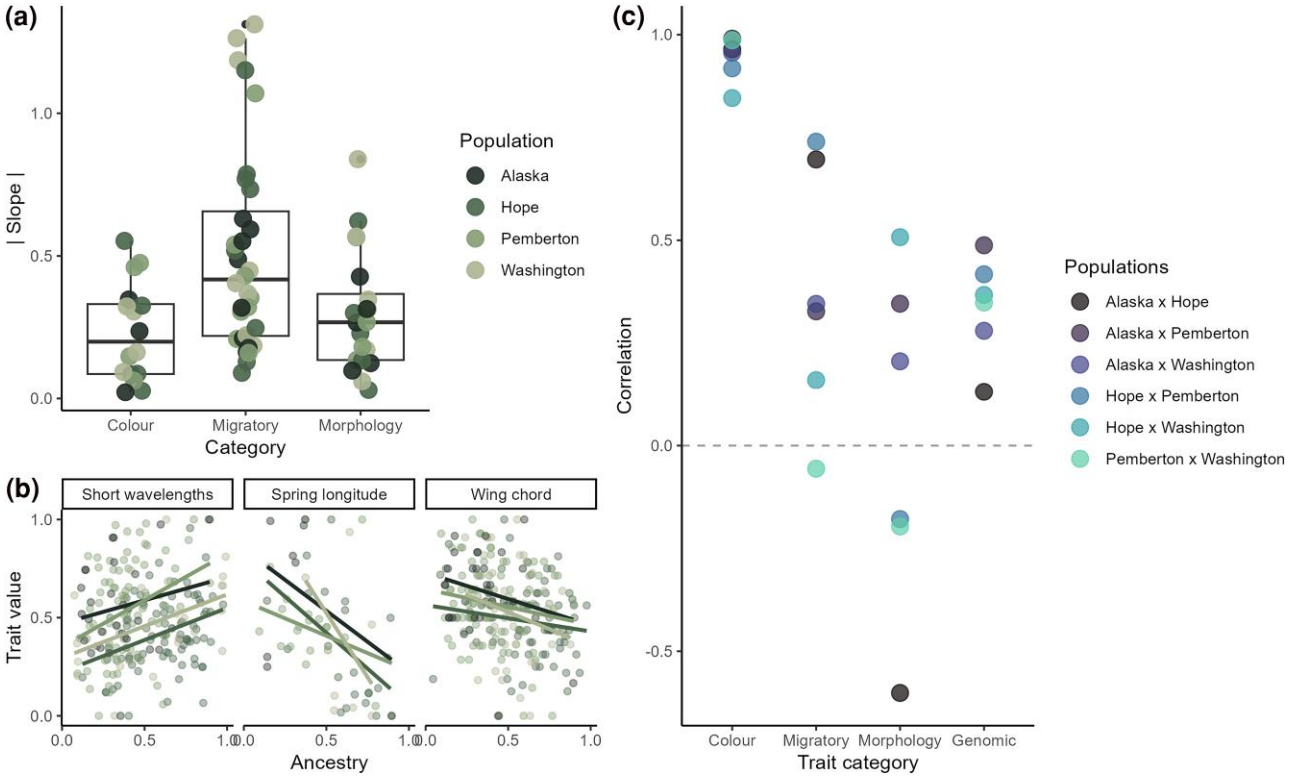

Fig. 7. Phenotypic repeatability across populations. a) Absolute values of slopes estimated between genome-wide ancestry and normalized trait values. Each point represents the slope for 1 trait, with color indicating population. Traits were divided into 3 categories: color (UV, long, medium, and short wavelengths), morphology (Kipp's distance, tail length, wing chord, tarsus length, and distal wing length), and migratory behaviors (fall longitude at 30°N, spring longitude at 30°N, overwintering site latitude and longitude, fall and spring migration distances, and timing). b) Relationships between genome-wide ancestry and phenotypes. Each point represents 1 individual while lines indicate linear slopes for each population. These slopes are examples of the values included in (a). The left panel shows a color trait (spectrophotometry at short wavelengths), the middle panel shows a migratory trait (spring longitude at 30°N), and the right panel shows a morphological trait (wing chord). c) Correlations between populations in relationships between trait values and genome-wide ancestry. Each point represents the correlation coefficient for a pair of populations between slopes relating phenotypes to genome-wide ancestry or genomic cline gradients.

evolved relatively recently, after the Swainson's thrush diverged from the rest of *Catharus*. Therefore, it is possible that these regions diverged as the subspecies were adapting to their distinct environments and migration strategies, which aligns with models for the evolution of ecological incompatibilities (Thompson et al. 2024). Incompatibilities can evolve via divergent selection acting on large genomic blocks formed by selection on multiple loci with high genetic linkage (Todesco et al. 2020).

Genomic structural features and subspecies evolutionary history contributed to the repeatable patterns across hybrid populations, as regions with low recombination rates and high interspecific divergence had the strongest signatures of selection and were least permissible to introgression. These relationships suggest that admixed genomic regions may be under widespread purifying selection, as introgressed variants tended to occur in regions of the genome where they are less likely to be linked to an incompatibility (Martin et al. 2019). This relationship to recombination rates has been observed in several hybrid populations, including in *Heliconius* butterflies (Martin et al. 2019), *Xiphophorus* fish (Langdon et al. 2022), and *Formica* ants (Nouhaud et al. 2022). Gene density did not emerge as a particularly important predictor of either selection or introgression, while F_{ST} did, suggesting that the presence of genes in a region may matter less than their identity. Karyotypes shape genomic structural features in avian genomes, as microchromosomes feature higher average recombination rates and gene densities than macrochromosomes (Ellegren 2013; Zhang et al. 2014). In the Swainson's thrush, microchromosomes were under relaxed selection and more permissible to introgression relative to macrochromosomes and the Z chromosome. The difference between the Z chromosome and microchromosomes is consistent with theory predicting reduced gene flow between diverging populations on sex chromosomes (Muirhead and Presgraves 2016; Lasne et al. 2017), while the difference between macrochromosomes and microchromosomes may be expected due to their differences in recombination rates (Butlin 2005; Tigano et al. 2022). Our results differ somewhat with a pattern of faster divergent evolution on microchromosomes than macrochromosomes in hummingbird species pairs (Henderson and Brelsford 2020), but align with higher introgression found on smaller chromosomes in an Amazonian antbird (Musher et al. 2024). In the Swainson's thrush, this could be driven in part by the concentration of loci under strong selection on a few macrochromosomes. While this could have been driven by structural genomic features, such as recombination, it is also possible that strong selection on a smaller number of loci on chromosomes 1 and 5 has led to indirect selection, and therefore coupling, in these regions (Barton 1983).

Our results cumulatively suggest that linkage among loci under selection has shaped landscapes of divergence and gene flow in the Swainson's thrush subspecies. The primary ecological difference between the subspecies—seasonal migration—is a syndrome, comprised of several interconnected behavioral, morphological, and physiological traits (Alerstam et al. 2003; Dingle 2006). Divergent evolution across this suite of traits likely involved selection acting on multiple targets with varying effect sizes, if the complex nature of seasonal migration meant that multiple genetic pathways could lead to migratory differences (Liedvogel et al. 2011; Justen and Delmore 2022). However, the easiest path to divergent evolution in response to selection may have been through blocks of loci in genetic linkage with each other. Supergenes, which are formed when sets of genes become genetically linked and experience selection as a single genomic block, can maintain adaptive divergence in the presence of gene flow (Twyford and Friedman 2015; Hager et al. 2022). The linkage is commonly maintained by structural variation, such as a large inversion or sets of smaller adjacent structural variants (Thompson and Jiggins 2014), which can lead to stronger selection within hybrid zones on structural variants (Zhang et al. 2023). A linked set of small inversions has been identified on chromosome 1 in Swainson's thrushes (Justen et al. 2024), and the block of loci under selection on chromosome 5, along with smaller sets of loci on chromosomes 2 and 3, are suggestive that some combination of structural variation and linked selection may be shaping species barriers.

On average, cline centers were shifted toward the inland subspecies, suggesting greater inland introgression. This pattern could have a demographic explanation, as the effective population size of the coastal species is much smaller than that of the inland species (Bay and Ruegg 2017). Because of this, the coastal subspecies could be harboring more weakly deleterious alleles. This pattern of a bias toward the more populous parental species has been observed in hybrid populations from several taxa (Moran et al. 2021). Alternatively, inland alleles could be favored by ecological selection in the hybrid zone, allowing them to introgress more easily. Consistent with this possibility, juvenile inland backcrosses survive migration at higher rates than coastal backcrosses (Blain et al. 2024).

Contrary to expectations, we did not find overrepresentation of misexpressed genes and phenotype-associated SNPs in regions with high cline gradients or reduced ancestry transitions. Transregulatory mechanisms, where gene expression depends on regulatory regions that are not physically near the focal gene (Hill et al. 2021), are important in producing patterns of misexpression in the Swainson's thrush (Louder et al. 2024). Therefore, if regulatory regions for misexpressed genes are under selection, they may not be captured by an analysis looking for linkage to the genes themselves (Albert et al. 2018). Alternatively, regulatory differences between the subspecies may instead be epigenetic, and therefore not show sequence-based signatures of selection (Boman et al. 2024). For both misexpressed genes and phenotype-associated SNPs, it is possible that most of these regions are under weak or no selection. To show clear patterns associated with selection, individual AIMs may need to be linked to alleles that are either under strong selection or associated with multiple selected alleles. Alternatively, the same alleles might be favored in hybrids of all backgrounds even if they are divergent between the subspecies. For example, the earlier migration timing associated with the inland subspecies or shorter routes associated with the coastal subspecies could be beneficial for all birds migrating from within the hybrid zone.

Repeatability of relationships between traits and genome-wide ancestry depended on the type of trait, with high repeatability in color, low repeatability in migratory behaviors, and intermediate repeatability in morphology and genomic cline gradients. This variation among trait categories could be due to differences in heritability. Migratory behaviors vary in their heritability estimates, with migratory orientation showing high heritability relative to migratory timing (Justen et al. 2024). Past estimates from Swainson's thrushes suggest that olive versus russet coloration does not have a strongly additive genetic basis (Delmore et al. 2016), although this contrasts

with the large effect color genes found in other vertebrates (Hoekstra 2006; Hubbard et al. 2010). Alternatively, the different patterns could arise from differences among populations in selective environments, or ancestry-dependent phenotypic plasticity, given that Swainson's thrushes in Alaska breed in a harsher climate and migrate much farther than those in Washington. For migratory traits, in particular, there were strong relationships between trait values and ancestry in some populations. This could reflect variation in selective environments, if fitness differences associated with intermediate versus parental migratory trait values are greater in some populations than others. If capacity for phenotypic plasticity has diverged between the coastal and inland subspecies, then variation in ancestry by trait relationships could also reflect different behavioral responses of the subspecies to these environments (Chevin and Lande 2011). Selection within hybrid zones is often structured by environmental gradients (Garroway et al. 2010; Taylor et al. 2014; Ryan et al. 2017), and this may be true of selection on some but not all traits that differentiate coastal and inland Swainson's thrushes. Color may be under similar selection across all hybrid populations, especially if it is important for mate selection, while selection on migratory traits might be environmentally dependent. Although the coastal and inland Swainson's thrushes readily hybridize, they differ in songs as well as their response to interspecific song playbacks (Ruegg et al. 2006; Ruegg 2008), suggesting that they may mate assortatively to some extent. Color is commonly used by birds to identify conspecifics and could be used alongside song for assortative mating if it is present (Roulin 2004). Traits involved in assortative mating can easily evolve linkage with other differentiating loci (Servedio et al. 2011), perhaps explaining the high repeatability of this trait with genomic ancestry.

Overall, patterns of selection on hybrid genomes were generally repeatable across 4 populations spanning the Swainson's thrush hybrid zone. This repeatability was most striking in large genomic blocks on chromosomes 1 and 5 that showed introgression in opposing directions and contained regions that have been previously linked to migratory traits. Repeated patterns were likely shaped by links among genome structural features, the evolutionary history of the subspecies, and selection and introgression in hybrids, with regions featuring low recombination and high subspecies differentiation tending to have stronger patterns of selection and reduced introgression. Open questions remain regarding the contribution of interactions among loci within and among selected regions as well as the relative importance of intrinsic and ecological selection. Nonetheless, locus-specific patterns across Swainson's thrush hybrid genomes are repeatably shaped by selection in the wild.

Methods

Sample Collection and Sequencing

Tissue samples were collected from adult birds from 4 populations within the hybrid zone spanning from southeastern Alaska to northern Washington during the breeding seasons between 2012 and 2023 (supplementary table S1, Supplementary Material online). Sampling was conducted using protocols approved by the Institutional Animal Care and Use Committee at Texas A&M 444 (IACUC 2019-0066) and permits obtained from Environment and Climate Change Canada 445 (SC-BC-2020-0016). A set of 5 inland subspecies individuals were sampled from an allopatric inland population in Quebec (supplementary table S1, Supplementary Material online). Additional tissue samples from the hybrid population in northern Washington were provided by the Burke Museum (supplementary table S1, Supplementary Material online). We obtained a final sample size of 299 hybrids, with 29 from Alaska, 122 from Pemberton, 106 from Hope, and 42 from Washington. DNA was extracted from blood using a standard phenol–chloroform protocol and libraries prepared for sequencing following methods described in Justen et al. (2024). Briefly, the protocol for library preparation was modified from Picelli et al. (2014) and Schumer et al. (2018), and uses Tn5 transposase enzyme to cut DNA, attaches identifying i7 and i5 indices to each fragment, and amplifies sequences with PCR. Genomes were sequenced to low coverage on the Illumina NovaSeq 6,000 (read depth range 1-13×; average 3.8×). We used a reference panel of 14 birds from each subspecies that were previously sequenced to high coverage (Louder et al. 2024). DNA for these birds was extracted using a standard phenol–chloroform protocol and libraries were prepared with DNA Flex Library Prep kits by Texas A&M's Institute for Genomic Sciences. Sequencing was performed on a NovaSeq 6000, with an average read depth of 18 × and a range of 14 to 39×.

Ancestry Informative Markers

We called genotypes at AIMs from low coverage whole genomes of hybrids using AncestryHMM, a hidden Markov model (HMM) based program that infers ancestry from read counts given subspecies allele frequencies and genetic distances between AIMs (Corbett-Detig and Nielsen 2017). Before running AncestryHMM, we (i) identified a set of AIMs, (ii) obtained inland and coastal reference panel allele frequencies, (iii) estimated genetic distances between AIMs, and (iv) estimated read counts at AIMs for low coverage whole genome sequences from hybrid populations.

We determined a set of AIMs using reference genomes from each subspecies and a reference panel of 14 coastal and 14 inland birds (Vertebrate Genomes Project 2021; Louder et al. 2024). We first obtained an initial list of putative AIMs by identifying fixed differences between the coastal and inland reference genomes. To generate a coastal pseudoreference genome that was collinear to the inland reference genome, we first simulated fastq files from the coastal reference genome fasta file using wgsim (parameters: -e 0 -N 113163516 -r 0 -R 1 −1 150 −2 150). We then aligned the simulated coastal fastq's to the inland reference genome with bwa and called variants using bcftools. We used bcftools to filter out indels, select homozygous sites only, apply quality filtering (-e QUAL < 30 || DP < 7 || DP > 60 || MQ < 40), and obtain a consensus sequence in fasta format that could be used as a coastal pseudoreference. We identified all sites that differed between the inland reference and coastal pseudoreference as an initial AIMs set of 10,280,598 loci.

To determine a final AIMs set, we identified loci from the initial AIMs set with an allele frequency difference of >0.5 between the coastal and inland reference panels. We aligned the reference panel sequences to both the coastal and inland genomes, retaining reads that mapped to both references. Using bcftools, we then called variants in the reference panel, extracted the loci in the initial AIMs set, and extracted allele frequencies for each coastal and inland birds. We then identified a final set of 930,466 AIMs by retaining loci on scaffolds

assigned to chromosomes where the reference allele frequency was higher in the inland than coastal reference samples by >0.5. This allele frequency cutoff was chosen to retain alleles that were substantially differentiated between the subspecies, while maintaining representation of regions distributed throughout the Swainson's thrush genome. Inland and coastal reference allele counts at these AIMs were extracted with bcftools.

Next, we estimated the genetic distance between AIMs using a recombination map estimated with LDhat (Auton and McVean 2007; https://github.com/auton1/LDhat) and the physical distances between sites. We converted values of rho to cM using an N_e for the inland subspecies of 570,338 (Bay and Ruegg 2017).

We estimated read counts from low coverage whole genome sequences using a pipeline modified from ancestryinfer (Schumer et al. 2020). First, sequences were aligned to both the coastal and inland reference genomes, with markers aligning to both genomes retained. We then called variants for each individual using bcftools and extracted read counts from the resulting vcf file. Read counts were combined with a reference panel allele frequencies and genetic distances to produce the input for AncestryHMM. We estimated ancestry at each marker with AncestryHMM with the following parameters: "-a 2 0.633 0.367 -p 0 −1000000 0.633 -p 1 −3000 0.367" (Corbett-Detig and Nielsen 2017). We ran AncestryHMM separately on 2 datasets, to allow for ploidy differences on sex chromosomes: (i) autosomes, including both sexes and (ii) Z chromosome, males only. We retained ancestry state calls with a posterior probability above 0.9. We then converted the ancestry states to a plink format file (ped). Using plink, we filtered out sites with a minor allele frequency <5%, missing data >25%, and applied linkage pruning (200 20 0.2) to obtain a set of 1488 AIMs.

Genomic Clines

Genomic cline estimation was performed separately for each population (Alaska, Pemberton, Hope, Washington) in R, using *bgc-hm* (Gompert et al. 2024). This program applies a logit-logistic model that estimates the probability that a particular variant was inherited from one, relative to the other, source species (Fitzpatrick 2013). Using the *glik* model, we estimated genomic clines based on the posterior probabilities of ancestry states at each AIM in the LD-pruned dataset along with reference allele counts and hybrid index. We estimated allele counts as the sum of inland alleles in the coastal and inland reference panel, based on ancestry state calls. To obtain a hybrid index, we calculated the mean of an individual's ancestry states in the LD-pruned autosomal dataset. Cline gradient estimates were considered outliers if their lower 90% confidence limits fell above 1. A cline gradient of 1 represents a slope for a particular locus that matches the genome-wide average, while a sleeper slope indicates a sharper than expected transition from 1 species ancestry to the other relative to genome-wide ancestry. Cline center estimates were considered introgression outliers if their 90% confidence limits were entirely above 0.5 (inland) or below 0.5 (coastal). Cline center estimates reflect the value of the hybrid index where the locus has an equal probability of inheritance from either the coastal or inland subspecies. Median estimates from HMM's were used as the point estimates for each population. We estimated cline asymmetry as $|centre - 0.5|$. Estimates of variability in cline gradients (σ_v) and cline centers (σ_c) were drawn from all loci and calculated by *bgc-hm* as the standard deviation in these parameters.

Ancestry Tracts

We identified ancestry tracts by determining continuous stretches of the same ancestry state within a chromosome for each individual. We retained ancestry tracts with a length >10 kb, then estimated the length and start position of each tract. Ancestry transitions were defined as the start sites of ancestry tracts. We estimated the number of transitions per individual for each population within 1 Mb windows. The number of ancestry transitions in a window varies nonlinearly with the number of AIMs in that window (supplementary fig. S1, Supplementary Material online). We fit an asymptotic model using *SSasympOrig* from the R package *nls*. We then simulated 95% confidence intervals with *predictNLS* from the R package *propogate* for each window and identified outliers as windows with ancestry transitions below the lower bound of the confidence interval. We then estimated the expected and observed overlap in outliers among populations across windows with at least 1 ancestry transition in all populations (see *Repeatability of outliers)*.

ML Phylogenies

To determine when the coastal and inland haploblocks on chromosomes 1 and 5 diverged, we built maximum likelihood phylogenies. First, we defined the haploblocks as the continuous regions with 2 or fewer ancestry tract start sites in the reference panel (46 to 59.5 Mb on chromosome 1 and 21 to 39 Mb on chromosome 5). We used a vcf file with high coverage whole genome sequences from 2 inland Swainson's thrushes, 2 coastal Swainson's thrushes, and 1 individual from each of *Catharus fuscescens*, *Catharus minimus*, *Catharus bicknelli*, *Catharus dryas*, *Catharus fuscater*, *Catharus mexicanus*, *Catharus aurantiirostris*, *Catharus frantzii*, *Catharus occidentalis*, *Catharus guttatus* (western), *Catharus guttatus* (eastern), and *Catharus gracilirostris*. Sequencing, alignment, and SNP calling methods can be found in Delmore et al. (2025). We extracted the haploblocks from the vcf file with bcftools, then filtered to retain only biallelic SNPs. We converted the haploblock vcf files to phylip format, retaining sites with SNPs for at least 8 individuals. We used IQ-Tree to first select a model of sequence evolution based on BIC (Kalyaanamoorthy et al. 2017), and then fit maximum likelihood phylogenies with 10,000 bootstrap replicates (Hoang et al. 2017; Minh et al. 2020).

Repeatability of Outliers

We evaluated the expected overlap across pairs, trios, and all 4 populations for loci or regions identified as cline gradient, cline asymmetry, or ancestry transition outliers. We estimated the expected pairwise population overlap in outliers as the proportion of loci (out of 2,187 AIMs or 4118 100 kb windows) in 1 population that were outliers multiplied by the number of outliers (the number of possible "draws") in the second population. To estimate expected overlap among trios of populations, we multiplied the proportion of loci expected to be overlapping outliers between 2 of the populations by the number of outliers in the third population. Expected overlap among all 4 populations was estimated as the proportion of loci expected to overlap among a trio of populations multiplied by the number of outliers in the fourth population. To statistically evaluate whether overlap in outlier loci between 2 or more populations was greater than expected by chance, we applied a permutation test that compares the χ^2

statistic of the observed data to a null distribution of simulated χ^2 statistics, following Yeaman et al. (2018).

D-statistics

We used D-statistics to estimate the extent of introgression from the coastal to inland subspecies in parapatric populations in British Columbia. To do this, we used the coastal and inland reference panel birds as well as a set of 5 inland birds sampled from Quebec (see above) as the allopatric inland population and a Hermit thrush sampled from Alaska as the outgroup. The Hermit thrush was sequenced to high coverage (23.5×). We first called variants using bcftools, and then filtered variant sites to retain only biallelic SNPs with a missing allele frequency <25% and a minor allele frequency >5%. We estimated Patterson's D-statistic genome-wide using the *Dtrios* function from Dsuite (Malinsky et al. 2021). We then estimated f_d and f_{dM} in nonoverlapping 100 SNP windows using the *Dinvestigate* function. For downstream analyses of f_d, we filtered out windows with a negative value for f_{dM}.

Relationships to Genome Structure

We obtained estimates of recombination rate and gene density as genome structure features. We used estimates of rho to obtain recombination rates in 100 kb windows (see *Ancestry informative markers*). We estimated gene density as the number of genes per 100 kb window, based on the inland genome annotation (Vertebrate Genomes Project 2021). We defined chromosomes <20 Mb in length as microchromosomes and above 40 Mb as macrochromosomes (Ellegren 2013).

We used cline gradients and ancestry transitions per individual within 100 kb windows as estimates of selection. For ancestry transitions, we excluded windows with fewer than 50 AIMs. We used cline asymmetry and f_d as measures of introgression. We then fit a series of linear models to evaluate whether measures of selection and introgression depend on recombination rate (cM/Mb—log-transformed), gene density (genes/100 kb), chromosome type (macrochromosome, microchromosome, or sex chromosome), or chromosome length (Mb). Before doing so, we removed 5 windows that were outliers for predictor variables (recombination rate > 3 cM/Mb, > 40 genes/100 kb). For models of cline gradients, ancestry transitions, and cline asymmetry, we included population as a random effect.

Relationships to Positive Selection and Subspecies Divergence

We estimated positive selection separately in coastal and inland individuals from the reference panel with RAiSD, using default parameters (Alachiotis and Pavlidis 2018). RAiSD estimates μ, a statistic that combines patterns of polymorphism, the site frequency spectrum, and linkage disequilibrium to evaluate evidence for positive selection in sliding SNP windows. We calculated the median μ for each subspecies in 100 kb windows, and then calculated $\log \mu_{inland}/\mu_{coastal}$ to produce a variable that ranged from positive selection in the coastal subspecies to positive selection in the inland subspecies. We then removed outlier windows that were ≥ 2 standard deviations from the mean. To estimate divergence between the subspecies, we used vcftools to calculate weighted Weir–Cockerham F_{ST} in 100 kb windows (Danecek et al. 2011), with the coastal and inland reference panel.

To estimate relationships of selection and introgression to positive selection, we fit a linear model with $\log \mu_{inland}/\mu_{coastal}$ as the predictor and cline centers as the response variable and linear models with the absolute value of that ratio ($|\log \mu_{inland}/\mu_{coastal}|$) as the predictor and each cline gradients, ancestry transitions, and f_d as response variables. We used linear models to evaluate whether F_{ST} predicted measures of selection (cline gradients and ancestry transitions) and introgression (cline asymmetry and f_d). Population was included as a random effect for models of cline metrics and ancestry transitions.

Phenotypes

We divided traits into 3 categories: (i) color (UV, long wavelengths, medium wavelengths, and short wavelengths), (ii) morphology (Kipp's distance, tail length, wing chord, tarsus length, and distal wing length), and (3) migratory behaviors (fall longitude at 30°N, spring longitude at 30°N, overwintering site latitude, overwintering site longitude, and fall migration distance and timing, spring migration distance and timing). For each trait, data were available for only a subset of the birds included in the study, depending on field data collection, with sample sizes per population ranging from 8 to 96 (supplementary table S2, Supplementary Material online). We measured color traits from back feathers spectrophotometrically by quantifying the relative stimulation of each of the 4 cones in the avian visual system. We used a Maya2000-pro spectrometer (Ocean Optics, Inc.) and a DH2000-DUV light source (output 190 to 250 0 nm; Ocean Optics, Inc). Morphological traits were measured in the field with digital calipers. To measure behavioral traits, birds were tagged with archival tags (light-level geolocators or GPS tags) during the breeding season then recaptured the following year (see details in Delmore and Irwin 2014; Justen et al. 2024).

To evaluate whether trait by ancestry relationships varied among trait categories, we first normalized each trait so that values fell between 0 and 1. Where y_i is the original trait estimate for each trait value i, the normalized trait value was estimated as: $\frac{y_i - \min(y)}{\max(y) - \min(y)}$. We then fit linear models for each trait and population with genome-wide ancestry as the predictor and normalized trait value as the response. We extracted slope estimates from those models, then used a linear mixed-effects model to test whether those slopes varied among trait categories, with trait and population included as random effects.

We then evaluated repeatability across populations in relationships between traits and genome-wide ancestry. To do this, for each trait category, we estimated correlations between pairs of populations in their trait by ancestry slopes estimated. We additionally estimated correlations between each population pair in cline gradient estimates. We used a linear mixed effects model to test whether correlations between populations depended on trait category, with the pair of populations included as a random effect.

Supplementary Material

Supplementary material is available at *Molecular Biology and Evolution* online.

Acknowledgments

Funding for this research came from an NSF CAREER grant to KED (IOS-2143004), startup funds from Texas A&M University. We greatly appreciate assistance in the field from members of the Delmore lab (especially Aeris Clarkson, Miranda Anderson, Hayley Madden, Scarlet Byron and Catherine Paul and several undergraduates), Todd Alleger, Zoe Crysler, John Klicka, and Kevin Epperly. We thank members

of the Delmore lab for discussions about clines and Swainson's thrushes. We thank Zach Gompert for help implementing *bgc-hm* and Molly Schumer for advice on using ancestry informative markers in the Swainson's thrush system.

Data Availability

Pipelines for calling ancestry informative markers and carrying out selection and introgression analyses are publicly available on GitHub: https://github.com/stephblain/thrush_ancestry_repeatability. All sequences have been uploaded to the NCBI Sequence Read Archive under BioProject numbers PRJNA979932, PRJNA1024534, and PRJNA1157204 (supplementary table S1, Supplementary Material online). Phenotypic data have been uploaded as a supplementary file.

References

Alachiotis N, Pavlidis P. RAiSD detects positive selection based on multiple signatures of a selective sweep and SNP vectors. *Commun Biol*. 2018:1(1):79. https://doi.org/10.1038/s42003-018-0085-8.

Albert FW, Bloom JS, Siegel J, Day L, Kruglyak L. Genetics of trans-regulatory variation in gene expression. *eLife*. 2018:7:e35471. https://doi.org/10.7554/eLife.35471.

Alerstam T, Hedenström A, Åkesson S. Long-distance migration: evolution and determinants. *Oikos*. 2003:103(2):247–260. https://doi.org/10.1034/j.1600-0706.2003.12559.x.

Auton A, McVean G. Recombination rate estimation in the presence of hotspots. *Genome Res*. 2007:17(8):1219–1246. https://doi.org/10.1101/gr.6386707.

Barton NH. Multilocus clines. *Evolution*. 1983:37(3):454–471. https://doi.org/10.2307/2408260.

Barton NH. The role of hybridization in evolution. *Mol Ecol*. 2001:10(3):551–568. https://doi.org/10.1046/j.1365-294x.2001.01216.x.

Barton NH, Hewitt GM. Adaptation, speciation and hybrid zones. *Nature*. 1989:341(6242):497–503. https://doi.org/10.1038/341497a0.

Bay RA, Ruegg K. Genomic islands of divergence or opportunities for introgression? *Proc R Soc B: Biol Sci*. 2017:284:20162414. https://doi.org/10.1098/rspb.2016.2414.

Blain SA, Justen HC, Easton WE, Delmore KE. Reduced hybrid survival in a migratory divide between songbirds. *Ecol Lett*. 2024:27(4):e14420. https://doi.org/10.1111/ele.14420.

Bolnick DI, Barrett RDH, Oke KB, Rennison DJ, Stuart YE. (Non)Parallel evolution. *Annu Rev Ecol Evol Syst*. 2018:49(1):303–330. https://doi.org/10.1146/annurev-ecolsys-110617-062240.

Boman J, Qvarnström A, Mugal CF. Regulatory and evolutionary impact of DNA methylation. *BMC Biol*. 2024:22(1):124. https://doi.org/10.1186/s12915-024-01920-2.

Butlin RK. Recombination and speciation. *Mol Ecol*. 2005:14(9):2621–2635. https://doi.org/10.1111/j.1365-294X.2005.02617.x.

Chaturvedi S, Lucas LK, Buerkle CA, Fordyce JA, Forister ML, Nice CC, Gompert Z. Recent hybrids recapitulate ancient hybrid outcomes. *Nat Commun*. 2020:11(1):2179. https://doi.org/10.1038/s41467-020-15641-x.

Chevin L, Lande R. Adaptation to marginal habitats by evolution of increased phenotypic plasticity. *Evolution*. 2011:24:1462–1476. https://doi.org/10.1111/j.1420-9101.2011.02279.x.

Corbett-Detig R, Nielsen R. A hidden Markov model approach for simultaneously estimating local ancestry and admixture time using next generation sequence data in samples of arbitrary ploidy. *PLoS Genet*. 2017:13(1):e1006529. https://doi.org/10.1371/journal.pgen.1006529.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, *et al*. The variant call format and VCFtools. *Bioinformatics*. 2011:27(15):2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Delmore KE, DaCosta JM, Winker K. Thrushes in love: extensive gene flow, with differential resistance and selection, obscures and reveals the evolutionary history of a songbird clade. *Mol Ecol*. 2025:e17635 . https://doi.org/10.1111/mec.17635.

Delmore KE, Irwin DE. Hybrid songbirds employ intermediate routes in a migratory divide. *Ecol Lett*. 2014:17(10):1211–1218. https://doi.org/10.1111/ele.12326.

Delmore KE, Toews DPL, Germain RR, Owens GL, Irwin DE. The genetics of seasonal migration and plumage color. *Curr Biol*. 2016:26(16):2167–2173. https://doi.org/10.1016/j.cub.2016.06.015.

Dingle H. Animal migration: is there a common migratory syndrome? *J Ornithol*. 2006:147(2):212–220. https://doi.org/10.1007/s10336-005-0052-2.

Ellegren H. The evolutionary genomics of birds. *Annu Rev Ecol Evol Syst*. 2013:44(1):239–259. https://doi.org/10.1146/annurev-ecolsys-110411-160327.

Firneno TJ, Semenov G, Dopman EB, Taylor SA, Larson EL, Gompert Z. Quantitative analyses of coupling in hybrid zones. *Cold Spring Harb Perspect Biol*. 2023:15(12):a041434. https://doi.org/10.1101/cshperspect.a041434.

Fitzpatrick BM. Alternative forms for genomic clines. *Ecol Evol*. 2013:3(7):1951–1966. https://doi.org/10.1002/ece3.609.

Garroway CJ, Bowman J, Cascaden TJ, Holloway GL, Mahan CG, Malcolm JR, Steele MA, Turner G, Wilson PJ. Climate change induced hybridization in flying squirrels. *Glob Chang Biol*. 2010:16(1):113–121. https://doi.org/10.1111/j.1365-2486.2009.01948.x.

Gavrilets S. Hybrid zones with Dobzhansky-type epistatic selection. *Evolution*. 1997:51(4):1027–1035. https://doi.org/10.2307/2411031.

Gompert Z, Buerkle CA. Bayesian estimation of genomic clines. *Mol Ecol*. 2011:20(10):2111–2127. https://doi.org/10.1111/j.1365-294X.2011.05074.x.

Gompert Z, Buerkle CA. bgc: Software for Bayesian estimation of genomic clines. *Mol Ecol Resour*. 2012:12(6):1168–1176. https://doi.org/10.1111/1755-0998.12009.x.

Gompert Z, DeRaad DA, Buerkle CA. A next generation of hierarchical Bayesian analyses of hybrid zones enables direct quantification of variation in introgression in R. *Ecol Evol*. 2024:14(11):e70548. https://doi.org/10.1002/ece3.70548.

Hager ER, Harringmeyer OS, Wooldridge TB, Theingi S, Gable JT, McFadden S, Neugeboren B, Turner KM, Jensen JD, Hoekstra HE. A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. *Science*. 2022:377(6604):399–405. https://doi.org/10.1126/science.abg0718.

Henderson EC, Brelsford A. Genomic differentiation across the speciation continuum in three hummingbird species pairs. *BMC Evol Biol*. 2020:20(1):113. https://doi.org/10.1186/s12862-020-01674-9.

Hill MS, Zande VP, Wittkopp PJ. Molecular and evolutionary processes generating variation in gene expression. *Nat Rev Genet*. 2021:22(4):203–215. https://doi.org/10.1038/s41576-020-00304-w.

Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol*. 2017:35(2):518–522. https://doi.org/10.1093/molbev/msx281.

Hoekstra HE. Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity (Edinb)*. 2006:97(3):222–234. https://doi.org/10.1038/sj.hdy.6800861.

Hubbard JK, Uy JAC, Hauber ME, Hoekstra HE, Safran RJ. Vertebrate pigmentation: from underlying genes to adaptive function. *Trends Genet*. 2010:26(5):231–239. https://doi.org/10.1016/j.tig.2010.02.002.

Hvala JA, Frayer ME, Payseur BA. Signatures of hybridization and speciation in genomic patterns of ancestry. *Evolution*. 2018:72(8):1540–1552. https://doi.org/10.1111/evo.13509.

Justen H, Delmore KE. The genetics of bird migration. *Curr Biol*. 2022:32(20):R1144–R1149. https://doi.org/10.1016/j.cub.2022.07.008.

Justen H, Lee-Yaw JA, Delmore KE. Reduced habitat suitability and landscape connectivity in a songbird migratory divide. *Glob Ecol Biogeogr*. 2021:30(10):2043–2056. https://doi.org/10.1111/geb.13367.

Justen HC, Easton WE, Delmore KE. Mapping seasonal migration in a songbird hybrid zone—heritability, genetic correlations, and genomic patterns linked to speciation. *Proc Nat Acad Sci USA*. 2024:121:e2313442121. https://doi.org/10.1073/pnas.2313442121.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Haeseler AV, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 2017:14(6):587–589. https://doi.org/10.1038/nmeth.4285.

Langdon QK, Powell DL, Kim B, Banerjee SM, Payne C, Dodge TO, Moran B, Fascinetto-Zago P, Schumer M. Predictability and parallelism in the contemporary evolution of hybrid genomes. *PLoS Genet*. 2022:18(1):e1009914. https://doi.org/10.1371/journal.pgen.1009914.

Lasne C, Sgrò CM, Connallon T. The relative contributions of the X chromosome and autosomes to local adaptation. *Genetics*. 2017:205(3):1285–1304. https://doi.org/10.1534/genetics.116.194670.

Liedvogel M, Åkesson S, Bensch S. The genetics of migration on the move. *Trends Ecol Evol*. 2011:26(11):561–569. https://doi.org/10.1016/j.tree.2011.07.009.

Louder MIM, Justen H, Kimmitt AA, Lawley KS, Turner LM, Dickman JD, Delmore KE. Gene regulation and speciation in a migratory divide between songbirds. *Nat Commun*. 2024:15(1):1–14. https://doi.org/10.1038/s41467-023-44352-2.

Malinsky M, Matschiner M, Svardal H. Dsuite—fast D-statistics and related admixture evidence from VCF files. *Mol Ecol Resour*. 2021:21(2):584–595. https://doi.org/10.1111/1755-0998.13265.

Martin SH, Davey JW, Salazar C, Jiggins CD. Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biol*. 2019:17(2):e2006288. https://doi.org/10.1371/journal.pbio.2006288.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Michael D, Haeseler AV, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020:37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015.

Moran BM, Payne C, Langdon Q, Powell DL, Brandvain Y, Schumer M. The genomic consequences of hybridization. *eLife*. 2021:10:1–33. https://doi.org/10.7554/eLife.69016.

Muirhead CA, Presgraves DC. Hybrid incompatibilities, local adaptation, and the genomic distribution of natural introgression between species. *Am Nat*. 2016:187(2):249–261. https://doi.org/10.1086/684583.

Musher LJ, Del-rio G, Marcondes RS, Brumfield RT, Bravo GA, Thom G. Geogenomic predictors of genetree heterogeneity explain phylogeographic and introgression history: a case study in an Amazonian bird (*Thamnophilus aethiops*). *Syst Biol*. 2024:73(1):36–52. https://doi.org/10.1093/sysbio/syad061.

Noor MAF, Feder JL. Speciation genetics: evolving approaches. *Nat Rev Genet*. 2006:7(11):851–861. https://doi.org/10.1038/nrg1968.

Nouhaud P, Martin SH, Portinha B, Sousa VC, Kulmuni J. Rapid and predictable genome evolution across three hybrid ant populations. *PLoS Biol*. 2022:20(12):1–20. https://doi.org/10.1371/journal.pbio.3001914.

Orr HA, Masly JP, Presgraves DC. Speciation genes. *Curr Opin Genet Dev*. 2004:14(6):675–679. https://doi.org/10.1016/j.gde.2004.08.009.

Picelli S, Bjorklund AK, Reinius B, Sargasser S, Winberg G, Sandberg R. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res*. 2014:24(12):2033–2040. https://doi.org/10.1101/gr.177881.114.

Roulin A. The evolution, maintenance and adaptive function of genetic colour polymorphism in birds. *Biol Rev Camb Philos Soc*. 2004:79(4):815–848. https://doi.org/10.1017/S1464793104006487.

Ruegg K. Genetic, morphological, and ecological characterization of a hybrid zone that spans a migratory divide. *Evolution*. 2008:62(2):452–466. https://doi.org/10.1111/j.1558-5646.2007.00263.x.

Ruegg K, Anderson EC, Slabbekoorn H. Differences in timing of migration and response to sexual signalling drive asymmetric hybridization across a migratory divide. *J Evol Biol*. 2012:25(9):1741–1750. https://doi.org/10.1111/j.1420-9101.2012.02554.x.

Ruegg K, Slabbekoorn H, Clegg S, Smith TB. Divergence in mating signals correlates with ecological variation in the migratory songbird, Swainson's thrush (*Catharus ustulatus*). *Mol Ecol*. 2006:15(11):3147–3156. https://doi.org/10.1111/j.1365-294X.2006.03011.x.

Ruegg KC, Smith TB. Not as the crow flies: a historical explanation for circuitous migration in Swainson's thrush (*Catharus ustulatus*). *Proc R Soc Biol Sci*. 2002:269:1375–1381. https://doi.org/10.1098/rspb.2002.2032.

Ryan SF, Fontaine MC, Scriber JM, Pfrender ME, O'Neil ST, Hellmann JJ. Patterns of divergence across the geographic and genomic landscape of a butterfly hybrid zone associated with a climatic gradient. *Mol Ecol*. 2017:26(18):4725–4742. https://doi.org/10.1111/mec.14236.

Satokangas I, Martin SH, Helanterä H, Saramäki J, Kulmuni J. Multi-locus interactions and the build-up of reproductive isolation. *Philos Trans R Soc B*. 2020:375(1806):20190543. https://doi.org/10.1098/rstb.2019.0543.

Schluter D. Evidence for ecological speciation and its alternative. *Science*. 2009:323(5915):737–741. https://doi.org/10.1126/science.1160006.

Schluter D, Rieseberg LH. Three problems in the genetics of speciation by selection. *Proc Natl Acad Sci U S A*. 2022:119(30):e2122153119. https://doi.org/10.1073/pnas.2122153119.

Schumer M, Powell DL, Corbett-Detig R. Versatile simulations of admixture and accurate local ancestry inference with mixnmatch and ancestryinfer. *Mol Ecol Resour*. 2020:20(4):1141–1151. https://doi.org/10.1111/1755-0998.13175.

Schumer M, Xu C, Powell DL, Durvasula A, Skov L, Holland C, Blazier JC, Sankararaman S, Andolfatto P, Rosenthal GG, *et al*. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*. 2018:3684:eaar3684. https://doi.org/10.1126/science.aar3684.

Servedio MR, Van Doorn GS, Kopp M, Frame AM, Nosil P. Magic traits in speciation: "magic" but not rare? *Trends Ecol Evol*. 2011:26(8):389–397. https://doi.org/10.1016/j.tree.2011.04.005.

Taylor SA, White TA, Hochachka WM, Ferretti V, Curry RL, Lovette I. Climate-mediated movement of an avian hybrid zone. *Curr Biol*. 2014:24(6):671–676. https://doi.org/10.1016/j.cub.2014.01.069.

Thompson KA, Brandvain Y, Coughlan JM, Delmore KE, Justen H, Linnen CR, Ortiz-Barrientos D, Rushworth CA, Schneemann H, Schumer M, *et al*. The ecology of hybrid incompatibilities. *Cold Spring Harb Perspect Biol*. 2024:16:a041440. https://doi.org/10.1101/cshperspect.a041440.

Thompson MJ, Jiggins CD. Supergenes and their role in evolution. *Heredity (Edinb)*. 2014:113(1):1–8. https://doi.org/10.1038/hdy.2014.20.

Tigano A, Khan R, Omer AD, Weisz D, Dudchenko O, Multani AS, Pathak S, Behringer RR, Aiden EL, Fisher H, *et al*. Chromosome size affects sequence divergence between species through the interplay of recombination and selection. *Evolution*. 2022:76(4):782–798. https://doi.org/10.1111/evo.14467.

Todesco M, Owens GL, Bercovich N, Légaré JS, Soudi S, Burge DO, Huang K, Ostevik KL, Drummond EB, Imerovski I, *et al*. Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature*. 2020:584(7822):602–607. https://doi.org/10.1038/s41586-020-2467-6.

Turner LM, White MA, Tautz D, Payseur BA. Genomic networks of hybrid sterility. *PLoS Genet*. 2014:10(2):e1004162. https://doi.org/10.1371/journal.pgen.1004162.

Twyford AD, Friedman J. Adaptive divergence in the monkey flower Mimulus guttatus is maintained by a chromosomal inversion. *Evolution*. 2015:69(6):1476–1486. https://doi.org/10.1111/evo.12663.

Vertebrate Genomes Project. Towards complete and error-free genome assemblies of all vertebrate species. *Nature*. 2021:592(7856):737–746. https://doi.org/10.1038/s41586-021-03451-0.

Westram AM, Faria R, Johannesson K, Butlin R. Using replicate hybrid zones to understand the genomic basis of adaptive divergence. *Mol Ecol*. 2021:30(15):3797–3814. https://doi.org/10.1111/mec.15861.

Yeaman S, Gerstein AC, Hodgins KA, Whitlock MC. Quantifying how constraints limit the diversity of viable routes to adaptation. *PLoS Genet*. 2018:14(10):1–25. https://doi.org/10.1371/journal.pgen.1007717.

Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW, Ödeen A, *et al*. Comparative genomics brings insights into avian genome evolution and adaptation. *Science*. 2014:346(6215):1311–1321. https://doi.org/10.1126/science.1251385.

Zhang L, Chaturvedi S, Nice CC, Lucas LK, Gompert Z. Population genomic evidence of selection on structural variants in a natural hybrid zone. *Mol Ecol*. 2023:32(6):1497–1514. https://doi.org/10.1111/mec.16469.