RESEARCH ARTICLE OPEN ACCESS

Combining Individual-Based Radio-Tracking With Whole-Genome Sequencing Data Reveals Candidate for Genetic Basis of Partial Migration in a Songbird

Matthias H. Weissensteiner[1] | Kira Delmore[2,3] | Valentina Peona[4,5] | Juan Sebastian Lugo Ramos[2,6] | Gregoire Arnaud[7] | Julio Blas[8] | Bruno Faivre[9] | Ivan Pokrovsky[10] | Martin Wikelski[10] | Jesko Partecke[10] | Miriam Liedvogel[1,2,11]

[1]Institute of Avian Research "Vogelwarte Helgoland", Wilhelmshaven, Germany | [2]Max Planck Research Group Behavioural Genomics, Max Planck Institute for Evolutionary Biology, Plön, Germany | [3]Department of Ecology Evolution and Environmental Biology, Columbia University, New York, New York, USA | [4]Vogelwarte, Sempach, Switzerland | [5]Department of Genetics and Bioinformatics, Swedish Natural History Museum, Stockholm, Sweden | [6]Neural Circuits and Evolution Laboratory, The Francis Crick Institute, London, UK | [7]Centre d'Ecologie Fonctionnelle et Evolutive (CEFE), Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France | [8]Department of Conservation Biology and Global Change, Estación Biológica de Doñana (EBD—CSIC), Sevilla, Spain | [9]UMR CNRS BioGéoSciences, Université de Bourgogne, Dijon, France | [10]Department of Migration, Max Planck Institute of Animal Behavior, Radolfzell, Germany | [11]Department of Biology and Environmental Sciences, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

Correspondence: Matthias H. Weissensteiner (mh.weissensteiner@gmail.com)

ABSTRACT

Partial migration is a phenomenon where migratory and resident individuals of the same species co-exist within a population, and has been linked to both intrinsic (e.g., genetic) as well as environmental factors. Here we investigated the genomic architecture of partial migration in the common blackbird, a songbird that comprises resident populations in the southern distribution range, partial migratory populations in central Europe, and exclusively migratory populations in northern and eastern Europe. We generated whole-genome sequencing data for 60 individuals, each of which was phenotyped for migratory behavior using radio-telemetry tracking. These individuals were sampled across the species' distribution range, including resident populations (Spain and France), obligate migrants (Russia), and a partial migratory population with equal numbers of migratory and resident individuals in Germany. We estimated genetic differentiation (F_{ST}) of single-nucleotide variants (SNVs) in 2.5 kb windows between all possible population and migratory phenotype combinations, and focused our characterization on birds from the partial migratory population in Germany. Despite overall low differentiation within the partial migratory German population, we identified several outlier regions with elevated differentiation on four distinct chromosomes. The region with the highest relative and absolute differentiation was located on chromosome 9, overlapping *PER2*, which has previously been shown to be involved in the control of the circadian rhythm across vertebrates. While this region showed high levels of differentiation, no fixed variant

---

- - - - - - - -

could be identified, supporting the notion that a complex phenotype such as migratory behavior is likely controlled by a large number of genetic loci.

1 | Introduction

Bird migration is one of the most fascinating and best-studied behavioral phenomena in the animal kingdom. While patterns of variation in migratory behavior and its ecological and evolutionary causes have been investigated in great detail, its genetic and physiological control remains understudied, limiting our understanding of the proximate control of migration, and its evolutionary potential (Helbig 1996; Plummer et al. 2015; Pulido 2011). Migration can vary drastically between species, among populations of the same species (e.g., Berthold et al. 1992), and among individuals within populations (Hegemann, Fudickar, and Nilsson 2019; Fudickar et al. 2013). In numerous species, including the common blackbird, *Turdus merula*, the propensity to migrate as well as the distance traveled differs with latitude, and many species include partial migratory populations, that is, populations where only some individuals migrate, while others stay year-round (Pulido 2011; Hegemann, Fudickar, and Nilsson 2019; Linek et al. 2021). Quantitative genetics analyses of selection and cross-breeding experiments under controlled laboratory conditions clearly showed a heritable component of migratory traits (Berthold et al. 1992; Helbig 1991; Helbig 1996; Pulido and Berthold 2010), but the genomic architecture underlying this behavior remains poorly understood. In order to investigate the molecular basis of a complex trait such as migration, the behavior needs to be accurately measured and characterized. Migratory flights may cover thousands of kilometers across and beyond continents, making it difficult to accurately record the focal behavior or trait in a natural setting. Much of our current understanding of the molecular basis of bird migration is based on phenotype characterization using indirect measures, most notably captive breeding schemes and quantitative genetics analyses on few study species (e.g., the blackcap, (Berthold and Querner 1981)), displacement experiments (Perdeck 1958; Thorup et al. 2007; Chernetsov, Kishkinev, and Mouritsen 2008), ringing recovery analyses (Mettler et al. 2013), and studies using stable isotope signatures wintering location (Bensch et al. 2009; Lundberg et al. 2017). Importantly, these measures are approximations (proxies) for the behavior in focus, and are mostly based on population-based averages, which especially in the case of partially migratory populations, where individuals within the same population exhibit different phenotypes, precludes the characterization of individual behavior. However, recent advances in tracking technology and weight reduction of tagging devices are increasingly permissive of the characterization of migration behavior at the individual level with increasing temporal and spatial resolution (Jetz et al. 2022). Much of this research has focused on characterizing variable orientation behavior across migratory divides, that is, areas where neighboring populations breed in close vicinity, but follow distinct migratory directions (e.g., Delmore and Irwin 2014; Delmore et al. 2016; Delmore et al. 2020; Sokolovskis et al. 2019). Studies on the molecular basis of other migratory traits, such as propensity to migrate especially in the context of partial migration (Hegemann, Fudickar, and Nilsson 2019; Hegemann, Marra, and Tieleman 2015), have so far been scarce; a transcriptomic study on a partial migratory population of the common blackbird detected four differentially expressed genes between migrants and residents which may play roles in determining the decision to migrate, or control physiological processes required for migration (Franchini et al. 2017).

Here, we focus on migratory behavior in common blackbirds (*Turdus merula*). Blackbirds are ideally suited for this because they show a clear pattern of migratory phenotype distribution (propensity to migrate) across their distribution range with year-round resident populations in Southern Europe, partial migrant populations in central Europe and exclusively migratory populations in Northern and Eastern Europe (Figure 1). In addition to that, the overwintering strategy of resident and migratory individuals within the partial migratory population in Germany is highly consistent (Partecke et al. *unpublished data*). Individual migratory strategies can be accurately assessed using a tracking routine that allows to remotely follow individual blackbirds in the wild over consecutive years (Fudickar et al. 2013). To understand the genomic architecture underlying migratory strategies in this system, we here (I) assembled a high-quality reference genome; (II) obtained behavioral phenotypes (i.e., migratory strategy) of individuals within a mixed population; (III) generated whole-genome sequencing data of these individuals in addition to individuals of year-round resident and fully migratory populations; and (IV) performed population genomic analyses to identify genomic candidate loci under selection. Our results show that while there is very little genetic differentiation between migratory phenotypes in the German partial migratory population–as expected in a fully panmictic population–there are a few divergent genomic regions that provide insights into the genetic architecture of the migratory phenotype, as they allow a direct association to individual behavioral phenotypes.

2 | Materials and Methods

2.1 | Sampling, Phenotyping, and Sequencing

We sampled adult male European blackbirds at four different study sites across the Eurasian breeding range, covering the species′ migratory phenotype spectrum: 10 individuals from a breeding population in Spain (N 37° 17′, W 6° 20′) and 10 individuals from a population in France (N 44° 16′, E 4° 43′) representing year-round residents; a partial migratory population in Germany (N 47° 46′, E 9° 2′) with 15 migratory and 15 resident birds individually phenotyped; and 10 migratory individuals breeding in Russia (N 55° 27′, E 37° 10′) (Table S1). To characterize migratory strategies of all individuals, we caught them with mist nets during spring and summer preceding autumn migration and deployed long-lasting radio transmitters (≤ 2.6 g delivered by Sparrow Systems, Fisher, IL, USA) using the leg-loop harness method as previously described (Fudickar et al. 2013). One to five stationary automated receivers (Sparrow Systems, Fisher, IL, USA) were deployed at each study site to continuously monitor the presence of individuals, and eventual departure dates (Crofoot et al. 2008; Kays et al. 2011). Each
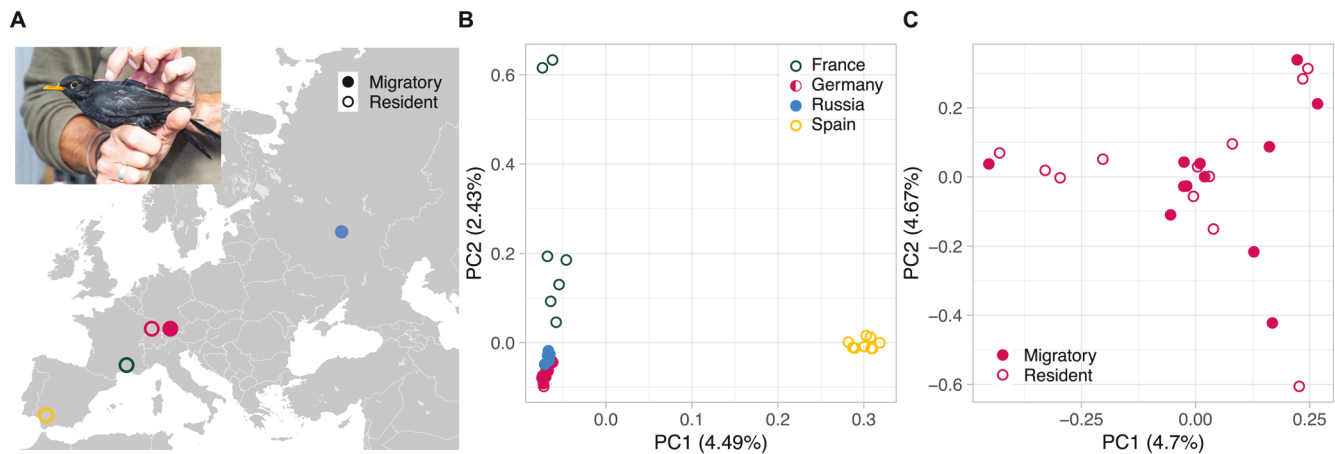
FIGURE 1 | Blackbird sampling and population structure. (A) Geographic location of blackbirds sampled: Spain (resident $n = 9$), France (resident $n = 7$), Germany (resident $n = 12$, migratory $n = 12$), and Russia (migratory $n = 9$). Circles represent individuals colored according to their population, open circles indicate resident, filled circles migratory phenotypes. The photograph in the upper left corner shows a male blackbird equipped with a radio transmitter (taken by Jesko Partecke). (B) Genome-wide principal component analysis (PCA) of SNV genotypes. Individuals from the Spanish population are clustered together, with distinct separation on the first principal component, while variation within the French population is captured mostly on the second principal component. Individuals from the German and Russian populations are tightly clustered, especially on the first principal component. (C) PCA of individuals from the German partial migratory populations with 12 resident (filled circles) and 12 migratory individuals (open circles). The lack of clear clusters illustrates that there is no population subdivision corresponding to the migratory phenotype.

automated receiver searched for a maximum of 16 frequencies every 60 s. Automated receivers were connected to H antennas (ATS, Isanti, MN, USA), mounted 3–12 m high. 24-h ARU monitoring enabled us to precisely determine individual departure events via a rapid decline of signal strength of the radio transmitters (Zúñiga et al. 2016). We used ARU data sightings and manual tracking to confirm the absence of an individual within a 2.5 km radius. Manual tracking was done via a combination of handheld H antenna (Andreas Wagener Telemetry Systems, Köln, DE) and Yaesu VR 500 receiver (Vertex Standard USA, Cypress, CA, USA). For the partial migratory German population, we also used a car-mounted Yagi-antenna (AF Antronics Inc., Urbana, IL, USA) and an airplane equipped with two H-antennas and two Biotrack receivers (Lotek, Newmarket, ON, Can) to confirm departure of an individual within a 20 km radius of the study site and thus validate the 2.5 km radius, which we used to define departure events in the other study areas.

Blood samples (ca. 50–100 μL) of all individuals were collected from the brachial vein and stored in 1 mL Queen's lysis buffer or 500 μL 95% EtOH; samples were stored in a −80°C freezer until further processing. High molecular weight DNA was extracted from blood following a standard salt extraction protocol. We prepared small insert size libraries using DNA from each individual and sequenced five samples per lane on NextSeq 500 with paired-end 150 bp reads.

2.2 | Reference Genome

We used an adult male from the German population, originating from our in-house breeding population and not characterized for migratory phenotype in the wild to generate a new reference genome. Blood (100 μL) was collected in 500 μL 95% EtOH, flash frozen in liquid nitrogen and stored at −80°C until further processing. High molecular weight genomic DNA was

extracted using the KingFisher Cell & Tissue DNA preparation Kit (Cat# 97030196). We used a combined Pacific Biosystems (PacBio) long-read sequencing and BioNano Genomics optical mapping strategy to generate a hybrid assembly of the blackbird reference genome. BioNano genome maps were generated using Direct Label and Stain (DLS) technology. Large insert size (15–20 kb) libraries were prepared for PacBio Sequel I Sequencing and subsequently sequenced on 20 PacBio SMRT cells. FALCON v1.9.0 and FALCON unzip v1.0.6 were used to generate haplotype phased contigs, which were then super-scaffolded using the BioNano Irys software hybrid scaffolding pipeline (Chin et al. 2016).

We then used SatsumaSynteny (Grabherr et al. 2010) to determine to which avian chromosome each scaffold in our final assembly corresponded to. Specifically, we aligned the blackbird scaffolds to the collared flycatcher (*Ficedula albicollis*) assembly (Ellegren et al. 2012) and organized scaffolds by chromosome and location. We validated the final assembly searching for ultra-conserved elements (UCEs) identified by Faircloth et al. (2012) (using whole genome alignments of the chicken and zebra finch). We downloaded the sequences for these elements and used NCBIs 'blastn' to determine how many were in our assembly (−evalue 1e-20 -perc_identity 90).

2.3 | Genome Annotation

We carried out a gene prediction to annotate and assign putative functions and protein domains using MAKER (Cantarel et al. 2008). Gene prediction was performed using as input a previously assembled blood transcriptome of blackbirds (18,219 assembled transcripts (Franchini et al. 2017); assembly source: https://datadryad.org/stash/dataset/doi:10.5061/dryad.tc722) and cDNA sequences from three avian species (zebra finch taeGut3.2.4, chicken Galgal4 and collared

flycatcher FicAlb_1.4) obtained from the Ensembl database. After four rounds of gene prediction with MAKER, we only kept gene models with an Annotation edit Distance (AED) less than 0.5 and more than 50aa in length. The functional annotation of the predicted protein sequences was done using blastp against a database of predicted proteins from ensemble (same species as above). Protein domains and GO ontologies were annotated using Interproscan.

2.4 | Repetitive Element Characterization and Annotation

To characterize the repetitive content of the blackbird genome, we produced a de novo repeat library of raw consensus sequences using RepeatModeler2 (with the -ltr_struct option) (Flynn et al. 2020). On top of RepeatModeler2, we also detected full-length LTR retrotransposons with LTR Harvest (Ellinghaus, Kurtz, and Willhoeft 2008) and LTR Digest (Steinbiss et al. 2009) as described in (Peona, Palacios-Gimenez, et al. 2021). The library of raw consensus sequences from RepeatModeler2 together with the avian transposable element library from Peona et al. (2021) was then used to annotate the blackbird genome with RepeatMasker v4.1.2 (Smit, Hubley, and Green 2015). The output of RepeatMasker (.align file) was further processed with the scripts calcDivergenceFromAlign.pl. from RepeatMasker to re-calculate the divergence of insertions from their consensus sequences taking CpG sites into account for subsequent plotting of the repeat landscape.

2.5 | Read Mapping and Variant Calling

Adaptor trimming of raw re-sequencing fastq files was performed with *bbduk* (sourceforge.net/projects/bbmap/). We then mapped fastq files of individuals to the blackbird reference assembly *bTurMer1* using *bwa-mem2* (Vasimuddin et al. 2019). Following sorting and indexing of alignment files, we removed duplicates with *samtools* (Danecek et al. 2011). To identify an initial set of SNVs per individual, we first used *bcftools mpileup*, specifying a minimum mapping quality of 20, and a minimum base quality of 20, followed by *bcftools call* with set to 2 (Danecek et al. 2011). The resulting gvcf files were merged per scaffold and the following filters were applied using *bcftools*: a SNV quality of above 20, a read depth of more than the total number of samples and less than twice the average read depth, mapping quality above 30, and RPBZ between −3 and 3. For sites with a read depth less than four, genotypes were set to missing. Further, we removed individuals with less than half of the genome-wide average read depth, and SNVs with more than 20% missingness. Finally, we also removed variants that overlapped with entries of our repeat annotation using *BEDtools* (Quinlan and Hall 2010).

2.6 | Population Genomic Analysis

The resulting SNV vcf file was used for all downstream analyses, including genome-wide F_{ST} scans, principal component analysis (PCA), and linkage-disequilibrium (LD) analyses. First, we converted the vcf file to bim format using *plink* (Purcell et al. 2007). We then used the R package *gdsfmt* (Zheng et al. 2012) to convert it to a geno file. Using this geno file we then performed the PCA and relatedness analysis using *SNPRelate* in R (Zheng et al. 2012). The latter enabled us to identify closely related individuals (parent-offspring and half-siblings) in our dataset, of which we then subsequently excluded the individual with the lower genome-wide average read depth from each pair. To calculate genomic differentiation (F_{ST}) in 2.5 kb windows, we used *popgenwindows.py* (https://github.com/simonhmartin/genomics_general), which outputs π for both populations of the comparison, as well as D_{xy} and F_{ST} (Hudson, Slatkin, and Maddison 1992). We chose this window size based on the assumption that a polygenic architecture underlies our trait of interest (migratory behavior) (Le Corre and Kremer 2012), therefore allowing a more fine-grained investigation of allele frequency changes of SNVs aids identifying genomic regions under selection, especially in a setting in which a presumably fully admixed population contains two divergent phenotypes. To discern population genetic processes that are common in all populations (e.g., background selection) from those that are of interest in the focal population comparison (i.e., selection on the behavioral phenotype), we calculated ΔF_{ST}'. Following Vijay et al. (2016), we first estimated F_{ST} for a comparison that is within the same phenotype (i.e., the two resident populations of Spain and France). We then Z-transformed and subtracted this estimate for each orthologous 2.5 kb window from the Z-transformed estimates of the between-phenotype comparisons (e.g., resident vs. migratory German individuals), and presumably retain a signal indicative of divergent selection. We also calculated per-SNV F_{ST} after Weir and Cockerham (1984) using *vcftools* (Danecek et al. 2011). To identify signatures of selection, we estimated Tajima's D in the same 2.5 kb windows used in the analyses above using *vcftools* (Danecek et al. 2011). To investigate the potential occurrence of an inversion on the blackbird equivalent of collared flycatcher chromosome 9 in, we calculated pairwise linkage disequilibrium (LD) among all SNVs present on that chromosome with *vcftools* (Danecek et al. 2011).

Downstream data handling and visualization have been performed with the *R* programming language, using the packages *ggplot2* and *tidyr* (Wickham, Vaughan, and Girlich 2024).

3 | Results

3.1 | Genome Assembly and Annotation

To produce a draft reference assembly, we generated 110.6 Gigabytes of PacBio long-read sequencing data from 21 SMRT cells, corresponding to a 45.27-fold genome coverage and achieving an average read length of 14 kb (10.5 kb average subread length). After assembling an initial draft reference (see Methods), we generated 4,643,798 BioNano single-molecule maps and assembled them into consensus maps to be then used to super-scaffold PacBio scaffolds resulting in a final reference genome of 1.01 Gb (1,012,066,560 bp) in length. This new reference assembly (*bTurMer1*, available under accession number GCA_046127255.1) consists of 125 scaffolds and with a scaffold N50 of 42.23 Mb. A total of 113 scaffolds mapped to the collared flycatcher genome (an average of four scaffolds per chromosome). The *de novo* gene annotation using both *in silico* and evidence-based approaches resulted in the identification

of 18,074 protein-coding genes and 19,060 transcripts. Results from the UCE analysis suggest this reference is nearly complete as more than 95% of the UCEs identified in amniotes were also present (2443/2560 UCEs).

3.2 | Repeat Library and Annotation

The *de novo* repeat library obtained with RepeatModeler2 contains a total of 430 consensus sequences (Data S1). As expected from the repeat composition of other birds (Peona, Blom, et al. 2021; Kapusta and Suh 2017; Bravo, Jonathan Schmitt, and Edwards 2021), the vast majority of consensus sequences are classified as LTR retrotransposons (197) or LINE Chicken Repeat 1 (103). The rest of the elements are classified as DNA transposons (41), SINEs (11) a few are satellite DNA monomers, while 88 sequences remain unclassified. The repeat annotation (Data S2 and S3) showed that the blackbird genome is ~11% repetitive percentage similar to many other birds (Kapusta and Suh 2017; Peona, Blom, et al. 2021) where the most common transposable elements are LINEs (3.73%) and LTR retrotransposons (3.64% of which only 8% are found to be full-length, Data S4). While the 20% of the blackbird-specific repeat library remains unclassified, these elements account for only 0.77% of the genome assembly. Finally, the blackbird genome shows an accumulation of LTR retrotransposons in ancient and recent times, while LINE retrotransposons show an accumulation of only older insertions (with a divergence from consensus sequence greater than 10%; Figure S1). Since the blackbird-specific repeat library has not been manually curated, precise calculation of divergence of insertions from their consensus sequences cannot be made.

3.3 | Population re-Sequencing

In total, we generated whole-genome re-sequencing data of 60 male individuals from 4 different European populations with divergent phenotypes (Spain = 10, France = 10, Germany = 30, and Russia = 10; Figure 1A). The higher number of individuals in the German population stems from the fact that we included migrants ($n = 15$) and residents ($n = 15$) from the focal partial migratory population to facilitate direct comparison of behavioral phenotypes of the same geographical region. Mean sequencing read depth among all individuals was 15.63X, and in an initial filtering step, we excluded eight individuals that showed an average read depth of less than half of the mean across all individuals ($< 7.82X$, Table S1). We then mapped sequence reads to the *bTurMer1* reference genome assembly and identified single-nucleotide variants (SNVs) in each individual. Merging and extensive filtering (see Methods) resulted in a final set of 21,408,162 variants, which were analyzed subsequently. Using these SNVs, we performed a relatedness analysis using the Method of Moments (Purcell et al. 2007) approach to identify closely related individuals. We found five pairs of individuals among the German and French samples that exhibited a kinship coefficient over 0.1 (Figure S2), which led us to exclude three individuals (one migratory individual from the German and two individuals from the French population, respectively). The total number of individuals was therefore 49 individuals (9 from Spain, 7 from France, 12 German resident individuals, 12 German migratory individuals, and 9 from Russia, Table S1).

3.4 | Population Structure

To visualize structure across blackbird populations in Europe, we performed principal component analysis (PCA) using genotypes of all filtered SNVs as input. Figure 1B shows that genomic variation separating the Spanish individuals from the rest of the European populations is captured by the first principal component, whereas the second principal component mainly illustrates variation within the French population, with German and Russian individuals tightly clustered together. Additional principal components did not exhibit further clustering according to populations or phenotypes (Figure S3). In a PCA using only individuals from the German population, we did not observe any clustering according to the behavioral phenotypes (Figure 1C).

3.5 | Genome-Wide Differentiation

To identify genomic regions that are associated with phenotypic differences in migration strategy, we performed F_{ST} outlier scans. For each population comparison, we calculated mean F_{ST} after Hudson, Slatkin, and Maddison (1992) in non-overlapping windows of 2.5 kb in size and a minimum of 10 SNVs. To obtain the baseline level of differentiation presumably not related to differences in migration strategy (i.e., a 'control' comparison), we first focused on the within-phenotype comparison of Spanish and French resident blackbirds. Mean genome-wide F_{ST} in this comparison was low with 0.04 (median 0.029) ranging from 0 to 0.878 per 2.5 kb window, yet the differentiation landscape was heterogeneous across chromosomes (see Figure 2A for Z-transformed F_{ST} estimates, Figure S4A for untransformed estimates). Next, we estimated differentiation by comparing migratory phenotypes in the German partial migratory population. To discern regions of elevated F_{ST} due to divergent selection from those due to common population processes (e.g., background selection), we Z-transformed F_{ST} estimates (F_{ST}') and subtracted the within-phenotype (Spain-France) F_{ST}' from the between-phenotype F_{ST}' to yield a measure of net differentiation ($\Delta F_{ST}'$). The genome-wide $\Delta F_{ST}'$ landscape shows very low levels of differentiation on average, with a few regions of distinctly elevated $\Delta F_{ST}'$ (Figure 2B, Figure S4B for absolute F_{ST}). Windows above the 99th percentile of $\Delta F_{ST}'$ were deemed as outliers. Overall, only seven $\Delta F_{ST}'$ outlier windows overlapped with F_{ST}' outlier windows of the 'control' comparison between both resident Spanish and French individuals. When five or more outlier windows occurred consecutively (ignoring windows with missing values)—indicative of selection potentially acting on haplotypes instead of single SNVs—we classified them as outlier clusters. In total, we found eight outlier clusters belonging to five regions on four chromosomes (Table 1).

The two clusters with highest overall net and absolute differentiation, as well as the most consecutive outlier windows were located on the equivalent of collared flycatcher chromosome 9
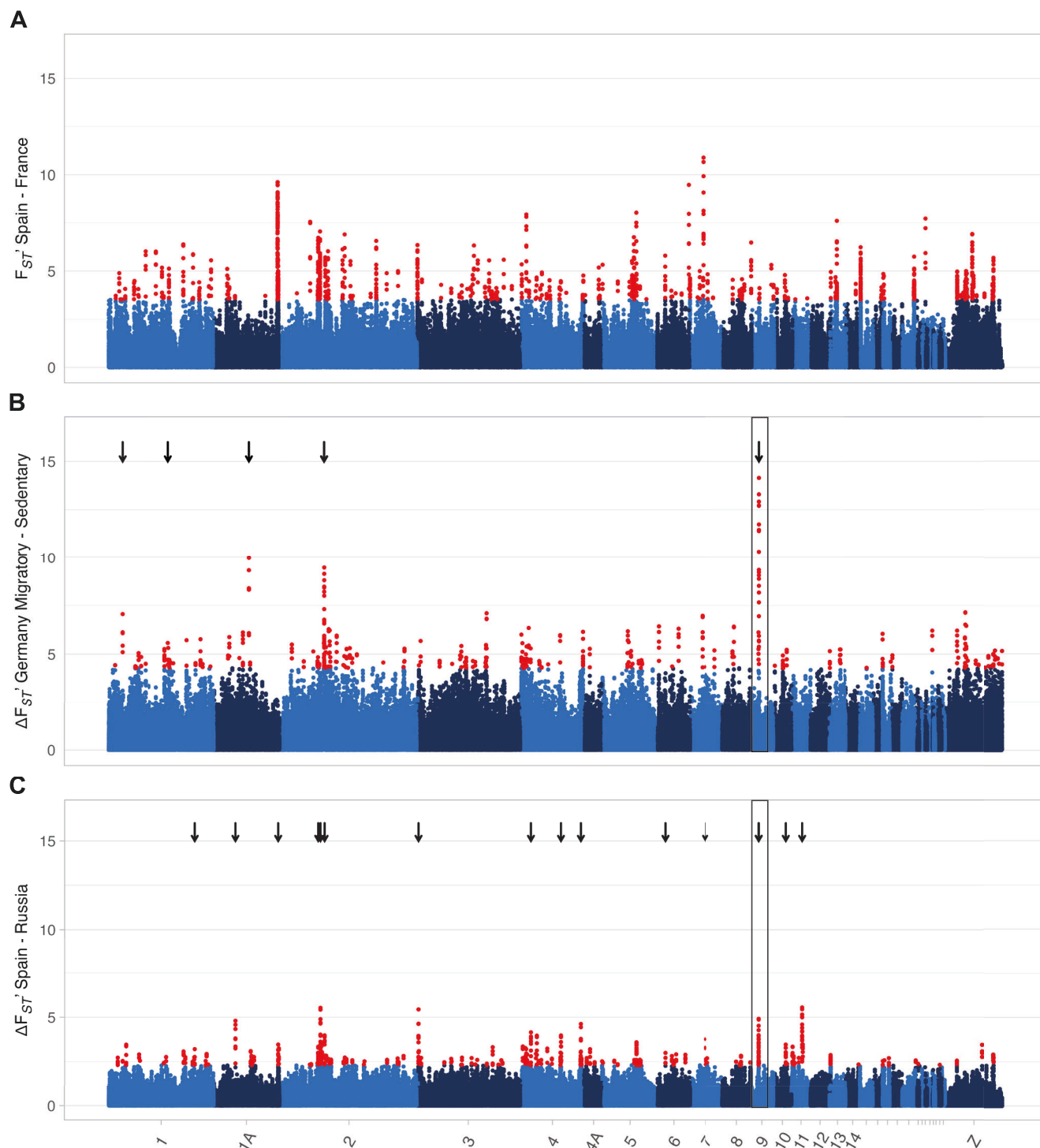
FIGURE 2 | Genome-wide differentiation scans. In (A), we compared geographically separated populations exhibiting the same migratory phenotype (Spain and France) and calculated F_{ST} in 2.5 kb windows. We then normalized these estimates by scaffold (F_{ST}') and plotted the running mean over five windows with different shades of blue corresponding to collared flycatcher chromosome models (respective chromosome number indicated on x-axis in C). Windows above the 99th percentile are shown in red. (B) Comparison between migratory and resident individuals within the partial migratory population (Germany). Net differentiation ($\Delta F_{ST}' = F_{ST\ \text{Spain-France}}' - F_{ST\ \text{GER migratory-GER resident}}'$) is plotted the same way as in (A). While the overall differentiation landscape is similarly heterogeneous compared to the within-phenotype comparison, there is one region of markedly increased differentiation on the equivalent of collared flycatcher chromosome 9 (indicated by the rectangle). $\Delta F_{ST}'$ outlier clusters, that is, five or more consecutive outlier windows are marked by black arrows. Note that only five arrows are visible due to the close proximity of clusters on chromosomes 1, 2, and 9. (C) $\Delta F_{ST}'$ for the between-phenotype comparison between Spain and Russia. A total of 73 outlier windows (in red) are shared between B and C, two of them also occurred in outlier cluster 1 of chromosome 9 in the between-phenotype comparison in the German population (indicated by the rectangle).

TABLE 1 | Summary of ΔF_{ST}' outlier clusters for the between-phenotype German comparison.

Cluster ID	Windows in cluster	Mean ΔFST'	Mean FST	Median FST	Mean Tajima's D Migratory	Mean Tajima's D resident	Chromosome
Super-Scaffold_100156_cluster401	7	11.664	0.136	0.111	−0.68	−1.688	9
Super-Scaffold_100156_cluster405	12	10.465	0.109	0.1	−0.531	−0.882	9
Super-Scaffold_100168_cluster14781	6	9.166	0.07	0.066	−0.254	−0.724	1A
Super-Scaffold_100237_cluster1039	7	8.107	0.066	0.065	−0.344	−0.387	2
Super-Scaffold_100237_cluster1044	7	8.286	0.068	0.059	−0.047	−0.14	2
Super-Scaffold_100237_cluster944	5	5.56	0.043	0.039	−0.014	−1.279	2
Super-Scaffold_100355_cluster12929	5	7.06	0.06	0.047	−0.268	−0.419	1
Super-Scaffold_100355_cluster775	5	5.573	0.052	0.051	−0.389	−0.341	1

(Table 1). This separation was also clearly visible on the haplotype structure (Figure 3A). To determine whether there are any fixed variants between the two phenotypes, we calculated F_{ST} per SNV after Weir and Cockerham (1984). Single-SNV values in the outlier cluster region on chromosome 9 corroborated the pattern observed in the ΔF_{ST}' estimates per 2.5-kb window (Figure 3B), however, we found no fixed ($F_{ST} = 1$) variant in that comparison. Notably, the largest outlier cluster region on chromosome 9 is partially overlapping the *PER2* gene (Figure 3B), and in this region we observed more heterozygous and homozygous SNVs for the alternative allele in migratory individuals (909 and 584, respectively) compared to resident ones (601 and 150, respectively).

To further investigate whether this outlier region might be under divergent selection, we estimated Tajima's *D* in 2.5 kb windows for both migratory and resident individuals in the partial migratory population in Germany. Genome-wide, Tajima's *D* estimates were on average negative with −0.602 and −0.593 for migratory and resident individuals, respectively (Figure 3C), with this slight difference (0.985-fold decrease in resident individuals) being significant (Kruskal–Wallis test, $p = 1.33^{-12}$). When focusing on windows belonging to outlier clusters of the across-phenotype comparison in the German population, we found significant differences in Tajima's *D* estimates in three outlier clusters (two of them belonging to the main differentiation peak on chromosome 9). Two clusters exhibited the most pronounced differences: cluster 1 on chromosome 9 with a mean Tajima's *D* value of −1.687 in resident, and −0.6801 in migratory individuals, respectively, which also exhibits the highest ΔF_{ST}' estimates, and cluster 1 on chromosome 2 (−0.013 and −1.278, respectively). While the overall negative values in both resident and migratory individuals suggest that selection might be acting on both phenotypes, these results might also indicate population expansion.

To expand beyond the within population comparison of resident and migratory individuals within the partial migratory German population, and to provide a broader perspective on the phenotypic contrast between migratory and resident individuals, we estimated ΔF_{ST}' (as described above) for the comparison between the Spanish (resident) and Russian (migratory) population (Figure 2C, Figure S4C for FST), and identified outliers and outlier clusters. We found 17 outlier clusters in total, spread over 10 chromosomes. Due to the substantial geographic distance between the two compared populations, the number of outlier clusters is expectedly increased (Table S2). While no outlier clusters were shared directly between this and the within-Germany comparison, we found a total of 73 outlier windows shared between comparisons, of which two from the Spain–Russia comparison were also part of outlier cluster 2 on chromosome 9 in the within Germany comparison.

3.6 | Inversion Detection

When estimating genetic differentiation between the Spanish and other populations, we noticed a peculiar pattern on the equivalent of flycatcher chromosome 9 (see Figure 4A for the Spain–Russia comparison): One of the two scaffolds that have been assigned to collared flycatcher chromosome 9 by synteny, Super-Scaffold_100189, exhibited markedly elevated F_{ST} estimates (4.95-fold compared to the genome-wide average and 5.98-fold to the other scaffold assigned to this chromosome). This pattern is observed when comparing the Spanish and other populations, and to a lesser degree in comparisons with the
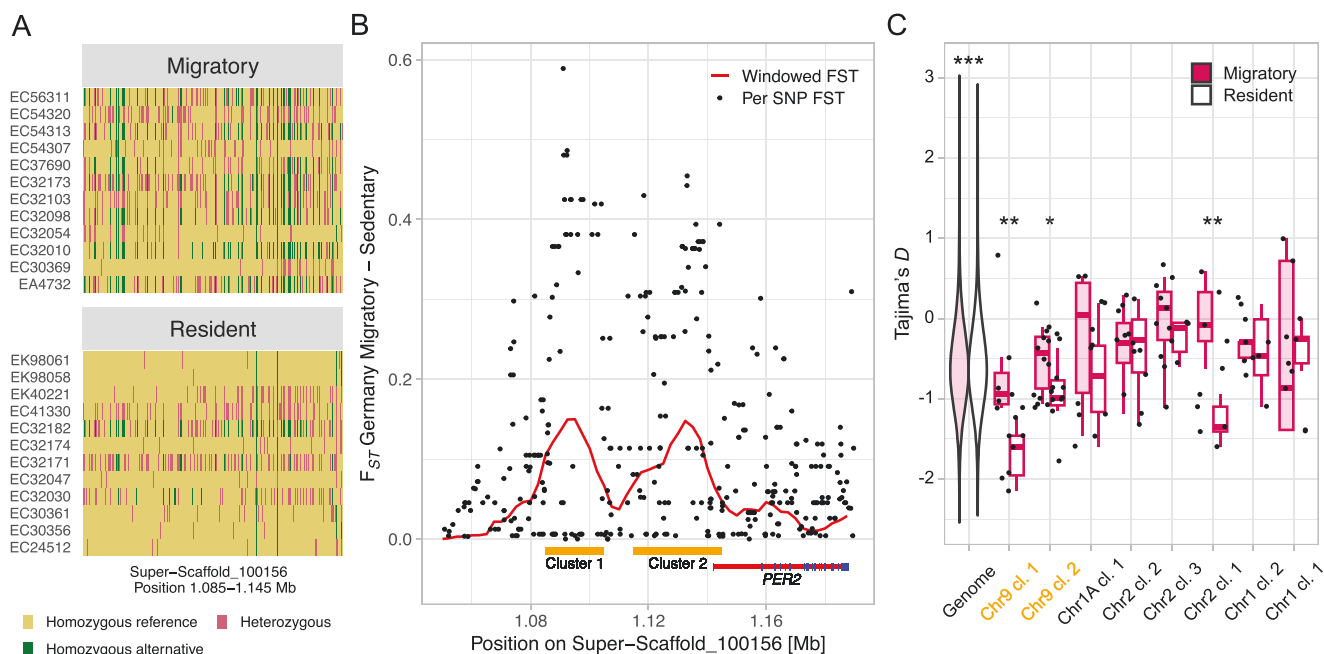
FIGURE 3 | Between-phenotype German comparison ΔF_{ST}' outlier cluster on the equivalent of flycatcher chromosome 9. (A) Haplotype plot of SNVs. Rows represent individuals, vertical tiles SNV genotypes colored according to genotype. While there are no variants fixed between the two migratory phenotypes, there is a clear separation of haplotypes visible. (B) Windowed (red line indicating the running mean) and per-SNV F_{ST} (black dots) estimates. Yellow bars indicate the two cluster regions with 7 and 12 consecutive outlier windows, respectively. The gene model of the *PER2* gene is shown in the lower right corner. (C) Tajima's D estimates in 2.5 kb windows of migratory phenotypes in the German population. The two violin plots indicate the genome-wide estimates for migratory (pink) and resident (white) phenotypes. Box- and jitter plots show estimates for each outlier cluster. Genome-wide, clusters 1 and 2 on chromosome 9 (orange as in B), and cluster 3 on chromosome 2 estimates are significantly different (Kruskal–Wallis test, $p = 1.33^{-12}$, 1.75^{-3}, 3.77^{-2}, and 9.02^{-3}, respectively).
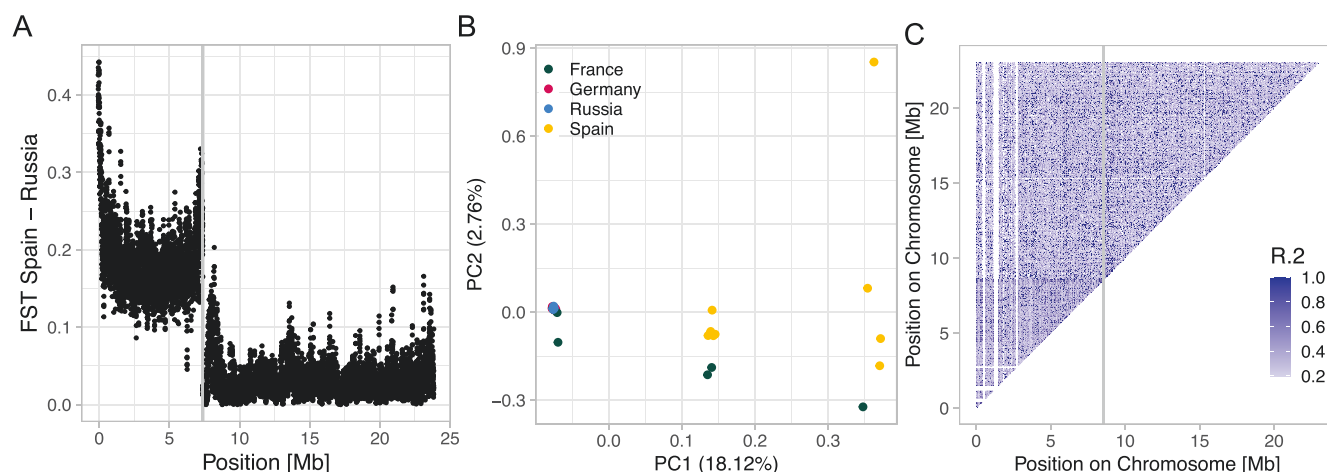
FIGURE 4 | Putative inversion on chromosome 9. (A) Running mean of F_{ST} between the Spanish and Russian population in 2.5 kb windows over chromosome 9. One of the two scaffolds assigned to this chromosome shows elevated differentiation compared to the other scaffold as well as the genome-wide average. The two scaffolds are separated by a vertical gray line. (B) PCA of individuals from all populations using only SNVs from the scaffold presumably harboring the inversion (Super-Scaffold_100189). PC1 separates individuals into three clearly distinct clusters, with the middle and right cluster containing only individuals from the Spanish and French populations. (C) Chromosome-wide linkage disequilibrium (LD) including individuals from all populations. The *x*- and *y*-axis represent physical position on chromosome 9, and tiles are colored according to LD between SNVs. Elevated LD is only observed within one scaffold, and does not extend beyond the scaffold boundary (indicated by the gray vertical line; the inverted scaffold is shown left of the gray line).

French population, but is neither visible in the within-Germany nor Germany-Russia comparison (Figure S5A,B). While the PCA using genome-wide SNVs showed more or less a pattern

expected from geographic isolation-by-distance (Figure 1B), results using SNVs only from Super-Scaffold_100189 showed a clear separation into three clusters not corresponding to

geographic structure (Figure 4B). Such a distinct pattern is indicative of an inversion segregating within a population (Huang et al. 2020; Ma and Amos 2012; Todesco et al. 2020), and when including individuals from the German and Russian populations, this pattern disappears (Figure S5C), suggesting that the inversion only segregates within the Spanish and French populations. To determine whether the putative inversion might extend beyond scaffold boundaries and reach into Super-Scaffold_100156, which harbors the main differentiation peak in the focal comparison, we estimated linkage disequilibrium (LD) over the entire chromosome (Figure 4C). Increased LD was confined to Super-Scaffold_100189, which also showed the PCA pattern indicative of an inversion (Figure 4B) and did not extend into the other scaffold.

4 | Discussion

Here we investigated the genetic basis of migratory behavior in a European songbird, the common blackbird, across the species' distribution range, using a newly assembled and annotated reference genome in combination with population whole-genome re-sequencing.

Overall population structure (Figure 1B) showed a pattern commonly seen in other pan-European species: a divergent Iberian population forming a distinct, separated cluster and individuals from populations at higher latitudes following an isolation-by-distance pattern and/or clustering more tightly together. Such a pattern is usually explained by a common biogeographic history following repeated glaciation cycles (Hewitt 2000; Vijay et al. 2016). Populations at higher latitudes therefore often show lower levels of genetic diversity due to postglacial range expansions, concordant with our observations in the German and Russian populations.

When focusing on the partial migratory German population to investigate the genetic basis of migratory behavior, we found no structure or clustering according to migratory phenotypes when using genome-wide information (Figure 1C), which suggests that the population is fully panmictic and presumably no assortative mating with respect to migratory strategy occurs. Within-population levels of genetic differentiation (both absolute and ΔF_{ST}') between migratory phenotypes were very low, a pattern also seen in other genomic comparisons of migratory phenotypes (e.g., Delmore et al. 2020, 2023 in blackcaps). However, among the outlier clusters identified, we found one region with pronounced contrasts in both absolute and relative differentiation in 2.5 kb windows, as well as per-SNV F_{ST}. This region, located on the equivalent of the collared flycatcher chromosome 9, overlaps with the *PER2* gene (Figure 3B), a member of the *clock* gene family and known to be influential in the regulation of the circadian clock (Cassone 2014). Besides influencing regulation and entrainment of the circadian rhythm, *clock* genes have also been shown to affect the circannual rhythm, orchestrating not only reproductive cycles, but also behavioral and physiological changes related to migratory behavior. In a previous study of the same partial migratory population, Franchini et al. (2017) have found *PER2* to be differentially expressed in blood samples between two groups of migratory individuals that differed in the timespan between sampling and onset of migration (8–9 days

vs. 14–18 days). Furthermore, Louder et al. (2024) have found another gene of the *clock* family, *PER3*, to be differentially expressed between migratory states in Swainson's thrushes in all brain regions. These results corroborate the important role of the *clock* genes in orchestrating and controlling migratory behavior, making a causal relationship between the differentiated region identified in our study and the propensity to migrate plausible. While we did not find any variants fully associated with the phenotype, that is, fixed SNVs, this genomic locus might still be relevant for the regulation of migratory behavior, as it can be expected that a complex behavioral phenotype is controlled by many or at least multiple variants (as shown in e.g., honey bees, Page, Rueppell, and Amdam 2012). While speculative, the fact that we found more homozygous SNVs for the alternative allele in that region potentially indicates an influence of allele number on the expression of *PER2*. Other outlier clusters were also located in or nearby genes of potential relevance for the migratory phenotype (e.g., the outlier cluster on chromosome 2 overlaps with the *TPK1* gene, which is involved in regulating metabolism). While these regions were highlighted as ΔF_{ST}' outlier clusters, their absolute differentiation is substantially lower compared to the outliers on chromosome 9 (see Table 1), suggesting that either selection acting on these regions might not have been as strong, or ΔF_{ST}' might have been artificially increased due to the very low overall differentiation.

Genome-wide estimates of Tajima's D were negative in both phenotypes of the German population, which is commonly interpreted as an indication of recent population expansion (Peart et al. 2020), and is concordant with findings in other songbird populations in higher latitudes (Poelstra et al. 2014). Although these genome-wide estimates were significantly lower in resident individuals, the difference was minute and likely not biologically relevant. In some of the ΔF_{ST}' outlier clusters however, particularly on chromosome 9, the decrease in Tajima's D in resident individuals was much more pronounced, which is possibly an indication of purifying selection (Tajima 1989; Jackson, Campos, and Zeng 2015).

On chromosome 9, we detected a putative inversion present on one of the two scaffolds assigned to this chromosome (different to the one harboring the F_{ST} outlier region). Our F_{ST} analysis suggests that the inversion segregates only in the Spanish and French populations (Figure 4A), which is supported by the PCA pattern when using individuals of all populations (Figure 4B). While the inversion does not seem to extend to the outlier region also present on the same chromosome (Figure 4C), its presence is peculiar. It is possible that there is a cline in the inversion frequency, however more sampling is necessary to support that. Segregating inversions are common in birds, with sometimes demonstrated phenotypic consequences (Sanchez-Donoso et al. 2022; Hooper and Price 2017; Knief et al. 2016), yet other times seemingly without any phenotypic or fitness effects at all. Since the assignment to the chromosome is in this case solely by synteny to other bird species, it is also possible that the scaffold harboring the inversion indeed belongs to another chromosome.

In conclusion, our results present a first step toward understanding the genetic architecture that leads to the partial migratory phenotype in a songbird species. While the region with elevated genetic differentiation on chromosome 9 overlapping

the *PER2* gene certainly does not provide proof for a genetic control of the migration behavior, it can be assumed that genetic variants present in that region contribute to the manifestation of migratory phenotypes, especially considering a polygenic control. The next step is now to take this knowledge to build an experimental framework that involves breeding and crossing of known phenotypes to tease apart the genetic underpinnings of this behavior.

Author Contributions

Matthias H. Weissensteiner: data curation (lead), formal analysis (lead), investigation (equal), validation (lead), visualization (lead), writing – original draft (equal). **Kira Delmore:** conceptualization (equal), formal analysis (supporting), investigation (supporting), writing – review and editing (supporting). **Valentina Peona:** formal analysis (supporting), visualization (supporting), writing – review and editing (supporting). **Juan Sebastian Lugo Ramos:** formal analysis (supporting), writing – review and editing (supporting). **Gregoire Arnaud:** methodology (supporting), writing – review and editing (supporting). **Julio Blas:** methodology (supporting), writing – review and editing (supporting). **Bruno Faivre:** methodology (supporting), writing – review and editing (supporting). **Ivan Pokrovsky:** methodology (supporting), writing – review and editing (supporting). **Martin Wikelski:** conceptualization (equal), investigation (equal), methodology (equal), project administration (equal). **Jesko Partecke:** conceptualization (equal), investigation (equal), methodology (lead), project administration (equal), writing – review and editing (supporting). **Miriam Liedvogel:** conceptualization (lead), funding acquisition (lead), investigation (lead), project administration (lead), supervision (equal), writing – original draft (equal).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Whole-genome sequencing data produced during this project is available at the European Nucleotide Archive under project ID PRJEB76551 (accessions for each individual in Table S1).

References

Bensch, S., M. Grahn, N. Müller, L. Gay, and S. Akesson. 2009. "Genetic, Morphological, and Feather Isotope Variation of Migratory Willow Warblers Show Gradual Divergence in a Ring." *Molecular Ecology* 18, no. 14: 3087–3096.

Berthold, P., A. J. Helbig, G. Mohr, and U. Querner. 1992. "Rapid Microevolution of Migratory Behaviour in a Wild Bird Species." *Nature* 360, no. 6405: 668–670.

Berthold, P., and U. Querner. 1981. "Genetic Basis of Migratory Behavior in European Warblers." *Science* 212, no. 4490: 77–79.

Bravo, G. A., C. J. Schmitt, and S. V. Edwards. 2021. "What Have we Learned From the First 500 Avian Genomes?" *Annual Review of Ecology, Evolution, and Systematics* 52: 611–639.

Cantarel, B. L., I. Korf, S. M. Robb, et al. 2008. "MAKER: An Easy-To-Use Annotation Pipeline Designed for Emerging Model Organism Genomes." *Genome Research* 18, no. 1: 188–196.

Cassone, V. M. 2014. "Avian Circadian Organization: A Chorus of Clocks." *Frontiers in Neuroendocrinology* 35, no. 1: 76–88.

Chernetsov, N., D. Kishkinev, and H. Mouritsen. 2008. "A Long-Distance Avian Migrant Compensates for Longitudinal Displacement During Spring Migration." *Current Biology: CB* 18, no. 3: 188–190.

Chin, C. S., P. Peluso, F. J. Sedlazeck, et al. 2016. "Phased Diploid Genome Assembly With Single-Molecule Real-Time Sequencing." *Nature Methods* 13, no. 12: 1050–1054.

Crofoot, M. C., I. C. Gilby, M. C. Wikelski, and R. W. Kays. 2008. "Interaction Location Outweighs the Competitive Advantage of Numerical Superiority in *Cebus Capucinus* Intergroup Contests." *Proceedings of the National Academy of Sciences of the United States of America* 105, no. 2: 577–581.

Danecek, P., A. Auton, G. Abecasis, et al. 2011. "The Variant Call Format and VCFtools." *Bioinformatics* 27, no. 15: 2156–2158.

Delmore, K., J. C. Illera, J. Perez-Tris, G. Segelbacher, J. S. Lugo Ramos, and G. Durieux. 2020. "The Evolutionary History and Genomics of European Blackcap Migration." *eLife* 9, e54462. April. https://doi.org/10.7554/eLife.54462.

Delmore, K. E., and D. E. Irwin. 2014. "Hybrid Songbirds Employ Intermediate Routes in a Migratory Divide." *Ecology Letters* 17, no. 10: 1211–1218.

Delmore, K. E., D. P. Toews, R. R. Germain, G. L. Owens, and D. E. Irwin. 2016. "The Genetics of Seasonal Migration and Plumage Color." *Current Biology: CB* 26, no. 16: 2167–2173.

Delmore, K. E., B. M. van Doren, K. Ullrich, T. Curk, H. P. van der Jeugd, and M. Liedvogel. 2023. "Structural Genomic Variation and Migratory Behavior in a Wild Songbird." *Evolution Letters* 7, no. 6: 401–412.

Ellegren, H., L. Smeds, R. Burri, P. I. Olason, N. Backström, and T. Kawakami. 2012. "The Genomic Landscape of Species Divergence in Ficedula Flycatchers." *Nature* 491, no. 7426: 756–760.

Ellinghaus, D., S. Kurtz, and U. Willhoeft. 2008. "LTRharvest, an Efficient and Flexible Software for de Novo Detection of LTR Retrotransposons." *BMC Bioinformatics* 9, no. January: 18.

Faircloth, B. C., J. E. McCormack, N. G. Crawford, M. G. Harvey, R. T. Brumfield, and T. C. Glenn. 2012. "Ultraconserved Elements Anchor Thousands of Genetic Markers Spanning Multiple Evolutionary Timescales." *Systematic Biology* 61, no. 5: 717–726.

Flynn, J. M., R. Hubley, C. Goubert, et al. 2020. "RepeatModeler2 for Automated Genomic Discovery of Transposable Element Families." *Proceedings of the National Academy of Sciences of the United States of America* 117, no. 17: 9451–9457.

Franchini, P., I. Irisarri, A. Fudickar, et al. 2017. "Animal Tracking Meets Migration Genomics: Transcriptomic Analysis of a Partially Migratory Bird Species." *Molecular Ecology* 26, no. 12: 3204–3216.

Fudickar, A. M., A. Schmidt, M. Hau, M. Quetting, and J. Partecke. 2013. "Female-Biased Obligate Strategies in a Partially Migrator y Population." *Journal of Animal Ecology* 82, no. 4: 863–871.

Grabherr, M. G., P. Russell, M. Meyer, et al. 2010. "Genome-Wide Synteny Through Highly Sensitive Sequence Alignment: Satsuma." *Bioinformatics* 26, no. 9: 1145–1151.

Hegemann, A., A. M. Fudickar, and J. Å. Nilsson. 2019. "A Physiological Perspective on the Ecology and Evolution of Partial Migration." *Journal of Ornithology* 160, no. 3: 893–905.

Hegemann, A., P. P. Marra, and B. I. Tieleman. 2015. "Causes and Consequences of Partial Migration in a Passerine Bird." *American Naturalist* 186, no. 4: 531–546.

Helbig, A. 1996. "Genetic Basis, Mode of Inheritance and Evolutionary Changes of Migratory Directions in Palaearctic Warblers (Aves: Sylviidae)." *Journal of Experimental Biology* 199, no. Pt 1: 49–55.

Helbig, A. J. 1991. "Inheritance of Migratory Direction in a Bird Species: A Cross-Breeding Experiment With SE- and SW-Migrating Blackcaps (*Sylvia atricapilla*)." *Behavioral Ecology and Sociobiology* 28, no. 1: 9–12.

Hewitt, G. 2000. "The Genetic Legacy of the Quaternary Ice Ages." *Nature* 405, no. 6789: 907–913.

Hooper, D. M., and T. D. Price. 2017. "Chromosomal Inversion Differences Correlate With Range Overlap in Passerine Birds." *Nature Ecology & Evolution* 1, no. 10: 1526–1534.

Huang, K., R. L. Andrew, G. L. Owens, K. L. Ostevik, and L. H. Rieseberg. 2020. "Multiple Chromosomal Inversions Contribute to Adaptive Divergence of a Dune Sunflower Ecotype." *Molecular Ecology* 29, no. 14: 2535–2549.

Hudson, R. R., M. Slatkin, and W. P. Maddison. 1992. "Estimation of Levels of Gene Flow From DNA Sequence Data." *Genetics* 132, no. 2: 583–589.

Jackson, B. C., J. L. Campos, and K. Zeng. 2015. "The Effects of Purifying Selection on Patterns of Genetic Differentiation Between *Drosophila melanogaster* Populations." *Heredity* 114, no. 2: 163–174.

Walter, J., G. Tertitski, R. Kays, U. Mueller, and M. Wikelski. 2022. "Biological Earth Observation With Animal Sensors." *Trends in Ecology & Evolution* 37, no. 8: 719–724.

Kapusta, A., and A. Suh. 2017. "Evolution of Bird Genomes-a Transposon's-Eye View." *Annals of the New York Academy of Sciences* 1389, no. 1: 164–185.

Kays, R., S. Tilak, M. Crofoot, et al. 2011. "Tracking Animal Location and Activity With an Automated Radio Telemetry System in a Tropical Rainforest." *Computer Journal* 54, no. 12: 1931–1948.

Knief, U., G. Hemmrich-Stanisak, M. Wittig, et al. 2016. "Fitness Consequences of Polymorphic Inversions in the Zebra Finch Genome." *Genome Biology* 17, no. 1: 199.

le Corre, V., and A. Kremer. 2012. "The Genetic Differentiation at Quantitative Trait Loci Under Local Adaptation." *Molecular Ecology* 21, no. 7: 1548–1566.

Linek, N., P. Brzęk, M. Phillip Gienapp, et al. 2021. "A Partial Migrant Relies Upon a Range-Wide Cue Set but Uses Population-Specific Weighting for Migratory Timing." *Movement Ecology* 9, no. 1: 63.

Louder, M. I. M., H. Justen, A. A. Kimmitt, et al. 2024. "Gene Regulation and Speciation in a Migratory Divide Between Songbirds." *Nature Communications* 15, no. 1: 98.

Lundberg, M., M. Liedvogel, K. Larson, et al. 2017. "Genetic Differences Between Willow Warbler Migratory Phenotypes Are Few and Cluster in Large Haplotype Blocks." *Evolution Letters* 1, no. 3: 155–168.

Ma, J., and C. I. Amos. 2012. "Investigation of Inversion Polymorphisms in the Human Genome Using Principal Components Analysis." *PLoS One* 7, no. 7: e40224.

Mettler, R., H. Martin Schaefer, N. Chernetsov, et al. 2013. "Contrasting Patterns of Genetic Differentiation Among Blackcaps (*Sylvia atricapilla*) With Divergent Migratory Orientations in Europe." *PLoS One* 8, no. 11: e81365.

Page, R. E., Jr., O. Rueppell, and G. V. Amdam. 2012. "Genetics of Reproduction and Regulation of Honeybee (*Apis mellifera* L.) Social Behavior." *Annual Review of Genetics* 46, no. August: 97–119.

Peart, C. R., S. Tusso, S. D. Pophaly, et al. 2020. "Determinants of Genetic Variation Across Eco-Evolutionary Scales in Pinnipeds." *Nature Ecology & Evolution* 4, no. 8: 1095–1104.

Peona, V., M. P. K. Blom, X. Luohao, et al. 2021a. "Identifying the Causes and Consequences of Assembly Gaps Using a Multiplatform Genome Assembly of a Bird-of-Paradise." *Molecular Ecology Resources* 21, no. 1: 263–286.

Peona, V., O. M. Palacios-Gimenez, J. Blommaert, et al. 2021b. "The Avian W Chromosome Is a Refugium for Endogenous Retroviruses With Likely Effects on Female-Biased Mutational Load and Genetic Incompatibilities." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 376: 20200186.

Perdeck, A. C. 1958. "Two Types of Orientation in Migrating Starlings, *Sturnus vulgaris* L., and Chaffinches, *Fringilla coelebs* L., as Revealed by Displacement Experiments." *Ardea* 55, no. 1–2: 1–2.

Plummer, K. E., G. M. Siriwardena, G. J. Conway, K. Risely, and M. P. Toms. 2015. "Is Supplementary Feeding in Gardens a Driver of Evolutionary Change in a Migratory Bird Species?" *Global Change Biology* 21, no. 12: 4353–4363.

Poelstra, J. W., N. Vijay, C. M. Bossu, et al. 2014. "The Genomic Landscape Underlying Phenotypic Integrity in the Face of Gene Flow in Crows." *Science* 344, no. 6190: 1410–1414.

Pulido, F. 2011. "Evolutionary Genetics of Partial Migration—The Threshold Model of Migration Revis(it)ed." *Oikos* 120, no. 12: 1776–1783.

Pulido, F., and P. Berthold. 2010. "Current Selection for Lower Migratory Activity Will Drive the Evolution of Residency in a Migratory Bird Population." *Proceedings of the National Academy of Sciences of the United States of America* 107, no. 16: 7341–7346.

Purcell, S., B. Neale, K. Todd-Brown, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81, no. 3: 559–575.

Quinlan, A. R., and I. M. Hall. 2010. "BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features." *Bioinformatics* 26, no. 6: 841–842.

Sanchez-Donoso, I., J. Sara Ravagni, D. Rodríguez-Teijeiro, et al. 2022. "Massive Genome Inversion Drives Coexistence of Divergent Morphs in Common Quails." *Current Biology: CB* 32, no. 2: 462–469.

Smit, A. F. A., R. Hubley, and P. Green. 2015. "RepeatMasker Open-4.0." http://www.repeatmasker.org.

Sokolovskis, K., M. Lundberg, M. Liedvogel, et al. 2019. "Phenotypic and Genetic Characterization of the East Siberian Willow Warbler (*Phylloscopus trochilus yakutensis* Ticehurst, 1935) in Relation to the European Subspecies." *Journal of Ornithology* 160, no. 3: 721–731.

Steinbiss, S., U. Willhoeft, G. Gremme, and S. Kurtz. 2009. "Fine-Grained Annotation and Classification of de Novo Predicted LTR Retrotransposons." *Nucleic Acids Research* 37, no. 21: 7002–7013.

Tajima, F. 1989. "Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism." *Genetics* 123, no. 3: 585–595.

Thorup, K., I.-A. Bisson, M. S. Bowlin, et al. 2007. "Evidence for a Navigational Map Stretching Across the Continental U.S. in a Migratory Songbird." *Proceedings of the National Academy of Sciences of the United States of America* 104, no. 46: 18115–18119.

Todesco, M., G. L. Owens, N. Bercovich, et al. 2020. "Massive Haplotypes Underlie Ecotypic Differentiation in Sunflowers." *Nature* 584, no. 7822: 602–607.

Vasimuddin, M., S. Misra, H. Li, and S. Aluru. 2019. "Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems." In *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 314–324. IEEE.

Vijay, N., C. M. Bossu, J. W. Poelstra, et al. 2016. "Evolution of Heterogeneous Genome Differentiation Across Multiple Contact Zones in a Crow Species Complex." *Nature Communications* 7: 13195.

Weir, B. S., and C. Clark Cockerham. 1984. "Estimating F-Statistics for the Analysis of Population Structure." *Evolution; International Journal of Organic Evolution* 38, no. 6: 1358–1370.

Wickham, H., D. Vaughan, and M. Girlich. 2024. "Tidyr: Tidy Messy Data. R Package Version 1.3.1." https://github.com/tidyverse/tidyr, https://tidyr.tidyverse.org.

Zheng, X., D. Levine, J. Shen, S. M. Gogarten, C. Laurie, and B. S. Weir. 2012. "A High-Performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data." *Bioinformatics* 28, no. 24: 3326–3328.

Zúñiga, D., J. Falconer, A. M. Fudickar, et al. 2016. "Abrupt Switch to Migratory Night Flight in a Wild Migratory Songbird." *Scientific Reports* 6: 34207.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.