



Original Article

# Variable Signatures of Selection Despite Conserved Recombination Landscapes Early in Speciation

Sheela P. Turbek<sup>✉</sup>, Georgy A. Semenov<sup>✉</sup>, Erik D. Enbody<sup>✉</sup>,  
Leonardo Campagna, and Scott A. Taylor

From the Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO 80309, USA (Turbek, Semenov, and Taylor); the Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden (Enbody); the Fuller Evolutionary Biology Program, Cornell Lab of Ornithology, Ithaca, NY, USA (Campagna); the Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA (Campagna)

Address correspondence to S. P. Turbek at the address above, or e-mail: [sheela.turbek@colorado.edu](mailto:sheela.turbek@colorado.edu).

Received May 11, 2021; First decision June 16, 2021; Accepted September 8, 2021.

Corresponding Editor: Arun Sethuraman

## Abstract

Recently diverged taxa often exhibit heterogeneous landscapes of genomic differentiation, characterized by regions of elevated differentiation on an otherwise homogeneous background. While divergence peaks are generally interpreted as regions responsible for reproductive isolation, they can also arise due to background selection, selective sweeps unrelated to speciation, and variation in recombination and mutation rates. To investigate the association between patterns of recombination and landscapes of genomic differentiation during the early stages of speciation, we generated fine-scale recombination maps for six southern capuchino seedeaters (*Sporophila*) and two subspecies of White Wagtail (*Motacilla alba*), two recent avian radiations in which divergent selection on pigmentation genes has likely generated peaks of differentiation. We compared these recombination maps to those of Collared (*Ficedula albicollis*) and Pied Flycatchers (*Ficedula hypoleuca*), non-sister taxa characterized by moderate genomic divergence and a heterogeneous landscape of genomic differentiation shaped in part by background selection. Although recombination landscapes were conserved within all three systems, we documented a weaker negative correlation between recombination rate and genomic differentiation in the recent radiations. All divergence peaks between capuchinos, wagtails, and flycatchers were located in regions with lower-than-average recombination rates, and most divergence peaks in capuchinos and flycatchers fell in regions of exceptionally reduced recombination. Thus, co-adapted allelic combinations in these regions may have been protected early in divergence, facilitating rapid diversification. Despite largely conserved recombination landscapes, divergence peaks are specific to each focal comparison in capuchinos, suggesting that regions of elevated differentiation have not been generated by variation in recombination rate alone.

**Subject area:** Molecular Adaptation and Selection

**Keywords:** heterogeneous genomic divergence, linkage disequilibrium, recent radiations, recombination rate variation, songbirds

During population divergence and speciation, the level of genetic differentiation among closely related taxa can vary widely across the genome, a pattern referred to as heterogeneous genomic divergence (Nosil et al. 2009). Regions of elevated differentiation are often thought to harbor barrier loci responsible for reproductive isolation, while regions of low differentiation are attributed to the homogenizing effects of gene flow or a lack of divergence (Turner et al. 2005). Accordingly, researchers routinely rely on genomic scans (e.g.,  $F_{ST}$  outlier tests) to identify loci whose genetic differentiation exceeds neutral expectations and that are therefore assumed to be under divergent selection (Nosil et al. 2008). However, a growing body of literature indicates that genomic landscapes of differentiation are much more complex and difficult to interpret than previously thought (Cruickshank and Hahn 2014; Burri et al. 2015; Stankowski et al. 2019). For example, background selection, selective sweeps unrelated to speciation, and variation in recombination and mutation rates can contribute to patterns of heterogeneous genomic divergence by altering levels of within-population diversity. These forces can thereby confound the findings of genomic scans and potentially yield high false positive rates when searching for speciation loci under divergent selection (Charlesworth et al. 1993; Ravinet et al. 2017; Booker et al. 2020).

While chromosomal rearrangements and their corresponding suppression of recombination in heterozygotes have long been known to promote genomic differentiation (Rieseberg 2001), the contribution of genome-wide variation in recombination rate to patterns of heterogeneous genomic divergence during speciation has received comparatively less attention (Butlin 2005; Roesti et al. 2012). Recently, the greater availability of whole-genome sequence data and advances in algorithmic methods for estimating recombination have provided the opportunity to investigate the chromosomal features that generate variation in patterns of recombination across the genome (Smukowski and Noor 2011). Studies have documented substantial variation in recombination rate among different genomic regions, according to factors such as chromosome type (e.g., large autosomes, microchromosomes, or sex chromosomes), distance to centromere and chromosome center, gene density, and GC content, and demonstrated that recombination estimates can differ among closely related taxa (Smukowski and Noor 2011; Berner and Roesti 2017; Stapley et al. 2017). In addition, these studies have revealed the presence of recombination hotspots, or regions that shuffle genetic variation at higher rates than the rest of the genome (Lichten and Goldman 1995; Singhal et al. 2015), emphasizing the need to consider variation in local recombination rate when interpreting patterns of elevated genomic differentiation.

In genomic areas of low recombination, the reduction in diversity at linked sites due to background selection and selective sweeps (i.e., linked selection) and the effects of drift can extend over physically larger regions than in areas of high recombination, elevating estimates of genomic differentiation between populations. Indeed, empirical studies in a variety of systems have detected a negative relationship between  $F_{ST}$  and recombination rate (Stephan et al. 1998; Keinan and Reich 2010; Renaut et al. 2013), supporting the idea that linked selection in regions of low recombination can generate heterogeneous differentiation landscapes even in the absence of selection against gene flow. However, few studies have investigated the association between patterns of recombination and landscapes of genomic differentiation during the earliest stages of the speciation process (Smukowski and Noor 2011). Often, incipient species are only differentiated by a handful of genomic regions (Martin et al. 2013; Marques et al. 2016). In several avian systems, in particular,

the few regions of elevated genomic differentiation between closely related taxa contain genes involved in plumage patterning and coloration, suggesting that divergent selection on plumage traits used in mate choice and territorial defense has played a key role in speciation (Poelstra et al. 2014; Toews et al. 2016; Campagna et al. 2017; Cooper and Uy 2017; Strykowski and Sorenson 2017; Semenov et al. 2021; Turbek et al. 2021). Nonetheless, in systems such as these, it is unknown whether regions containing pigmentation genes are associated with particularly high, low, or average recombination rates, and, consequently, the role of recombination in generating peaks of differentiation associated with reproductive isolation early in divergence remains unclear.

To further investigate the interplay between recombination rate and divergent selection at various stages of the speciation process, we generated fine-scale recombination maps for southern capuchino seedeaters (*Sporophila*) and White Wagtails (*Motacilla alba*), two recent avian radiations with well-characterized landscapes of genomic differentiation and a history of ongoing or recent gene flow. We compared the patterns observed in these radiations to those detected in Collared (*Ficedula albicollis*) and Pied Flycatchers (*Ficedula hypoleuca*), non-sister taxa that exhibit moderate genomic divergence but still occasionally interbreed (Burri et al. 2015). These three systems were chosen because there is ongoing gene flow within each taxon and their landscapes of genomic differentiation have been shaped by known selection pressures (Burri et al. 2015; Hejase et al. 2020; Semenov et al. 2021). Collared and Pied Flycatchers differ in male plumage coloration, song, and various life-history traits, and hybridize at low frequencies in areas of sympatry in eastern Central Europe and the Baltic Islands of Gotland and Öland (Qvarnström et al. 2010; Ellegren et al. 2012). While female F1 hybrids are entirely sterile, male hybrids experience reduced fitness as a result of sexual selection against intermediate plumage patterns (Svedin et al. 2008).

In contrast to *Ficedula* flycatchers, southern capuchino seedeaters constitute one of the most recent and rapid avian radiations and comprise ten predominantly sympatric species that are ecologically and genomically similar, yet exhibit striking differences in male plumage patterning and song (Campagna et al. 2012, 2017). Of the few genomic regions that differentiate the species, many contain pigmentation genes involved in the melanogenesis pathway (Campagna et al. 2017) and are associated with recent, species-specific selective sweeps (Hejase et al. 2020). In addition, two syntopic capuchino species (*S. hypoxantha* and *S. iberaensis*) mate assortatively in Iberá National Park in Corrientes, Argentina, where pre-mating isolation appears to be based on plumage coloration and song (Turbek et al. 2021). Similar to capuchino seedeaters, White Wagtails (*Motacilla alba*) are a recent radiation consisting of nine subspecies characterized by variation in plumage and body size, but little genomic divergence (Harris et al. 2018; Semenov et al. 2018). Two subspecies, *M. a. alba* and *M. a. personata*, form a hybrid zone in central Siberia, where social pairs mate assortatively by head plumage (Semenov et al. 2017), a trait with a relatively simple genetic basis that maps to two small autosomal regions of elevated genomic differentiation (Semenov et al. 2021). Thus, divergent mate preferences, combined with selection on genes encoding plumage pigmentation, appear to underlie the maintenance of phenotypic differentiation in both capuchino seedeaters and White Wagtails.

The genomic landscapes of differentiation between *Ficedula* flycatchers, capuchinos, and wagtails are highly heterogeneous, with regions of elevated and reduced divergence (Ellegren et al. 2012; Campagna et al. 2017; Semenov et al. 2021; Turbek et al.

2021). However, while capuchinos and wagtails are characterized by an extremely low level of background genomic differentiation [capuchinos: genome-wide  $F_{ST} = 0.008$  (Campagna et al. 2017), wagtails: genome-wide  $F_{ST} = 0.046$  (Semenov et al. 2021)] punctuated by few narrow divergence peaks [capuchinos: mean peak width = ~243 kb (Campagna et al. 2017), wagtails: mean peak width = ~175 kb (Semenov et al. 2021)], Collared and Pied Flycatchers exhibit a much higher background of differentiation (autosomal  $F_{ST} = 0.350$ ) and have comparatively wide divergence peaks (mean peak width = 625 kb) that are more abundant and scattered throughout the genome (Ellegren et al. 2012). Most of these  $F_{ST}$  peaks are found repeatedly in the same genomic regions among populations within *Ficedula* species and among independent lineages of *Ficedula* flycatchers, suggesting that some of the more prominent features of the heterogeneous landscape of genomic differentiation between Collared and Pied Flycatchers evolved as a result of background selection (Burri et al. 2015), though selective sweeps due to positive selection in regions of low recombination have also played a role (Chase et al. 2021). The three systems also differ in their potential for admixture, which occurs more frequently in genomic regions with high recombination rates (Schumer et al. 2018; Martin et al. 2019). Wagtails and capuchinos are largely undifferentiated across the genome and have a high potential for gene flow as evidenced by field observations, admixture and ABBA-BABA tests (Semenov et al. 2021), and demographic modeling (Turbek et al. 2021), while Pied and Collared Flycatchers experience a detectable, yet relatively low, rate of introgression (Wiley et al. 2009; Ellegren et al. 2012). By comparing *Ficedula* flycatchers to the recent capuchino and wagtail radiations, we can therefore examine the degree of conservation of recombination landscapes among closely related species and assess the location of known divergence peaks relative to recombination rates at various stages of speciation. Background selection is thought to explain accentuated genomic differentiation among more divergent taxa, while positive selection may play a more prominent role in generating regions of elevated differentiation early in divergence (Burri 2017); however, few studies have evaluated this hypothesis by comparing landscapes of heterogeneous genomic differentiation throughout the speciation process (but see Burri et al. 2015; Van Doren et al. 2017; Delmore et al. 2018).

Given that recombination breaks up favorable allele combinations, potentially imposing fitness costs early in divergence, we expected to observe a statistical association between regions of low recombination and high  $F_{ST}$  in all three systems. While this pattern is well-established later in the speciation process due to the accumulating effects of background selection (Cruickshank and Hahn 2014; Burri 2017), and should therefore be supported in *Ficedula* flycatchers, it may not be detectable in the earliest stages of speciation. If the time to divergence in capuchino seedeaters and White Wagtails is too recent for this pattern to be detected, we expected little to no connection between recombination rate and genomic differentiation in capuchinos and wagtails, in contrast to *Ficedula* flycatchers. The genetic architecture of phenotypic differentiation in both capuchinos and wagtails is relatively simple, involving only a few narrow genomic regions, and could potentially persist regardless of broader (e.g., chromosome-scale) recombination patterns. Finally,  $F_{ST}$  outliers in capuchino seedeaters and White Wagtails could fall in regions of high recombination. While recombination breaks up co-adapted genetic combinations, it also generates new variation upon which natural selection can act (Ortiz-Barrientos et al. 2016) and can increase the effectiveness of selection by reducing the linkage disequilibrium between loci under selection (i.e., Hill–Robertson

interference), thereby facilitating adaptation in natural populations (Hill and Robertson 1966; Comeron et al. 2008). High recombination in regions containing pigmentation genes could thus generate novel associations among alleles and/or elements in their regulatory network, leading to new plumage patterns, and make it easier for natural selection to operate on individual loci that contribute to phenotypic diversity. Therefore, while genes underlying plumage patterning could occur throughout the genome, pigmentation genes located in genomic regions of high recombination may be disproportionately found in divergence peaks.

## Materials and Methods

### Sampling

To calculate recombination rates for each system, we used published whole-genome sequencing data from southern capuchino seedeaters (Campagna et al. 2017; Turbek et al. 2021), White Wagtails (Semenov et al. 2021), and *Ficedula* flycatchers (Burri et al. 2015). The capuchino dataset included 123 individuals of 10 southern capuchino seedeater species and 4 individuals of two outgroup species: the Tawny-bellied Seedeater (*S. hypoxantha*;  $n = 28$ ), the Iberá Seedeater (*S. iberensis*;  $n = 21$ ), the Black-bellied Seedeater (*S. melanogaster*;  $n = 12$ ), the Black-and-tawny Seedeater (*S. nigrorufa*;  $n = 12$ ), the Marsh Seedeater (*S. palustris*;  $n = 12$ ), the Pearly-bellied Seedeater (*S. pileata*;  $n = 12$ ), the Chestnut Seedeater (*S. cinnamomea*;  $n = 3$ ), the Rufous-rumped Seedeater (*S. hypochroma*;  $n = 4$ ), the Dark-throated Seedeater (*S. ruficollis*;  $n = 15$ ), the Copper Seedeater (*S. bouvreuil*;  $n = 4$ ), and the outgroups Chestnut-bellied Seedeater (*S. castaneiventris*;  $n = 2$ ) and Ruddy-breasted Seedeater (*S. minuta*;  $n = 2$ ; Table S1). The Tawny-bellied Seedeater and Iberá Seedeater datasets included a combination of males and females, while only males were included for the other capuchino species. For White Wagtails, we examined sequences from ten males of *M. a. alba* and ten males of *M. a. personata* that were sampled from allopatric populations in western Siberia and Uzbekistan, respectively, as well as five individuals of the Forest Wagtail (*Dendronanthus indicus*) and four individuals of the Gray Wagtail (*Motacilla cinerea*) as outgroups (Table S1). The *Ficedula* dataset contained ten males each of the Collared (*F. albicollis*) and Pied Flycatcher (*F. hypoleuca*) that were sampled in a zone of sympatry in the Czech Republic, along with single individuals of two outgroup species, the Red-breasted Flycatcher (*F. parva*) and the Snowy-browed Flycatcher (*F. hyperythra*) (Table S1).

### Whole-Genome Sequencing and Variant Calling

Raw sequencing reads were processed as described in Turbek et al. (2021) for capuchinos and *Ficedula* flycatchers and Semenov et al. (2021) for wagtails. Briefly, we used AdapterRemoval (capuchinos: v.2.1.7, flycatchers: v.2.3.1) to trim capuchino and *Ficedula* reads with a Phred quality score below 10, remove adapter sequences, and merge overlapping paired-end reads (Schubert et al. 2016). We aligned the capuchino reads to a reference genome of *S. hypoxantha* (Campagna et al. 2017) and the *Ficedula* reads to the chromosome-level Collared Flycatcher genome assembly FicAlb1.5 (Kawakami et al. 2014) using the very sensitive, local option in Bowtie2 (capuchinos: v.2.3.4, flycatchers: v.2.4.2) (Langmead and Salzberg 2012). For wagtails, we used the default settings in Trimmomatic v.0.39 to trim poor-quality reads and remove Illumina adapters (Bolger et al. 2014). We then used a chromosome-level assembly of *M. alba alba* (NCBI BioProject PRJNA662706) to align the reads

using bwa mem (v.0.7.17-r1188) within the Sentieon wrapper (Freed et al. 2017).

Following alignment, we used SAMtools (v.1.7) to sort and index the capuchino and flycatcher data and Picard Tools (capuchinos: v.2.17.10, flycatchers: v.2.23.8) to mark PCR duplicates, realign around indels, and fix mate pairs (Li et al. 2009; Broad Institute 2018). Variant calling was performed with HaplotypeCaller in GATK (capuchinos: v.3.8.0, flycatchers: v.4.1.9) (McKenna et al. 2010). We merged the resulting G.VCF files with CombineGVCFs and genotyped the data with GenotypeGVCFs. For wagtails, we ran duplicate removal, HaplotypeCaller (GATK 4.1), and GenotypeGVCFs (GATK 4.1) using the Sentieon wrapper tools (Freed et al. 2017). Next, we used VariantFiltration in GATK to filter out variants based on the following parameters: QD < 2.0, FS > 60.0, MQ < 30.0, MQRankSum < -12.5 and ReadPosRankSum < -8.0 for capuchinos and flycatchers and QD < 2.0, FS > 60.0, MQ < 40.0, and MQRankSum < -12.5 for wagtails. Finally, we used VCFtools (capuchinos: v.0.1.13, wagtails and flycatchers: v.0.1.15) to remove individuals with more than 20% missing data (3 *S. hypoxantha*, 3 *S. nigrorufa*, 2 *S. palustris*, 2 *S. pileata*) and exclude variants that had a minor allele frequency less than 0.05, more than 10% missing data, or were not biallelic (Danecek et al. 2011). We additionally filtered out sites that had a quality score below 20 and mean depth of coverage less than 4 or greater than 75 (for wagtails) and sites with a mean depth of coverage less than 2 or greater than 50 (for capuchinos and flycatchers). This pipeline yielded 10 638 161 SNPs after filtering for capuchinos, 37 215 097 SNPs for wagtails, and 15 072 062 SNPs for *Ficedula* flycatchers. Although we followed different pipelines to call variants for each system, it is unlikely that this affected our downstream analyses, as our comparisons are made within, rather than between, species complexes.

To avoid conducting a large number of pairwise comparisons, we restricted the capuchino dataset to six species and also removed species that lacked the minimum necessary number of individuals to calculate recombination rate (we discarded *S. cinnamomea*, *S. hypochroma*, *S. ruficollis*, and *S. bouvreuil* for these reasons). We created separate VCF files for each species in order to estimate recombination rate (for the ingroups) or infer the ancestral allele at each variable site (for the outgroups). Although the wagtail and *Ficedula* references genomes were assembled to chromosome level, our reference for *S. hypoxantha* was not a chromosome-level assembly. We therefore estimated recombination rate for each capuchino scaffold individually, reoriented and ordered the scaffolds according to their alignment to the zebra finch assembly (*Taeniopygia guttata*-3.2.4), removing scaffolds that had a cumulative matching sum <5 kb, and merged the LDhelmet output for scaffolds assigned to each chromosome.

### Recombination Rate Estimation

We computed recombination rate over 50-kilobase (kb) windows using LDhelmet v.1.10 (Chan et al. 2012). LDhelmet infers recombination rates by examining patterns of linkage disequilibrium (LD) across phased haplotypes and requires 1) an estimate of  $\theta$ , the population mutation rate, 2) a set of phased haplotypes, 3) knowledge of the ancestral allele at each segregating site, and 4) a 4 × 4 mutation matrix containing counts of mutation types (Chan et al. 2012). We estimated the population mutation rate for each system by calculating Watterson's  $\theta$  per-chromosome using a custom Python script modified from Shanfelter et al. (2019) (<https://github.com/georgysemenov/wagtails>). The script used the R package PopGenome (Pfeifer et al. 2014) to estimate Watterson's  $\theta$  over 2-kb

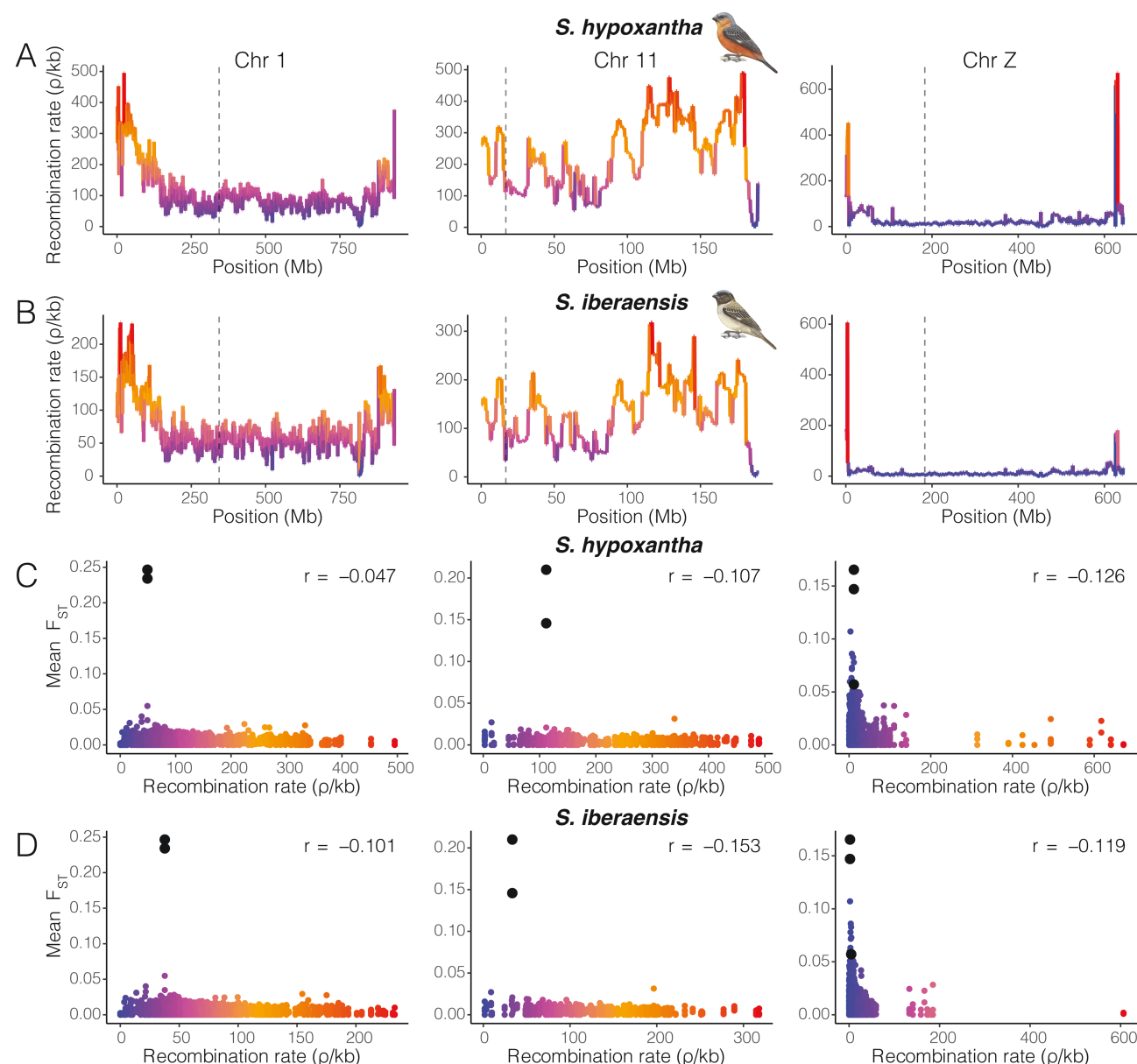
windows with a sliding window of 1 kb and averaged all windows together for each chromosome. Given that estimates of Watterson's  $\theta$  were very similar among taxa within each system ( $\pm 0.001$ ), we used the same population mutation rate estimate within capuchinos, wagtails, and *Ficedula* flycatchers. A previous study found that small changes in  $\theta$  have little influence on the resulting recombination estimates generated in LDhelmet (Stukenbrock and Dutheil 2018). To infer haplotypes, we phased each chromosome (or scaffold for capuchinos) independently with read-aware phasing in ShapeIt2, which uses reads that span at least two heterozygous sites (i.e., phase informative reads) to improve phasing accuracy (Delaneau et al. 2013). ShapeIt2 requires a minimum of ten individuals for phasing. Given that three individuals of *S. nigrorufa* had more than 20% missing data and were excluded from the VCF, we included one individual of *hypoxantha* with those of *S. nigrorufa* for the phasing step, as capuchino seedeaters exhibit little neutral genomic structure, and subsequently removed this *S. hypoxantha* individual from the phased VCF for downstream analyses. For capuchino species with more than ten individuals following filtering (*S. hypoxantha*, *S. iberaensis*, and *S. melanogaster*), we excluded birds with more than 10% missing data prior to phasing each scaffold, as recommended to improve phasing accuracy. All wagtail and flycatcher individuals had low rates of missing data and were therefore retained in the phasing analysis.

We used two outgroup species per system to infer the ancestral allele at each variable site. If the outgroup taxa were homozygous for the same allele at a particular locus, the nucleotide carried by the outgroups was assigned a prior probability of 0.91, while the remaining three nucleotides were assigned prior probabilities of 0.03 to account for uncertainty in ancestral allele reconstruction, as in Singhal et al. (2015). For all other sites (for which a polymorphism was segregating among the outgroup species), each nucleotide was assigned a prior probability derived from the overall frequency of that particular nucleotide on the chromosome (Singhal et al. 2015). We estimated mutation matrices for each species by quantifying the total number of each type of mutation away from the ancestral allele at all positions for which an ancestral allele could be inferred.

**Table 1.** Pearson's correlation ( $r$ ) between the recombination landscapes of closely related taxa of capuchino seedeaters, White Wagtails, and *Ficedula* flycatchers. Each row refers to a chromosome that contained at least one divergence peak between the two species. Adjusted  $P < 0.05$  (Benjamini and Hochberg method) for all comparisons

| Comparison                                     | Chr | Pearson's $r$ |
|--|-----|---------------|
| <i>S. hypoxantha</i> vs. <i>S. iberaensis</i>  | 1   | 0.88          |
| <i>S. hypoxantha</i> vs. <i>S. iberaensis</i>  | 11  | 0.90          |
| <i>S. hypoxantha</i> vs. <i>S. iberaensis</i>  | Z   | 0.72          |
| <i>S. nigrorufa</i> vs. <i>S. melanogaster</i> | 1   | 0.76          |
| <i>S. nigrorufa</i> vs. <i>S. melanogaster</i> | 1A  | 0.88          |
| <i>S. nigrorufa</i> vs. <i>S. melanogaster</i> | 11  | 0.84          |
| <i>S. nigrorufa</i> vs. <i>S. melanogaster</i> | 20  | 0.85          |
| <i>S. nigrorufa</i> vs. <i>S. melanogaster</i> | Z   | 0.69          |
| <i>M. a. alba</i> vs. <i>M. a. personata</i>   | 1A  | 0.87          |
| <i>M. a. alba</i> vs. <i>M. a. personata</i>   | 20  | 0.70          |
| <i>F. albicollis</i> vs. <i>F. hypoleuca</i>   | 1   | 0.82          |
| <i>F. albicollis</i> vs. <i>F. hypoleuca</i>   | 1A  | 0.86          |
| <i>F. albicollis</i> vs. <i>F. hypoleuca</i>   | 11  | 0.80          |
| <i>F. albicollis</i> vs. <i>F. hypoleuca</i>   | 20  | 0.62          |
| <i>F. albicollis</i> vs. <i>F. hypoleuca</i>   | Z   | 0.79          |



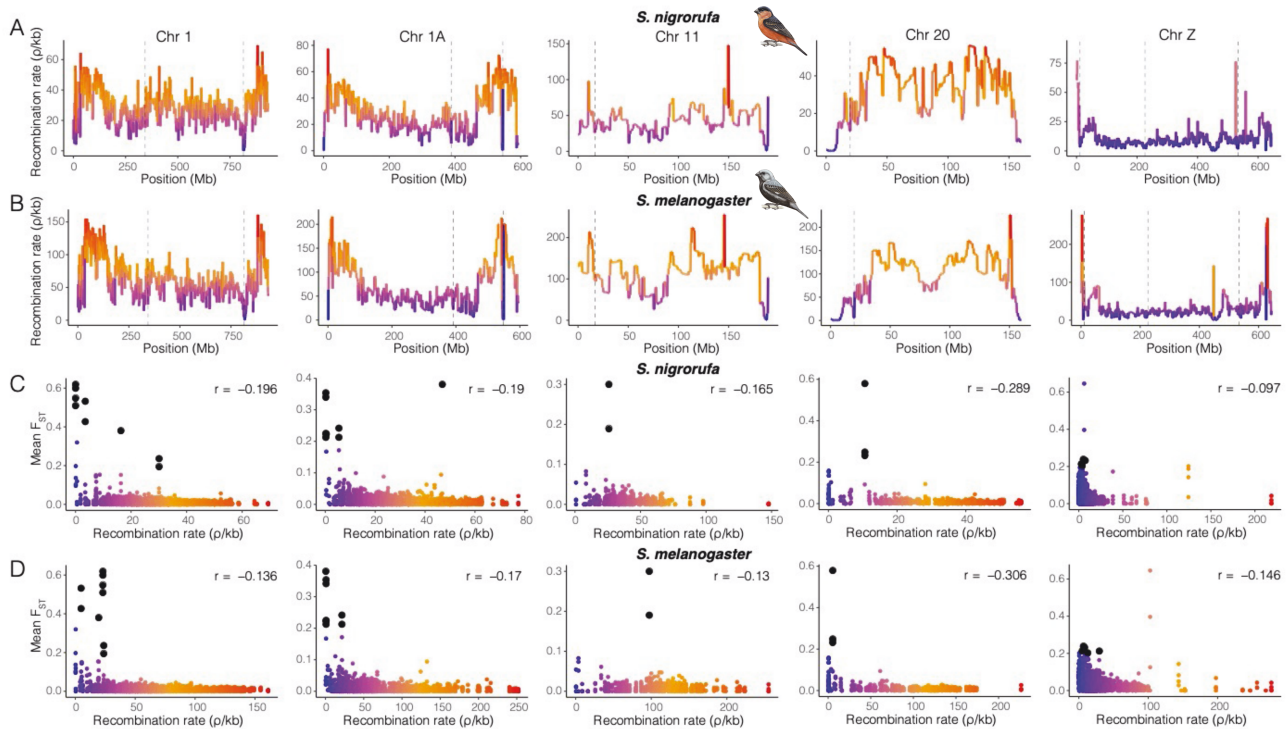


**Figure 1.** (A, B) Fine-scale recombination maps across chromosomes 1, 11, and Z in (A) *S. hypoxantha* and (B) *S. iberaensis*. Dashed gray lines indicate genomic regions containing divergence peaks. Recombination estimates were averaged over 100-kb windows and are reported in  $\rho$ /bp, where  $\rho$  is the population-scaled recombination rate ( $4N_e r$ ) and  $r$  is the per-generation mutation rate (i.e., the probability of a recombination event occurring during meiosis). (C, D) The relationship between mean  $F_{ST}$ , calculated over 25-kb windows, and recombination rate in (C) *S. hypoxantha* and (D) *S. iberaensis*. Windows containing divergence peaks are depicted as black circles and Pearson's correlation coefficient between mean  $F_{ST}$  and recombination rate is shown in the upper right corner of each plot. All plots are colored by recombination rate.

Ancestral allele probabilities and mutation matrices were generated using a custom Perl script modified from Shanfelter et al. (2019).

We converted the VCF file for each ingroup species into SNP sequence and SNP position files with the flag “-ldhelmet” in VCFtools. For each chromosome of interest, we used the *find\_confs* module in LDhelmet with a window size of 50 SNPs (-w 50) to create one haplotype configuration file. Next, we generated one likelihood lookup table per chromosome with the *table\_gen* command using the recommended grid of  $p$  values (-r 0.0 0.1 10.0 1.0 100.0) and our estimate of Watterson's  $\theta$  for each system (-t; capuchinos: 0.01, wagtails: 0.0045, flycatchers: 0.004). Additionally, we ran the *pade* module to generate Padé coefficient tables with our estimate of Watterson's  $\theta$  (-t) and the recommended 11 coefficients (-x11).

LDhelmet uses a reversible-jump Markov chain Monte Carlo (rjMCMC) scheme to approximate the posterior distribution of recombination rates, as the prior on the recombination map is a step function with an unknown number of change points (Chan et al. 2012). For the *rjmc* step, we followed Singhal et al. (2015) and ran the Markov chain for the recommended 1 000 000 iterations, discarding 100 000 iterations as burn-in, to estimate recombination rate for each chromosome (or scaffold in capuchinos) individually over the recommended window size of 50 SNPs (-w 50 -burn\_in 100000 -n 1000000). The *rjmc* procedure also requires specification of the block penalty, a penalty assigned to the likelihood when the recombination rate changes across the genome. A previous study that used LDhelmet to generate recombination maps in birds found



**Figure 2.** (A, B) Fine-scale recombination maps across chromosomes 1, 1A, 11, 20, and Z in (A) *S. nigrorufa* and (B) *S. melanogaster*. Dashed gray lines indicate genomic regions containing divergence peaks. Details as in Figure 1. (C, D) The relationship between mean  $F_{ST}$ , calculated over 25-kb windows, and recombination rate in (C) *S. nigrorufa* and (D) *S. melanogaster*. Other details as in Figure 1.

that a block penalty of 5 maximizes the power to detect hotspots, while a block penalty of 100 generates the most accurate recombination maps over longer distances (Singhal et al. 2015). As we were interested in broader-scale patterns of recombination, we ran each chromosome with a block penalty of 100 (-b 100) in all three systems. Finally, we extracted population-scaled recombination rates using the *post\_to\_text* command in LDhelmet. Recombination rates are reported in p/bp, where  $\rho$  is the population-scaled recombination rate ( $4N_e r$ ) and  $r$  refers to the per-generation recombination rate (i.e., the probability of a recombination event occurring during meiosis).

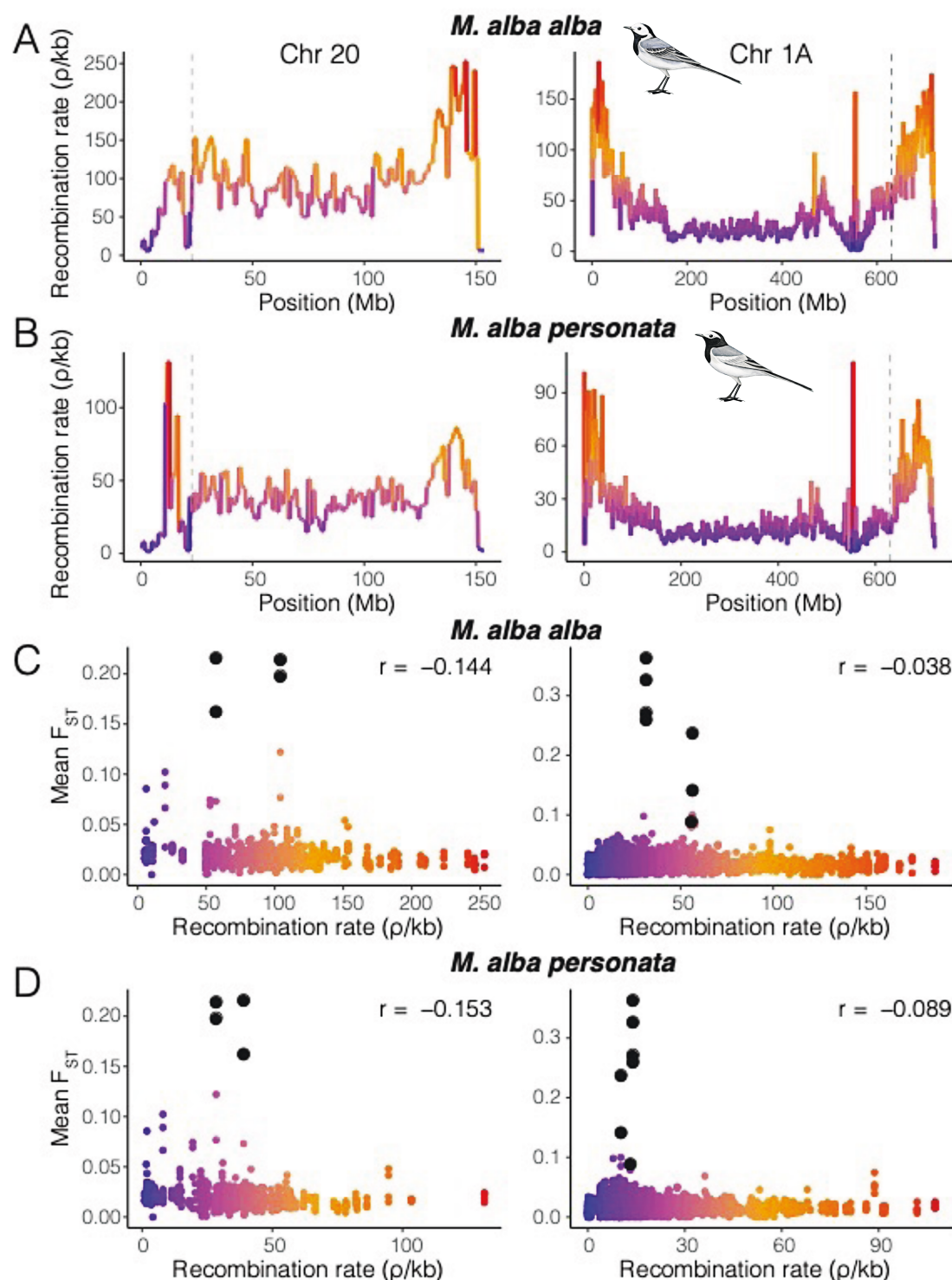
We estimated recombination rates for chromosomes containing the main divergence peaks between the focal species (1, 1A, 11, 20, and Z in capuchinos and 1A and 20 in wagtails). For *Ficedula* flycatchers, we generated recombination maps for all five of the analyzed chromosomes in the focal species, as each of the five chromosomes contains a divergence peak between Pied and Collared Flycatchers (Ellegren et al. 2012). Given that we do not make explicit comparisons between capuchinos, wagtails, and flycatchers, we did not determine synteny among the species complexes. However, chromosomal synteny is high among songbirds (Ellegren 2010), and the analyzed chromosomes are therefore likely similar across the three systems.

### Statistical Analyses

For each chromosome, we filtered out recombination estimates between SNPs that were located greater than 5 kb apart, as SNP density was likely too low to have the necessary power to detect recombination events, and averaged recombination rate ( $\rho$ ) over 100-kb windows with a 100-kb step using the R package *WindowScanR* (Tavares 2021). In addition, we used VCFtools to calculate  $F_{ST}$  over 25-kb windows between the species pairs within each system. We

note that excluding invariant sites has been shown to bias estimates of nucleotide diversity statistics (Korunes and Samuk 2021); however, it is unclear how the use of variant-only SNP data may affect windowed estimates of  $F_{ST}$  or  $\rho$ . To examine the extent to which recombination landscapes are conserved within systems, we calculated Pearson's correlation coefficient ( $r$ ) between the windowed  $\rho$  estimates of species pairs and adjusted the resulting  $p$ -values for multiple comparisons with the Benjamini and Hochberg method (Benjamini and Hochberg 1995; Pike 2011). We also computed Pearson's  $r$  between recombination rate and  $F_{ST}$  for each chromosome of interest to compare the association between regions of elevated genomic differentiation and patterns of recombination in *Ficedula* flycatchers and the recent capuchino and wagtail radiations. We note that the correlation coefficients and associated  $P$  values are affected by pseudoreplication, as loci are genetically linked and therefore not independent; however, we included the correlations in the study to illustrate and summarize the genome-wide patterns.

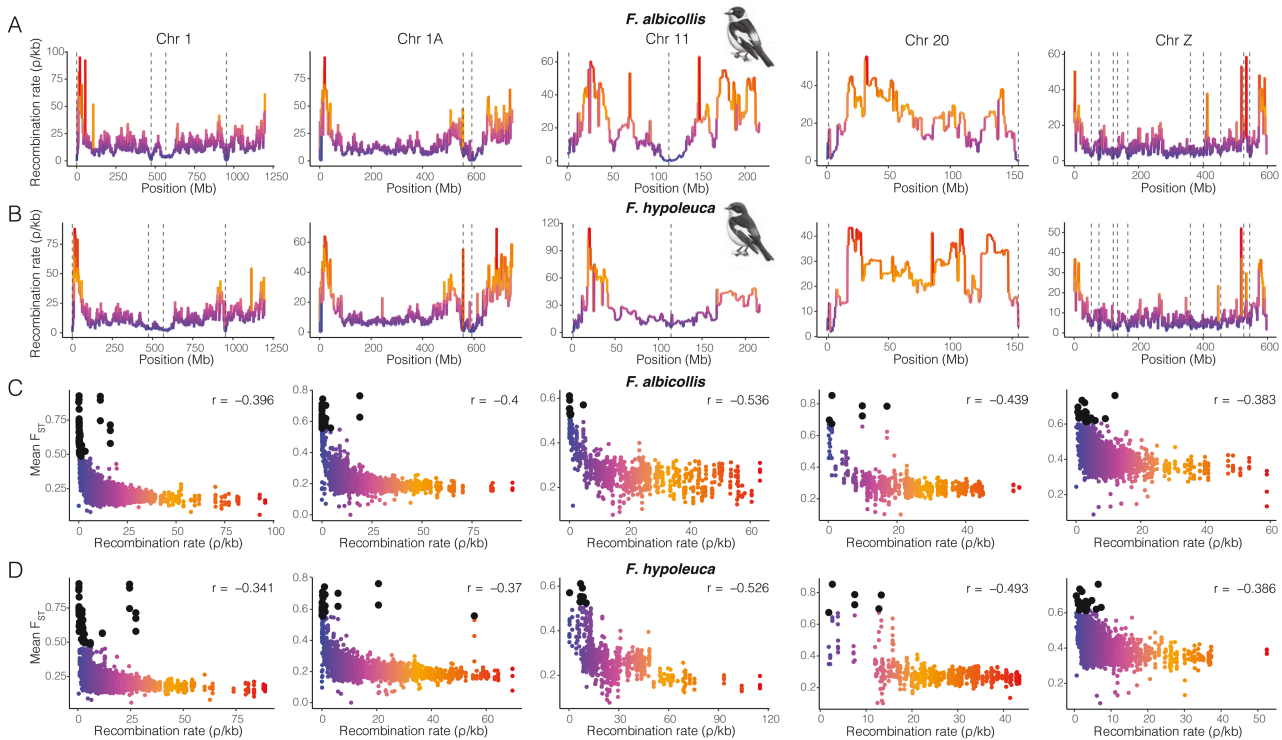
To determine whether divergence peaks in capuchino seedeaters and White Wagtails that are known to be under divergent selection fall in regions of exceptionally high or low recombination, we carried out randomization tests for *M. a. alba* versus *M. a. personata* and the capuchino species that exhibited the highest (*S. nigrorufa* vs. *S. melanogaster*;  $F_{ST} = 0.008$ ) and lowest (*S. hypoxantha* vs. *S. iberaensis*;  $F_{ST} = 0.005$ ) genomic differentiation (mean genome-wide  $F_{ST}$  calculated for 25-kb windows). These two capuchino comparisons represent the extremes of geographic overlap within the radiation, as *S. hypoxantha* and *S. iberaensis* are syntopic while *S. nigrorufa* and *S. melanogaster* are found in allopatry, and likely reflect the current probability of gene flow among the species. For each chromosome, we calculated mean  $\rho$  within the divergence peaks and randomly chose a region of the same size (i.e., the same



**Figure 3.** (A, B) Fine-scale recombination maps across chromosomes 20 and 1A in (A) *M. a. alba* and (B) *M. a. personata*. Dashed gray lines indicate genomic regions containing divergence peaks. Details as in Figure 1. (C, D) The relationship between mean  $F_{ST}$ , calculated over 25-kb windows, and recombination rate in (C) *M. a. alba* and (D) *M. a. personata*. Other details as in Figure 1.

number of windows as the divergence peaks) 10 000 times, with replacement. We computed mean  $\rho$  for the randomly selected region and calculated the two-tailed  $P$  value as  $2 \times$  the number of times the mean  $\rho$  of the randomly selected region fell below the mean observed  $\rho$  within the divergence peaks, divided by the number of

replicates. To determine the statistical significance of each test, we adjusted our  $P$  values for multiple comparisons with the Benjamini and Hochberg correction. The randomization tests examined the null hypothesis that the average  $\rho$  of the divergence peak was equal to the chromosome-wide average  $\rho$ . Divergence peaks were defined



**Figure 4.** (A, B) Fine-scale recombination maps across chromosomes 1, 1A, 11, 20, and Z in (A) *F. albicollis* and (B) *F. hypoleuca*. Dashed gray lines indicate genomic regions containing divergence peaks. Details as in Figure 1. (C, D) The relationship between mean  $F_{ST}$  calculated over 25-kb windows, and recombination rate in (C) *F. albicollis* and (D) *F. hypoleuca*. Other details as in Figure 1. Drawings courtesy of Dan Zetterström.

as described in Campagna et al. (2017) for capuchino seedeaters and Semenov et al. (2021) for White Wagtails. We also repeated the randomization tests for *Ficedula* flycatchers, where peaks of divergence have been produced through a combination of background selection and positive selection in regions of low recombination (Burri et al. 2015; Chase et al. 2021), to compare the patterns observed in each system. Divergence peaks in *Ficedula* flycatchers were defined as 25-kb windows in the top 1% of the  $F_{ST}$  distribution for each of the relevant chromosomes (see Figures S1 and S2 for examples of divergence peaks). All analyses were carried out in R v.3.6.2 (R Core Team 2018).

## Results

### Closely Related Species are Characterized by Similar Recombination Landscapes

On average, recombination landscapes were highly correlated within capuchino seedeaters, wagtails, and *Ficedula* flycatchers, indicating that recombination rates are fairly conserved among closely related avian species (Table 1, Figure S3). Correlation coefficients for wagtails, *Ficedula* flycatchers, and the capuchino species exhibiting the highest and lowest genomic differentiation are provided in Table 1, while correlation coefficients among all capuchino species are provided in Figure S4. Pearson's correlation across chromosomes ranged from 0.37 to 0.96 in capuchino seedeaters (five chromosomes), 0.70 to 0.87 in White Wagtails (two chromosomes), and 0.62 to 0.86 in *Ficedula* flycatchers (five chromosomes). Removing a single outlier chromosome in *S. pileata* (Chr 11) increased the lower bound of the range to 0.63 in capuchinos. Although divergence peaks are specific to each focal comparison of capuchino species (i.e., not all

capuchinos exhibit the same divergence peaks; see example in Figure S2) (Campagna et al. 2017), recombination rate was similar among capuchino species (Figures S3 and S4). Males and females can differ in their fine-scale recombination patterns (Sardell and Kirkpatrick 2020); however, *S. hypoxantha* and *S. iberaensis* exhibited similar recombination estimates to the other capuchino species despite the fact that our sampling included a combination of males and females (Figure S3).

### Weak Genome-wide Correlation Between Recombination and Differentiation Early in Speciation

Given the longer time since divergence between Collared and Pied Flycatchers, we expected to detect an association between regions of low recombination and high  $F_{ST}$  in *Ficedula* flycatchers. However, it is unclear if this pattern is detectable in the early stages of the speciation process. *Ficedula* flycatchers were characterized by a much higher background level of genomic differentiation than capuchinos and wagtails across all chromosomes of interest (mean genome-wide  $F_{ST}$ : *S. hypoxantha* vs. *iberaensis*: 0.005, *S. nigrorufa* vs. *S. melanogaster*: 0.008, wagtails: 0.020, *Ficedula* flycatchers: 0.235). We observed a negative correlation between mean  $F_{ST}$  and recombination rate in all contrasts (Figures 1–4). However, this negative relationship was weaker in capuchino seedeaters (Figures 1 and 2; *S. hypoxantha*:  $r = -0.047$  to  $-0.126$ , *S. iberaensis*:  $r = -0.101$  to  $-0.153$ , *S. nigrorufa*:  $r = -0.097$  to  $-0.289$ , *S. melanogaster*:  $r = -0.13$  to  $-0.306$ ) and wagtails (Figure 3; *M. a. alba*:  $r = -0.038$  to  $-0.144$ , *M. a. personata*:  $r = -0.089$  to  $-0.153$ ) than in *Ficedula* flycatchers (Figure 4; *F. albicollis*:  $r = -0.383$  to  $-0.536$ , *F. hypoleuca*:  $r = -0.341$  to  $-0.526$ ). We note that  $\rho$  and  $F_{ST}$  can only co-vary to the degree



**Table 2.** Results of randomization tests to determine if divergence peaks fell in regions with higher or lower recombination rates ( $\rho$ ) than expected by chance. For each chromosome, 10 000 regions with the same number of windows as the divergence peaks were randomly sampled with replacement. The two-tailed  $P$  value reports  $2 \times$  the proportion of randomly selected regions within the chromosome that had lower recombination rates than that of the divergence peak. Significant results (adjusted  $P < 0.05$ ; Benjamini and Hochberg method) are shown in bold

| Species                | Chr | Mean rho of peaks | Mean rho of distribution | Adjusted two-tailed $P$ |
|------------------------|-----|-------------------|--------------------------|-------------------------|
| <i>S. hypoxantha</i>   | 1   | 49.07             | 163.04                   | <b>0.006</b>            |
| <i>S. hypoxantha</i>   | 11  | 111.83            | 163.16                   | 0.382                   |
| <i>S. hypoxantha</i>   | Z   | 11.65             | 146.50                   | <b>&lt;0.0001</b>       |
| <i>S. iberaensis</i>   | 1   | 37.75             | 90.97                    | <b>0.012</b>            |
| <i>S. iberaensis</i>   | 11  | 33.43             | 90.96                    | <b>0.005</b>            |
| <i>S. iberaensis</i>   | Z   | 2.96              | 83.93                    | <b>&lt;0.0001</b>       |
| <i>S. nigrorufa</i>    | 1   | 9.31              | 29.03                    | <b>&lt;0.0001</b>       |
| <i>S. nigrorufa</i>    | 1A  | 7.41              | 28.39                    | <b>&lt;0.0001</b>       |
| <i>S. nigrorufa</i>    | 11  | 25.91             | 41.30                    | 0.056                   |
| <i>S. nigrorufa</i>    | 20  | 10.62             | 30.00                    | <b>0.022</b>            |
| <i>S. nigrorufa</i>    | Z   | 4.52              | 12.16                    | <b>0.005</b>            |
| <i>S. melanogaster</i> | 1   | 18.78             | 66.41                    | <b>&lt;0.0001</b>       |
| <i>S. melanogaster</i> | 1A  | 5.98              | 73.31                    | <b>&lt;0.0001</b>       |
| <i>S. melanogaster</i> | 11  | 97.46             | 118.21                   | 0.257                   |
| <i>S. melanogaster</i> | 20  | 5.50              | 91.30                    | <b>0.004</b>            |
| <i>S. melanogaster</i> | Z   | 12.52             | 34.25                    | <b>0.006</b>            |
| <i>M. a. alba</i>      | 1A  | 42.04             | 42.82                    | 0.953                   |
| <i>M. a. alba</i>      | 20  | 80.61             | 96.55                    | 0.551                   |
| <i>M. a. personata</i> | 1A  | 12.76             | 19.91                    | 0.257                   |
| <i>M. a. personata</i> | 20  | 33.74             | 37.96                    | 0.670                   |
| <i>F. albicollis</i>   | 1   | 2.10              | 15.50                    | <b>&lt;0.0001</b>       |
| <i>F. albicollis</i>   | 1A  | 1.85              | 17.26                    | <b>&lt;0.0001</b>       |
| <i>F. albicollis</i>   | 11  | 0.66              | 22.76                    | <b>&lt;0.0001</b>       |
| <i>F. albicollis</i>   | 20  | 6.54              | 23.04                    | <b>0.0004</b>           |
| <i>F. albicollis</i>   | Z   | 3.47              | 9.45                     | <b>&lt;0.0001</b>       |
| <i>F. hypoleuca</i>    | 1   | 5.23              | 14.08                    | <b>&lt;0.0001</b>       |
| <i>F. hypoleuca</i>    | 1A  | 3.97              | 15.59                    | <b>&lt;0.0001</b>       |
| <i>F. hypoleuca</i>    | 11  | 7.24              | 26.91                    | <b>&lt;0.0001</b>       |
| <i>F. hypoleuca</i>    | 20  | 7.60              | 25.97                    | <b>0.0004</b>           |
| <i>F. hypoleuca</i>    | Z   | 2.81              | 8.34                     | <b>&lt;0.0001</b>       |

that effective population size ( $N_e$ ) varies across the genome, as the population-scaled recombination rate ( $\rho$ ) depends on  $N_e$ .

### Many Divergence Peaks Fell in Regions of Low Recombination

Although we expected  $F_{ST}$  outliers to fall in regions of low recombination in *Ficedula* flycatchers due to the accumulating effects of background selection, high recombination in regions containing pigmentation genes could help explain the phenotypic diversity observed in capuchinos and wagtails by generating novel associations among alleles and/or elements in their regulatory network. Randomization tests revealed that divergence peaks were located in regions of significantly low recombination in 5 of 6 (83%) analyzed chromosome/peak combinations between *S. hypoxantha* and *S. iberaensis*, 8 of 10 (80%) comparisons between *S. nigrorufa* and *S. melanogaster*, 0 of 4 (0%) comparisons between wagtail subspecies, and all 10 (100%) comparisons between Collared and Pied Flycatchers. (Table 2). While the comparisons between wagtail subspecies were not statistically significant, mean  $\rho$  of the divergence peaks was consistently lower than mean  $\rho$  of the randomly selected regions in all comparisons (Table 2), suggesting that peaks of differentiation tend to fall in regions on the lower end of the recombination spectrum in all three systems. The average difference in recombination rate between the divergence peaks and randomly selected regions was 56.21 p/kb

for capuchinos, 7.02 p/kb for wagtails, and 13.74 p/kb for *Ficedula* flycatchers.

### Discussion

Genetic recombination plays a fundamental role in the evolution of sexually reproducing organisms by generating novel allelic combinations that selection can act upon and breaking up associations between loci, thereby allowing selection to remove deleterious mutations that may otherwise accumulate in populations. Nonetheless, we know little about the ways in which broad-scale variation in recombination rate contributes to patterns of heterogenous genomic divergence early in the speciation process. With the development of high-throughput genomic sequencing technology, researchers are now able to take advantage of high-density genomic data to estimate fine-scale recombination rates from patterns of linkage disequilibrium in non-model organisms. Using LD-based methods to assess the location of divergence peaks relative to the recombination landscape at various stages of the speciation process, we detected similarities and differences between recent radiations and more divergent taxa.

Within systems, we found that recombination rate was highly correlated among closely related taxa. The conservation of recombination rate across the genome is perhaps unsurprising because genomically similar organisms likely share features that correlate

with patterns of recombination, such as gene density and GC content (Dumont and Payseur 2008; Smukowski and Noor 2011), and chromosomal synteny is high among songbirds (Ellegren 2010). However, in some mammalian taxa, recombination landscapes and the position of hotspots vary substantially among closely related species and subspecies [e.g., between subspecies of mice (Smagulova et al. 2016) and humans and chimpanzees (Auton et al. 2012)]. This discrepancy has been attributed to the rapid evolution of *PRDM9* (Auton et al. 2012; Smagulova et al. 2016), a gene that determines meiotic recombination hotspots in mice and apes but is absent from other vertebrate species, including birds. The observed similarity between the recombination landscapes of capuchino seedeaters, wagtails, and *Ficedula* flycatchers is therefore consistent with the conservation of hotspot locations reported in birds and other taxa that lack a *PRDM9* gene (Lam and Keeney 2015; Singhal et al. 2015; Kawakami et al. 2017). In capuchino seedeaters, recombination landscapes were remarkably comparable among species despite the fact that particular divergence peaks are present between some species but not others (Campagna et al. 2017), suggesting that regions of elevated differentiation have not been generated by variation in recombination rate alone. Drastic decreases in diversity (e.g., recurrent selective sweeps) can compromise the ability of LD-based methods to accurately infer recombination rates (Chan et al. 2012), and many of the divergence peaks in capuchino seedeaters are associated with recent species-specific selective sweeps (Hejase et al. 2020). However, LDhelmet is fairly robust to soft selective sweeps, like those observed in capuchinos (Chan et al. 2012; Hejase et al. 2020). Given the similarity between the recombination maps produced for different capuchino species, it is unlikely that positive selection significantly influenced the resulting recombination estimates.

In contrast to Collared and Pied Flycatchers, which exhibited a relatively strong negative correlation between mean  $F_{ST}$  and recombination rate across all analyzed chromosomes, we found a much weaker relationship between the two variables in capuchinos and wagtails. Negative associations between  $F_{ST}$  and recombination rate have been reported in a variety of taxa (Stephan et al. 1998; Keinan and Reich 2010; Renaut et al. 2013), as the reduction in intraspecific variation due to linked selection can extend over larger distances in regions of low recombination, thereby elevating estimates of genomic differentiation. Given that the pattern of heterogeneous genomic divergence in *Ficedula* flycatchers is thought to have evolved due to a combination of background selection and selective sweeps in areas of reduced recombination (Burri et al. 2015; Chase et al. 2021), we expected to find a strong negative correlation between mean  $F_{ST}$  and recombination rate in Collared and Pied Flycatchers. In capuchino seedeaters and White Wagtails, on the other hand, the few regions of elevated differentiation between closely related taxa are associated with plumage patterning and are thought to have resulted from selective sweeps. We likely did not find a strong relationship between chromosome-wide recombination rate and genomic differentiation in the recent radiations given the short time since divergence between the species, and therefore limited influence of background selection in generating the heterogeneous landscapes of differentiation.

Divergence peaks may lie in low or high recombination regions regardless of the strength of association between genome-wide  $F_{ST}$  and recombination rate. Capuchino seedeaters and White Wagtails are notable for their extensive phenotypic diversity, as both groups have rapidly radiated to form species or subspecies characterized by remarkable variation in plumage patterning. In capuchinos, the reshuffling of standing genetic variation among closely related species

is thought to have facilitated phenotypic diversification and rapid speciation (Turbek et al. 2021). High recombination rates in regions containing pigmentation genes could potentially allow the rapid generation of novel plumage patterns, for example through the reshuffling of enhancers that regulate gene expression (Wallbank et al. 2016). Although we only documented a weak negative correlation between recombination rate and mean  $F_{ST}$  in capuchinos and wagtails on a broad scale, we found that a large portion of the identified divergence peaks between capuchino species were located in regions of exceptionally low recombination. The pattern observed in capuchino seedeaters was similar to that found in *Ficedula* flycatchers but was decoupled from chromosome-level associations between divergence and recombination in capuchinos. In addition, the mean recombination estimates of all divergence peaks in capuchinos, wagtails, and flycatchers fell on the lower end of the distribution of recombination estimates for their respective chromosomes.

Chromosomal inversions in capuchinos and wagtails could maintain divergence peaks regardless of patterns of recombination within taxa due to a lack of recombination in hybrid individuals. However, Campagna et al. (2017) did not detect structural variants that coincided with peaks of differentiation. In addition, the narrow regions of elevated differentiation in wagtails readily recombine in hybrids (Semenov et al. 2021). Thus, it is unlikely that divergence peaks exist within large chromosomal inversions in either system.

The finding that divergence peaks tended to occur in regions of suppressed recombination early in divergence supports the results of previous theoretical and empirical studies suggesting that adaptation and divergence may be favored in regions of low recombination when populations are subject to genetic exchange (Nachman and Payseur 2012; Marques et al. 2016; Aeschbacher et al. 2017; Han et al. 2017; Samuk et al. 2017; Martin et al. 2019). Recombination can impede speciation with gene flow (which is likely ongoing in all three systems) by breaking apart favorable allelic combinations and homogenizing divergent populations. Thus, regions responsible for phenotypic differences are less likely to be broken up by gene flow early in divergence if they are located in areas of low recombination, allowing speciation to proceed despite genetic exchange.

During population divergence and speciation, numerous factors, including selection against gene flow, background selection, selective sweeps, and variation in recombination and mutation rates, contribute to heterogeneous genomic differentiation (Ravinet et al. 2017; Semenov et al. 2019; Stankowski et al. 2019). By considering variation in local recombination rate near peaks of divergence produced through known forces in the early stages of the speciation process, we demonstrate that pigmentation genes that underlie reproductive isolation tend to occur in regions of low recombination. Co-adapted genetic combinations in these regions were likely protected from the effects of gene flow early in divergence, facilitating rapid diversification.

## Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

## Funding

S.P.T. was supported by a National Science Foundation (NSF) Graduate Research Fellowship; G.A.S. was supported by an NSF grant (DEB 1928891) to S.A.T.; Both S.P.T. and G.A.S. were supported by Ecological, Evolutionary, and Conservation Genomics awards from the American Genetic Association.

## Acknowledgments

We thank Luís Fábio Silveira and the Museu de Zoologia da Universidade de São Paulo for providing access to the capuchino outgroup samples and Per Alström for the wagtail outgroup samples. This work utilized resources from the University of Colorado Boulder Research Computing Group, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University. In addition, we acknowledge support from the National Genomics Infrastructure in Genomics Production Stockholm funded by Science for Life Laboratory and SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science for assistance with parallel sequencing and access to the UPPMAX computational infrastructure.

## Data Availability

The sequencing data used in this study are publicly available as NCBI BioProject PRJNA382416, NCBI BioProject PRJNA690099, and EMBL-EBI European Nucleotide Archive accession PRJEB7359.

## References

- Aeschbacher S, Selby JP, Willis JH, Coop G. 2017. Population-genomic inference of the strength and timing of selection against gene flow. *Proc Natl Acad Sci U S A*. 114:7061–7066.
- Auton A, Fledel-Alon A, Pfeifer S, Venn O, Ségurel L, Street T, Leffler EM, Bowden R, Aneas I, Broxholme J, et al. 2012. A fine-scale chimpanzee genetic map from population sequencing. *Science*. 336:193–198.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B*. 57:289–300.
- Berner D, Roesti M. 2017. Genomics of adaptive divergence with chromosome-scale heterogeneity in crossover rate. *Mol Ecol*. 26:6351–6369.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30:2114–2120.
- Booker TR, Yeaman S, Whitlock MC. 2020. Variation in recombination rate affects detection of outliers in genome scans under neutrality. *Mol Ecol*. 29:4274–4279.
- Broad Institute. 2018. Picard Toolkit. Available from: <http://broadinstitute.github.io/picard>. Accessed April 6, 2020.
- Burri R. 2017. Interpreting differentiation landscapes in the light of long-term linked selection. *Evol Lett*. 1:118–131.
- Burri R, Nater A, Kawakami T, Mugal CF, Olason PI, Smeds L, Suh A, Dutoit L, Bureš S, Garamszegi LZ, et al. 2015. Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula flycatchers*. *Genome Res*. 25:1656–1665.
- Butlin RK. 2005. Recombination and speciation. *Mol Ecol*. 14:2621–2635.
- Campagna L, Benites P, Loughheed SC, Lijtmaer DA, Di Giacomo AS, Eaton MD, Tubaro PL. 2012. Rapid phenotypic evolution during incipient speciation in a continental avian radiation. *Proc Biol Sci*. 279:1847–1856.
- Campagna L, Repenning M, Silveira LF, Fontana CS, Tubaro PL, Lovette IJ. 2017. Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Sci Adv*. 3:e1602404.
- Chan AH, Jenkins PA, Song YS. 2012. Genome-wide fine-scale recombination rate variation in *Drosophila melanogaster*. *PLoS Genet*. 8:e1003090.
- Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics*. 134:1289–1303.
- Chase MA, Ellegren H, Mugal CF. 2021. Positive selection plays a major role in shaping signatures of differentiation across the genomic landscape of two independent *Ficedula* flycatcher species pairs. *Evolution*. 75:2179–2196.
- Cameron JM, Williford A, Kliman RM. 2008. The Hill-Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. *Heredity (Edinb)*. 100:19–31.
- Cooper EA, Uy JAC. 2017. Genomic evidence for convergent evolution of a key trait underlying divergence in island birds. *Mol Ecol*. 26:3760–3774.
- Cruikshank TE, Hahn MW. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol*. 23:3133–3157.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al.; 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics*. 27:2156–2158.
- Delaneau O, Zagury JF, Marchini J. 2013. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods*. 10:5–6.
- Delmore KE, Lugo Ramos JS, Van Doren BM, Lundberg M, Bensch S, Irwin DE, Liedvogel M. 2018. Comparative analysis examining patterns of genomic differentiation across multiple episodes of population divergence in birds. *Evol Lett*. 2:76–87.
- Dumont BL, Payseur BA. 2008. Evolution of the genomic rate of recombination in mammals. *Evolution*. 62:276–294.
- Ellegren H. 2010. Evolutionary stasis: the stable chromosomes of birds. *Trends Ecol Evol*. 25:283–291.
- Ellegren H, Smeds L, Burri R, Olason PI, Backström N, Kawakami T, Künstner A, Mäkinen H, Nadachowska-Brzyska K, Qvarnström A, et al. 2012. The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature*. 491:756–760.
- Freed D, Aldana R, Weber JA, Edwards JS. 2017. The Sentieon Genomics Tools-A fast and accurate solution to variant calling from next-generation sequence data. *BioRxiv* 115717.
- Han F, Lamichaney S, Grant BR, Grant PR, Andersson L, Webster MT. 2017. Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Res*. 27:1004–1015.
- Harris RB, Alström P, Ödeen A, Leaché AD. 2018. Discordance between genomic divergence and phenotypic variation in a rapidly evolving avian genus (*Motacilla*). *Mol Phylogenet Evol*. 120:183–195.
- Hejase HA, Salman-Minkov A, Campagna L, Hubisz MJ, Lovette IJ, Gronau I, Siepel A. 2020. Genomic islands of differentiation in a rapid avian radiation have been driven by recent selective sweeps. *Proc Natl Acad Sci U S A*. 117:30554–30565.
- Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet Res*. 8:269–294.
- Kawakami T, Mugal CF, Suh A, Nater A, Burri R, Smeds L, Ellegren H. 2017. Whole-genome patterns of linkage disequilibrium across flycatcher populations clarify the causes and consequences of fine-scale recombination rate variation in birds. *Mol Ecol*. 26:4158–4172.
- Kawakami T, Smeds L, Backström N, Husby A, Qvarnström A, Mugal CF, Olason P, Ellegren H. 2014. A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. *Mol Ecol*. 23:4035–4058.
- Keinan A, Reich D. 2010. Human population differentiation is strongly correlated with local recombination rate. *PLoS Genet*. 6:e1000886.
- Korunes KL, Samuk K. 2021. pixy: Unbiased estimation of nucleotide diversity and divergence in the presence of missing data. *Mol Ecol Resour*. 21:1359–1368.
- Lam I, Keeney S. 2015. Nonparadoxical evolutionary stability of the recombination initiation landscape in yeast. *Science*. 350:932–937.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 9:357–359.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/Map format and SAMtools. *Bioinformatics*. 25:2078–2079.
- Lichten M, Goldman AS. 1995. Meiotic recombination hotspots. *Annu Rev Genet*. 29:423–444.
- Marques DA, Lucek K, Meier JI, Mwaiko S, Wagner CE, Excoffier L, Seehausen O. 2016. Genomics of rapid incipient speciation in sympatric threespine stickleback. *PLoS Genet*. 12:e1005887.
- Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, Blaxter M, Manica A, Mallet J, Jiggins CD. 2013. Genome-wide evidence

- for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* 23:1817–1828.
- Martin SH, Davey JW, Salazar C, Jiggins CD. 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biol.* 17:e2006288.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297–1303.
- Nachman MW, Payseur BA. 2012. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. *Philos Trans R Soc Lond B Biol Sci.* 367:409–421.
- Nosil P, Egan SP, Funk DJ. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: “isolation by adaptation” and multiple roles for divergent selection. *Evolution.* 62:316–336.
- Nosil P, Funk DJ, Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. *Mol Ecol.* 18:375–402.
- Ortiz-Barrientos D, Engelstädter J, Rieseberg LH. 2016. Recombination rate evolution and the origin of species. *Trends Ecol Evol.* 31:226–236.
- Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. 2014. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol.* 31:1929–1936.
- Pike N. 2011. Using false discovery rates for multiple comparisons in ecology and evolution. *Methods Ecol. Evol.* 2:278–282.
- Poelstra JW, Vijay N, Bossu CM, Lantz H, Ryll B, Müller I, Baglione V, Unneberg P, Wikelski M, Grabherr MG, et al. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science.* 344:1410–1414.
- Qvarnström A, Rice AM, Ellegren H. 2010. Speciation in *Ficedula* flycatchers. *Philos Trans R Soc Lond B Biol Sci.* 365:1841–1852.
- R Core Team. 2018. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, Rafajlović M, Noor MAF, Mehlig B, Westram AM. 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *J Evol Biol.* 30:1450–1477.
- Renaut S, Grassa CJ, Yeaman S, Moyers BT, Lai Z, Kane NC, Bowers JE, Burke JM, Rieseberg LH. 2013. Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nat Commun.* 4:1827.
- Rieseberg LH. 2001. Chromosomal rearrangements and speciation. *Trends Ecol Evol.* 16:351–358.
- Roesti M., Hendry AP, Salzburger W, Berner D. 2012. Genome divergence during evolutionary diversification as revealed in replicate lake–stream stickleback population pairs. *Mol. Ecol.* 21:2852–2862.
- Samuk K, Owens GL, Delmore KE, Miller SE, Rennison DJ, Schluter D. 2017. Gene flow and selection interact to promote adaptive divergence in regions of low recombination. *Mol Ecol.* 26:4378–4390.
- Sardell JM, Kirkpatrick M. 2020. Sex differences in the recombination landscape. *Am Nat.* 195:361–379.
- Schubert M, Lindgreen S, Orlando L. 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res Notes.* 9:88.
- Schumer M, Xu C, Powell DL, Durvasula A, Skov L, Holland C, Blazier JC, Sankararaman S, Andolfatto P, Rosenthal GG, et al. 2018. Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science.* 360:656–660.
- Semenov GA, Koblik EA, Red'kin YA, Badyaev AV. 2018. Extensive phenotypic diversification coexists with little genetic divergence and a lack of population structure in the White Wagtail subspecies complex (*Motacilla alba*). *J Evol Biol.* 31:1093–1108.
- Semenov GA, Linck E, Enbody ED, Harris RB, Khaydarov DR, Alström P, Andersson L, Taylor SA. 2021. Asymmetric introgression reveals the genetic architecture of a plumage trait. *Nat Commun.* 12:1019.
- Semenov GA, Safran RJ, Smith CCR, Turbek SP, Mullen SP, Flaxman SM. 2019. Unifying theoretical and empirical perspectives on genomic differentiation. *Trends Ecol Evol.* 34:987–995.
- Semenov GA, Scordato ESC, Khaydarov DR, Smith CCR, Kane NC, Safran RJ. 2017. Effects of assortative mate choice on the genomic and morphological structure of a hybrid zone between two bird subspecies. *Mol Ecol.* 26:6430–6444.
- Shanfelter AF, Archambeault SL, White MA. 2019. Divergent fine-scale recombination landscapes between a freshwater and marine population of threespine stickleback fish. *Genome Biol. Evol.* 11:1552.
- Singhal S, Leffler EM, Sannareddy K, Turner I, Venn O, Hooper DM, Strand AI, Li Q, Raney B, Balakrishnan CN, et al. 2015. Stable recombination hotspots in birds. *Science.* 350:928–932.
- Smagulova F, Brick K, Pu Y, Camerini-Otero RD, Petukhova GV. 2016. The evolutionary turnover of recombination hot spots contributes to speciation in mice. *Genes Dev.* 30:266–280.
- Smukowski CS, Noor MA. 2011. Recombination rate variation in closely related species. *Heredity (Edinb).* 107:496–508.
- Stankowski S, Chase MA, Fuiten AM, Rodrigues MF, Ralph PL, Streisfeld MA. 2019. Widespread selection and gene flow shape the genomic landscape during a radiation of monkeyflowers. *PLoS Biol.* 17:e3000391.
- Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Philos Trans R Soc B Biol Sci.* 372:20160455.
- Stephan W, Xing L, Kirby DA, Braverman JM. 1998. A test of the background selection hypothesis based on nucleotide data from *Drosophila ananassae*. *Proc Natl Acad Sci U S A.* 95:5649–5654.
- Stryjewski KE, Sorenson MD. 2017. Mosaic genome evolution in a recent and rapid avian radiation. *Nat Ecol Evol.* 1:1912–1922.
- Stukenbrock EH, Duthel JY. 2018. Fine-Scale recombination maps of fungal plant pathogens reveal dynamic recombination landscapes and intragenic hotspots. *Genetics.* 208:1209–1229.
- Svedin N, Wiley C, Veen T, Gustafsson L, Qvarnström A. 2008. Natural and sexual selection against hybrid flycatchers. *Proc Biol Sci.* 275:735–744.
- Tavares H. 2021. windowscanr: apply functions using sliding windows. R package version 0.1.
- Toews DP, Taylor SA, Vallender R, Brelsford A, Butcher BG, Messer PW, Lovette IJ. 2016. Plumage genes and little else distinguish the genomes of hybridizing Warblers. *Curr Biol.* 26:2313–2318.
- Turbek SP, Browne M, Di Giacomo AS, Kopuchian C, Hochachka WM, Estalles C, Lijtmaer DA, Tubaro PL, Silveira LF, Lovette IJ, Safran RJ, Taylor SA, Campagna L. 2021. Rapid speciation via the evolution of pre-mating isolation in the Iberá Seedeater. *Science.* 371:eabc0256.
- Turner TL, Hahn MW, Nuzhdin SV. 2005. Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biol.* 3:e285.
- Van Doren BM, Campagna L, Helm B, Illera JC, Lovette IJ, Liedvogel M. 2017. Correlated patterns of genetic diversity and differentiation across an avian family. *Mol Ecol.* 26:3982–3997.
- Wallbank RW, Baxter SW, Pardo-Diaz C, Hanly JJ, Martin SH, Mallet J, Dasmahapatra KK, Salazar C, Joron M, Nadeau N, et al. 2016. Evolutionary novelty in a butterfly wing pattern through enhancer shuffling. *PLoS Biol.* 14:e1002353.
- Wiley C, Qvarnström A, Andersson G, Borge T, Saetre GP. 2009. Postzygotic isolation over multiple generations of hybrid descendants in a natural hybrid zone: how well do single-generation estimates reflect reproductive isolation? *Evolution.* 63:1731–1739.