Deciphering the historical recombination landscape of a migratory songbird, the Eurasian blackcap

Dissertation

in fulfilment of the requirements for the degree

"Doctor rerum naturalium"

of the Faculty of Mathematics and Natural Sciences

at the Christian Albrechts University of Kiel

Submitted by

Karen Bascón-Cardozo

Max Planck Institute for Evolutionary Biology

Plön, November 2022

First examiner: Prof. Dr. Miriam Liedvogel

Second examiner: Prof. Dr. Eva H. Stukenbrock

Date of oral examination: 24th of February, 2023

"Great things are not done by impulse, but by a series of small things brought together"

V.Van Gogh

Abstract

Meiotic recombination involves the exchange of genetic information between parental chromosomes generating new combinations of alleles in the offspring. Recombination is a crucial determinant for reproductive success, proper chromosomal segregation, genetic variability, adaptation rates, and influencing the efficacy of selection. The historical recombination landscape comprises all recombination events captured in DNA sequences among individuals over many generations. Recombination involves different molecular mechanisms and recombining events are not spread randomly across the genome. The placement of recombination events in most mammals is defined by the protein PRDM9, a key player in the initiation of recombination dynamics. However, some organismal groups, such as birds, have lost this protein enabling them with unique recombining properties and conferring an apparent stasis in the positioning of recombination events. However, recent studies in species lacking PRDM9 demonstrated that recombination rates differ widely across the genome and between species. The variation in recombination rates, its causes, and its impact on evolutionary and behavioural processes remain poorly understood, particularly in natural bird populations. Here, I studied the variation of the historical recombination landscape in the Eurasian blackcap, *Sylvia atricapilla,* and the interplay of recombination rate variation with specific genomic and epigenetic features at different scales: across the genome, at inter- and intraspecific levels.

The Eurasian blackcap is an excellent model for bird migration and evolutionary research due to its vast repertoire of migratory phenotypes across a wide geographical range. The initial characterisation of recombination rates at the genome level revealed a heterogeneous recombination landscape in this species. Specific genomic features such as nucleotide diversity, GC content, CpG islands, and retrotransposons contribute to this variation. The analysis including annotation features confirmed that most recombination events are prevalent in regulatory regions, particularly cis-regulatory regions. Additionally, the work presented here reveals that the density of CpG islands, location, and methylation patterns are essential determinants in shaping the recombination landscape in blackcaps. Increased methylation levels of shared CpG sites among different tissues were associated with low recombining regions, particularly in shorter chromosomes. Our analyses within the framework of TE characterisation suggested that the association with recombination rates could be influenced by TE family, location, and age.

By including the Garden warbler (*Sylvia borin*), the closest sister taxon of blackcaps, I assessed the variation in recombination rates at inter-specific scales. This analysis showed a genome-wide correlation between recombination maps of both species. However, the intra-chromosomal patterns revealed both conserved and divergent recombining regions between species. These results highlight that the recombination landscape in the blackcap is not entirely stable, particularly at finer scales. This led me to expand the historical recombination landscape evaluation to the population level.

This project presents an in-depth evaluation of recombination variation among phenotypically distinct migratory blackcap populations and reveals variable conservation of recombination maps at the intraspecific level. The comparison of continental populations with island populations revealed the most striking differences. Interestingly, the correlation level between recombination maps could reflect the genetic structure and demographic history among populations. The conducted population genomic analysis supported recombination as an essential determinant in population differentiation. There is a positive association between recombination divergence and population differentiation. Additionally, the interplay between recombination rates with nucleotide diversity and population divergence suggests that recombination contributes to maintaining genetic variation and influences the efficacy of linked selection which appears to shape genomic variation and population differentiation in blackcaps.

This thesis provides a more profound understanding of recombination rate variation at different scales. The findings presented here highlight that recombination is a variable trait that may impact the course of the evolution of genomes, populations, and species.

Kurzfassung

Bei der meiotischen Rekombination werden genetische Informationen zwischen den elterlichen Chromosomen ausgetauscht, wodurch neue Allelkombinationen in den Nachkommen entstehen. Die Rekombination ist ein entscheidender Faktor für den Fortpflanzungserfolg, die korrekte chromosomale Segregation, die genetische Variabilität, die Anpassungsraten und die Wirksamkeit der Selektion. Die historische Rekombinationslandschaft umfasst alle Rekombinationsereignisse, die in DNA-Sequenzen zwischen Individuen über viele Generationen hinweg erfasst wurden. Rekombination involviert verschiedene molekulare Mechanismen, und die Rekombinationsereignisse sind nicht zufällig über das Genom verteilt. Das Auftreten von Rekombinationsereignissen an bestimmten Stellen im Genom wird bei den meisten Säugetieren durch das Protein PRDM9 bestimmt, welches eine Schlüsselrolle bei der Initiierung der Rekombinationsdynamik spielt. Einige Organismengruppen, wie z. B. Vögel, haben dieses Protein jedoch verloren, wodurch sie über einzigartige Rekombinationsmerkmale verfügen und die Positionierung der Rekombinationsereignisse scheinbar stagniert. Jüngste Studien an Arten, denen PRDM9 fehlt, haben jedoch gezeigt, dass die Rekombinationsraten im gesamten Genom und zwischen den Arten stark variieren. Die Variation der Rekombinationsraten, ihre Ursachen und ihre Auswirkungen auf Evolutions- und Verhaltensprozesse sind nach wie vor kaum verstanden, insbesondere in wilden Vogelpopulationen. In dieser Arbeit habe ich die Variation der historischen Rekombinationslandschaft bei der Mönchsgrasmücke (*Sylvia atricapilla*) und das Zusammenspiel der Rekombinationsratenvariation mit spezifischen genomischen und epigenetischen Merkmalen auf verschiedenen Ebenen untersucht: im gesamten Genom sowie auf inter- und intraspezifischer Ebene.

Die Mönchsgrasmücke ist ein hervorragendes Modell für die Vogelzug- und Evolutionsforschung, da sie über ein großes Repertoire an Zugphänotypen in einem weiten geografischen Bereich verfügt. Die erste Charakterisierung der Rekombinationsraten auf Genomebene ergab eine heterogene Rekombinationslandschaft bei dieser Art. Spezifische genomische Merkmale wie Nukleotiddiversität, GC-Gehalt, CpG-Inseln und Retrotransposons tragen zu dieser Variation bei. Die Analyse unter Einbeziehung von Annotationsmerkmalen bestätigte, dass die meisten Rekombinationsereignisse in regulatorischen Regionen, insbesondere in cis-regulatorischen Regionen, vorkommen. Darüber hinaus zeigt die hier vorgestellte Arbeit, dass die Dichte der CpG-Inseln, die Lage und die Methylierungsmuster wesentliche Determinanten für die Gestaltung der

Rekombinationslandschaft in Mönchsgrasmücken sind. Erhöhte Methylierungsniveaus gemeinsamer CpG-Inseln in verschiedenen Geweben wurden mit Regionen mit geringer Rekombination in Verbindung gebracht, insbesondere in kürzeren Chromosomen. Unsere Analysen im Rahmen der TE-Charakterisierung deuteten darauf hin, dass deren Assoziation mit den Rekombinationsraten durch die TE-Familie, die Position und dasAlter beeinflusst werden könnte.

Unter Einbeziehung der Gartengrasmücke (*Sylvia borin*), dem nächsten Schwestertaxon der Mönchsgrasmücke, habe ich die Variation der Rekombinationsraten auf interspezifischer Ebene untersucht. Diese Analyse zeigte eine genomweite Korrelation zwischen den Rekombinationskarten der beiden Arten. Die intrachromosomalen Muster zeigten jedoch sowohl konservierte als auch divergierende Rekombinationsregionen zwischen den Arten. Diese Ergebnisse zeigen, dass die Rekombinationslandschaft der Mönchsgrasmücke nicht völlig stabil ist, insbesondere auf feineren Skalen. Dies veranlasste mich, die Auswertung der historischen Rekombinationslandschaft auf die Populationsebene auszuweiten.

In diesem Projekt wird eine eingehende Bewertung der Rekombinationsvariation zwischen phänotypisch unterschiedlich zieenden Mönchsgrasmückenpopulationen vorgenommen und die unterschiedliche Erhaltung von Rekombinationskarten auf intraspezifischer Ebene aufgezeigt. Der Vergleich von Kontinentalpopulationen mit Inselpopulationen ergab die auffälligsten Unterschiede. Interessanterweise könnte das Korrelationsniveau zwischen Rekombinationskarten die genetische Struktur und die demografische Geschichte der Populationen widerspiegeln. Die durchgeführten Populationsgenomanalysen belegen, dass die Rekombination eine wesentliche Determinante der Populationsdifferenzierung ist. Es besteht ein positiver Zusammenhang zwischen Rekombinationsdivergenz und Populationsdifferenzierung. Die Wechselwirkung zwischen Rekombinationsraten, Nukleotiddiversität und Populationsdivergenz deutet darauf hin, dass die Rekombination zur Aufrechterhaltung der genetischen Variation beiträgt und die Wirksamkeit der damit verbundenen Selektion beeinflusst, die offenbar die genomische Variation und die Populationsdifferenzierung bei Mönchsgrasmücken beeinflusst.

Die vorliegende Arbeit ermöglicht ein tieferes Verständnis der Variation der Rekombinationsrate auf verschiedenen Ebenen. Die hier vorgestellten Ergebnisse zeigen, dass die Rekombination ein variables Merkmal ist, das den Verlauf der Evolution von Genomen, Populationen und Arten beeinflussen kann.

Table of Contents

CHAPTER I

General Introduction

Understanding the evolution of complex traits involves disentangling the dynamic of intricate systems which contribute to the adaptation of individuals, populations, and species over time. Diversity is a major factor that enables organisms to cope with harsh conditions and adapt to fluctuating environments. One of the primary sources of genetic diversity is recombining genetic material and creating new phenotypes in the progeny. Studying the processes that generate genetic variability and adaptation is one way to broaden our understanding of behavioural and evolutionary traits.

1.1 Recombination: the process

Sexual reproduction in eukaryotes produces genetically diverse individuals, which involves meiosis; the division of cells where diploid germ cells differentiate into haploid gametes. Meiotic division involves a single round of DNA replication of parental chromosomes followed by two rounds of chromosome segregation. Before the first meiotic division (during prophase-I, see Fig. 1), homologous chromosomes from different parental origins duplicate the genetic material and receive programmed DNA double-strand breaks (DSBs). To repair and complete the gap induced by DSBs, sister chromatids of one chromosome search for their homologous chromosome to pair in these specific regions known as 'chiasmata'. Chromosome pairing is followed by the formation of a protein structure called the 'synaptonemal complex' (SC) (Fawcett, 1956; Moses, 1956). Here is where the process of recombination occurs, DNA information is shuffled and swapped, creating a new combination of loci and alleles in the daughter cells. DSBs can be resolved in two main ways, either as cross-overs (COs) resulting in the reciprocal exchange of genetic material (DNA) or non-crossovers, the non-reciprocal exchange of DNA, which can result in gene-conversion (GC) (reviewed in Lam & Keeney, 2015; Lenormand et al., 2016) (Fig. 1). Mainly, COs ensure the proper homolog segregation, while generating new haplotype combinations that increase genome diversity in the progeny.

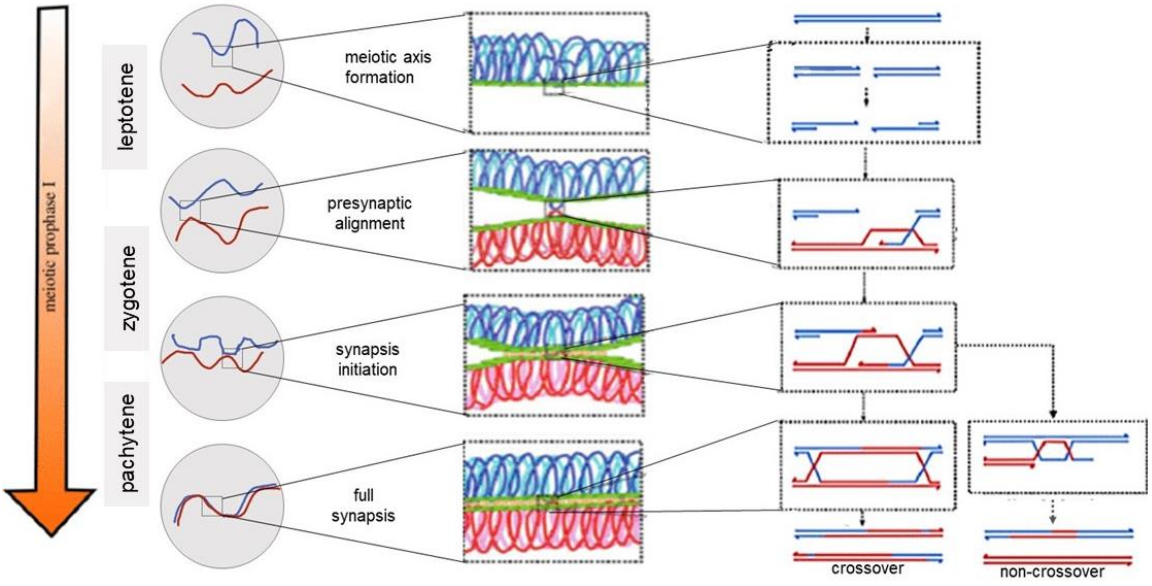


Figure 1. Mechanisms of meiotic recombination during meiosis. The various stages of recombination that occur during prophase I are illustrated and include global interactions across entire chromosomes and local interactions at the single-molecule level. Blue and red strands represent homologous chromosomes, and the meiotic axis and synaptonemal complex are coloured green and orange, respectively. Figure and legend adapted from Morgan et al.(2017).

1.2 Different approaches to measure recombination

The places where recombination events take place are not randomly distributed along the genome. CO interference ensures a specific distance between two different COs along a chromosome (Muller,1916; Sturtevant, 1915). At least one recombination event (CO) per homologous chromosome pair per meiosis is essential for proper chromosome segregation; some vertebrates can also show more COs per chromosome (Calderón & Pigozzi, 2006; del Priore & Pigozzi, 2020).

Across the genome, regions where the majority of recombination events cluster are called "hot spots" or high recombining regions (HRR), and regions where recombination is suppressed are "cold spots" or low recombining regions (LRR). The evolution of high and low recombination rates will have different causes and implications according to the organism, genome structure, molecular dynamics, and genetic regulation (reviewed in Peñalba & Wolf, 2020; Stapley et al., 2017).

Studying the recombination process involves disentangling different scales from diverse perspectives. Recombination landscapes or maps capture the distribution of recombination events (COs) and how often haplotypes are shuffled in different parts of the genome. The characterisation

of recombination maps is usually based on two main estimates: the number or frequency of COs (at different scales: in a given meiosis, across generations, between genes or loci) and the location of COs in the genome. Direct and indirect methods exist to infer recombination landscapes according to the scale and focus of interest. Based on that, recombination is typically expressed as either recombination frequency or recombination rates represented in centimorgan per megabase (in cM/Mb) per generation.

Direct methods include cytological approaches tracking meiotic processes through immunostaining assays on single meiotic cells. The frequency of recombination events (CO) is estimated by counting the number of recombination nodules (Anderson et al., 2003) present during crossing-over (pachytene stages, Fig. 1) or the number of chiasmata (Henderson & Edwards, 1968; Pollock & Fechheimer, 1978). Another approach is staining specific meiotic proteins such as the mismatch repair protein MLH1on SCs and counting foci to estimate COs frequency and distribution (Anderson et al., 1999) across different chromosomes. These methods estimate the recombination frequency of an individual as the average number of crossovers among cells. More advanced method have been developed in the last few years, such as using chromatin immunoprecipitation and targeting specific recombinases (e.g. RAD51 and DMC1) followed by sequencing (ChIP-seq) to map meiotic DNA double-strand breaks (Smagulova et al., 2011). Direct methods can provide insight into the recombination process based on the cellular context: the different stages of recombination, the mechanisms controlling crossover frequency and interference, and the effects of chromosomal translocations or polymorphic inversions.

Indirect or genomic methods are more suitable for assessing recombination at bigger scales and include *pedigree* approaches where parent and offspring genotypes are compared to evaluate the transmission of alleles (DNA markers) on chromosomes between generations. The frequency or fraction of recombination between loci or a pair of markers is calculated and converted to a genetic distance (or genetic map) measured in centiMorgan (cM). One centiMorgan (cM) is equal to a 1 % chance that two markers on a chromosome will become separated from one another due to a recombination event during meiosis. The generation of a linkage map enables the examination of recombination rate variation along chromosomes. With the pedigree-approach, it is possible to differentiate recombination rates between sexes and even among individuals of the parental generation (Johnston et al., 2016; Weng et al., 2019). However, this approach is more suitable for species with short generation times and reared in captivity or within a long-term controlled

breeding system. Another method to quantify the frequency of COs based on haploid gametes is *sperm typing* which will yield recombination rates specific to the male sex. These two approaches assess recombination at local scales but may not capture the historical recombination landscape of a focal species and the resolution among populations across different generations. The population-scaled recombination approach is suitable for this purpose and to study non-model species and natural populations from which pedigree data is challenging to get in natural conditions.

The *population-scaled* approach, also known as the LD-based approach, relies on linkage disequilibrium (LD) or coalescence following the rationale that recombination is the main force to break the non-random physical linkage between loci. This approach detects the number of recombination events that have occurred in the history of a sample across many generations, yielding sex- and time-averaged historical recombination rates from unrelated individuals (Spence & Song, 2019). This is represented as $\rho = 2cN_e r$, where the population recombination rate (ρ) is equal to the recombination rate per base-pair per generation (r) scaled by the effective population size (N_e) and multiplied by the ploidy of the genome (c). LD-based inference provides a genome-wide, high-resolution map with a small number of individuals and is widely used in population genetics. It is essential to consider that population-scale recombination rate approaches rely on some assumptions such as constant population size (Ne), mutation rate, and the absence of population migration and genetic drift (Chapman & Thompson, 2001; Dapper & Payseur, 2018). However, these assumptions are not always met, especially in natural populations, which can generate a considerable distortion in the recombination estimation (Dapper & Payseur, 2018). New programs have been developed in the last years to alleviate those potential biases, including the population-specific demography for the recombination inference (Hermann et al., 2019; Spence & Song, 2019).

1.3 Recombination variability at different scales and possible causes

At broad (genome-wide) and fine scales recombination rates and their distribution can vary widely across eukaryotes, closely related taxa, populations, individuals, and within the genome. Recombination variation can be the cause or consequence of several phenomena. But why do recombination rates vary, and what factors influence this variation? Here I explain some of them including different scales.

a) Variation across taxonomic groups

Among different factors influencing historical recombination rate variation between other groups of organisms, some of the most important are: i) **Genome size and composition;** initially, it has been suggested that larger genomes have lower recombination rates. Stapley et al., 2017 found that the genome-wide recombination rate is lower in larger plant genomes but there is not enough evidence in animals and fungi to suggest a decrease in recombination rate with genome size. This relation will also depend on the composition and architecture of the genome. Large genomes usually contain a higher density of repetitive elements and sequences, which is generally associated negatively with recombination rates (Kent et al., 2017; Tiley & Burleigh, 2015). ii) The **genome organization** (including karyotype) may also influence recombination variation. Following the assumption that at least one CO is required for proper segregation, increased genome-wide recombination rates may result from the larger number of chromosomes in a specific genome (Fledel-Alon et al., 2009). Perhaps that could explain why recombination rates are higher, for example, in birds (2n=~80 chromosomes) compared to humans (2n=46). Nevertheless, not only the number but also the size of chromosomes may be correlated with variations in recombination rates. Organisms such as birds, some reptiles, and plants contain small chromosomes with increased recombination rates (Bascón-Cardozo et al., 2022a; Brazier & Glémin, 2022; Schield et al., 2020; Smeds et al., 2016). In contrast, mammals with lower recombination rates have more evenly sized chromosomes (e.g. humans). iii) **Molecular mechanisms** influencing recombination dynamics contribute to recombination rate variation. Most mammals possess the PR-domain containing protein 9 (PRDM9), which directs DSB formation and recombination hotspots in mice and humans (Baudat et al., 2010; Berg et al., 2011; Myers et al., 2010; Parvanov et al., 2010). *Prdm9* evolves rapidly resulting in less conserved recombination landscapes in these organisms. However, other mechanisms exist for organisms that lack this protein (e.g. Drosophila, yeast, dogs, birds, and most plants). In these organisms, recombination events are located in euchromatic regions associated with regulatory elements such as promoters suggesting a more conserved location of recombination events, like in birds and yeast (Singhal et al., 2015; Tsai et al., 2010). Notably, the variation in recombination dynamics suggests that recombination may evolve and play different roles according to the organism and its specific genomic mechanisms.

b) Variation among and within species

Recombination rates among species are more conserved at broad scales (genome-wide), while differences may be more evident at fine or local scales even in closely related species(Auton et al., 2012(Auton et al., 2012; Kawakami et al., 2017; Peñalba et al., 2020; Ross et al., 2015; Stukenbrock & Dutheil, 2018; van Oers et al., 2014, see chapter II Bascón-Cardozo et al., 2022a). Recombination rates can vary between species according to their life history traits, generation time (Burt & Bell, 1987)), longevity (Jaramillo-Correa et al., 2010), as well as rearing conditions, for example, divergent recombination rates are seen in domesticated compared to wild species (Dreissig et al., 2019; Moyers et al., 2018; Ross-Ibarra, 2004; Schwarzkopf et al., 2020). Different behavioural traits could result in contrasting recombination rates between species. For example, in insects, where social bees diverged in recombination rates with solitary bees (Jones et al., 2019).

At the population scale, specific features for the focal population may contribute to recombination evolution and variation, such as environmental and demographic factors, nucleotide diversity (Begun & Aquadro, 1992, see chapter II and V Bascón-Cardozo et al., 2022a,b), the degree of gene flow, introgression, genetic drift, selection (and linked selection) and also population divergence (Burri et al., 2015; Chase et al., 2021; Martin et al., 2019, see chapter III Bascón-Cardozo et al., 2022b). For example, smaller Ne is associated with lower recombination rates (Charlesworth, 2009). However, a negative association was also reported, including indirect measurements of Ne (Burt & Bell, 1987) or population density (Price & Bantock, 1975). Selection in changing environments can result in increased recombination rates due to an overabundance of deleterious allelic combinations when they become less advantageous at a certain time point (Otto & Michalakis, 1998; reviewed in Dapper & Payseur, 2017). Abiotic factors such as temperature (Dreissig et al., 2019; Jackson et al., 2015; Lloyd et al., 2018; Phillips et al., 2015; Zhang et al., 2017) and precipitation (Dreissig et al., 2019) may also trigger plasticity in genetic modifiers generating divergent recombination rates within and between species and also among individuals.

c) Variation between and within individuals

Recombination rates can diverge even between individuals. Recombination landscapes differ between sexes in several eukaryotes, a phenomenon called heterochiasmy (Sardell & Kirkpatrick, 2020; Wang et al., 2016). Other factors, such as aging, could also contribute to individual variation (Speed, 1977), as recombination rates increase with maternal age in humans, cattle, and plants (Li et al., 2017; Martin et al., 2015; Wang et al., 2016). Moreover, pathogen infections and the

coevolution with parasites were reported to increase recombination frequencies in the host (Fischer & Schmid-Hempel, 2005; Kerstes et al., 2012), following the Red Queen hypothesis (Salathé et al., 2009).

d) Variation across the genome and among chromosomes

The frequency and distribution of recombination events are usually heterogeneous across the genome in eukaryotes. Many factors influence this variation, such as the organization and architecture of the genome (Schield et al., 2020; Smeds et al., 2016; see chapter II: Bascón-Cardozo et al., 2022a). Structural chromosomal features such as centromeres and telomeres where recombination is typically suppressed or increased (Choo, 1998; Subramanian et al., 2019). Other regions are characterised by high linkage and therefore experience suppressed recombination due to structural variants or genomic rearrangements: Chromosomal translocations, inversions where recombination is reduced in individuals that are heterozygous for the inversion (Ishigohoka et al., 2021), duplications, and supergenes (i.e. a cluster of tightly linked genes that regulate a specific phenotype) (Thompson & Jiggins, 2014). Moreover, specific genomic features such as genic and intergenic regions, CpG islands, tandem repeats, and transposable elements (TEs) will affect historical recombination rate variation (see chapter II, Bascón-Cardozo et al., 2022a). Genomic regions under different selective pressures and Ne will also affect local recombination patterns (Gossmann et al., 2011). Given the tight association between recombination and epigenetic regulation, histone modification and methylation patterns will additionally play a role in shaping the recombination landscape across the genome (Buard et al., 2009; Mirouze et al., 2012, see chapter III).

1.4 Historical recombination in the context of evolution and population genomics: the implications of variable recombination rate

Most asexual organisms have shorter lifespans (Engelstädter, 2008), highlighting the significant role that meiotic recombination plays in sexual organisms. Recombination is a crucial component in gene and genome evolution; for example, it reduces genome size by enabling the removal of LTR retrotransposons, alters codon usage by GC-biased gene conversion, and contributes to gene duplication and loss (Tiley & Burleigh, 2015). Recombination also contributes to maintaining genes that are important for specific behavioural traits (Kent et al., 2012). In population genetics, recombination is a key component in the adaptation and evolution of populations and species

(Felsenstein, 1974; Grandaubert et al., 2019; Samuk et al., 2020). Hence, besides mutation, selection, genetic drift, and migration, recombination is one of the pillar parameters considered by empirical and theoretical evolutionary biologists.

More specifically, recombination can have direct and indirect advantages. The immediate benefits involve the correct chromosomal segregation in most eukaryotes, which increases reproductive success and fitness in the offspring (Kong et al., 2004), and generates new combinations of alleles and phenotypes. The indirect advantages relate to recombination breaking up the linkage between loci on haplotypes (LD). Fisher and Muller were the first to discover recombination's evolutionary advantages in finite populations (Fisher, 1930; Muller, 1964). Muller proposed that in the absence of recombination, harmful mutations accumulate irreversibly in finite populations, effect known as Muller's ratchet (Muller, 1964). Later on, it was proposed that recombination can act directly on loci under selection by speeding up the fixation of advantageous mutations (Crow & Kimura, 1965). Theoretical models continued to be developed, including the interaction of genetic drift, selection, and LD. Then, the concept of Hill-Robertson interference (HRi) (Hill & Robertson, 1966) was described, which states that linkage between loci under selection (positive or negative) will reduce the overall effectiveness of selection and thus fitness (Fig. 2A). Recombination reduces this interference by i) disrupting the linkage between advantageous alleles from deleterious mutations, ii) bringing together advantageous alleles from different individuals in a population and potentially facilitating adaptation, and iii) enabling natural selection to promote the fixation of beneficial mutations and removal of deleterious ones with greater efficacy (Tiley & Burleigh, 2015) (see Fig. 2).

Recombination may either increase or decrease the time to fixation of the advantageous alleles depending on the selection regime and the type of disequilibrium: If a population is initially in negative LD (or negative epistasis) (e.g. excess of intermediate genotypes and a scarcity of extremes), the advantage of recombination is increased as it reduces LD, generates more extreme phenotypes, increases the genetic variance in a population and builds optimal combinations of alleles (reviewed in Dapper & Payseur, 2017). Whereas, if a population initiates with positive LD (in equilibrium or positive epistasis) between beneficial alleles, the advantages of recombination are decreased (Otto et al., 1994). In this scenario, recombination can disrupt advantageous or adaptative combinations of alleles (Dapper & Payseur, 2017), an effect referred as recombination load (Charlesworth & Barton, 1996; Otto & Lenormand, 2002). A good strategy to maintain adaptative loci would be to reduce recombination in specific chromosomal stretches through

chromosomal rearrangements where recombination is typically suppressed. Adaptive allele combinations may be preserved within inversions when multiple rearrangements are maintained within a population in homogeneous environments (Kirkpatrick & Barton, 2006). Recombination reduction or suppression also has important consequences for shaping the evolution of reproductive isolation and speciation (Kirkpatrick & Barton, 2006) as well as patterns of introgression and divergence in the genome among populations (Martin et al., 2019; Samuk et al., 2017; Schumer et al., 2018).

Another indirect advantage of recombination relies on increasing or maintaining genetic variance in the population by affecting the strength of linked selection and reducing the loss of neutral diversity linked to the loci that are under either positive (Genetic Hitchhiking, selective sweep) or negative selection (Background selection) (Begun & Aquadro, 1992; Charlesworth et al., 1993; Cutter & Payseur, 2013) (Fig. 2B, 2C).

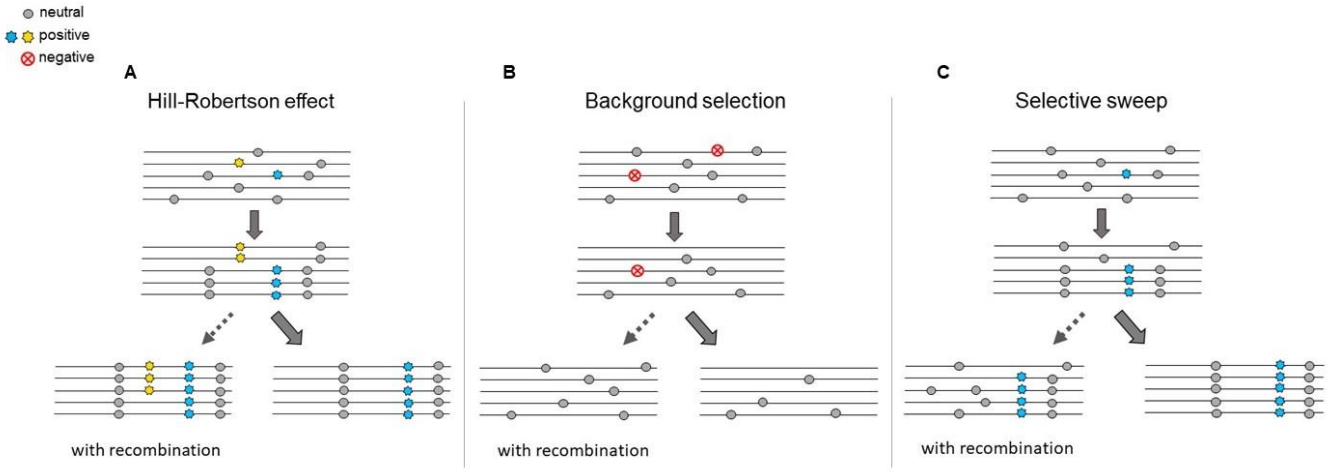


Figure 2. Schematic representation of A) Hill-Robertson interference: advantageous mutations arise on the population at distinct loci, interfere with each other, and prevent each other's fixation resulting in the loss of one advantageous mutation and fixation of the other one. If recombination occurs (dotted arrow), the advantageous mutations decouple, eventually reaching fixation (adapted from Barroso et al., 2020). **B)** Background selection: Reduction of neutral diversity linked to deleterious mutations under purifying selection. If recombination occurs, the linkage is broken, the deleterious mutations are eventually removed, and neutral diversity is maintained. **C)** Selective sweeps: an advantageous mutation arises in the population and increases in frequency, which leads to the reduction of genetic diversity at linked neutral loci. If recombination occurs, the advantageous mutation is decoupled from neutral sites, and diversity is maintained. Black lines represent individual genomes. Neutral sites are denoted as grey circles, advantageous mutations as light blue and yellow stars, and deleterious mutations as circles with a cross in red.

1.5 Avian genome and recombination in birds

Specific characteristics makes avian genomes attractive to study, such as the small genome size, the presence of macro- and micro-chromosomes, and sex heterogamety where males are homogametic (ZZ) and females heterogametic (ZW). Birds usually have less variation in chromosome number (2n=76–80), indicating karyotype stability compared to other vertebrates (Ellegren, 2013; Kawakami et al., 2014). The compact genomes of birds possess less repetitive elements, shorter introns and intragenic regions, and higher gene density, particularly in micro-chromosomes, compared to other taxa such as reptiles and mammals (Zhang et al., 2014). Across the evolution of avian genomes, intra-chromosomal rearrangements are more predominant than inter-chromosomal rearrangements (Skinner & Griffin, 2012), especially chromosomal inversions and translocations were identified (Völker et al., 2010). Break-points of chromosomal rearrangements and copy number variants (CNVs) appear in association with elevated recombination rates supporting the role of recombination in the genome evolution of birds either as a cause or a consequence of structural genomic rearrangement (Kawakami et al., 2014; Völker et al., 2010). Additionally, the evolution of sex chromosomes involved several periods of recombination suppression (Xu et al., 2019).

Recombination rates in avian genomes are presumably conserved at broad scales (Singhal et al., 2015). Yet, genome-wide differences were also reported among species like chicken, showing higher rates than zebra finch (Backström et al., 2010). At fine scales (e.g. intrachromosomal patterns), the divergence in the location and frequency of recombination among and within species is more evident. Heterochiasmy was reported, where females which carry the WZ chromosomes show low recombination rates and smoother patterns compared to males with typically higher recombination rates (Kawakami et al., 2014). However, that is not a general pattern, and some species don't show significant differences between sexes (Calderón & Pigozzi, 2006; Stapley et al., 2008; van Oers et al., 2014).

Among species, the recombination maps reveal characteristics specific to some species. For example, recombination maps of zebra finches (*Taeniopygia guttata*) usually show deserts of recombination in the middle of chromosomes (Backström et al., 2010) which are not observed in other passerine species, like the flycatcher (*Ficedula albicollis*) (Kawakami et al., 2014). A comparative analysis revealed that the genetic map of the zebra finch appears more similar to other passerines like the superb fairy-wren (*Malurus Cyaneus)* and great tit *(Parus major*), whereas the flycatcher, is

more similar to the distantly related chicken (*Gallus gallus*) (Peñalba et al., 2020). Although, other characteristics in recombination maps are more conserved across avian lineages, such as higher rates in micro-chromosomes compared to macro-chromosomes (Backström et al., 2010; Kawakami et al., 2014; Singhal et al., 2015, see chapter II Bascón-Cardozo et al., 2022a). The sex chromosomes typically lack recombination except for the Pseudoautosomal region (PAR) in the homogametic sex, i.e. males in birds are ZZ (see chapter II, Bascón-Cardozo et al., 2022; Ponnikas et al., 2022; Smeds et al., 2014).

In birds, like in most PRDM9-deficient organisms, the majority of recombination events are enriched in regulatory regions and presumably associated with open chromatin regions. This opens the exciting opportunity to study the role of recombination and the possible causes and implications of its variation in these organisms. Specific genomic features are associated with variation in recombination rates across the genome (Kawakami et al., 2017; Kent et al., 2017; Peñalba et al., 2020; Singhal et al., 2015; see chapter II Bascón-Cardozo et al., 2022). In a population genetic context, genome-wide variation in diversity and genetic differentiation can be explained by variation in recombination rates (Henderson & Brelsford, 2020; see chapter IV) as reported in some passerines (e.g. flycatchers, stonechats, blackcap). Long-term reductions in recombination rates result in genomic regions with elevated differentiation ("differentiation islands") that are shared among lineages (Burri et al., 2015; Chase et al., 2021; Chase & Mugal, 2022; Van Doren et al., 2017). Additionally, a recent study revealed that evolutionary changes in the recombination rate between flycatcher species drive signatures of linked selection (Chase & Mugal, 2022).

1.6 Our study system: the iconic blackcap

The Eurasian blackcap (*Sylvia atricapilla*) is a small passerine. Blackcaps have a wide-breeding distribution throughout Europe, including areas of Norway and Russia (Fig. 3A), and the breeding season takes place between mid-April and August. The blackcap is a nocturnal, solitary, and seasonal migrant. The majority of blackcaps migrate from breeding grounds to overwinter in the western and central areas of the Mediterranean Sea and north of the African continent (Helbig et al., 1989)(see Fig. 3A). Given the broad distribution, blackcaps encompass an entire spectrum of migratory phenotypes from non-migrant (or residents) populations on the continent (Iberian Peninsula) and Atlantic islands to migratory populations following long, medium, and short distances. Additionally, migrant populations vary in orientation, taking either south-eastern (SE) or

south-western (SW) routes and generating a contact zone where populations with different migratory strategies breed adjacent to one another (migratory divide) (Helbig et al., 1989; Helbig, 1991a, Berthold, et al., 1990). Since the 1960s, a new north-west (NW) migratory route from central European breeding-grounds to over-wintering in Great Britain and Ireland was discovered (Berthold et al., 1992; Helbig, 1991a).

The vast repertoire of migratory phenotypes that blackcaps display has been shown to have a genetic component. In support of this, cross-breeding experiments have shown that hybrids instinctively display intermediate orientations to the parents (Helbig et al., 1989; Helbig, 1991b; Helbig, 1994). Divergent chromosomal regions potentially involved in controlling migratory traits have been identified in some passerine species (Delmore et al., 2016; Lundberg et al., 2017; Toews et al., 2019). A recent work reported a repeat-rich region part of LTR retrotransposons associated with migratory phenotypes in the willow warbler (*Phylloscopus trochilus*) (Caballero-López et al., 2022). In blackcaps, Delmore et al. (2020) documented regions under positive selection containing genes with functions related to the transition from migration to residency. The same study suggests that cis-regulatory changes are important for this transition. However, a clear understanding of the genomic makeup controlling variability in migratory traits is still in ongoing process.

What do we know about the population genomics of the blackcap?

Previous studies revealed that blackcap populations began to diverge approx. 30.000 years ago (ya), resulting in populations exhibiting variable migratory phenotypes from migrants with variable distances and orientations to residents distributed in the continent and different islands (Fig. 3A) (Delmore et al., 2020).

Demography data show that blackcap populations started to diverge around the last glacial period, indicative of the evolution of migratory phenotypes corresponding environmental changes. After the split, continental populations (migrant and resident) expanded and increased Ne, whereas all island resident population show decrease in effective population size, which corresponds to colonisation events in which small number of individuals were involved (Fig. 3B) (Delmore et al., 2020, Ishigohoka et al. in preparation).

Even though blackcap populations exhibit low genetic differentiation, population structure can be detected (Delmore et al., 2020; Bours et al., in preparation). Visualization of the general population structure through PCA analysis indicated a genetic clustering of continental populations (migrants

and residents) and island populations that are close to the continent (Mallorca and Crete), separated from Macaronesian island populations (Azores, Cape Verde, Madeira, and the Canary Islands) in PC1 (Fig. 3C). The island populations of the Azores and Cape Verde cluster closely, suggesting higher genetic similarity to each other in comparison to the rest of the populations, a similar pattern can be found between Madeira and the Canary Islands. The resident population of Mallorca separates from the rest of the migrants and resident populations in PC2 (Fig. 3C). Finally, PC3 separates continental residents and the Crete population from each other as well as from the rest of the migrant populations and Mallorca.

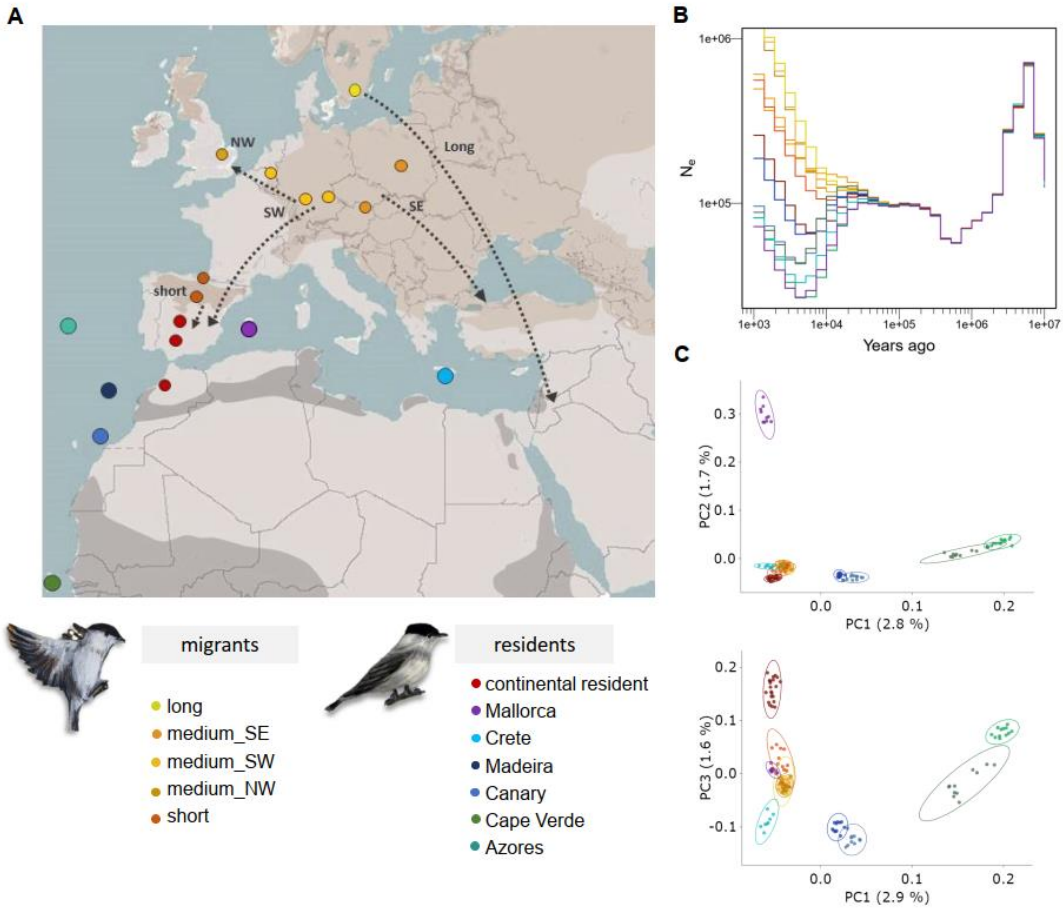


Figure 3. Blackcap distribution, population structure, and demography. A) Sites corresponding to populations with different migratory phenotypes. Arrows show the distance and direction of migration. In brown shadow, the breeding range and the dark-gray shadow correspond to the overwinter (non-breeding) range (BirdLife International 2022). **B)** Effective population size fluctuation by time is estimated for all populations. **C)** Principle Component Analysis (PCA) representing population structure.

1.7 Scope of the thesis

Recombination landscapes are heterogeneous and not as highly conserved as previously thought. However, some questions remain: how is the variation of recombination rates in birds? And what are the causes and implications of this variation? This thesis covers some of those questions and aims to characterise the historical recombination landscape in the blackcap at different scales and to assess the interplay of recombination rate variation with genomic, population, and epigenetic features. This is addressed in the following three chapters:

Chapter II: Fine-scale map reveals highly variable recombination rates associated with genomic features in the European blackcap

In the second chapter, I initially assessed how the recombination rate varies across the genome. I characterised the historical recombination landscape across the genome and between chromosomes. To determine how genomic features influence recombination rate variation, I evaluated the association of recombination rates with specific genomic and annotation features. Finally, by comparing recombination maps between the blackcap and its closest sister taxa, the Garden warbler, I assessed the conservation of recombination rates at broad- and fine-scale between these species.

Chapter III: The relationship of CpG methylation shared among tissues with recombination rates across the blackcap genome

In the third chapter, I explored the relationship of epigenetic variation, particularly DNA methylation, with recombination rates including different blackcap populations. I evaluated the methylation patterns of CpG sites shared among tissues within high and low recombining regions across the genome. This chapter adds a layer of information and complements the characterisation of the recombination landscape and genomic features in the Eurasian blackcap.

Chapter IV: Historical recombination maps diverge between blackcap populations with distinct migratory phenotype

In the fourth chapter, I extended the recombination landscape characterisation to the population scale by evaluating the variation of recombination rates among blackcap populations with different migratory phenotypes. I additionally assessed the interplay of intra-specific recombination rate variation with nucleotide diversity, population divergence patterns, and potential signs of selection.

1.8 References

Anderson, L. K., Doyle, G. G., Brigham, B., Carter, J., Hooker, K. D., Lai, A., Rice, M., & Stack, S. M. (2003). High-Resolution Crossover Maps for Each Bivalent of Zea mays Using Recombination Nodules. *Genetics*, *165*(2), 849–865. https://doi.org/10.1093/genetics/165.2.849

Anderson, L. K., Reeves, A., Webb, L. M., & Ashley, T. (1999). Distribution of crossing over on mouse synaptonemal complexes using immunofluorescent localization of MLH1 protein. *Genetics*, *151*(4), 1569–1579. https://doi.org/10.1093/genetics/151.4.1569

Auton, A., Fledel-Alon, A., Pfeifer, S., Venn, O., Ségurel, L., Street, T., Leffler, E. M., Bowden, R., Aneas, I., Broxholme, J., Humburg, P., Iqbal, Z., Lunter, G., Maller, J., Hernandez, R. D., Melton, C., Venkat, A., Nobrega, M. A., Bontrop, R., … McVean, G. (2012). A Fine-Scale Chimpanzee Genetic Map from Population Sequencing. *Science*, *336*(6078), 193–198. https://doi.org/10.1126/science.1216872

Backström, N., Forstmeier, W., Schielzeth, H., Mellenius, H., Nam, K., Bolund, E., Webster, M. T., Öst, T., Schneider, M., Kempenaers, B., & Ellegren, H. (2010). The recombination landscape of the zebra finch Taeniopygia guttata genome. *Genome Research*, *20*(4), 485–495. https://doi.org/10.1101/gr.101410.109

Barroso, G. V., Moutinho, A. F., & Dutheil, J. Y. (2020). A Population Genomics Lexicon. In J. Y. Dutheil (Ed.), *Statistical Population Genomics* (pp. 3–17). Springer US. https://doi.org/10.1007/978-1-0716-0199-0_1

Bascón-Cardozo, K., Bours, A., Manthey, G., Pruisscher, P., Durieux, G., Dutheil, J., Odenthal-Hesse, L., & Liedvogel, M. (2022a). *Fine-scale map reveals highly variable recombination rates associated with genomic features in the European blackcap* [Preprint]. Preprints. https://doi.org/10.22541/au.165423614.49331155/v1

Bascón-Cardozo, K., Bours, A., Ishigohoka, J., Odenthal-Hesse, L., & Liedvogel, M. (2022b). Historical recombination maps diverge between Eurasian blackcap populations with distinct migratory strategies. https://doi.org/10.22541/au.166790167.72861799/v1

Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., & de Massy, B. (2010). PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science (New York, N.Y.)*, *327*(5967), 836–840. https://doi.org/10.1126/science.1183439

Begun, D. J., & Aquadro, C. F. (1992). Levels of naturally occurring DNA polymorphism correlate with recombination rates in D. melanogaster. *Nature*, *356*(6369), 6369. https://doi.org/10.1038/356519a0

Berg, I. L., Neumann, R., Sarbajna, S., Odenthal-Hesse, L., Butler, N. J., & Jeffreys, A. J. (2011). Variants of the protein PRDM9 differentially regulate a set of human meiotic recombination hotspots highly active in African populations. *Proceedings of the National Academy of Sciences*, *108*(30), 12378–12383. https://doi.org/10.1073/pnas.1109531108

Berthold, P., Querner, U. & Schlenker, R. (1990) *Die Mönchsgrasmücke*. Die Neue Brehm-Bücherei

Berthold, P., Helbig, A. J., Mohr, G., & Querner, U. (1992). Rapid microevolution of migratory behaviour in a wild bird species. *Nature*, *360*(6405), 6405. https://doi.org/10.1038/360668a0

Brazier, T., & Glémin, S. (2022). Diversity and determinants of recombination landscapes in flowering plants. *PLOS Genetics*, *18*(8), e1010141. https://doi.org/10.1371/journal.pgen.1010141

Buard, J., Barthès, P., Grey, C., & de Massy, B. (2009). Distinct histone modifications define initiation and repair of meiotic recombination in the mouse. *The EMBO Journal*, *28*(17), 2616–2624. https://doi.org/10.1038/emboj.2009.207

Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S.-A., Sætre, G.-P., Török, J., & Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Research*, *25*(11), 1656–1665. https://doi.org/10.1101/gr.196485.115

Burt, A., & Bell, G. (1987). Mammalian chiasma frequencies as a test of two theories of recombination. *Nature*, *326*(6115), 6115. https://doi.org/10.1038/326803a0

Caballero-López, V., Lundberg, M., Sokolovskis, K., & Bensch, S. (2022). Transposable elements mark a repeat-rich region associated with migratory phenotypes of willow warblers (Phylloscopus trochilus). *Molecular Ecology*, *31*(4), 1128–1141. https://doi.org/10.1111/mec.16292

Calderón, P. L., & Pigozzi, M. I. (2006). MLH1-focus mapping in birds shows equal recombination between sexes and diversity of crossover patterns. *Chromosome Research*, *14*(6), 605–612. https://doi.org/10.1007/s10577-006-1059-0

Chapman, N. H., & Thompson, E. A. (2001). 25 Linkage disequilibrium mapping: The role of population history, size, and structure. In *Advances in Genetics* (Vol. 42, pp. 413–437). Academic Press. https://doi.org/10.1016/S0065-2660(01)42034-7

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 3. https://doi.org/10.1038/nrg2526

Charlesworth, B., & Barton, N. H. (1996). Recombination load associated with selection for increased recombination. *Genetics Research*, *67*(1), 27–41. https://doi.org/10.1017/S0016672300033450

Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, *134*(4), 1289–1303. https://doi.org/10.1093/genetics/134.4.1289

Chase, M. A., Ellegren, H., & Mugal, C. F. (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*(9), 2179–2196. https://doi.org/10.1111/evo.14234

Chase, M. A., & Mugal, C. F. (2022). The role of recombination dynamics in shaping signatures of direct and indirect selection across the Ficedula flycatcher genome [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2022.08.11.503468

Choo, K. H. A. (1998). Why Is the Centromere So Cold? *Genome Research*, *8*(2), 81–82. https://doi.org/10.1101/gr.8.2.81

Crow, J. F., & Kimura, M. (1965). Evolution in Sexual and Asexual Populations. *The American Naturalist*. https://doi.org/10.1086/282389

Cutter, A. D., & Payseur, B. A. (2013). Genomic signatures of selection at linked sites: Unifying the disparity among species. *Nature Reviews Genetics*, *14*(4), 4. https://doi.org/10.1038/nrg3425

Dapper, A. L., & Payseur, B. A. (2017). Connecting theory and data to understand recombination rate evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160469. https://doi.org/10.1098/rstb.2016.0469

Dapper, A. L., & Payseur, B. A. (2018). Effects of Demographic History on the Detection of Recombination Hotspots from Linkage Disequilibrium. *Molecular Biology and Evolution*, *35*(2), 335–353. https://doi.org/10.1093/molbev/msx272

del Priore, L., & Pigozzi, M. I. (2020). MLH1 focus mapping in the guinea fowl (Numida meleagris) give insights into the crossover landscapes in birds. *PLoS ONE*, *15*(10), e0240245. https://doi.org/10.1371/journal.pone.0240245

Delmore, K. E., Toews, D. P. L., Germain, R. R., Owens, G. L., & Irwin, D. E. (2016). The Genetics of Seasonal Migration and Plumage Color. *Current Biology*, *26*(16), 2167–2173. https://doi.org/10.1016/j.cub.2016.06.015

Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Lugo Ramos, J. S., Durieux, G., Ishigohoka, J., & Liedvogel, M. (2020). The evolutionary history and genomics of European blackcap migration. *ELife*, *9*, e54462. https://doi.org/10.7554/eLife.54462

Dreissig, S., Mascher, M., & Heckmann, S. (2019). Variation in Recombination Rate Is Shaped by Domestication and Environmental Conditions in Barley. *Molecular Biology and Evolution*, *36*(9), 2029–2039. https://doi.org/10.1093/molbev/msz141

Ellegren, H. (2013). The Evolutionary Genomics of Birds. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 239–259. https://doi.org/10.1146/annurev-ecolsys-110411-160327

Engelstädter, J. (2008). Constraints on the evolution of asexual reproduction. *BioEssays*, *30*(11–12), 1138–1150. https://doi.org/10.1002/bies.20833

Fawcett, D. W. (1956). The fine structure of chromosomes in the meiotic prophase of vertebrate spermatocytes. *The Journal of Biophysical and Biochemical Cytology*, *2*(4), 403–406. https://doi.org/10.1083/jcb.2.4.403

Felsenstein, J. (1974). THE EVOLUTIONARY ADVANTAGE OF RECOMBINATION. *Genetics*, *78*(2), 737–756. https://doi.org/10.1093/genetics/78.2.737

Fischer, O., & Schmid-Hempel, P. (2005). Selection by parasites may increase host recombination frequency. *Biology Letters*, *1*(2), 193–195. https://doi.org/10.1098/rsbl.2005.0296

Fisher, R. A. (1930). *The genetical theory of natural selection*. Рипол Классик.

Fledel-Alon, A., Wilson, D. J., Broman, K., Wen, X., Ober, C., Coop, G., & Przeworski, M. (2009). Broad-Scale Recombination Patterns Underlying Proper Disjunction in Humans. *PLOS Genetics*, *5*(9), e1000658. https://doi.org/10.1371/journal.pgen.1000658

Gossmann, T. I., Woolfit, M., & Eyre-Walker, A. (2011). Quantifying the Variation in the Effective Population Size Within a Genome. *Genetics*, *189*(4), 1389–1402. https://doi.org/10.1534/genetics.111.132654

Grandaubert, J., Dutheil, J. Y., & Stukenbrock, E. H. (2019). The genomic determinants of adaptive evolution in a fungal pathogen. *Evolution Letters*, *3*(3), 299–312. https://doi.org/10.1002/evl3.117

Helbig, A. J., Berthold, P., & Wiltschko, W. (1989). Migratory Orientation of Blackcaps (Sylvia atricapilla): Population-specific Shifts of Direction during the Autumn. *Ethology*, *82*(4), 307–315. https://doi.org/10.1111/j.1439-0310.1989.tb00510.x

Helbig, A. J. (1991a). SE- and SW-migrating Blackcap (Sylvia atricapilla) populations in Central Europe: Orientation of birds in the contact zone. *Journal of Evolutionary Biology*, *4*(4), 657–670. https://doi.org/10.1046/j.1420-9101.1991.4040657.x

Helbig, A. J. (1991b). 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*(1), 9–12. https://doi.org/10.1007/BF00172133

Helbig, A. J. (1994). Genetic base and evolutionary change of migratory directions in a european passerine migrant sylvia atricapilla. *Ostrich*,*65*(2), 151-159.

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

Henderson, S. A., & Edwards, R. G. (1968). Chiasma Frequency and Maternal Age in Mammals. *Nature*, *218*(5136), 5136. https://doi.org/10.1038/218022a0

Hermann, P., Heissl, A., Tiemann-Boege, I., & Futschik, A. (2019). LDJump: Estimating variable recombination rates from population genetic data. *Molecular Ecology Resources*, *19*(3), 623–638. https://doi.org/10.1111/1755-0998.12994

Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetics Research*, *8*(3), 269–294. https://doi.org/10.1017/S0016672300010156

Ishigohoka, J., Bascón-Cardozo, K., Bours, A., Fuß, J., Rhie, A., Mountcastle, J., Haase, B., Chow, W., Collins, J., Howe, K., Uliano-Silva, M., Fedrigo, O., Jarvis, E. D., Pérez-Tris, J., Illera, J. C., & Liedvogel, M. (2021). *Recombination suppression and selection affect local ancestries in genomes of a migratory songbird* [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2021.12.22.473882

Jackson, S., Nielsen, D. M., & Singh, N. D. (2015). Increased exposure to acute thermal stress is associated with a non-linear increase in recombination frequency and an independent linear decrease in fitness in Drosophila. *BMC Evolutionary Biology*, *15*(1), 175. https://doi.org/10.1186/s12862-015-0452-8

Jaramillo-Correa, J. P., Verdú, M., & González-Martínez, S. C. (2010). The contribution of recombination to heterozygosity differs among plant evolutionary lineages and life-forms. *BMC Evolutionary Biology*, *10*(1), 22. https://doi.org/10.1186/1471-2148-10-22

Johnston, S. E., Bérénos, C., Slate, J., & Pemberton, J. M. (2016). Conserved Genetic Architecture Underlying Individual Recombination Rate Variation in a Wild Population of Soay Sheep (Ovis aries). *Genetics*, *203*(1), 583–598. https://doi.org/10.1534/genetics.115.185553

Jones, J. C., Wallberg, A., Christmas, M. J., Kapheim, K. M., & Webster, M. T. (2019). Extreme Differences in Recombination Rate between the Genomes of a Solitary and a Social Bee. *Molecular Biology and Evolution*, *36*(10), 2277–2291. https://doi.org/10.1093/molbev/msz130

Kawakami, T., Mugal, C. F., 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. *Molecular Ecology*, *26*(16), 4158–4172. https://doi.org/10.1111/mec.14197

Kawakami, T., Smeds, L., Backström, N., Husby, A., Qvarnström, A., Mugal, C. F., 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. *Molecular Ecology*, *23*(16), 4035–4058. https://doi.org/10.1111/mec.12810

Kent, C. F., Minaei, S., Harpur, B. A., & Zayed, A. (2012). Recombination is associated with the evolution of genome structure and worker behavior in honey bees. *Proceedings of the National Academy of Sciences*, *109*(44), 18012–18017. https://doi.org/10.1073/pnas.1208094109

Kent, T. V., Uzunović, J., & Wright, S. I. (2017). Coevolution between transposable elements and recombination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160458. https://doi.org/10.1098/rstb.2016.0458

Kerstes, N. A., Bérénos, C., Schmid-Hempel, P., & Wegner, K. M. (2012). Antagonistic experimental coevolution with a parasite increases host recombination frequency. *BMC Evolutionary Biology*, *12*(1), 18. https://doi.org/10.1186/1471-2148-12-18

Kirkpatrick, M., & Barton, N. (2006). Chromosome Inversions, Local Adaptation and Speciation. *Genetics*, *173*(1), 419–434. https://doi.org/10.1534/genetics.105.047985

Kong, A., Barnard, J., Gudbjartsson, D. F., Thorleifsson, G., Jonsdottir, G., Sigurdardottir, S., Richardsson, B., Jonsdottir, J., Thorgeirsson, T., Frigge, M. L., Lamb, N. E., Sherman, S., Gulcher, J. R., & Stefansson, K. (2004). Recombination rate and reproductive success in humans. *Nature Genetics*, *36*(11), 1203–1206. https://doi.org/10.1038/ng1445

Lam, I., & Keeney, S. (2015). Mechanism and Regulation of Meiotic Recombination Initiation. *Cold Spring Harbor Perspectives in Biology*, *7*(1). https://doi.org/10.1101/cshperspect.a016634

Lenormand, T., Engelstädter, J., Johnston, S. E., Wijnker, E., & Haag, C. R. (2016). Evolutionary mysteries in meiosis. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2016.0001

Li, F., De Storme, N., & Geelen, D. (2017). Dynamics of male meiotic recombination frequency during plant development using Fluorescent Tagged Lines in Arabidopsis thaliana. *Scientific Reports*, *7*(1), 1. https://doi.org/10.1038/srep42535

Lloyd, A., Morgan, C., H. Franklin, F. C., & Bomblies, K. (2018). Plasticity of Meiotic Recombination Rates in Response to Temperature in Arabidopsis. *Genetics*, *208*(4), 1409–1420. https://doi.org/10.1534/genetics.117.300588

Lundberg, M., Liedvogel, M., Larson, K., Sigeman, H., Grahn, M., Wright, A., Åkesson, S., & Bensch, S. (2017). Genetic differences between willow warbler migratory phenotypes are few and cluster in large haplotype blocks. *Evolution Letters*, *1*(3), 155–168. https://doi.org/10.1002/evl3.15

Martin, H. C., Christ, R., Hussin, J. G., O'Connell, J., Gordon, S., Mbarek, H., Hottenga, J.-J., McAloney, K., Willemsen, G., Gasparini, P., Pirastu, N., Montgomery, G. W., Navarro, P., Soranzo, N., Toniolo, D., Vitart, V., Wilson, J. F., Marchini, J., Boomsma, D. I., … Donnelly, P. (2015). Multicohort analysis of the maternal age effect on recombination. *Nature Communications*, *6*(1), 1. https://doi.org/10.1038/ncomms8846

Martin, S. H., Davey, J. W., Salazar, C., & Jiggins, C. D. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLOS Biology*, *17*(2), e2006288. https://doi.org/10.1371/journal.pbio.2006288

Mirouze, M., Lieberman-Lazarovich, M., Aversano, R., Bucher, E., Nicolet, J., Reinders, J., & Paszkowski, J. (2012). Loss of DNA methylation affects the recombination landscape in Arabidopsis. *Proceedings of the National Academy of Sciences*, *109*(15), 5880–5885. https://doi.org/10.1073/pnas.1120841109

Morgan, C. H., Zhang, H., & Bomblies, K. (2017). Are the effects of elevated temperature on meiotic recombination and thermotolerance linked via the axis and synaptonemal complex? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1736), 20160470. https://doi.org/10.1098/rstb.2016.0470

Moses, M. J. (1956). Chromosomal Structures in Crayfish Spermatocytes. *The Journal of Biophysical and Biochemical Cytology*, *2*(2), 215–218.

Moyers, B. T., Morrell, P. L., & McKay, J. K. (2018). Genetic Costs of Domestication and Improvement. *Journal of Heredity*, *109*(2), 103–116. https://doi.org/10.1093/jhered/esx069

Muller, H. J. (1916). The Mechanism of Crossing-Over. *THE AMERICAN NATURALIST*, 29.

Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *1*(1), 2–9. https://doi.org/10.1016/0027-5107(64)90047-8

Myers, S., Bowden, R., Tumian, A., Bontrop, R. E., Freeman, C., MacFie, T. S., McVean, G., & Donnelly, P. (2010). Drive Against Hotspot Motifs in Primates Implicates the PRDM9 Gene in Meiotic Recombination. *Science*, *327*(5967), 876–879. https://doi.org/10.1126/science.1182363

Otto, S. P., Feldman, M. W., & Christiansen, F. B. (1994). Some Advantages and Disadvantages of Recombination. In S. A. Levin (Ed.), *Frontiers in Mathematical Biology* (pp. 198–211). Springer. https://doi.org/10.1007/978-3-642-50124-1_12

Otto, S. P., & Lenormand, T. (2002). Resolving the paradox of sex and recombination. *Nature Reviews Genetics*, *3*(4), 252–261. https://doi.org/10.1038/nrg761

Otto, S. P., & Michalakis, Y. (1998). The evolution of recombination in changing environments. *Trends in Ecology & Evolution*, *13*(4), 145–151. https://doi.org/10.1016/S0169-5347(97)01260-3

Parvanov, E. D., Petkov, P. M., & Paigen, K. (2010). Prdm9 Controls Activation of Mammalian Recombination Hotspots. *Science*, *327*(5967), 835–835. https://doi.org/10.1126/science.1181495

Peñalba, J. V., Deng, Y., Fang, Q., Joseph, L., Moritz, C., & Cockburn, A. (2020). Genome of an iconic Australian bird: High-quality assembly and linkage map of the superb fairy-wren (Malurus cyaneus). *Molecular Ecology Resources*, *20*(2), 560–578. https://doi.org/10.1111/1755-0998.13124

Peñalba, J. V., & Wolf, J. B. W. (2020). From molecules to populations: Appreciating and estimating recombination rate variation. *Nature Reviews Genetics*, *21*(8), 8. https://doi.org/10.1038/s41576-020-0240-1

Phillips, D., Jenkins, G., Macaulay, M., Nibau, C., Wnetrzak, J., Fallding, D., Colas, I., Oakey, H., Waugh, R., & Ramsay, L. (2015). The effect of temperature on the male and female recombination landscape of barley. *New Phytologist*, *208*(2), 421–429. https://doi.org/10.1111/nph.13548

Pollock, D. L., & Fechheimer, N. S. (1978). The chromosomes of cockerels (Gallus domesticus) during meiosis. *Cytogenetic and Genome Research*, *21*(5), 267–281. https://doi.org/10.1159/000130903

Ponnikas, S., Sigeman, H., Lundberg, M., & Hansson, B. (2022). Extreme variation in recombination rate and genetic diversity along the Sylvioidea neo-sex chromosome. *Molecular Ecology*, *31*(13), 3566–3583. https://doi.org/10.1111/mec.16532

Price, D. J., & Bantock, C. R. (1975). Marginal Populations of Cepaea nemoralis (L.) on the Brendon Hills, England. II. Variation in Chiasma Frequency. *Evolution*, *29*(2), 278–286. https://doi.org/10.2307/2407216

Ross, C. R., DeFelice, D. S., Hunt, G. J., Ihle, K. E., Amdam, G. V., & Rueppell, O. (2015). Genomic correlates of recombination rate and its variability across eight recombination maps in the western honey bee (Apis mellifera L.). *BMC Genomics*, *16*(1), 107. https://doi.org/10.1186/s12864-015-1281-2

Ross-Ibarra, J. (2004). The Evolution of Recombination under Domestication: A Test of Two Hypotheses. *The American Naturalist*, *163*(1), 105–112. https://doi.org/10.1086/380606

Salathé, M., Kouyos, R. D., & Bonhoeffer, S. (2009). On the Causes of Selection for Recombination Underlying the Red Queen Hypothesis. *The American Naturalist*, *174*(S1), S31–S42. https://doi.org/10.1086/599085

Samuk, K., Manzano-Winkler, B., Ritz, K. R., & Noor, M. A. F. (2020). Natural Selection Shapes Variation in Genome-wide Recombination Rate in Drosophila pseudoobscura. *Current Biology*, *30*(8), 1517-1528.e6. https://doi.org/10.1016/j.cub.2020.03.053

Samuk, K., Owens, G. L., Delmore, K. E., Miller, S. E., Rennison, D. J., & Schluter, D. (2017). Gene flow and selection interact to promote adaptive divergence in regions of low recombination. *Molecular Ecology*, *26*(17), 4378–4390. https://doi.org/10.1111/mec.14226

Sardell, J. M., & Kirkpatrick, M. (2020). Sex Differences in the Recombination Landscape. *The American Naturalist*, *195*(2), 361–379. https://doi.org/10.1086/704943

Schield, D. R., Pasquesi, G. I. M., Perry, B. W., Adams, R. H., Nikolakis, Z. L., Westfall, A. K., Orton, R. W., Meik, J. M., Mackessy, S. P., & Castoe, T. A. (2020). Snake Recombination Landscapes Are Concentrated in Functional Regions despite PRDM9. *Molecular Biology and Evolution*, *37*(5), 1272–1294. https://doi.org/10.1093/molbev/msaa003

Schumer, M., Xu, C., Powell, D. L., Durvasula, A., Skov, L., Holland, C., Blazier, J. C., Sankararaman, S., Andolfatto, P., Rosenthal, G. G., & Przeworski, M. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, *360*(6389), 656–660. https://doi.org/10.1126/science.aar3684

Schwarzkopf, E. J., Motamayor, J. C., & Cornejo, O. E. (2020). Genetic differentiation and intrinsic genomic features explain variation in recombination hotspots among cocoa tree populations. *BMC Genomics*, *21*(1), 332. https://doi.org/10.1186/s12864-020-6746-2

Singhal, S., Leffler, E. M., Sannareddy, K., Turner, I., Venn, O., Hooper, D. M., Strand, A. I., Li, Q., Raney, B., Balakrishnan, C. N., Griffith, S. C., McVean, G., & Przeworski, M. (2015). Stable recombination hotspots in birds. *Science (New York, N.Y.)*, *350*(6263), 928–932. https://doi.org/10.1126/science.aad0843

Skinner, B. M., & Griffin, D. K. (2012). Intrachromosomal rearrangements in avian genome evolution: Evidence for regions prone to breakpoints. *Heredity*, *108*(1), 1. https://doi.org/10.1038/hdy.2011.99

Smagulova, F., Gregoretti, I. V., Brick, K., Khil, P., Camerini-Otero, R. D., & Petukhova, G. V. (2011). Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. *Nature*, *472*(7343), 7343. https://doi.org/10.1038/nature09869

Smeds, L., Kawakami, T., Burri, R., Bolivar, P., Husby, A., Qvarnström, A., Uebbing, S., & Ellegren, H. (2014). Genomic identification and characterization of the pseudoautosomal region in highly differentiated avian sex chromosomes. *Nature Communications*, *5*(1), 5448. https://doi.org/10.1038/ncomms6448

Smeds, L., Mugal, C. F., Qvarnström, A., & Ellegren, H. (2016). High-Resolution Mapping of Crossover and Non-crossover Recombination Events by Whole-Genome Re-sequencing of an Avian Pedigree. *PLOS Genetics*, *12*(5), e1006044. https://doi.org/10.1371/journal.pgen.1006044

Speed, R. M. (1977). The effects of ageing on the meiotic chromosomes of male and female mice. *Chromosoma*, *64*(3), 241–254. https://doi.org/10.1007/BF00328080

Spence, J. P., & Song, Y. S. (2019). Inference and analysis of population-specific fine-scale recombination maps across 26 diverse human populations. *Science Advances*, *5*(10), eaaw9206. https://doi.org/10.1126/sciadv.aaw9206

Stapley, J., Birkhead, T. R., Burke, T., & Slate, J. (2008). A Linkage Map of the Zebra Finch Taeniopygia guttata Provides New Insights Into Avian Genome Evolution. *Genetics*, *179*(1), 651–667. https://doi.org/10.1534/genetics.107.086264

Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation in recombination frequency and distribution across eukaryotes: Patterns and processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160455. https://doi.org/10.1098/rstb.2016.0455

Stukenbrock, E. H., & Dutheil, J. Y. (2018). Fine-Scale Recombination Maps of Fungal Plant Pathogens Reveal Dynamic Recombination Landscapes and Intragenic Hotspots. *Genetics*, *208*(3), 1209–1229. https://doi.org/10.1534/genetics.117.300502

Sturtevant, A. H. (1915). The behavior of the chromosomes as studied through linkage. *Zeitschrift Für Induktive Abstammungs- Und Vererbungslehre*, *13*(1), 234–287. https://doi.org/10.1007/BF01792906

Subramanian, V. V., Zhu, X., Markowitz, T. E., Vale-Silva, L. A., San-Segundo, P. A., Hollingsworth, N. M., Keeney, S., & Hochwagen, A. (2019). Persistent DNA-break potential near telomeres increases

initiation of meiotic recombination on short chromosomes. *Nature Communications*, *10*(1), 970. https://doi.org/10.1038/s41467-019-08875-x

Thompson, M. J., & Jiggins, C. D. (2014). Supergenes and their role in evolution. *Heredity*, *113*(1), 1. https://doi.org/10.1038/hdy.2014.20

Tiley, G. P., & Burleigh, J. G. (2015). The relationship of recombination rate, genome structure, and patterns of molecular evolution across angiosperms. *BMC Evolutionary Biology*, *15*(1), 194. https://doi.org/10.1186/s12862-015-0473-3

Toews, D. P. L., Taylor, S. A., Streby, H. M., Kramer, G. R., & Lovette, I. J. (2019). Selection on VPS13A linked to migration in a songbird. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(37), 18272–18274. https://doi.org/10.1073/pnas.1909186116

Tsai, I. J., Burt, A., & Koufopanou, V. (2010). Conservation of recombination hotspots in yeast. *Proceedings of the National Academy of Sciences*, *107*(17), 7847–7852. https://doi.org/10.1073/pnas.0908774107

Van Doren, B. M., Campagna, L., Helm, B., Illera, J. C., Lovette, I. J., & Liedvogel, M. (2017). Correlated patterns of genetic diversity and differentiation across an avian family. *Molecular Ecology*, *26*(15), 3982–3997. https://doi.org/10.1111/mec.14083

van Oers, K., Santure, A. W., De Cauwer, I., van Bers, N. E., Crooijmans, R. P., Sheldon, B. C., Visser, M. E., Slate, J., & Groenen, M. A. (2014). Replicated high-density genetic maps of two great tit populations reveal fine-scale genomic departures from sex-equal recombination rates. *Heredity*, *112*(3), 307–316. https://doi.org/10.1038/hdy.2013.107

Völker, M., Backström, N., Skinner, B. M., Langley, E. J., Bunzey, S. K., Ellegren, H., & Griffin, D. K. (2010). Copy number variation, chromosome rearrangement, and their association with recombination during avian evolution. *Genome Research*, *20*(4), 503–511. https://doi.org/10.1101/gr.103663.109

Wang, Z., Shen, B., Jiang, J., Li, J., & Ma, L. (2016). Effect of sex, age and genetics on crossover interference in cattle. *Scientific Reports*, *6*(1), 1. https://doi.org/10.1038/srep37698

Weng, Z., Wolc, A., Su, H., Fernando, R. L., Dekkers, J. C. M., Arango, J., Settar, P., Fulton, J. E., O'Sullivan, N. P., & Garrick, D. J. (2019). Identification of recombination hotspots and quantitative trait loci for recombination rate in layer chickens. *Journal of Animal Science and Biotechnology*, *10*(1), 20. https://doi.org/10.1186/s40104-019-0332-y

Xu, L., Auer, G., Peona, V., Suh, A., Deng, Y., Feng, S., Zhang, G., Blom, M. P. K., Christidis, L., Prost, S., Irestedt, M., & Zhou, Q. (2019). Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds. *Nature Ecology & Evolution*, *3*(5), 834–844. https://doi.org/10.1038/s41559-019-0850-1

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., Storz, J. F., Antunes, A., Greenwold, M. J., Meredith, R. W., Ödeen, A., Cui, J., Zhou, Q., Xu, L., Pan, H., Wang, Z., Jin, L., Zhang, P., Hu, H., … Wang, J. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science (New York, N.Y.)*, *346*(6215), 1311–1320. https://doi.org/10.1126/science.1251385

Zhang, K., Wu, X.-C., Zheng, D.-Q., & Petes, T. D. (2017). Effects of Temperature on the Meiotic Recombination Landscape of the Yeast Saccharomyces cerevisiae. *MBio*, *8*(6), e02099-17. https://doi.org/10.1128/mBio.02099-17

CHAPTER II

Fine-scale map reveals highly variable recombination rates associated with genomic features in the European blackcap

Karen Bascón-Cardozo[1#], Andrea Bours[1], Georg Manthey[1,2], Gillian Durieux[1], Julien Y. Dutheil[1,4], Peter Pruisscher[1,5], Linda Odenthal-Hesse*[1#] and Miriam Liedvogel*[1,2#]

Affiliations:

[1] Max Planck Institute for Evolutionary Biology, MPRG Behavioural Genomics, Plön, Germany
[2] Institute of Avian Research, Wilhelmshaven, Germany
[4] Institute of Evolutionary Science of Montpellier (ISEM), University of Montpellier, Montpellier, France
[5] Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden

* equal contribution

Submitted manuscript: This manuscript was submitted to Genome Biology and Evolution (GBE) in August 2022 and it is currently revised based on comments received on October 18, 2022.

Published preprint: https://www.authorea.com/doi/full/10.22541/au.165423614.49331155
DOI:10.22541/au.165423614.49331155/v1

2.1 Abstract

Recombination is responsible for breaking up haplotypes, influencing genetic variability, and the efficacy of selection. Bird genomes lack the protein PRDM9, a key determinant of recombination dynamics in most metazoans. The historical recombination maps in birds show an apparent stasis in the positioning of recombination events. This highly conserved recombination pattern over long timescales may constrain the evolution of recombination in birds, but extensive variation in recombination rate across the genome and between different species has been reported. Here, we characterise a fine-scale historical recombination map of an iconic migratory songbird, the European blackcap (*Sylvia atricapilla*) using a LD-based approach which accounts for population demography. Our results revealed variable recombination rates among and within chromosomes, which associate positively with nucleotide diversity and GC content, and negatively with chromosome size. The recombination rates increased significantly at regulatory regions and not necessarily at gene bodies. CpG islands associated strongly with recombination rates; however, their specific position and local DNA methylation patterns likely influenced this relationship. The association with retrotransposons varied according to specific family and location. Our results also provide evidence of a heterogeneous conservation of recombination maps between the blackcap and its closest sister taxon, the garden warbler at the intra-chromosomal level. These findings highlight the considerable variability of recombination rates at different scales and the role of specific genomic features in shaping this variation. This study opens the possibility of further investigating the impact of recombination on specific population-genomic features.

Keywords

historical recombination rates, bird genome, transposable elements, CpG islands, regulatory elements

2.2 Introduction

Meiotic recombination reshuffles parental genetic material, thus creates new combinations of alleles providing the primary source of genetic variation. Recombination impacts the efficacy and strength of selection at linked loci under Hill-Roberson interference (Hill and Robertson 1966) (Betancourt et al. 2009; Hickey and Golding 2018) by breaking up non-random associations between loci across the genome, and thus both increasing the probability of advantageous alleles to fix and impeding the propagation of deleterious alleles. Hence, recombination is a key player in influencing the rate of genetic diversity, introgression, differentiation, and subsequently speciation (Butlin 2005; Nachman and Payseur 2012; Mugal et al. 2013; Kawakami et al. 2017; Martin et al. 2019). In a behavioural context, recombination affects the evolution and regulation of genes involved in specific adaptative behaviours (Kent et al., 2012).

Recombination events occur during meiosis-I when specific regions of a chromosome receive programmed double-strand breaks (DSBs), which are then repaired via exchange of genetic material either as non-reciprocal gene-conversion events or reciprocal crossing-over (CO). Both recombination outcomes influence allele dynamics and genetic diversity. Recombination rates are inferred from the accumulation of COs recombination events over long evolutionary timescales, and vary widely across different scales: between taxa, species, populations, sexes, and along the genome (Kong et al. 2010; Stapley et al. 2017; Peñalba and Wolf 2020). Most recombination events are not randomly distributed along the genome and evolve differently depending on the organism, genome structure, and its mechanisms of genetic regulation (Stapley et al. 2017).

In the majority of mammals and most metazoans, CO positions are determined in *trans* by the meiotic methyltransferase PR-domain containing protein 9 (PRDM9), a protein that binds specific nucleotide motifs via a zinc finger-DNA-binding domain (Baudat et al. 2010; Parvanov et al. 2010; Berg et al. 2011). In these organisms, recombination events cluster into intergenic regions away from functional elements (Brick et al. 2012). The genome-wide location of recombination events, often called recombination maps, changes continuously, due to a combination of rapid evolution of the zinc-fingers that confer specificity to DNA motifs and subsequent erosion of the motifs via biased gene conversion (Baudat et al. 2010; Myers et al. 2010; Parvanov et al. 2010). In contrast, many PRDM9 deficient species like some plants, yeast, canids, platypus, crocodiles, some species of fish, and birds cluster past recombination events in locations with existing open-chromatin marks, such as promoters or functional regions where the transcription machinery has greater

access to the DNA (Choi et al. 2013; Lam and Keeney 2015) At least in yeasts and birds, the positioning of recombination events appears static over long evolutionary timescales, (Lam and Keeney 2015; Singhal et al. 2015).

Among vertebrate genomes lacking functional PRDM9 orthologs, birds are an iconic model for studying evolution. Bird genomes are compact, containing a high density of coding regions, especially in micro-chromosomes, and only few repetitive elements (Zhang et al. 2014). Historical recombination maps from different bird species have been characterised at both broad and fine-scale using cytological methods (Calderón and Pigozzi 2006; Malinovskaya et al. 2018; del Priore and Pigozzi 2020) or by inferring historical population recombination rates based on Linkage Disequilibrium (LD) patterns using coalescent models (McVean et al. 2002). High recombination rates were localised at genomic features such as Transcription Start Sites (TSS), Transcription Stop Sites (TES), GC rich sequences including CpG islands (CpGi), and specific Transposable Elements (TEs) (Backström et al. 2010; Kawakami et al. 2014; Singhal et al. 2015; Kawakami et al. 2017; Weng et al. 2019). The association between recombination rates and TEs, however, remains puzzling. A negative association was reported between recombination and TEs (Rizzon et al. 2002; Shen et al. 2017), while evidence of positive association was found in another system (Kawakami et al. 2017).

Despite the apparent conservation of recombination maps in birds' genomes, recombination rates differ widely along the genome and between different (sub) species (Backström et al., 2010, Singhal et al., 2015, Kawakami, 2017, Peñalba et al., 2020). From an evolutionary and ecological perspective, it is crucial to understand the causes and consequences of this variability. Here we characterise the recombination map of the European blackcap *Sylvia atricapilla*. This migratory songbird species is a powerful model to investigate genomic features in an evolutionary framework, due to its wide-range of migratory phenotypes that have been shown to be heritable (e.g. Berthold et al., 1992; Helbig, 1991; Liedvogel et al., 2011; Merlin & Liedvogel, 2019). Understanding the genomic architecture and recombination landscape are instrumental in characterising the genetic basis of a complex behaviour such as seasonal migration.

Most analyses of historical bird recombination were based on domesticated or inbred lines (Backström et al., 2010; Calderón & Pigozzi, 2006; Mugal, et al., 2016), with only few carried out on natural populations (Kawakami et al. 2017; Peñalba et al. 2020). Additionally, these characterisations were based mainly on genomes with low coverage, compromising the resolution

of recombination inference, due to low mapping efficiencies especially on telomeric and other repetitive sequences (Ho et al. 2020). Very importantly, the underlying population demography can heavily influence the accuracy of recombination inference at finer scales (Kamm et al. 2016a; Dapper and Payseur 2018) and population-scaled recombination methods used previously did not account for this. Our study is the first to overcome many of these limitations in an iconic avian system using whole-genome resequencing (WGS) data of wild-caught individuals mapped to a high-quality chromosome-level genome, assembled with a three-pronged approach, including scaffolding with optical maps, long-range, and in-depth short-read sequencing (Ishigohoka et al., 2021). Further, we apply an LD-based approach accounting for past demographic fluctuations to alleviate potential bias in the inference of the population-scaled recombination rates.

Our study aims to (i) characterise the population-scaled historical recombination map for the blackcap, (ii) assess the variability of recombination rates across the genome, and (iii) analyse the variation in recombination rate across the genomes with respect to specific genomic features such as chromosome size, GC content, density of genes and CpGi, functional elements, and different families of retrotransposons. Finally, we (iv) evaluate the conservation of the recombination map in a comparative framework, including data from its closest sister species, the garden warbler, *Sylvia borin*. Using European blackcaps as focal species to characterise fine-scale recombination rates and their association with genomic features not only allows an unprecedented resolution in a wild bird species, but also generates a deeper understanding of this iconic species, and inform mapping studies to characterise the genetic basis of traits such as migration.

2.3 Results

Inter- and intra-chromosomal recombination rate variation

The population-scaled recombination rate average was 5.9 cM/Mb across the European blackcap genome (as shown in Table S1), The fine-scale characterisation of recombination in the blackcap revealed a heterogeneous pattern with variation in recombination rates among and within chromosomes (Fig. 1A), where rates varied substantially between autosomes, with extremes as low as 0.1 and as high as 30 cM/Mb (Table S1). The recombination rate significantly increased with decreasing chromosome size (Fig. 1B, Kendall's tau (r_τ) =-0.61, p=4.7e-8). Avian genomes contain macro and micro-chromosomes, and particularly micro-chromosomes (<20 Mb) had almost three-

fold higher average recombination rates of 13.9 cM/Mb, Table S1) compared to macro-chromosomes with an average rate of 4.8 cM/Mb (Welch's two-sample *t*-tests, p=2.8e-5). In most micro-chromosomes, recombination was elevated along the entire chromosome (Fig. 1A, C, Table S1, SD ± 11.06), yet two of the shortest chromosomes, chromosomes 31 and 32, had particularly low recombination rates (Fig. 1B, Table S1). These chromosomes also have lower SNP densities compared to all other autosomes (8.3 and 3.1 variants/kb on average for chromosomes 31 and 32, respectively, while the minimum average of all other autosomes was equal to 14.3 variants/kb). Macro-chromosomes displayed a central recombination desert and elevated recombination rates towards the distal regions of chromosomes, followed by a sharp drop at the very ends (Fig. 1A).

The lowest average recombination rate was seen on sex chromosome Z (1.15 cM/Mb, Fig. 1A**.** Table S1). At the same time, this chromosome displays a 65 kb interval with the overall highest recombination rate (163.17 cM/Mb) at one chromosome end (Fig. 1A), located within a 800 kb block of elevated average recombination rate (89.5 cM/Mb).

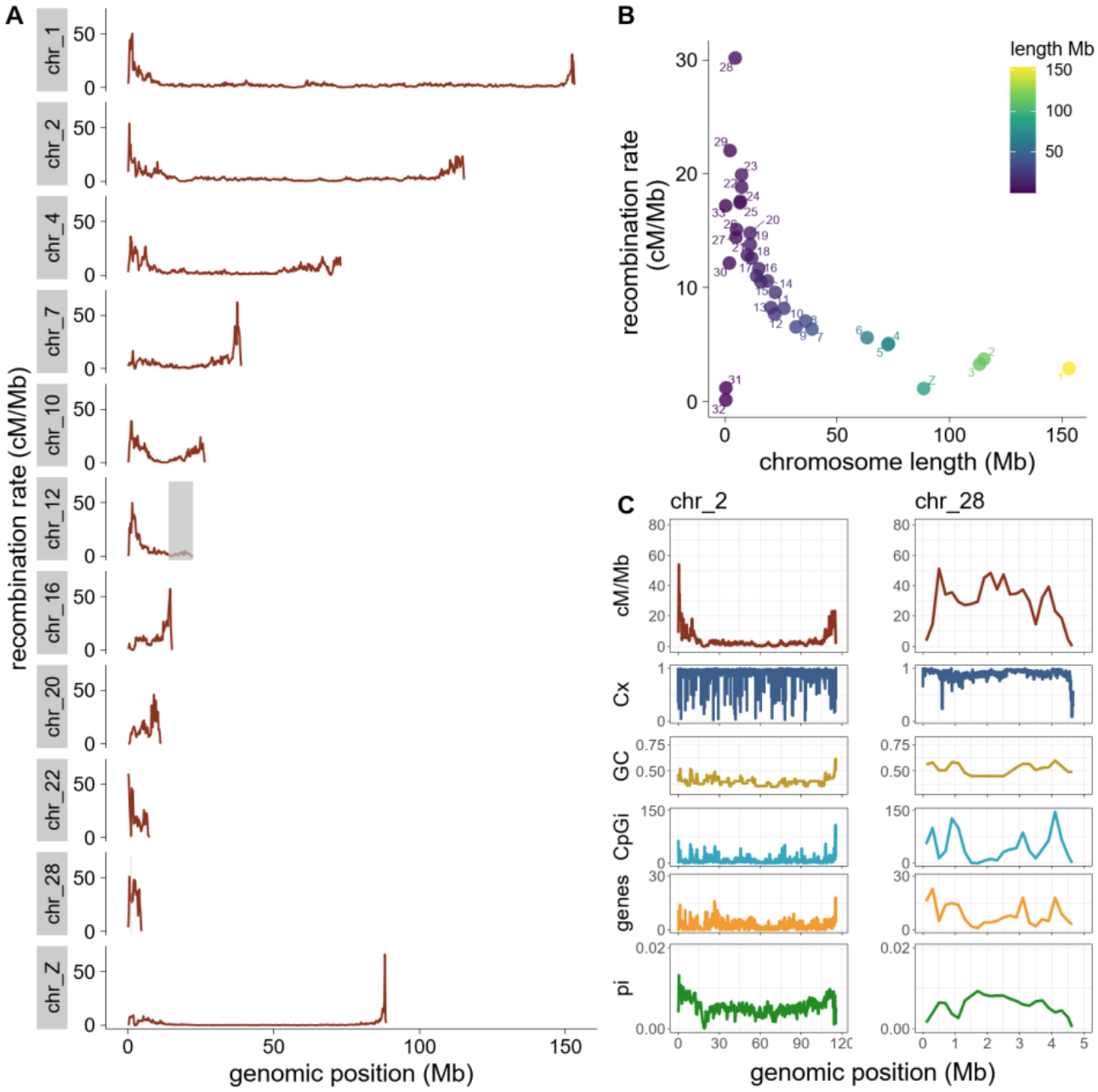


Figure 1 Historical recombination map across the genome and within chromosomes in the European blackcap A Recombination rates calculated in 200kb windows across genomic positions in a subset of chromosomes representative for different size. **B** The relationship between chromosome size and recombination rates measured by Kendall's rank correlation coefficient (τ =-0.6, p=4.7e-08). **C** Distribution of recombination rates and genomic features (from top to bottom: complexity Cx, GC content, CpGi density, gene density, and nucleotide diversity as pi) calculated in 200kb windows across exemplified chromosomes 2 and 28.

Recombination rates are strongly associated with GC-content and CpG islands

We found recombination rates strongly associated with GC content (r_τ = 0.4, p<0.0001) and CpGi (r_τ = 0.3, p<0.0001) density genome-wide. These relationships were also significant after performing partial correlations and when using different window sizes (Fig. 2A, Supplementary Table S3). Additionally, micro-chromosomes had significantly higher GC content (r = -0.84, p<0.001) and density of CpGi (r = -0.68, p<0.001) (for example see chromosome 2, 28, Fig1C, Table S2) compared to macro-chromosomes.

Considering that CpGi may play an important role in recombination dynamics and transcription regulation, we investigated the intra-chromosomal distribution of CpGi density and recombination rate further. Both showed similar patterns across most chromosomes, and we observed peaks in recombination rates overlapping with CpGi dense regions (chromosome 2, Fig. 1C, Fig. S2). in some chromosomes, however, we saw contrasting patterns where regions of low recombination were instead CpGi-rich (chromosome 28, Fig. 1C**,** Fig. S2). In order to decipher whether recombining rates were increased within CpGi or relatively abundant in adjacent regions, we took recombination rates as a function of distance from the nearest CpGi into account. We defined the starting point of each CpGi as zero and calculated recombination rates in 5 kb windows spanning a total of 200 kb regions upstream and downstream. This analysis revealed that recombination increased within the entire CpGi track, flanked by an interval of decreased recombination rates in the adjacent up and downstream regions (Fig. 2B).

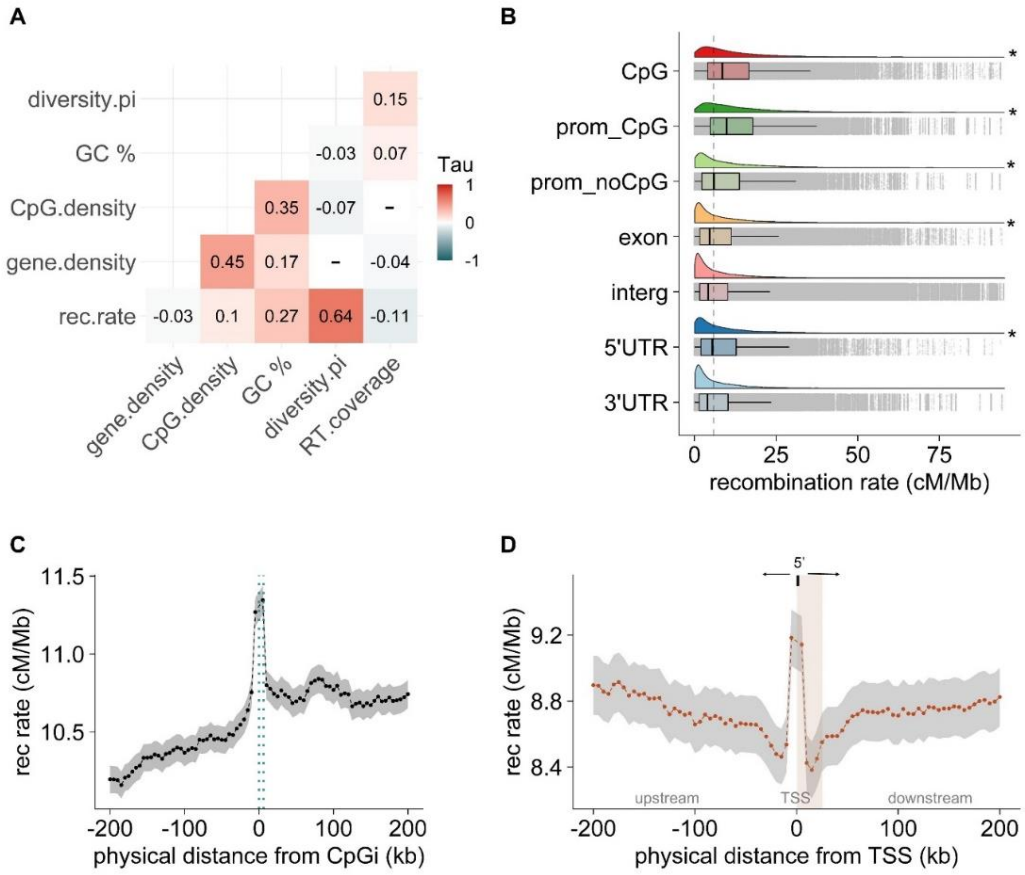


Figure 2. Recombination variation is associated with genomic features and functional elements. A Genome-wide partial Kendall's rank correlations between recombination rates and genomic features calculated in 200kb windows. All correlations are significant (p <0,001), except those marked with (-)**. B** Cloud-plots and boxplots showing median and average recombination rate in different annotation categories. The dashed line denotes the average genome-wide recombination rate (5.8 cM/Mb). Features with significantly higher recombination median compared to the genome-wide average recombination rate are marked with * (Wilcoxon rank-sum test, p<0.001). **C** Average recombination rate as a function of distance to the nearest CpG island. Dotted lines indicate start and end of CpG islands. **D** Average recombination rates as a function of distance to the nearest Transcription Start Site (TSS), the orange shadow denotes mRNA of the genes annotated. Dots represent the average recombination rate for each 5kb window and the grey shadow shows 95% confidence interval.

High recombination rates in regulatory regions but not necessarily in gene bodies

Recombination rates were positively correlated with gene density in genome-wide pairwise comparison (r_τ = 0.2, p<0.001, Supplementary Table S3). However, after controlling for variables such as GC content and CpG density with partial correlation the association became slightly

negative (r_τ = -0.03, p<0.05, Fig. 2A). Shorter chromosomes had significantly higher gene density (r = -0.96, p<0.001). The intra-chromosomal distribution of gene density and recombination rates revealed variable patterns with high recombining regions in low gene density in some chromosomes (for example, see chromosome 28, Fig. 1C, and chromosomes16 and 18 Fig. S3). The relationship of recombination with genes or functional elements may not only depend on the density of genes but also on the location and specific regulatory motifs surrounding them and controlling their expression. These included *cis*-regulatory elements such as promoters and 5'prime regions, which were significantly enriched for higher recombination rates (Fig. 2C, Wilcoxon rank test corrected p<2e-16). In contrast, recombination rates within intergenic regions and 3'UTR regions were significantly lower than the genome-wide recombination average (Fig. 2C, Wilcoxon test p<0,01). Since recombination rates in organism lacking PRDM9 are usually elevated in TSS, we evaluate if the same occurs in the blackcap genome. Hence, we took recombination rates as a function of distance from TSS and we confirmed the significant increase of recombination rates in TSS compared to surrounding regions. In fully annotated genes, a peak of recombination in TSS was followed by a sharp decline in recombination rate in the adjacent downstream region, including gene bodies and 3-UTR regions. (Fig. 2D). Similarly, recombination rates in promoters with CpGi were significantly higher compared to those in promoters without CpGi, or CpGi in general (Fig. 2C, Wilcoxon rank test corrected, p<2e-16). In line with these results, CpGi and gene density showed a strong genome-wide association (Fig. 2A).

High recombination patterns coincide with unique sequences across the genome

We inferred complexity (Cx) as the uniqueness of sequences across the genome (ranging from 1 representing unique sequences to 0 for repeated sequences). Genome-wide complexity was 0.9, and intra-chromosomal distribution of Cx along shorter chromosomes displayed elevated patterns and few dips, whereas large chromosomes showed a distinctly more heterogeneous complexity pattern (Fig. 1C, Fig S1). Some regions with elevated recombination rates overlapped with areas of high complexity (or uniqueness) (Fig. 1C, Fig. S1). We observe drops of complexity in the distal regions in most chromosomes, where recombination rates also decline drastically (Fig. S1). Other regions of low complexity coincide with regions containing few CpGi and low gene density (chromosome 28**,** Fig. 1C). Nevertheless, the between-chromosomes comparisons reveal that complexity overall associated positively with chromosome size (r_τ = 0.6, p = 1.1e-07) and negatively with recombination rate (r_τ = -0.31, p = 0.01).

LTRs and non-LTRs elements associate differently with recombination rates

When we evaluated the association between recombination rates and coverage of retrotransposons (RTs) at the genome-wide level in 200-kb windows, recombination rates were not significantly correlated with RT coverage (Table S3). In the partial correlation, the coefficient became negative (r_τ = -0.09, p<0.0001, Fig. 2A). To clarify and further characterise the complex relationship between RTs and recombination, we analysed different RT families separately, LTR, LINE and SINE, as they have distinct transposition mechanisms and evolutionary histories. The association of recombination rate and RT coverage varied with each family: LTR and SINES were negatively associated with recombination rate (r_τ = -0.064, p = 5e-11 for LINE, r_τ = 0.1, p = 2.2e-11 for SINE, Fig. 3A), and LINES showed a weak but significant positive association pattern (r_τ = 0.052, p = 6.8e-08, Fig. 3A). The characterisation of RT families across different chromosomes revealed high coverage of LTR and LINE (non-LTR) elements, particularly at the distal parts of some macro-chromosomes where recombination rates are also elevated (Fig. S4, S5). In contrast, SINE (non-LTR) elements were predominantly found in the central part of chromosomes (Fig. S6). Within chromosomes, LTRs were additionally located within or near regions where recombination is suppressed.

To assess whether the presence of RTs contributes to the variation of recombination rates in specific genomic regions, we calculated and compared the median and average recombination rate of CpGi, promoters, genes, and intergenic regions with RTs, including the three families, and without them. Recombination rate was significantly higher in CpGi overlapping with LINEs and SINEs compared to CpGi alone (Wilcoxon rank-sum test, p<2.2e-16 for LINE and p<0.05 for SINE). A significantly elevated recombination rate was also observed in genes containing LTR and LINE elements than in those without these two families of RTs (Wilcoxon rank-sum test, p<2.2e-16). The recombination rate within intergenic regions and promoters associated only with LINEs was significantly higher compared to intergenic regions and promoters alone (Wilcoxon rank-sum test, p<2.2e-16) (Fig. 3B).

We also investigated each RT family's average recombination rate overlapping with different genomic features. We found that overall, all investigated features associated with LINE elements showed significantly elevated recombination compared to the same features overlapping the other two families of RTs (Wilcoxon rank-sum test, p<0.0001, Fig. 3B), except in CpGi, where the average recombination rate of CpGi with SINEs and LINEs were not significantly different

(Wilcoxon rank-sum test, p = 1, Fig. 3B). LINE and SINE elements located within CpG and promoters had increased recombination rates compared to the same elements located within genes and intergenic regions (Wilcoxon rank-sum test, p<2.2e-16, Fig. 3B). In contrast, LTRs in genes showed significantly higher recombination rates than LTRs located within intergenic regions, promoters, and CpGi, where recombination rates were generally very low (Wilcoxon rank-sum test, p<2.2e-16, Fig. 3B).

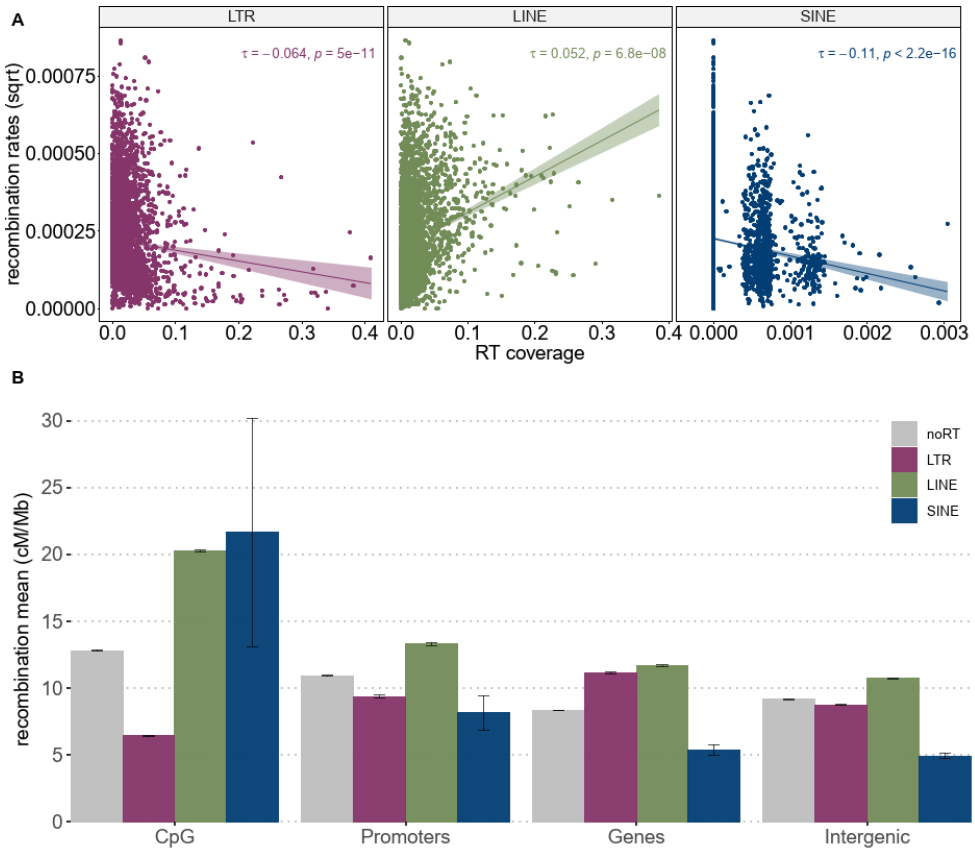


Figure 3. Association between recombination rates and retrotransposons. **A** Correlation between recombination rates (square root transformed) and the coverage of three types of RT in 200kb windows measured with Kendall's tau. **B** Bars plots showing the mean recombination rate of different annotation features intersecting with different types of retrotransposons and without RTs (grey). Errors bars represent 95% confidence intervals. Recombination average in genomic features with RTs are significantly different to genomic features without RTs (Wilcoxon rank-sum test, p<0.05). The average recombination rate of each RT at different genomic features are significantly different (Wilcoxon rank-sum test, p<0.001), except for SINE in genes and intergenic regions.

Heterogeneous conservation of recombination rates between closely related sister species

To assess the similarity of recombination maps of the blackcap with its closest sister species, the garden warbler, recombination rates in this species were estimated using the same pipeline as for the blackcap. Genome-wide average recombination rate in the garden warbler was 2.07 cM/Mb. Similar to the recombination map in the blackcap, micro-chromosomes had much higher recombination rates (5.13 cM/Mb on average) compared to macro-chromosomes with an average of 1.67 cM/Mb, which also showed deserts in the middle and peaks of recombination at chromosome ends.

We compared recombination rates between both species in 50 kb windows to analyse the conservation of these apparent similarities. We found a significant correlation of recombination rate genome-wide ($r_\tau = 0.60$, p = 2.2e-16), which increased with wider window sizes (Table S4). We also analysed the similarity at the level of individual chromosomes, except in some of the micro-chromosomes, 28, 29, 31, 32, where the SNP density from at least one species was not sufficient to perform a statistical comparison (Fig. S7). We observed a large variation in the degree of interspecies correlation between chromosomes. Recombination maps of autosomal chromosomes 2, 4, 6, 9, 11, 13, 19, 20, 21, 23, and 25 showed high similarity (Fig. 4A, Fig. S7, S8). Additionally, recombination rates along chromosome Z are highly similar between species ($r_\tau = 0.47$, p<2.2e-16). Lower correlation coefficients and thus less conservation between species was observed at chromosomes 8, 10, 15, 18, 22, and 27 (Fig. 4B, Fig. S7, S8).

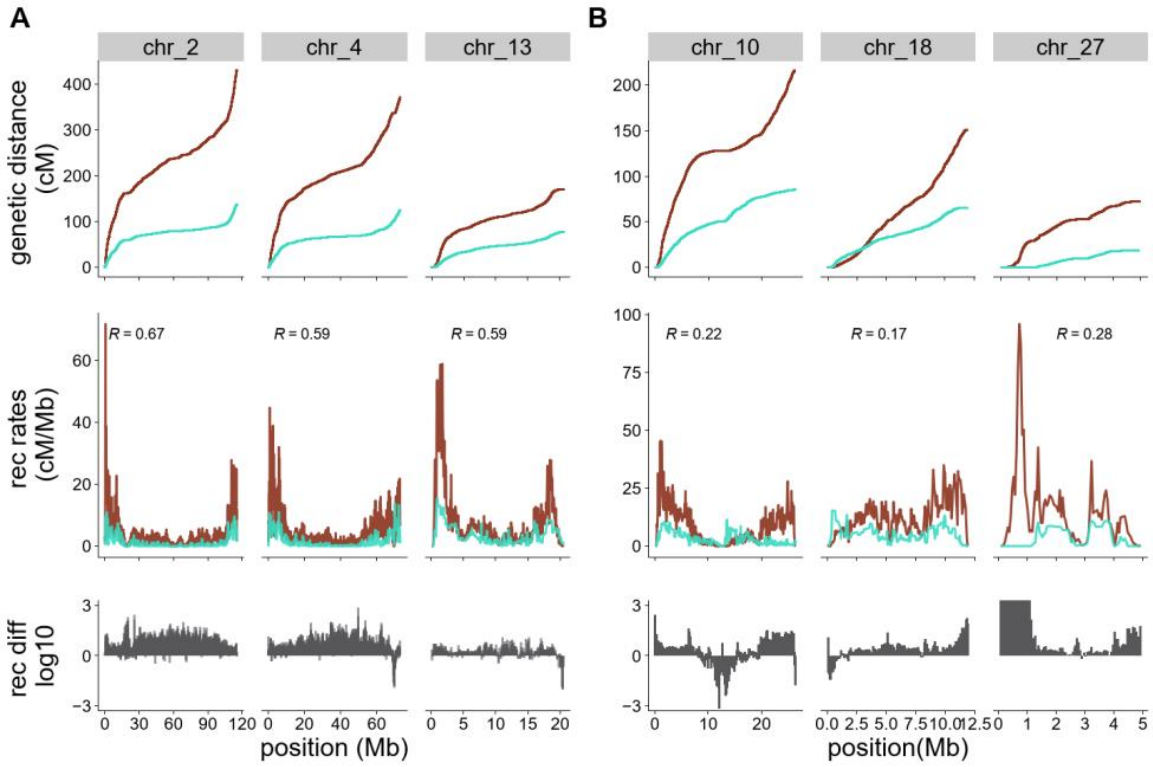


Figure 4. **Historical recombination map comparison between closely related sister species.** Recombination rates and genetic map in the blackcap (dark-red) and the garden warbler (cyan). Different chromosomes showing **A** higher similarity and **B** lower similarity of recombination rates between both species. The first row shows the genetic distance in cM, the middle row shows the distribution of recombination rates including the correlation coefficients measured by Kendall's rank correlation. All comparisons are significant $p<0.001$. The third row shows the normalized recombination rate difference calculated between both species across different chromosomes.

2.4 Discussion

Genomic and intra-chromosomal recombination rate variation

We show that historical recombination rates in the blackcap genome are highly variable both between and within chromosomes. Our fine-scale recombination rates were inferred from a natural population of blackcaps explicitly accounting for its demographic history. Average recombination rate was higher in blackcaps (5.8 cM/Mb) compared to zebra finch (1.5 cM/Mb) and collared flycatcher (3.69 cM/Mb) (Backström et al. 2010; Singhal et al. 2015; Kawakami et al. 2017). However, our results also call for caution when comparing absolute values since recombination maps were estimated using different approaches and differed parameters specific for each dataset (such as block penalties of 10 or 100 and bigger windows sizes of 100 or 200 kb). To allow for a more detailed comparison, we estimated recombination rates of a closely related sister species, the garden warbler, using the same approach with identical parameters, and we found an average recombination rate of 2.07 cM/Mb, which falls within the same order of magnitude as reported for other avian species. The difference in average recombination rates represents distinct evolutionary histories or different rates of COs in the two species, and probably different marker densities, since a smaller sample size of garden warblers was available. Accurate recombination estimation depends on the number of variable sites (SNPs) that can be analysed and the parameters used for the inference. The quality of our reference genome allowed for a high-quality assembly where SNPs could be confidently inferred and SNPs density remained high after hard-filtering.

Most of the chromosomes across the garden warbler genome show lower recombination rates compared to the blackcap, but common patterns, such as elevated recombination rates at the ends and deserts of recombination in the middle of macro-chromosomes are observed. This patterning, particularly suppression at the chromosomal centre, was also characterised in zebra finches (Backström et al., 2010; Peñalba et al., 2020; Singhal et al., 2015), and may reflect the position of the centromeres. Historical recombination maps were reported to be heterogeneous also in other passerine species (Kawakami et al., 2017; Peñalba et al., 2020; Singhal et al., 2015). The increased recombination towards chromosome ends also appears to be a general pattern, reported in many bird species (Singhal et al. 2015; Pigozzi 2016; Kawakami et al. 2017; del Priore and Pigozzi 2020), diverse organisms lacking PRDM9 (Campbell et al., 2016; Haenel et al., 2018; Higgins et al., 2012; Lam & Keeney, 2015; Yelina et al., 2015), and organisms with PRDM9, including humans (Campbell et al., 2016). A spatiotemporal theory explaining the increase in chromosome arms was

reported by Higgins et al. (2012) and Haenel et al. (2018) suggesting that recombination initiation occurs at the distal part of the chromosomes, where DSB and crossover events occur more often in comparison to the chromosome centres where euchromatin is present. Additionally, Haenel et al. (2018) propose that the high CO rates are given primarily in the chromosomes' peripheral or distal region. It was proposed by Thacker et al. (2014) that chromosome homology pairing initiated at the telomeres is slower, resulting in prolonged DSB activity on these chromosomes. Supporting this idea, Subramanian et al., 2019, revealed that in 100 kb regions flanking telomeres, continuously received DSBs even after Synaptonemal Complex (SC) components down-regulate interstitial recombination initiation in yeast. The high marker density of the blackcap reference genome allows for detecting such increased recombination rates at the chromosomal ends, even though these regions are of low genome complexity, overcoming limitations of previous studies which often had insufficient resolution, at chromosome ends, and other highly repetitive regions as they lacked the high-resolution genomic scaffold provided by whole-chromosome optical map.

The Z chromosome showed a differentiated pattern compared to the autosomes. The historical recombination rates were low along the chromosome except for a specific region at the end. This region has been identified and described before in other bird species like the flycatcher (Smeds et al. 2014; Kawakami et al. 2017) as the pseudoautosomal region (PAR) characteristic of sex chromosomes. This region represents obligatory synapsis and crossover between sex chromosomes where homologous recombination occurs to maintain the correct segregation of chromosomes at the meiotic spindle (Mohandas et al. 1992).

Differential recombination rate according to chromosome size

Genome-wide recombination rates are higher in birds than in mammals, reptiles, and other organisms (Stapley et al. 2017; Malinovskaya et al. 2018). Recombination rate variation is influenced by the genomic architecture as well as chromosomal organisation. Given that segregation requires an obligatory crossover (CO) for each chromosome pair, a larger number of recombination events are needed in genomes with higher numbers of chromosomes (Fledel-Alon et al. 2009). Birds appear to have more than one crossover per chromosomal pair, as reported in cytological studies that estimate CO from the number of foci of the MLH1 protein. Three or more MLH1 foci were seen in chicken macro-chromosomes, up to six in the Japanese quail and 1.5 or 2 in the zebra finch, on average (Calderón and Pigozzi 2006; del Priore and Pigozzi 2020).

Our results suggest that high recombination rates in birds can be sufficiently explained as stemming primarily from regions of highly elevated recombination in micro-chromosomes with a higher density of functional elements. Elevated recombination rates have been reported in birds and other genomes with shorter chromosomes such as fungi, plants (e.g., *Arabidopsis thaliana*), insects (e.g., honeybee, bumblebee), and snakes (Haenel et al. 2018; Schield et al. 2020). Cytological studies illustrated an increased density of CO in shorter chromosomes, which explains the high recombination based on the ratio of CO numbers to the chromosome length (Malinovskaya et al. 2018). Genome-wide complexity, defined as sequence uniqueness or non-repetitiveness, was higher in blackcaps (0.9) compared to humans (0.80) and mice (0.78), where high complexity was also found in functional elements and captured transposon insertions and copy number variation (Pirogov et al., 2019). In our data, the intra-chromosomal distribution showed generally elevated complexity patterns in micro-chromosomes in contrast to macro-chromosomes, yet the inter-chromosomal association between low complexity and high recombination remained positive. This correlation could be influenced by the estimation of complexity over the entire chromosome, probably not reflecting fine-scale fluctuations of complexity, especially within macro-chromosomes. Moreover, high recombining features like promoters and regions upstream of genes are abundant in tandem repeats in avian genomes, probably lowering the complexity value in those regions. Differences between shorter and longer chromosomes, are consistent with micro-chromosomes having a higher density of coding regions and fewer repetitive elements compared to macro-chromosomes (Consortium 2004; Waters et al. 2021). We also found regions within chromosomes showing low complexity, high abundance of LTRs, and low SNP diversity, which could hint at the presence of centromeres, as the human genome's centromeres were located in relatively long stretches of very low complexity (Pirogov et al., 2019).

Recombination variation associated with genomic and annotation features

Variation in recombination rates was significantly associated with specific genomic features. We confirmed a positive correlation between recombination and GC content, similar to other studies in birds (Singhal et al. 2015; Peñalba et al. 2020) and mammals (Galtier et al., 2001; Smukowski & Noor, 2011). It is not yet understood whether high GC content is a cause or a consequence of elevated levels of recombination, as most evidence of GC-bias is indirect and comes from observation of GC enrichment at sites with elevated recombination rates. Therefore high GC content could promote recombination initiation, and recombination promotes GC enrichment via

GC-biased gene conversion (gBGC), resulting from DSB repair favouring G:C over A:T nucleotides in highly recombining regions. Direct evidence for GC-biased gene conversion mechanism has thus far only been demonstrated in humans, where it is generated in a PRDM9 dependent manner, but independently of crossover (Odenthal-Hesse et al. 2014).

Gene density was positively correlated with recombination rates, and slightly negatively correlated in partial correlations. This association might not only be influenced by the density or proportion of genes, but also by their location and *cis*- or *trans*- factors regulating their mode of expression as suggested previously (Peñalba et al. 2020). Indeed, most recombination events did not occur within genes but instead regulatory regions such as TSS, promoters, and 5'UTR regions. Our findings are similar to those in other organisms lacking PRDM9, including other bird species but also canids, yeast, and plants (Auton et al. 2013; Choi et al. 2013; Lam and Keeney 2015; Singhal et al. 2015; Kawakami et al. 2017). In contrast, to our findings, most recombination events in PRDM9-dependent organisms like mammals (e.g., mouse) were instead concentrated in intergenic regions or away from functional elements. Interestingly, in mouse knockouts for PRDM9, recombination events appear at default open chromatin regions, such as promoter-like features associated with H3K4me3 peaks (Brick et al. 2012; Mihola et al. 2019). Our data offer additional support to the idea that recombination events predominantly occur in functional or regulatory regions in the absence of PRDM9.

From an evolutionary perspective, if recombination primarily reshuffles functional elements in genic regions in organisms lacking PRDM9, the role of recombination might be pivotal in the adaptation and maintenance of variability of functional elements in the genome. In support of this hypothesis, recombination rates in blackcaps were significantly associated with nucleotide diversity. Similar correlations have previously been found in yeast (Tsai et al. 2010), three breeds of chicken (Mugal et al. 2013), and other passerine species (Kawakami et al. 2017). The positive correlation is consistent with the expectations of genomic polymorphism being shaped by linked selection (Burri et al. 2015), particularly background selection where low recombining regions lead to low diversity.

Our data confirm that CpGi are key players in shaping recombination rate variation, with CpGi being strongly associated with recombination rates and gene density. Historical recombination events in blackcaps located primarily in regulatory features containing CpGi, further supporting CpGi as predictors of high recombination. Since CpGi are usually associated with DNA hypomethylation at gene promoters, gene expression, and transcription start sites (Jung and Pfeifer

2013; Angeloni and Bogdanovic 2021), recombination may be linked to regulatory regions and potentially to chromatin modifiers such as H3K4me3 and low levels of DNA methylation (Choi et al. 2013). Perhaps the contradictory patterns we found for some chromosomes with regions of high recombination and low CpGi (or the opposite) could be explained by CpGi located outside functional elements or inactive promoters, where DNA is generally hypermethylated, and recombination may be suppressed. Additionally, the presence of retrotransposons may alter methylation patterns of CpGi and adjacent regions. Our results indicate that the recombination landscape is not shaped solely by the location of functional elements but may also be influenced by their methylation status.

The differential association between recombination rates and retrotransposons

Our results show evidence of various evolutionary scenarios explaining the differential correlation of recombination with different types of RTs. Similar to our results, a negative association between recombination and retrotransposons was found in insects (e.g. *Drosophila*) and cotton (Rizzon et al. 2002; Shen et al. 2017). However, in the flycatcher and zebra finch (Kawakami et al. 2017) a positive association was reported. Among different factors, such as the mechanism of replication, insertion sites affinity, selective pressures, the methylation patterns, the age of RTs (Bourque et al., 2018; Kent et al., 2017; Sultana et al., 2017), as well as the type of RTs can heavily influence the direction of this association as our results demonstrate. We found evidence for a bidirectional correlation between recombination and non-LTR elements, with LINE positively and SINE negatively associated with recombination.

The positive association with LINEs could be explained by the fact that RTs can transcribe themselves by targeting upstream regions of genes transcribed by RNA pol III. This suggests that LINE elements, would predominantly be found in open chromatin stretches including functional genetic elements (e.g., promoters), similar to the recombination machinery. Moreover, RTs seem to have an affinity for AT-rich motifs (Abrusán et al., 2008; Lerat et al., 2002), which were associated with elevated recombination in birds and with nucleosome-depleted regions in plants and yeast (Jansen and Verstrepen 2011; Choi et al. 2013). This could reflect a scenario where RTs play beneficial roles in the regulation of gene expression or promoter dispersion (reviewed in Bourque et al., 2018). For instance, LINE-like retrotransposons form and maintain telomeres in Drosophila, replacing the role of the telomerase enzyme (Pardue and DeBaryshe 2011).

Interestingly, in the blackcap genome, we also found particularly high distributions of RTs in the distal part of some chromosomes, which is a pattern also reported in rice (Tian et al. 2009).

Our results demonstrate that the presence of LINEs is associated with the increased local recombination rate in specific genomic features such as promoters, genes, and CpGi, where the recombination rate is already elevated. However, LINEs were also found in intergenic regions where recombination is generally lower, though the average recombination rate in intergenic regions containing these elements was significantly lower than the other features. Here, intergenic regions contain introns and other genomic sequences located out of genes, which could interact with LINEs at some extent. Insert-site bias may also explain the low correlation coefficient, as it is widespread in LINEs, particularly in the CR1 subclass which is characteristic for birds. In contrast to L1 elements in mammals, high densities of CR1 were found not only in AT-rich, but also GC-rich regions where selection might have more efficiently acted on removing insertions from these sites, which can influence the association with recombination rates and the differential distribution of LINE families according to the age. At least in the chicken genome, young TEs were more predominant found in regions with elevated GC content compared to older families of TEs (Abrusán et al. 2008). Similarly in flycatchers, young retrotransposon subfamilies were also located in high recombination regions (Kawakami et al. 2017).

On the other hand, recombination rates increase with decreasing SINE coverage. SINE elements prefer GC sequences (Smit 1999), apparent in our data as the recombination rate in CpGi with SINE was significantly higher compared to other annotation features or CpGi without RTs. This suggests that SINE elements may partially be associated with increased recombination rates in GC-rich sequences and CpGi. The negative association of SINEs with recombination rates may point to a regulatory role via chromatin repression, further supported by our observation that recombination rates in promoters, genes and intergenic regions with SINEs were significantly lower than without these elements. However, methylation patterns and age might influence this interaction drastically, as younger SINE subfamilies show higher methylation compared to old subfamilies (Pehrsson et al., 2019). Moreover, in humans, young Alu elements are randomly distributed along the genome, while old elements, whose methylation patterns were fixed, were predominantly found in GC-rich sequences (Smit 1999; Pehrsson et al. 2019). Hence, the methylation patterns of different subfamilies of SINEs could be contributing to recombination variation.

We observed a weak negative association of recombination with LTRs and a high coverage distribution of these elements in regions where recombination is low or even suppressed, potentially indicative of centromeres (Fig. S5). In agreement with that, Rizzon et al., 2002 reported LTRs enriched in low gene density, GC-content sequences, and low recombining regions like centromeric and pericentromeric regions. Similarly, a study in chicken and rice genomes revealed a negative correlation between LTR densities and recombination and gene densities (Tian et al. 2009; Mason et al. 2016). Even though we found a higher average recombination rate in genes overlapping LTRs compared to genes without them, recombination in promoters and CpG, did not increase with the presence of these elements. This suggests a possible scenario where either recombination is boosting selection against RTs insertion and further expansion on the genome (Dolgin and Charlesworth 2008), or RTs are potentially playing a role in repetitive element maintenance, and contribute to recombination suppression.

The genome-wide distribution of RTs, their specific family, mechanisms of replication, and insertion sites have an important consequence on the variation of recombination rates. It remains unclear, whether RTs are removed from highly recombining regions through purifying selection, or whether they are accumulated in non-recombining regions instead, where selection is weaker (Dolgin and Charlesworth 2008), potentially contributing to recombination suppression (reviewed in Kent et al., 2017). Specific families of RTs in highly recombining regions may play beneficial roles in the regulation of gene expression or promoters dispersion (reviewed in Bourque et al., 2018, Pardue & DeBaryshe, 2011), and selection might not act against them.

In a broader context, our study not only characterises the fine-scale population historical recombination map accounting for past demographic fluctuations (and thus alleviates potential bias in the inference of recombination rates in a wild avian system), but also provides an important toolset to focally study sources of selection and other evolutionary forces shaping variation in behavioural traits and populations dynamics. Moreover, the identification of conserved and non-conserved regions across the recombination maps in very closely related species highlights the importance and potential to expand our characterisation of recombination analysis from this focal resident population to additional populations and species using variable demographic data, selective pressures, and species-specific mutation rates. With the advances in sequencing technology and the availability of a growing number of high-quality bird genomes, the approach applied here should facilitate the characterisation of comparable recombination maps in the future.

2.5 Materials and Methods

Samples, reference genome, and annotation

We used whole-genome resequencing data of 19 resident male blackcaps collected from two locations in Spain (Cazalla de la Sierra n=11, Gibraltar n=8) that were previously reported as part of a population genomics analysis (Delmore et al. 2020). Individual sequences were mapped against our chromosomal-level reference genome of a female from a resident population in Tarifa, Spain. The primary and alternate haplotypes were assembled using the Vertebrate Genomes Project (VGP) pipeline v1.5 (Ishigohoka et al. 2021; Rhie et al. 2021) and then made available under NCBI BioProject PRJNA558064, accession numbers GCA_009819655.1 and GCA_009819715.1 (see Ishigohoka et al., 2021 for more details).

Genome annotation was carried out using MAKER (Campbell et al., 2014). We ran RepeatMasker 4.0.9 (Smit et al., 2015) to soft-mask the genome assembly, using a library of transposable elements from the collared flycatcher (Suh et al. 2018) and blue-capped cordon bleu (Boman et al. 2019) as input. We *de novo* assembled a transcriptome from RNA-Seq data of 3 different brain regions (two blackcap individuals for each brain region) using Trinity (Grabherr et al. 2011). This data was combined with an ISOSeq based transcriptome of the brain and reproductive organs from one male and one female blackcap to serve as evidence for gene prediction. Additionally, cDNA and protein sequence of three avian species (zebra finch, chicken, and flycatchers) from Ensembl and TrEMBL as well as manually curated avian cDNA and protein sequence from RefSeq and Swiss-Prot were provided as further evidence. We used the *ab initio* gene predictors Augustus (Keller et al. 2011) (species chicken) and SNAP (Korf 2004) to refine the gene predictions. SNAP was trained iteratively on the output of MAKER to improve its accuracy. In the final annotation, we included only gene predictions with Annotation Edit distance (AED) ≤ 0.5. We performed functional annotation using BlastP with the Swiss-Prot database and InterProScan with the Pfam application on the predicted protein sequences.

Mapping and variant calling

For mapping and variant calling we used a modified pipeline based on the GATK best practices for short-read sequencing. First, the quality of adapter bases in the sequenced reads of the 19 samples was set to 2 with Picard MarkIlluminaAdapters, followed by paired-end mapping to the reference using BWA mem (Li 2013 May 26). Then, we marked reads as duplicates based on the

mapping position and insert size since we wanted to keep unmapped read mates and supplementary reads for completeness. To evaluate how well the samples mapped to the reference, we performed a quality control of the generated read alignments (BAM files) using QualiMap version 2.2.1 (Okonechnikov et al. 2016); Picard CollectMultipleMetrics, CollectRawWgsMetrics and CollectWgsMetrics; and MultiQC version 1.8 (Ewels et al. 2016). We used GATK version 4.1.6.0 HaplotypeCaller to call single-nucleotide polymorphic sites (SNPs) per sample and output them as gVCF files. Afterwards, the 19 gVCF files were combined using GATK CombineGVCFs. From the gVCF file with 19 samples, we called variants with GenotypeGVCFs. SNPs were selected using GATK SelectVariants, and hard-filtered with the following criteria: QD < 2.5; FS > 45.0; SOR > 3.0; MG < 40; MQRankSum < -12.5; ReadPosRankSum < -8.0. Finally, we selected 19,917,215 variants by applying a Genotype Quality filter of 20, a minimum depth of 5, a maximum depth of 60, and a missingness of 0.7, with VCFtools (Danecek et al. 2011). Furthermore, we removed singletons (minimum allele frequency > 0.03) using VCFtools and GATK SelectVariants, as non-reproducible singletons would include most randomly distributed sequencing errors that could bias linkage disequilibrium analyses. We extracted scaffolds and chromosomes by taking only bi-allelic sites using GATK SelectVariants. We only considered reads mapping to chromosomes for the recombination estimation and further analysis. In addition to the 19 individual blackcap resequencing data, we also included resequencing data for five individuals (four males, one female) of the closest sister species, the garden warbler. The VCFs with hard-filtered SNPs were taken from Ishigohoka et al., 2021.We further filtered SNPs as previously described yielding 11,386,509 SNPs for further analysis.

Inferring historical recombination rates across the blackcap genome

We used Pyrho to estimate population-specific recombination rates per generation and characterised their distribution across the genome (Kamm et al., 2016; Spence & Song, 2019). Pyrho relies on a composite-likelihood approach taking population-specific demography into account to compute lookup tables and calculate the optimal parameters for the estimation. The program estimates and scales the per-generation per-base recombination rate (r) using the effective population sizes (Ne) and mutation rate. Demography for the focal continental resident blackcap population was inferred using MSMC2 as part of a population genetic analysis on blackcaps across the species' distribution range (Delmore et al., 2020). We input unphased genotypes per individual in a VCF format and under the assumption that closely related bird species had similar mutation

rates, we used the mutation rate of a related species, the collared flycatcher ($4.6x10^{-09}$ site/generation) (Smeds, Qvarnström, et al. 2016). After running 'hyperparam' in Pyrho, we found a block penalty of 20 and a window size of 50 kb to give optimal resolution in the recombination rates estimation for our dataset. The estimates of Pyrho (scaled-recombination per site per generation rate) were converted to cM/Mb using a custom Unix script, which we also used to output the linkage genetic map (genetic distance in cM).

We first calculated recombination rates in non-overlapping windows of different sizes (50-kb, 100-kb, 200-kb and 1-Mb) to compare resolution at different window sizes. We used a custom python script to calculate the average rates weighted by the physical distance between each pair of sites where recombination was estimated. We used the 'ggplot2' package in R software to plot recombination rates against the physical distance(R Core Team 2021, Wickham, 2016).

Testing for association between recombination rates and genome complexity

To assess the association of recombination rates with reference genome complexity, a measure of the uniqueness of sequences across the genome, we used "Macle", a program that estimates the match complexity of strings in sequences genome-wide (Pirogov et al. 2019). This program outputs values of 1 for all unique sequences and values as low as 0 for sequences repeated multiple times. Complexity was initially calculated for the whole genome and each chromosome separately. We then calculated complexity maps across the genome using different overlapping window sizes (10, 50, and 200 kb). The best resolution was achieved for a window size of 10-kb, which was then chosen for the final estimation of the genome complexity for each chromosome.

Characterising recombination rate variation with respect to genomic features

To evaluate the genome-wide association between recombination rates and selected genomic features, we calculated the weighted averaged GC-content in 200-kb and 1-Mb non-overlapping windows considering the length of each sequence where GC was inferred. CpG islands, defined as DNA stretches with elevated CG content (usually greater than 50%), and frequent absence of DNA methylation (Jung & Pfeifer, 2013), were identified using a distance-base algorithm implemented in CpGcluster v1.0 (Hackenberg et al. 2006) setting a minimum length of at least 50 bp and maximal P-value of 1E-5. The density and coverage of CpGi and genes were calculated by counting the number of elements and calculating their fraction within 200-kb, and 1-Mb non-overlapping windows using bedtools 'annotate' (Quinlan 2014). We calculated nucleotide diversity

(π) using "all sites" VCF and varying tools from the "genomics general" toolkit (Martin et al. 2020). We used parseVCFs.py (https://github.com/simonhmartin/genomics_general/VCF_processing release 0.4) to filter the "all sites" VCF similarly to the variant VCF, with min-depth >5 and max <60. After which we ran popgenWindows.py (https://github.com/simonhmartin/genomics_general release 0.4), with a window size of 200-kb and ensuring a minimum 20-kb called in each window. As our missingness equivalent we set the proportion of individuals with at least 20kb per window to be 0.7 (-minData). Considering recombination rates are usually exponentially distributed, Kendal's rank correlation test with a significance threshold of 5% was used for all the pairwise and partial correlations in 200-kb and 1-Mb windows performed in R with the 'ggcorrplot' (**https://github.com/kassambara/**ggcorrplot) and "ppcor" package (Kim 2015) respectively. In partial correlation analyses, we controlled for nonpredictor variables by including them as potential confounding variables.

For the inter-chromosomal comparisons, we first estimated the GC content of each chromosome separately using 'Geecee' in EMBOSS (Rice et al. 2000), and we calculated the density of genes and CpGi as the number of elements divided by the length of each chromosome. Then, these variables were correlated with the weighted average recombination rate per chromosome in cM/Mb, which is equivalent to the genetic map distance divided by the chromosome length in Mb. In addition, the high quality of our reference genome with chromosome resolution allowed us to measure the association between chromosome length (measured in Mb and log-transformed) with recombination rates and genomic features separately using Pearson's correlation coefficient.

To investigate the variation of recombination rates with respect to specific genomic functional elements in the blackcap genome, we assigned different annotation categories: exons, intergenic regions, the untranslated upstream and downstream regions surrounding the mRNA (5'UTR, 3'UTR), Transcription Start Sites (first position of annotated mRNA), promoters (defined as 2 kb upstream of TSS) with and without CpGi. Then, recombination rates (in cM/Mb) were intersected with all genomic features using bedtools "intersects" and the average recombination rate was calculated for each annotation category. We performed Wilcoxon rank-sum test with multiple test groups and visualised with the 'ggplot2' package in R. Additionally, we calculated recombination rates as a function of distance from Transcription Start Sites (TSS) and CpGi. We used the bedtools and assigned 0 as the first position to the closest gene and CpGi, respectively. Next, we calculated

the average recombination rate in 5-kb non-overlapping windows spanning 200-kb upstream and downstream, considering the direction of genes for the case of TSS. The same was performed for CpGi. We plotted and performed statistics using R and the 'ggplot2' package.

Characterising the association between recombination rates and retrotransposons

To evaluated the association between recombination rates and retrotransposons (RTs), we utilised the annotation of RTs described in Bours et al., 2022. In brief, the prediction of transposons was calculated using RepeatModeler 1.0.11 (Bao et al. 2015), LTRs were predicted using LTRharvest (Ellinghaus et al. 2008) and LTR-related HMM (Hidden Markov Model) (Mistry et al. 2021). Transposons were then annotated with RepeatMasker 4.1.0 (Smit et al., 2015) using manually curated repeat libraries of two bird species.

We classified RTs into three families: (i) Long Terminal Repeats (LTR) and two subclasses of non-LTR retrotransposons: (ii) Long-interspersed nuclear elements (LINEs) and (iii) short interspersed transposable elements (SINEs). We then calculated the density and coverage for all RT together, and for each family separately, using 200 kb windows and measuring the correlation between recombination rates (square root transformed) using Kendall's rank correlation test in R. The coverage was used for the analysis since a number of RTs in the automated pipeline are fragmented and taking the density could skew the association of these RTs.

To assess whether the presence of RTs is associated with recombination rate variation in specific annotation features, we calculated and compared the median and average recombination rates of CpGi, genes, promoters, and intergenic regions with (overlapping) and without (non-overlapping) any of the three families of RTs. We additionally compared the average recombination rate of each RT family overlapping with different genomic features to characterise the association between TEs and recombination rates along the genome. We conducted unpaired two-sample Wilcoxon rank-sum tests (Mann-Whitney test) for each genomic feature and each RT family separately.

Evaluating the conservation of recombination maps between the blackcap and garden warbler

To evaluate the conservation of historical recombination rates, we compared the recombination maps of the blackcap with its closest sister species, the garden warbler. To estimate recombination rates in the garden warbler, we used Pyrho following the same pipeline as for the blackcap (Materials and methods 2.2) on an unphased VCF file with SNPs of five individuals. We took the

inference of garden warbler demography from Ishigohoka et al. (2021), and compared recombination rates between both species in 50, 100, 200 kb and 1 Mb non-overlapping windows. We performed Kendall's rank correlation test genome-wide and within chromosomes using cor.test function in R and ggplot2 for visualisation. Finally, we calculated the recombination rate ratio between both species by subtracting the log-transformed recombination rate calculated in 50-kb windows.

Acknowledgements

We would like to thank all members of the MPRG Behavioural Genomics and the Department of Evolutionary Genetics for constructive discussions during lab seminars; Bernhard Haubold, Elena Damm, and Kristian Ullrich for helpful discussions and suggestions.

This work was supported by the German Academic Exchange Service (DAAD scholarship to KBC grant 57381412), the Max Planck Society (MPRG grant MFFALIMN0001 to ML), and the DFG (project Z02 within SFB 1372 – Magnetoreception and Navigation in Vertebrates to ML).

Author Contributions

ML and LOH designed and jointly supervised the study. Mapping and variant calling was done by AB, genome annotation was generated by GM together with GD, and RT annotation was done by PP. KBC performed recombination rate estimates with input from JYD. Characterisation of genomic features, genome complexity and their respective associations with recombination rate were carried out by KBC with support from LOH and ML. KBC wrote the manuscript with input from LOH and ML. All authors provided comments on earlier versions of the manuscript.

Data Availability

The primary and alternate haplotype assemblies for the European blackcap can be found under NCBI BioProject PRJNA558064, accession numbers GCA_009819655.1 and GCA_009819715.1 (Ishigohoka et al. 2021). The final dataset and all scripts used for the analyses will be uploaded to Dryad upon acceptance of the manuscript, the full genome annotation will be deposited at ENA https://www.ebi.ac.uk/ena.

2.6 References

Abrusán G, Krambeck H-J, Junier T, Giordano J, Warburton PE. 2008. Biased Distributions and Decay of Long Interspersed Nuclear Elements in the Chicken Genome. Genetics. 178(1):573–581. doi:10.1534/genetics.106.061861.

Angeloni A, Bogdanovic O. 2021. Sequence determinants, function, and evolution of CpG islands. Biochemical Society Transactions. 49(3):1109–1119. doi:10.1042/BST20200695.

Auton A, Li YR, Kidd J, Oliveira K, Nadel J, Holloway JK, Hayward JJ, Cohen PE, Greally JM, Wang J, et al. 2013. Genetic Recombination Is Targeted towards Gene Promoter Regions in Dogs. PLOS Genetics. 9(12):e1003984. doi:10.1371/journal.pgen.1003984.

Backström N, Forstmeier W, Schielzeth H, Mellenius H, Nam K, Bolund E, Webster MT, Öst T, Schneider M, Kempenaers B, et al. 2010. The recombination landscape of the zebra finch Taeniopygia guttata genome. Genome Res. 20(4):485–495. doi:10.1101/gr.101410.109.

Bao W, Kojima KK, Kohany O. 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. Mobile DNA. 6(1):11. doi:10.1186/s13100-015-0041-9.

Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B. 2010. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. Science. 327(5967):836–840. doi:10.1126/science.1183439.

Berg IL, Neumann R, Sarbajna S, Odenthal-Hesse L, Butler NJ, Jeffreys AJ. 2011. Variants of the protein PRDM9 differentially regulate a set of human meiotic recombination hotspots highly active in African populations. Proceedings of the National Academy of Sciences. 108(30):12378–12383. doi:10.1073/pnas.1109531108.

Berthold P, Helbig AJ, Mohr G, Querner U. 1992. Rapid microevolution of migratory behaviour in a wild bird species. Nature. 360(6405):668–670. doi:10.1038/360668a0.

Betancourt AJ, Welch JJ, Charlesworth B. 2009. Reduced Effectiveness of Selection Caused by a Lack of Recombination. Current Biology. 19(8):655–660. doi:10.1016/j.cub.2009.02.039.

Boman J, Frankl-Vilches C, da Silva dos Santos M, de Oliveira EHC, Gahr M, Suh A. 2019. The Genome of Blue-Capped Cordon-Bleu Uncovers Hidden Diversity of LTR Retrotransposons in Zebra Finch. Genes. 10(4):301. doi:10.3390/genes10040301.

Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, Imbeault M, Izsvák Z, Levin HL, Macfarlan TS, et al. 2018. Ten things you should know about transposable elements. Genome Biol. 19(1):199. doi:10.1186/s13059-018-1577-z.

Bours, A., Pruisscher, P., Bascón-Cardozo, K., Odenthal-Hesse, L., & Liedvogel, M. (2022). *The blackcap (Sylvia atricapilla) genome reveals a species-specific accumulation of LTR retrotransposons* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-2166848/v1

Brick K, Smagulova F, Khil P, Camerini-Otero RD, Petukhova GV. 2012. Genetic recombination is directed away from functional genomic elements in mice. Nature. 485(7400):642–645. doi:10.1038/nature11089.

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(11):1656–1665. doi:10.1101/gr.196485.115.

Butlin RK. 2005. Recombination and speciation. Molecular Ecology. 14(9):2621–2635. doi:10.1111/j.1365-294X.2005.02617.x.

Calderón PL, Pigozzi MI. 2006. MLH1-focus mapping in birds shows equal recombination between sexes and diversity of crossover patterns. Chromosome Res. 14(6):605–612. doi:10.1007/s10577-006-1059-0.

Campbell CL, Bhérer C, Morrow BE, Boyko AR, Auton A. 2016. A Pedigree-Based Map of Recombination in the Domestic Dog Genome. G3 (Bethesda). 6(11):3517–3524. doi:10.1534/g3.116.034678.

Campbell MS, Holt C, Moore B, Yandell M. 2014. Genome Annotation and Curation Using MAKER and MAKER-P. Current Protocols in Bioinformatics. 48(1):4.11.1-4.11.39. doi:10.1002/0471250953.bi0411s48.

Choi K, Zhao X, Kelly KA, Venn O, Higgins JD, Yelina NE, Hardcastle TJ, Ziolkowski PA, Copenhaver GP, Franklin FCH, et al. 2013. Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. Nat Genet. 45(11):1327–1336. doi:10.1038/ng.2766.

Consortium ICGS. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 432(7018):695–716. doi:10.1038/nature03154.

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

Dapper AL, Payseur BA. 2018. Effects of Demographic History on the Detection of Recombination Hotspots from Linkage Disequilibrium. Molecular Biology and Evolution. 35(2):335–353. doi:10.1093/molbev/msx272.

Delmore K, Illera JC, Pérez-Tris J, Segelbacher G, Lugo Ramos JS, Durieux G, Ishigohoka J, Liedvogel M. 2020. The evolutionary history and genomics of European blackcap migration. Scordato E, Wittkopp PJ, editors. eLife. 9:e54462. doi:10.7554/eLife.54462.

Dolgin ES, Charlesworth B. 2008. The Effects of Recombination Rate on the Distribution and Abundance of Transposable Elements. Genetics. 178(4):2169–2177. doi:10.1534/genetics.107.082743.

Ellinghaus D, Kurtz S, Willhoeft U. 2008. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinformatics. 9(1):18. doi:10.1186/1471-2105-9-18.

Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 32(19):3047–3048. doi:10.1093/bioinformatics/btw354.

Fledel-Alon A, Wilson DJ, Broman K, Wen X, Ober C, Coop G, Przeworski M. 2009. Broad-Scale Recombination Patterns Underlying Proper Disjunction in Humans. PLOS Genetics. 5(9):e1000658. doi:10.1371/journal.pgen.1000658.

Galtier N, Piganeau G, Mouchiroud D, Duret L. 2001. GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. Genetics. 159(2):907–911.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 29(7):644–652. doi:10.1038/nbt.1883.

Hackenberg M, Previti C, Luque-Escamilla PL, Carpena P, Martínez-Aroza J, Oliver JL. 2006. CpGcluster: a distance-based algorithm for CpG-island detection. BMC Bioinformatics. 7(1):446. doi:10.1186/1471-2105-7-446.

Haenel Q, Laurentino TG, Roesti M, Berner D. 2018. Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. Molecular Ecology. 27(11):2477–2497. doi:10.1111/mec.14699.

Helbig AJ. 1991. Inheritance of migratory direction in a bird species: a cross-breeding experiment with SE- and SW-migrating blackcaps (Sylvia atricapilla). Behav Ecol Sociobiol. 28(1):9–12. doi:10.1007/BF00172133.

Hickey DA, Golding GB. 2018. The advantage of recombination when selection is acting at many genetic Loci. Journal of Theoretical Biology. 442:123–128. doi:10.1016/j.jtbi.2018.01.018.

Higgins JD, Perry RM, Barakate A, Ramsay L, Waugh R, Halpin C, Armstrong SJ, Franklin FCH. 2012. Spatiotemporal asymmetry of the meiotic program underlies the predominantly distal distribution of meiotic crossovers in barley. Plant Cell. 24(10):4096–4109. doi:10.1105/tpc.112.102483.

Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. Genetics Research. 8(3):269–294. doi:10.1017/S0016672300010156.

Ho SS, Urban AE, Mills RE. 2020. Structural variation in the sequencing era. Nat Rev Genet. 21(3):171–189. doi:10.1038/s41576-019-0180-9.

Ishigohoka J, Bascón-Cardozo K, Bours A, Fuß J, Rhie A, Mountcastle J, Haase B, Chow W, Collins J, Howe K, et al. 2021. Recombination suppression and selection affect local ancestries in genomes

of a migratory songbird. Evolutionary Biology. unpublished data, http://biorxiv.org/lookup/doi/10.1101/2021.12.22.473882, last accessed 2022 Jan 24.
Jansen A, Verstrepen KJ. 2011. Nucleosome positioning in Saccharomyces cerevisiae. Microbiol Mol Biol Rev. 75(2):301–320. doi:10.1128/MMBR.00046-10.
Jung M, Pfeifer GP. 2013. CpG Islands. In: Maloy S, Hughes K, editors. Brenner's Encyclopedia of Genetics (Second Edition). San Diego: Academic Press. p. 205–207. [accessed 2022 Feb 15]. https://www.sciencedirect.com/science/article/pii/B9780123749840003491.
Kamm JA, Spence JP, Chan J, Song YS. 2016a. Two-Locus Likelihoods Under Variable Population Size and Fine-Scale Recombination Rate Estimation. Genetics. 203(3):1381–1399. doi:10.1534/genetics.115.184820.
Kamm JA, Spence JP, Chan J, Song YS. 2016b. Two-Locus Likelihoods Under Variable Population Size and Fine-Scale Recombination Rate Estimation. Genetics. 203(3):1381–1399. doi:10.1534/genetics.115.184820.
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. Molecular Ecology. 26(16):4158–4172. doi:10.1111/mec.14197.
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. Molecular Ecology. 23(16):4035–4058. doi:10.1111/mec.12810.
Keller O, Kollmar M, Stanke M, Waack S. 2011. A novel hybrid gene prediction method employing protein multiple sequence alignments. Bioinformatics. 27(6):757–763. doi:10.1093/bioinformatics/btr010.
Kent CF, Minaei S, Harpur BA, Zayed A. 2012. Recombination is associated with the evolution of genome structure and worker behavior in honey bees. PNAS. 109(44):18012–18017. doi:10.1073/pnas.1208094109.
Kent TV, Uzunović J, Wright SI. 2017. Coevolution between transposable elements and recombination. Philosophical Transactions of the Royal Society B: Biological Sciences. 372(1736):20160458. doi:10.1098/rstb.2016.0458.
Kim S. 2015. ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. Commun Stat Appl Methods. 22(6):665–674. doi:10.5351/CSAM.2015.22.6.665.
Kong A, Thorleifsson G, Gudbjartsson DF, Masson G, Sigurdsson A, Jonasdottir Aslaug, Walters GB, Jonasdottir Adalbjorg, Gylfason A, Kristinsson KT, et al. 2010. Fine-scale recombination rate differences between sexes, populations and individuals. Nature. 467(7319):1099–1103. doi:10.1038/nature09525.
Korf I. 2004. Gene finding in novel genomes. BMC Bioinformatics. 5(1):59. doi:10.1186/1471-2105-5-59.
Lam I, Keeney S. 2015. Non-paradoxical evolutionary stability of the recombination initiation landscape in yeast. Science. 350(6263):932–937. doi:10.1126/science.aad0814.
Lerat E, Capy P, Biémont C. 2002. Codon Usage by Transposable Elements and Their Host Genes in Five Species. J Mol Evol. 54(5):625–637. doi:10.1007/s00239-001-0059-0.
Li H. 2013 May 26. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:13033997 [q-bio]. [accessed 2022 Mar 8]. http://arxiv.org/abs/1303.3997.
Liedvogel, M., Åkesson, S., & Bensch, S. 2011. The genetics of migration on the move. *Trends in ecology & evolution*,*26*(11), 561-569.
Malinovskaya L, Shnaider E, Borodin P, Torgasheva A. 2018. Karyotypes and recombination patterns of the Common Swift (Apus apus Linnaeus, 1758) and Eurasian Hobby (Falco subbuteo Linnaeus, 1758). Avian Research. 9(1):4. doi:10.1186/s40657-018-0096-7.
Martin SH, Davey JW, Salazar C, Jiggins CD. 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLOS Biology. 17(2):e2006288. doi:10.1371/journal.pbio.2006288.

Martin SH, Singh KS, Gordon IJ, Omufwoko KS, Collins S, Warren IA, Munby H, Brattström O, Traut W, Martins DJ, et al. 2020. Whole-chromosome hitchhiking driven by a male-killing endosymbiont. Malik HS, editor. PLoS Biol. 18(2):e3000610. doi:10.1371/journal.pbio.3000610.

Mason AS, Fulton JE, Hocking PM, Burt DW. 2016. A new look at the LTR retrotransposon content of the chicken genome. BMC Genomics. 17(1):688. doi:10.1186/s12864-016-3043-1.

McVean G, Awadalla P, Fearnhead P. 2002. A coalescent-based method for detecting and estimating recombination from gene sequences. Genetics. 160(3):1231–1241.

Merlin C, Liedvogel M. 2019. The genetics and epigenetics of animal migration and orientation: birds, butterflies and beyond. el Jundi B, Kelber A, Webb B, editors. Journal of Experimental Biology. 222(Suppl_1):jeb191890. doi:10.1242/jeb.191890.

Mihola O, Pratto F, Brick K, Linhartova E, Kobets T, Flachs P, Baker CL, Sedlacek R, Paigen K, Petkov PM, et al. 2019. Histone methyltransferase PRDM9 is not essential for meiosis in male mice. Genome Res. 29(7):1078–1086. doi:10.1101/gr.244426.118.

Mohandas TK, Speed RM, Passage MB, Yen PH, Chandley AC, Shapiro LJ. 1992. Role of the pseudoautosomal region in sex-chromosome pairing during male meiosis: meiotic studies in a man with a deletion of distal Xp. Am J Hum Genet. 51(3):526–533.

Mugal CF, Nabholz B, Ellegren H. 2013. Genome-wide analysis in chicken reveals that local levels of genetic diversity are mainly governed by the rate of recombination. BMC Genomics. 14(1):86. doi:10.1186/1471-2164-14-86.

Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C, MacFie TS, McVean G, Donnelly P. 2010. Drive Against Hotspot Motifs in Primates Implicates the PRDM9 Gene in Meiotic Recombination. Science. 327(5967):876–879. doi:10.1126/science.1182363.

Nachman MW, Payseur BA. 2012. Recombination rate variation and speciation: theoretical predictions and empirical results from rabbits and mice. Philosophical Transactions of the Royal Society B: Biological Sciences. 367(1587):409–421. doi:10.1098/rstb.2011.0249.

Odenthal-Hesse L, Berg IL, Veselis A, Jeffreys AJ, May CA. 2014. Transmission Distortion Affecting Human Noncrossover but Not Crossover Recombination: A Hidden Source of Meiotic Drive. PLOS Genetics. 10(2):11.

Okonechnikov K, Conesa A, García-Alcalde F. 2016. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics. 32(2):292–294. doi:10.1093/bioinformatics/btv566.

Pardue M-L, DeBaryshe PG. 2011. Retrotransposons that maintain chromosome ends. Proceedings of the National Academy of Sciences. 108(51):20317–20324. doi:10.1073/pnas.1100278108.

Parvanov ED, Petkov PM, Paigen K. 2010. Prdm9 Controls Activation of Mammalian Recombination Hotspots. Science. 327(5967):835–835. doi:10.1126/science.1181495.

Pehrsson EC, Choudhary MNK, Sundaram V, Wang T. 2019. The epigenomic landscape of transposable elements across normal human development and anatomy. Nat Commun. 10(1):5640. doi:10.1038/s41467-019-13555-x.

Peñalba JV, Deng Y, Fang Q, Joseph L, Moritz C, Cockburn A. 2020. Genome of an iconic Australian bird: High-quality assembly and linkage map of the superb fairy-wren (Malurus cyaneus). Molecular Ecology Resources. 20(2):560–578. doi:10.1111/1755-0998.13124.

Peñalba JV, Wolf JBW. 2020. From molecules to populations: appreciating and estimating recombination rate variation. Nat Rev Genet. 21(8):476–492. doi:10.1038/s41576-020-0240-1.

Pigozzi MI. 2016. The Chromosomes of Birds during Meiosis. Cytogenet Genome Res. 150(2):128–138. doi:10.1159/000453541.

Pirogov A, Pfaffelhuber P, Haubold B. High-complexity regions in mammalian genomes are enriched for developmental genes. :7.

del Priore L, Pigozzi MI. 2020. MLH1 focus mapping in the guinea fowl (Numida meleagris) give insights into the crossover landscapes in birds. PLoS One. 15(10):e0240245. doi:10.1371/journal.pone.0240245.

Quinlan AR. 2014. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Current Protocols in Bioinformatics. 47(1):11.12.1-11.12.34. doi:10.1002/0471250953.bi1112s47.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria. URL https://www.R-project.org/.

Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Fungtammasan A, Kim J, et al. 2021. Towards complete and error-free genome assemblies of all vertebrate species. Nature. 592(7856):737–746. doi:10.1038/s41586-021-03451-0.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: The European Molecular Biology Open Software Suite. Trends in Genetics. 16(6):276–277. doi:10.1016/S0168-9525(00)02024-2.

Rizzon C, Marais G, Gouy M, Biémont C. 2002. Recombination Rate and the Distribution of Transposable Elements in the Drosophila melanogaster Genome. Genome Res. 12(3):400–407. doi:10.1101/gr.210802.

Schield DR, Pasquesi GIM, Perry BW, Adams RH, Nikolakis ZL, Westfall AK, Orton RW, Meik JM, Mackessy SP, Castoe TA. 2020. Snake Recombination Landscapes Are Concentrated in Functional Regions despite PRDM9. Su B, editor. Molecular Biology and Evolution. 37(5):1272–1294. doi:10.1093/molbev/msaa003.

Shen C, Li X, Zhang R, Lin Z. 2017. Genome-wide recombination rate variation in a recombination map of cotton. Wang K, editor. PLoS ONE. 12(11):e0188682. doi:10.1371/journal.pone.0188682.

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(6263):928–932. doi:10.1126/science.aad0843.

Smeds L, Kawakami T, Burri R, Bolivar P, Husby A, Qvarnström A, Uebbing S, Ellegren H. 2014. Genomic identification and characterization of the pseudoautosomal region in highly differentiated avian sex chromosomes. Nat Commun. 5(1):5448. doi:10.1038/ncomms6448.

Smeds L, Mugal CF, Qvarnström A, Ellegren H. 2016. High-Resolution Mapping of Crossover and Non-crossover Recombination Events by Whole-Genome Re-sequencing of an Avian Pedigree. Henderson I, editor. PLoS Genet. 12(5):e1006044. doi:10.1371/journal.pgen.1006044.

Smeds L, Qvarnström A, Ellegren H. 2016. Direct estimate of the rate of germline mutation in a bird. Genome Res. 26(9):1211–1218. doi:10.1101/gr.204669.116.

Smit AF. 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. Current Opinion in Genetics & Development. 9(6):657–663. doi:10.1016/S0959-437X(99)00031-3.

Smit, A, Hubley, R & Green, P. 2015. RepeatMasker Open-4.0. 2013-2015 http://www.repeatmasker.org

Smukowski CS, Noor M a. F. 2011. Recombination rate variation in closely related species. Heredity. 107(6):496–508. doi:10.1038/hdy.2011.44.

Spence JP, Song YS. 2019. Inference and analysis of population-specific fine-scale recombination maps across 26 diverse human populations. Science Advances. 5(10):eaaw9206. doi:10.1126/sciadv.aaw9206.

Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM. 2017. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. Philosophical Transactions of the Royal Society B: Biological Sciences. 372(1736):20160455. doi:10.1098/rstb.2016.0455.

Subramanian VV, Zhu X, Markowitz TE, Vale-Silva LA, San-Segundo PA, Hollingsworth NM, Keeney S, Hochwagen A. 2019. Persistent DNA-break potential near telomeres increases initiation of meiotic recombination on short chromosomes. Nat Commun. 10(1):970. doi:10.1038/s41467-019-08875-x.

Suh A, Smeds L, Ellegren H. 2018. Abundant recent activity of retrovirus-like retrotransposons within and among flycatcher species implies a rich source of structural variation in songbird genomes. Molecular Ecology. 27(1):99–111. doi:10.1111/mec.14439.

Sultana T, Zamborlini A, Cristofari G, Lesage P. 2017. Integration site selection by retroviruses and transposable elements in eukaryotes. Nat Rev Genet. 18(5):292–308. doi:10.1038/nrg.2017.7.

Thacker D, Mohibullah N, Zhu X, Keeney S. 2014. Homologue engagement controls meiotic DNA break number and distribution. Nature. 510(7504):241–246. doi:10.1038/nature13120.

Tian Z, Rizzon C, Du J, Zhu L, Bennetzen JL, Jackson SA, Gaut BS, Ma J. 2009. Do genetic recombination and gene density shape the pattern of DNA elimination in rice long terminal repeat retrotransposons? Genome Res. 19(12):2221–2230. doi:10.1101/gr.083899.108.

Tsai IJ, Burt A, Koufopanou V. 2010. Conservation of recombination hotspots in yeast. Proc Natl Acad Sci USA. 107(17):7847–7852. doi:10.1073/pnas.0908774107.

Waters PD, Patel HR, Ruiz-Herrera A, Álvarez-González L, Lister NC, Simakov O, Ezaz T, Kaur P, Frere C, Grützner F, et al. 2021. Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. Proc Natl Acad Sci USA. 118(45):e2112494118. doi:10.1073/pnas.2112494118.

Weng Z, Wolc A, Su H, Fernando RL, Dekkers JCM, Arango J, Settar P, Fulton JE, O'Sullivan NP, Garrick DJ. 2019. Identification of recombination hotspots and quantitative trait loci for recombination rate in layer chickens. Journal of Animal Science and Biotechnology. 10(1):20. doi:10.1186/s40104-019-0332-y.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer.

Yelina NE, Lambing C, Hardcastle TJ, Zhao X, Santos B, Henderson IR. 2015. DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in Arabidopsis. Genes Dev. 29(20):2183–2202. doi:10.1101/gad.270876.115.

Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C, Storz JF, Antunes A, Greenwold MJ, Meredith RW, et al. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. Science. 346(6215):1311–1320. doi:10.1126/science.1251385.

2.7 Supplementary material

TableS1. Chromosomes lengths, weighted average, and SD recombination rate per chromosome for the European blackcap and the Garden warbler. Recombination rates were inferred using Pyrho with a block penalty of 20 and window size 50.

		European Blackcap			Garden warbler		
Chromosome	Length (Mb)	Genetic Map (cM)	Rec mean (cM/Mb)	Rec SD (cM/Mb)	Genetic Map (cM)	Rec mean (cM/Mb)	Rec SD (cM/Mb)
1	153,2	446,3	2,9	6,2	144,5	0,94	1,8
2	115,3	428,9	3,7	7,3	137,8E	1,19	2,3
3	113,4	372,6	3,3	6,5	108,5	0,96	2,2
4	73	370,5	5,1	7,8	124,6	1,71	2,8
5	72,6	363,1	5	8,6	117,2	1,61E	2,8
6	63,4	355,5	5,6	8	107,2	1,69E	2,4
7	38,9	246,7	6,4	8,7	79,5	2,05E	2,6
8	36	254,5	7,1	7,9	78,7	2,19	2,3
9	31,7	206,9	6,5	10,8	90,9	2,87	3,0
10	26,4	215,8	8,2	9,1	85,6	3,24	2,8
11	22,3	214,3	9,6	9,1	87,3	3,91	2,7
12	22,2	170,6	7,7	11,6	81,6	3,67	3,1
13	20,6	170	8,3	11,1	76,9	3,74	3,1
14	19,1	202,4	10,6	9,7	73,8	3,87	3,0
15	16,1	168,7	10,5	11,4	74	4,62	1,6
16	15,2	177,7	11,7	12,3	71,4	4,69	2,4
17	14,2	156,3	11	10,2	66,8	4,71	3,0
18	12	151	12,6	8,7	65,5	5,46	3,3
19	11,4	156,9	13,8	10,8	68,7	6,05	3,0
20	11,2	166,5	14,8	11,2	59,1	5,26	2,7
21	10,1	129,3	12,9	9,5	60,1	5,99	2,7
22	7,6	142,7	18,8	14,5	35,5	4,71	2,9
23	7,5	148,8	19,9	13,5	53,3	7,14	3,1
24	6,9	120,7	17,6	8,8	52,3	7,64	3,3
25	6,8	118,3	17,5	13,6	48,1	7,15	3,7
26	5,1	76,2	15,1	11,1	25,3	5,01	4,1
27	5	72	14,4	15,5	18,4	3,70	3,9
28	4,7	140,2	30,2	18,7	18,7	4,03	5,3
29	2,2	48,4	22	17,6	10,3	4,71	4,0
30	2	24,5	12,2	11,5	7,6	3,83	2,1
31	0,6	0,7	1,2	1,8	0,04	3,8	0,4
32	0,4	0,1	0,1	0,4	0,03	0,06	0
33	0,3	5,1	17,2	10,5	0,59	2	0
Z	88,6	102	1,2	7	11,6	0,13	0,4

Table S2. Intra-chromosomal comparison. Pearson correlations between chromosome length and the average recombination rates per chromosome with CpG density, Gene density, and GC content calculated for each chromosome. Kendall's tau correlation between chromosome length and the average recombination rates per chromosome with complexity (Cx).

Inter-chromosomal comparisons: Pearson's correlation coefficients				
	CpG density	**Gene density (logT)**	**GC content**	**Complexity** Kendall tau
Chr_Length(logT)	-0,68 ***	-0,96 ***	-0,84***	0,6***
Rec cM/Mb	0,68 ***	0,53 **	0,66***	-0,31*

***p<0.001,**p<0.01

Table S3. Genome-wide pairwise correlations and partial correlations between recombination rates with genomic features calculated in 200 kb and 1 Mb windows.

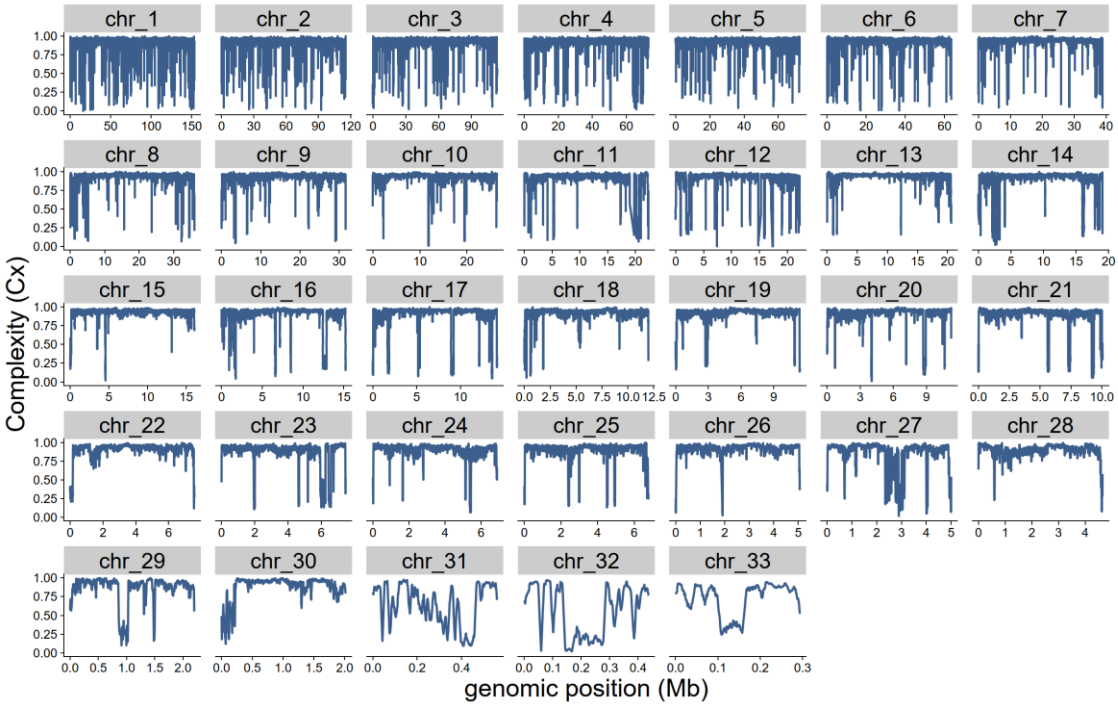
Kendall's tau correlation coefficients				
	Pairwise correlation		**Partial correlation**	
	200kb	**1 Mb**	**200kb**	**1 Mb**
CpG density	0,3***	0,4***	0,1***	0.13***
Gene density	0,2***	0,3***	-0,03*	ns
GC content	0,4***	0,4 ***	0,27***	0,27***
RT density	0,1***	0,1***	ns	0,12***
RT coverage	ns	ns	-0,09***	0,1***'

***p<0.001, **p<0,01, *p<0,05, ns= non-significant p>0.05

Table S4. Recombination rates comparison between the blackcap and garden warbler genome estimated in different window size.

Window size	**Kendall's tau coefficient**
50 kb	0,6083 ***
100 kb	0,6274 ***
200 kb	0,6469 ***
1 Mb	0,7115 ***

***p<0.001



FigureS1. Complexity distribution calculated in 10 kb overlapping windows across all chromosomes.

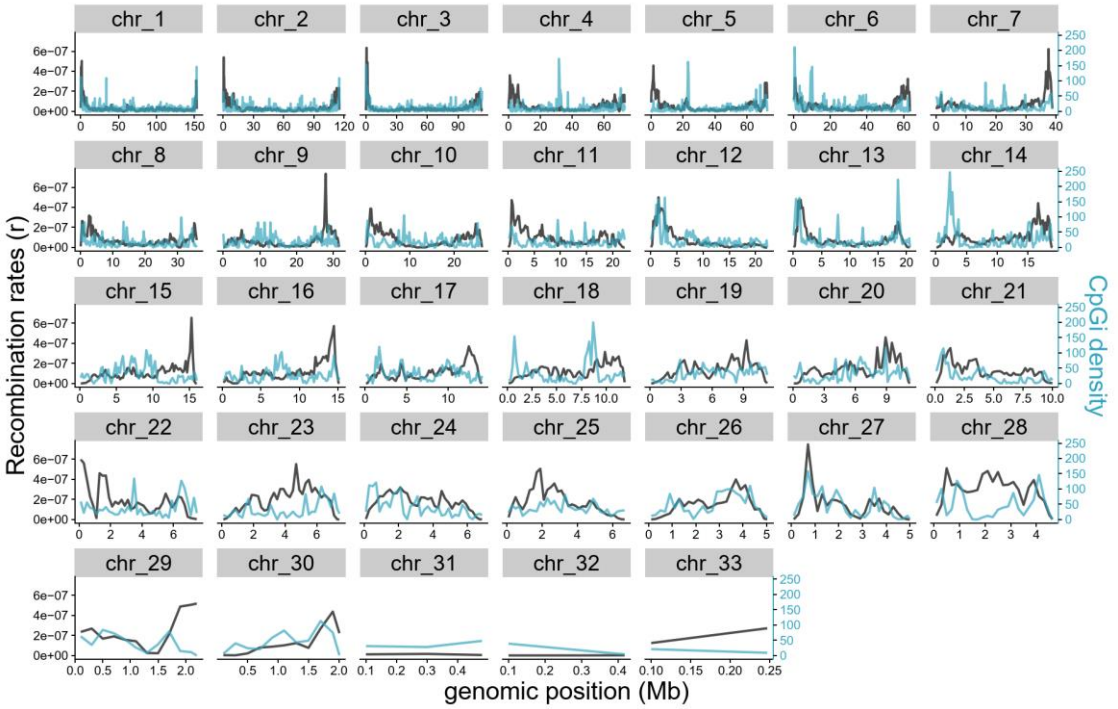


Figure S2. Distribution of CpG density and recombination rates in 200 kb windows across all chromosomes.

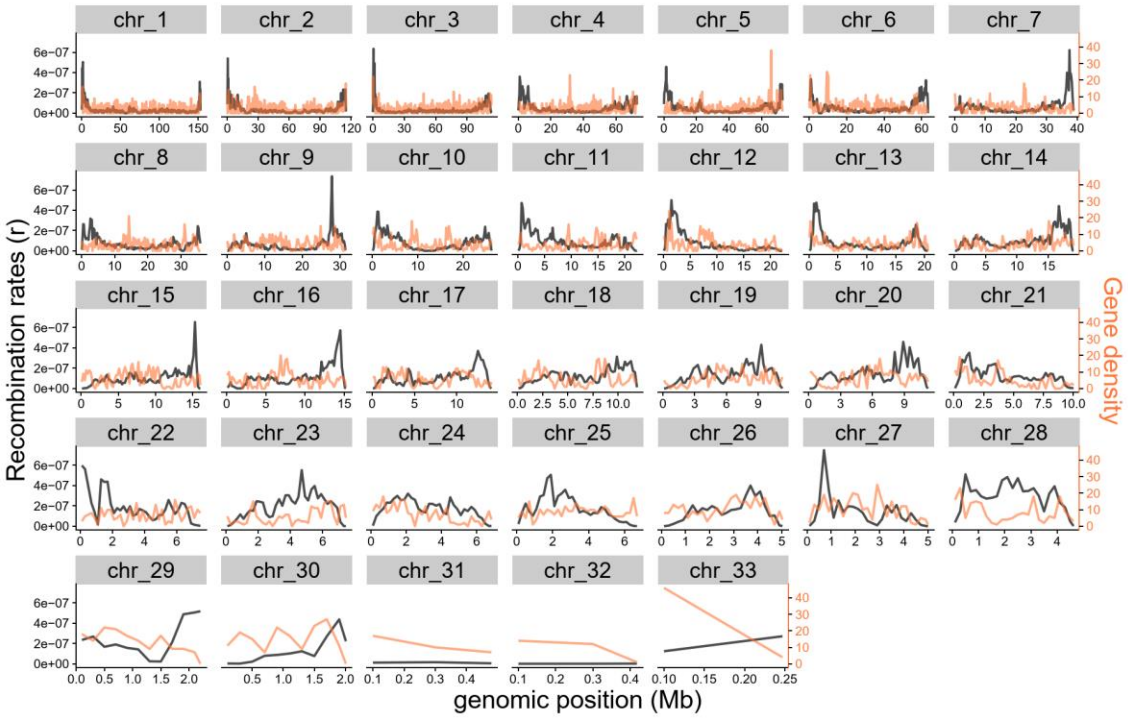


Figure S3 Distribution of gene density and recombination rates in 200 kb windows across all chromosomes.

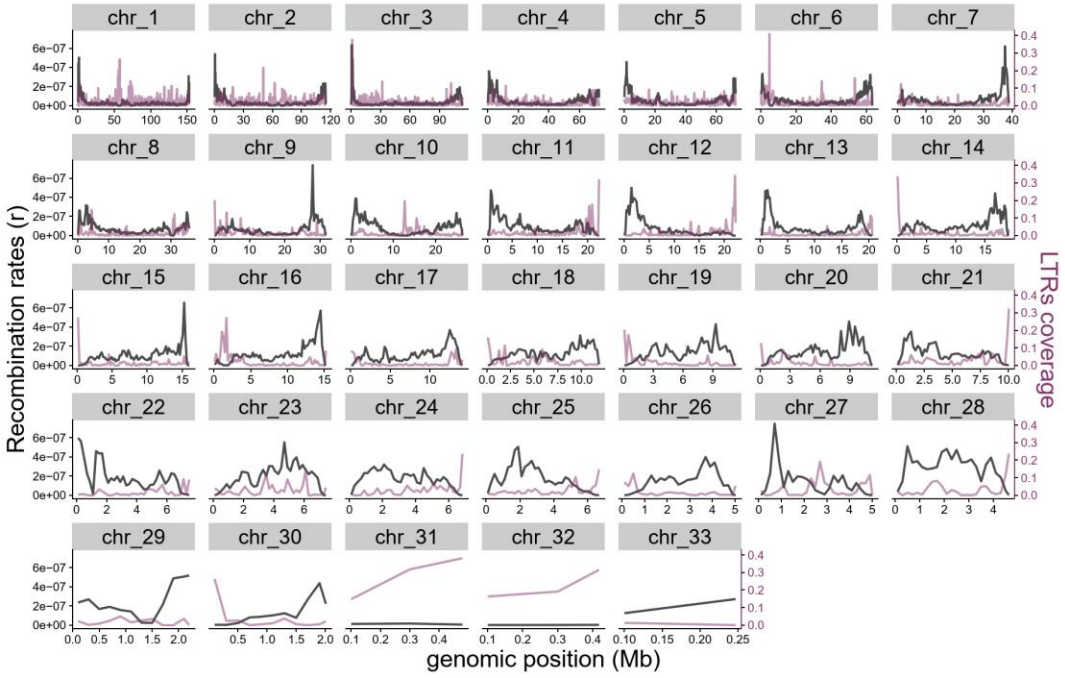


Figure S4 Distribution of LTR retrotransposons coverage and recombination rates in 200 kb windows across all chromosomes.

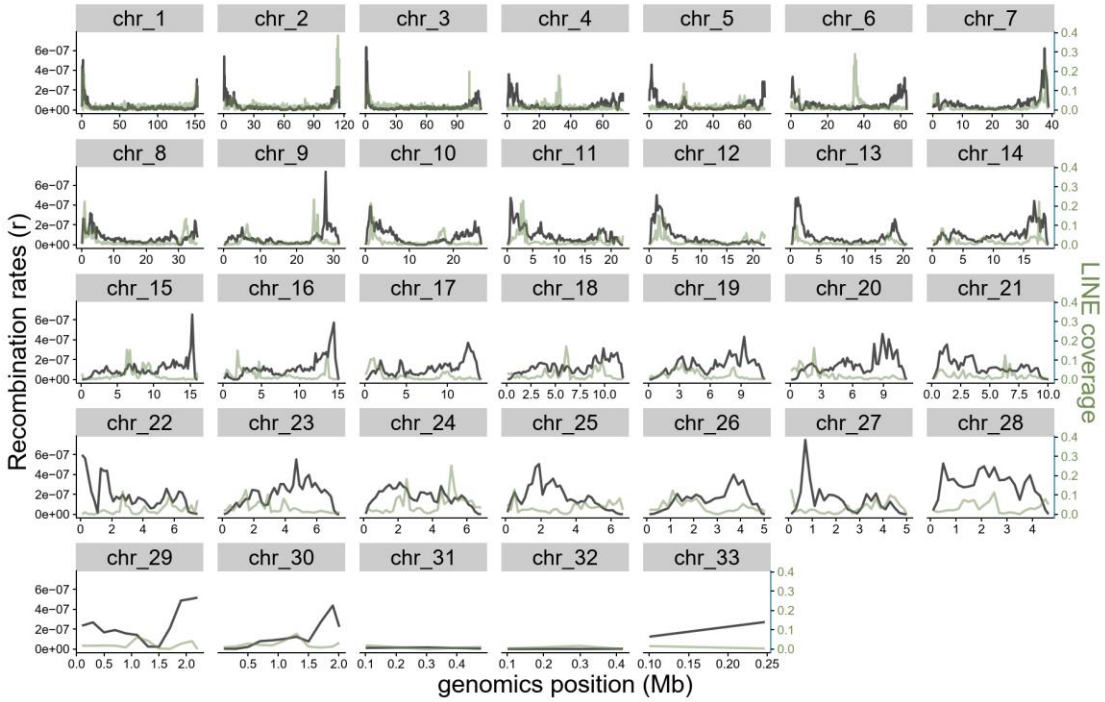


Figure S5 Distribution of LINE retrotransposons coverage and recombination rates in 200 kb windows across all chromosomes.

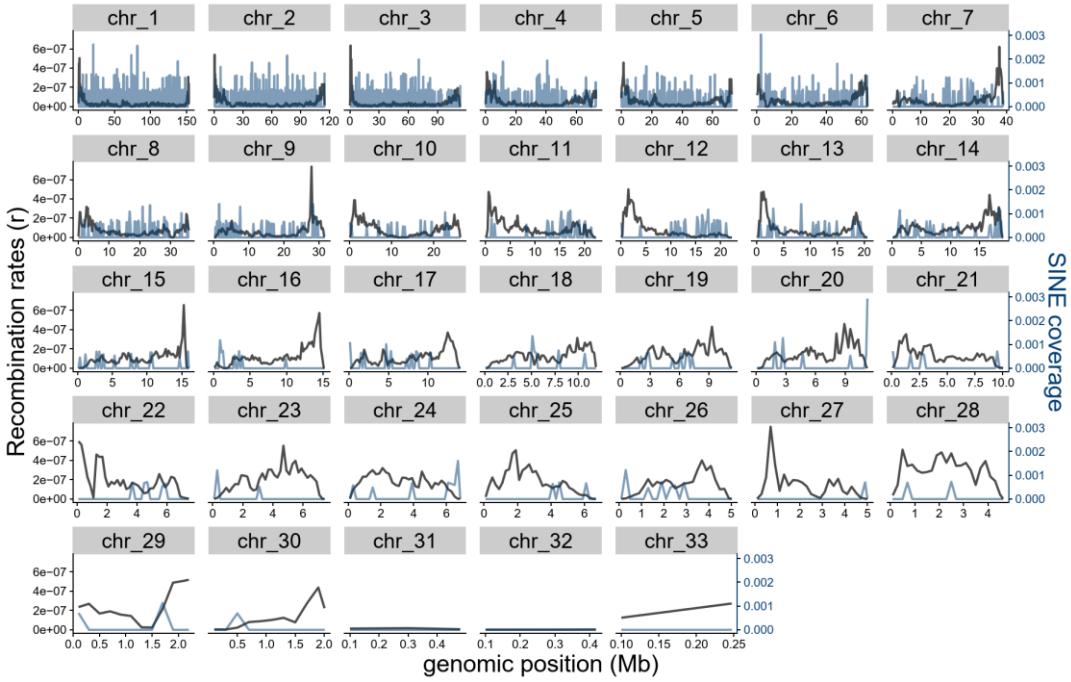


Figure S6 Distribution of SINE retrotransposons coverage and recombination rates in 200 kb windows across all chromosomes.

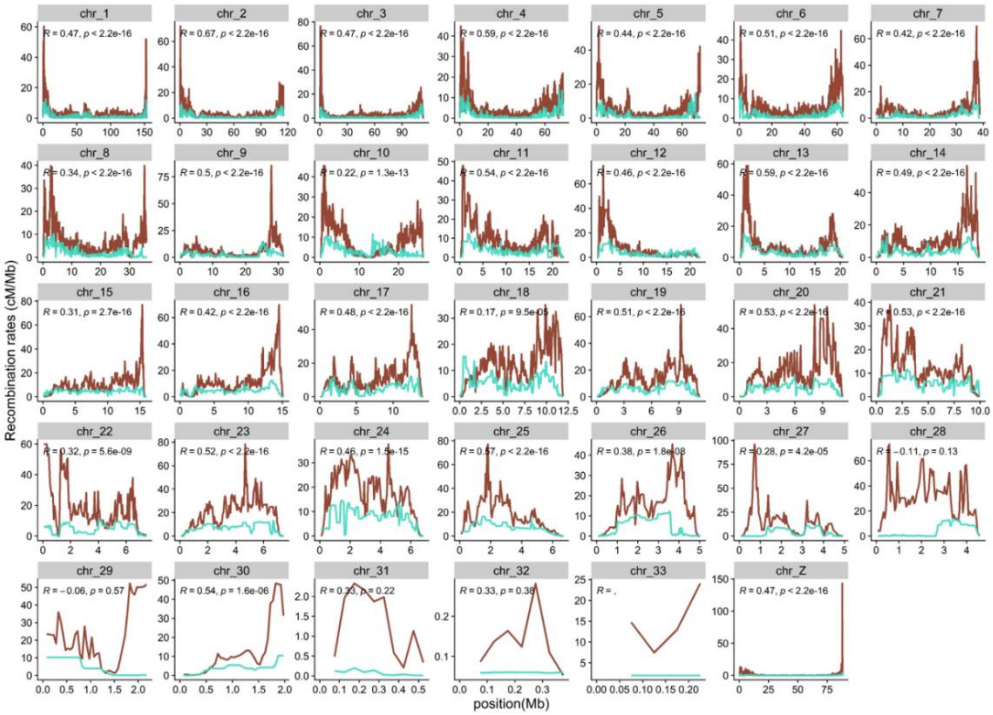


Fig S7 Distribution of recombination rates calculated in 50 kb non-overlapping windows in the Blackcap and Garden warbler across all chromosomes. Kendall non-parametric correlation coefficients reported for each chromosome.

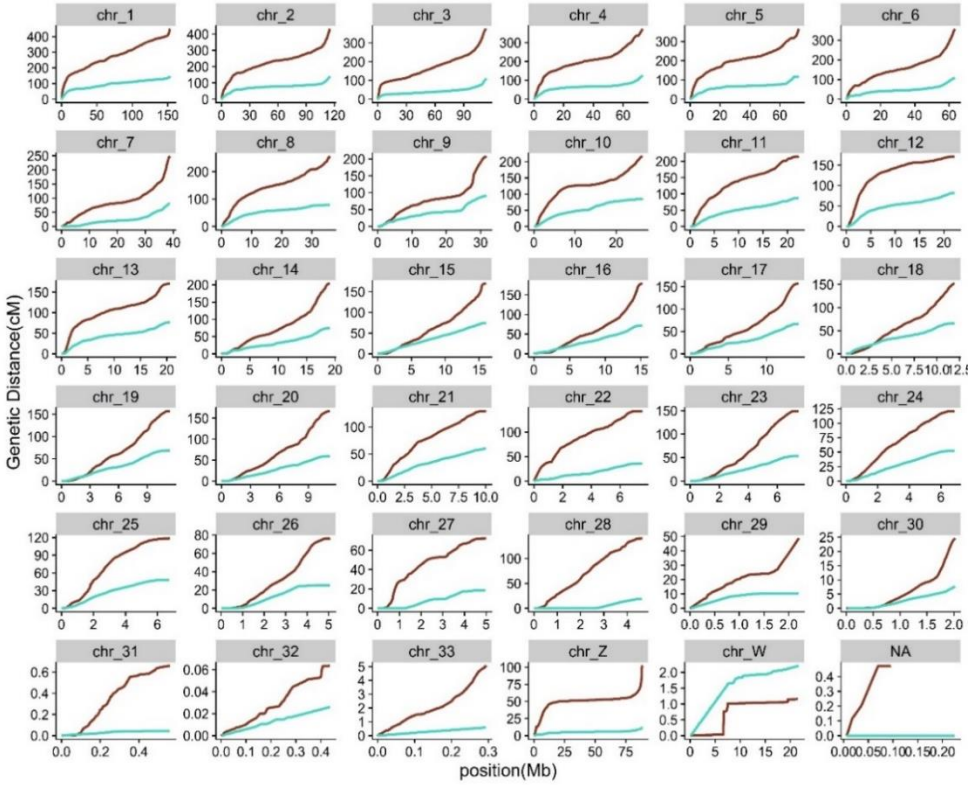


Figure S8 Genetic map of the blackcap (dark red) and garden warbler (cyan) across all chromosomes.

CHAPTER III

The relationship of CpG methylation shared among tissues with recombination rates across the blackcap genome

Karen Bascón-Cardozo[1], Britta Meyer[1], Linda Odenthal-Hesse*[1,2] and Miriam Liedvogel*[1]

Affiliations:

[1] Max Planck Institute for Evolutionary Biology, MPRG Behavioural Genomics, Plön, Germany
[2] Institute of Avian Research, Wilhelmshaven, Germany

*equal contribution

3.1 Abstract

Avian genomes lack the protein PRDM9, and most recombination events occur in regions where the chromatin allows access to the transcription machinery. Specific genomic features are tightly associated with historical recombination rates, such as GC content, CpG islands, promoter regions, and Transcription Start Sites. The interplay of CpG islands with recombination rates may be influenced by the density, the location across the genome and epigenetic regulation. This chapter aims to characterise the association of CpG methylation using reduced representation bisulfite sequencing (RRBS) from different tissues. We assessed recombination rate variation at different CpG methylation stages. We additionally evaluated CpG methylation patterns at regulatory regions and regions with a high density of transposable elements. Our results revealed that recombination rates increased significantly at intermediate CpG methylation percentages. HRR were associated with lower methylation percentage only in micro-chromosomes and Z chromosomes. However, we saw a clear pattern where methylation percentage drops at TSS where recombination is significantly higher. We evidenced some regions with high CpGi density and low recombination rates coinciding with hypermethylated CpGs. Intrachromosomal patterns also revealed low recombination patterns in regions where CpG methylation is elevated, and there is an enrichment of TEs. Our results support the idea that epigenetic regulation, particularly CpG methylation, might play an important role in shaping the fluctuation and evolution of the recombination landscape and specific genomic features.

3.2 Introduction

Understanding genome dynamics is important to disentangle evolutionary and behavioural processes. All the molecular factors and changes that have the potential to regulate the genome activity and expression independently of DNA sequence per-se is studied by epigenetics. This encompasses diverse processes, one of the most important and well-described is DNA methylation which involves the recruitment of specific enzymes (methyltransferases) to methylate cytosines residues in CpG context: Cytosine is followed by G dinucleotide separated by a phosphate group, and non-CpG context: CHG, and CHH sequence contexts, where H = A, T, or C. High levels of DNA methylation (hypermethylation) are typically associated with chromatin compaction (heterochromatin), repression of gene expression, sex chromosomes inactivation, and silencing of transposable elements (TEs) (reviewed in Allis & Jenuwein, 2016; Deniz et al., 2019; Jones, 2012; Kent et al., 2017; Lim et al., 2019; Zamudio et al., 2015). Whereas low levels of DNA methylation (hypomethylation, depletion of methylation and /or acetylation) lead to a more open chromatin structure (euchromatin), enabling accessibility for a diverse of processes such as replication, transcription, DNA repair as well as meiotic recombination (Zhong et al., 2021).

Epigenetic modifications such as post-transcriptional modification have been reported to activate or suppress the activity of recombination hotspots in mammals (Buard et al., 2009). Most importantly, in organisms whose recombination dynamics are functional in the absence of the protein PRDM9, the majority of recombination events are prevalent in Transcription Start Sites (TSS), cis-regulatory and functional elements (Bascón-Cardozo et al., 2022a, Kawakami et al., 2017; Lam & Keeney, 2015; Singhal et al., 2015) which are presumably euchromatic and hypomethylated. For example, the acquisition of DNA methylation in the form of H3K9me2 marks (the hallmark of constitutive heterochromatin) has been reported to silence euchromatic crossover hotspots in the plant *Arabidopsis thaliana* (Yelina et al., 2015). Similarly, within fungi, DNA methylation strongly suppresses meiotic crossovers in fission yeast (Ellermeier et al., 2010). Hence, DNA methylation and chromatin modifications may have an important influence on the recombination landscape of these organisms.

The historical population-scaled recombination rates in non-PRDM9 dependent organisms, particularly avian genomes, are highly associated with CpG islands (CpGi) and GC content (Bascón-Cardozo et al., 2022a; Kawakami et al., 2017; Singhal et al., 2015). In birds, almost 70% of CpG sites are methylated, except for CpG sites around TSS and promoters (Derks et al., 2016).

CpGi are stretches of DNA with high CG composition, predominantly not methylated, and usually associated with transcription and regulatory regions such as promoters and TSS (Deaton & Bird, 2011); hence, they are tightly related to high recombination rates. In a previous study (see chapter II, Bascón-Cardozo et al., 2022a), we found greater recombination frequency at promoters with CpGi compared to promoters alone in the Eurasian blackcap. However, we also evidenced discrepancies at fine scales where some intrachromosomal regions with high CpG islands density coincide with reduced recombination rates. Those findings lead us to speculate that the association between CpGi and recombination may not only depend on the presence or density of CpGi but also on their methylation profile.

To disentangle that and complement our understanding of recombination and CpGi methylation, here we assessed DNA methylation patterns in the blackcap using reduced representation bisulfite sequencing (RRBS) data. We include tissues that belong to the three germ layers (ectoderm, mesoderm, and endoderm) and therefore encompass a more comprehensive range of CpG methylation. We ask if CpG methylation patterns are associated with differential recombination rates: Is recombination prevalent at low methylation stages? Are hypomethylated shared CpG sites among tissues associated with high recombining regions or regulatory regions where recombination tends to be high (e.g., TSS)? We hypothesise that recombination rates will be higher at low methylation percentages of CpG-shared sites. More specifically, hypermethylated sites will associate with Low Recombining Regions (LRR) and hypomethylated with High recombining regions (HRR). Regions where we found a relatively high density of CpGi and low recombination rates will potentially be associated with CpG hypermethylated sites. Finally, we look into the association of CpG methylation and transposable elements (TEs), expecting that high methylation patterns will be related to a high abundance of TEs.

3.3 Materials and Methods

Samples information

Two male and two female blackcaps, *Sylvia atricapilla*, from a long-distance migratory population (Berthold, 1973; Fusani et al., 2014) were caught at the Pape Ornithological Station, Latvia (56°9'48"N, 21°1'35"E), between the end of August and the beginning of September 2011. Birds were tested in a behavioural experiment, as detailed in Fusani et al. 2014 with the retina as focal

tissue and stored at -80° C after the termination of the investigation. The organs: brain, lung, muscle, and heart were dissected of all four individuals, and extracted genomic DNA for Reduced Representation Bisulfite Sequencing (RRBS) as in Klughammer et al., 2022. Briefly, DNA extraction was performed using a standard phenol-chloroform extraction protocol, and RRBS was performed using 100 ng of genomic DNA. The efficiency of bisulfite conversion independent of CpG context was assessed by adding methylated and unmethylated spike-in controls at a concentration of 0.1%. DNA was digested using the restriction enzymes MspI and TaqI in combination. After library enrichment cycles using qPCR and PCR, libraries were sequenced on Illumina HiSeq 3000/4000 machines using the 50 or 60-bp single-read setup. Samples were processed, and the unaligned bam files were provided for each tissue and bird separately (four tissue samples per individual), with a total of 16 individual samples.

Methylation calls processing

The unaligned bam files provided were initially converted to fastq files. We then aligned the reads using the blackcap reference genome previously prepared or modified in bismark (Krueger & Andrews, 2011) with the option 'bismark_genome_preparation'. After that, we applied the 'bismark_methylation_extractor' where we cut-off 2bp from the 5' and 3' ends of the reads with the options '--ignore 2 --ignore_3prime 2' and we generated the general and the cytosine report. To merge the OT and OB strands, we applied a custom python script 'merge_CpG.py' obtained from https://github.com/rcristofari/penguin-tools/blob/master using the CpG report as input.

MethylKit was used for the methylation analysis (Akalin et al., 2012). We input the CpG report data with default parameters only considering methylation in the CpG context. Reads with less than tenfold coverage were filtered out. We further filtered bases with coverage above than 99.9th percentile of coverage in each sample. We visualized the methylation calls and performed some statistics using 'getMethlationStats' and 'getCoverageStats.'

To identify shared methylation sites that cover all samples from different tissues, we used the 'unite' function in methylkit. CpGs covered with at least 1 sample per group were returned. Then, we performed pairwise correlation and clustering analysis among samples. Finally, we calculated the methylation percentage by taking the numbers of Cs and coverage for each site in all the samples.

Evaluating recombination rates in shared methylation sites

To evaluate how variable genome-wide recombination rates are at different methylation levels, we intersect the shared methylation sites with recombination rates from all the chromosomes. We used the population-scaled recombination rates per site per generation for the continental resident population, which was used initially to characterise the recombination landscape and its association with specific genomic features in chapter II (Bascón-Cardozo et al. 2022a). We also used recombination map of the medium South East (medium-SE) population since the samples from the current study come from the same location and it is very likely that those individuals possessed the same migratory phenotype. For further details in the recombination rate inference see chapter II and the subsequent chapter IV which characterises the population-specific recombination maps (Bascón-Cardozo et al. 2022a; Bascón-Cardozo et al. 2022b). All the methylation sites (CpGs) were classified in levels and we calculated and compared the average genome-wide recombination rate for each methylation level. We performed the Wilcoxon rank test for the statistical comparisons.

Test for methylation comparison in high and low recombining regions

We identified High recombining regions (HRR) as recombination rates four times higher than the genome-wide average recombination rate and low recombining regions (LRR) as four times lower than the average genome-wide recombination rate. We calculated these regions for different blackcap populations, and we took only HRR and LRR that were shared between continental populations (medium distance, short and continental residents) (for further details in the inference of recombination maps see the subsequent chapter IV, Bascón-Cardozo et al. 2022b). To compare methylation patterns in HRR and LRR, we intersected the high and low recombining regions with the CpGs and their respective methylation percentages. We Initially compared the methylation percentage average between HRR and LRR genome-wide, then we contrasted macro and micro-chromosomes separately. The sex chromosome W was excluded from the analysis, and the Z chromosome was analysed separately. We performed the Wilcoxon rank test for all the comparisons in R.

Measuring the correlation between recombination and CpG methylation in windows

We summarized the methylation information in 100kb non-overlapping windows using the option 'tileMethylCounts' in the methylkit for each sample, setting a minimum coverage threshold of 3 reads per cytosine per sample initially. Subsequently, during the calculation in windows, we filtered

out regions with less than ten bases per region. To get windows covering all the samples, we used the 'unit' function and then we calculated the percentage of methylation for each sample in each window.

For the genome-wide analysis in windows, we intersected the methylation percentage of all the samples in 100 kb non-overlapping windows with recombination rates inferred for continental resident populations and medium_SE migrant population and calculated in the same window size. We performed pairwise correlations measured with Kendall's Tau correlation coefficient. Additionally, we included medium-SW and one island population (Crete) to evaluate the consistency of results among recombination maps. We plot the intra-chromosomal distribution of recombination rates and methylation percentage for all the CpGs. To complement our results, we calculated CpGi density in 100 kb and plotted the intrachromosomal distribution for visualisation. CpGi were previously identified in chapter II (Bascón-Cardozo et al., 2022a). We also included TEs coverage in 100kb windows and measured the correlation with methylation calculated in windows using Kendall's Tau correlation coefficient.

Methylation as a function of distance from Transcription Start Sites

Using the genome annotation, we assess the methylation percentage as a function of distance from Transcription Start Sites (TSS), defined as the first position of annotated mRNAs. Using bedtools, we assigned 0 as the first position to the closest gene, and we calculated the average methylation percentage in 5-kb non-overlapping windows spanning 200 kb upstream and downstream, considering the direction of the genes.

Data processing, analysis, and visualization

All the statistical analyses were performed using R software (R Core Team 2021). The data processing and figures were created using 'dplyr', 'ggpubr', 'tidyr', 'ggplot2' libraries.

3.4 Results

We obtained a total of 17.064 shared methylation sites (CpG positions, CpGs) across all autosomes covering the entire dataset (i.e., four different tissues for four individuals); 631 CpGs were identified for the sex chromosome Z.

CpGs methylation were correlated between tissues ($R>=0.6$, $p<0.001$) and the samples of each tissue cluster according to the germ layers: brain (ectoderm), muscle and heart (mesoderm), and lung (endoderm), with the exception of one sample from muscle clustering closer to the lung samples (Fig.S1).

High genome-wide recombination rates at intermediate methylation levels

The genome-wide average recombination rates calculated across a CpG methylation range revealed that the recombination rate was significantly higher at intermediate levels of methylation (between 30-50% and 60-70%) compared to very high or low methylation levels (0-10%, 90-100%) (Wilcoxon, $p<0.001$) (Fig. 1A). This result was more evident when methylation and recombination rates were analysed in 100 kb non-overlapping windows (Fig. S2). We obtained consistent results when we included recombination maps from other populations in the comparison (Fig. S2).

The genome-wide correlation between recombination rates and methylation percentage calculated in 100 kb windows was weak but positive for all the samples from different tissues ($p=> 0.17$, $p<0.001$). This result was consistent using recombination maps from different populations.

HRR associated with low methylation percentage in micro-chromosomes

Our data revealed a differential pattern in the comparison between methylation percentage in HRR and LRR. CpGs enriched in HRR showed a slightly lower methylation average in micro-chromosomes compared to LRR ($p<0.05$, Fig. 1B). In chromosome Z, this was more evident (Fig. 1B). The opposite pattern was observed in macro-chromosomes, HRR showed a higher methylation percentage than LRR ($p<0.01$, Fig. 1B). The comparison of methylation in HRR and LRR for the whole genome was not significant.

From 3361 CpGs which intersected with HRR and LRR, 1091 CpGs intersected with HRR and 1750 CpGs with LRR. CpGs in LRR were usually greater in number than CpGs in HRR across the methylation range, except for 30-60% methylation level, where the amount of CpGs enriched in HRR was slightly greater (Fig. S3). Generally, a considerable proportion of CpGs within HRR and

LRR were found in low methylation levels (0-10%) as well as high methylated levels (90-100%) in the case of LRR (Fig. S3).

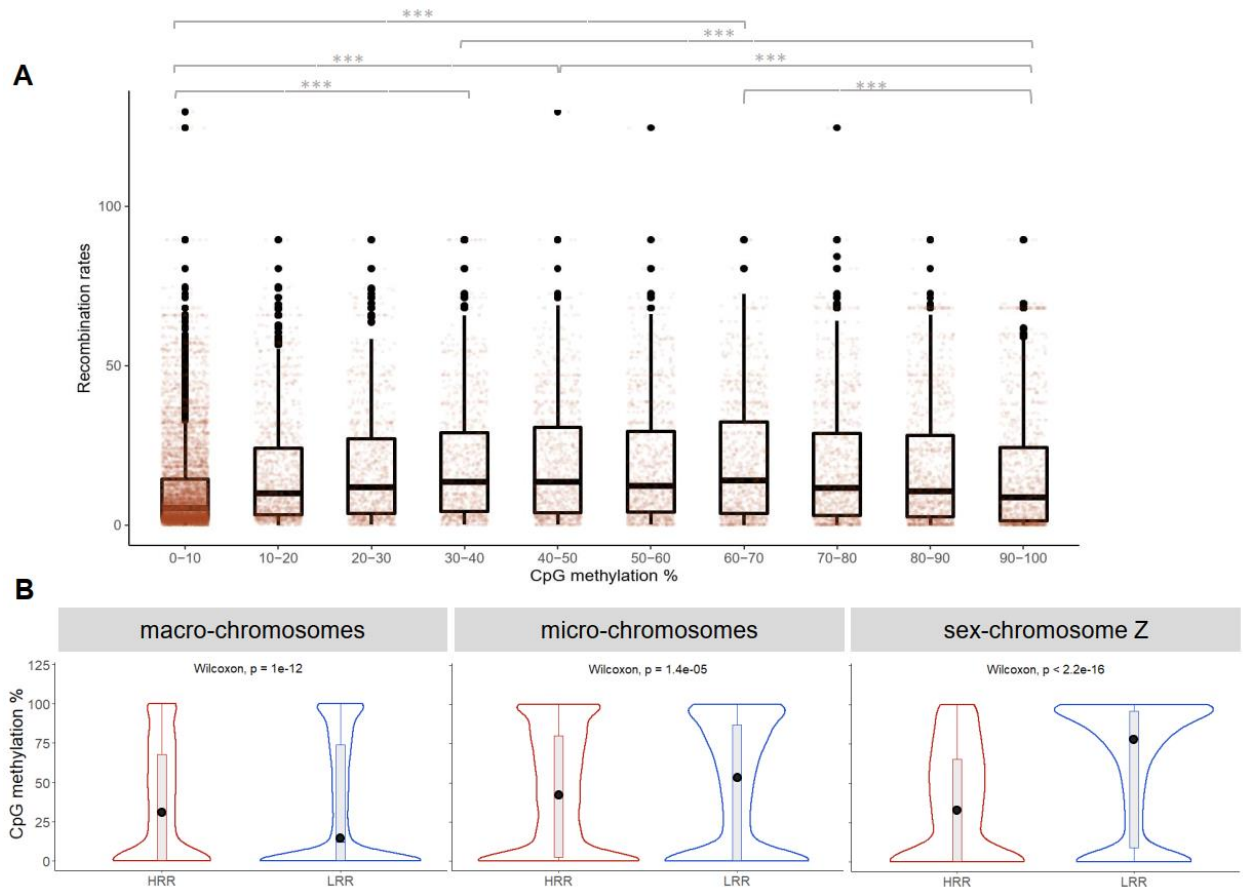


Figure 1. **Methylation patterns and recombination rates across the whole genome and within high and low recombining regions. A)** Boxplots showing the genome-wide recombination rate comparison at different methylation percentage levels. Only significant comparisons showed with asterisks *** $p<0.001$; the rest of the comparisons are not significant. **B)** Violin plots contrasting CpG methylation between High Recombining Regions (HRR in red) and Low Recombining Regions (LRR in blue) in macro- and micro-chromosomes and the Z chromosome. Dots represent the mean CpG methylation percentage for each group.

A relatively high amount of CpGi with low recombination rates is potentially hypermethylated

Despite the weak positive genome-wide correlation between recombination rates and methylation percentage calculated in 100 kb windows, the intra-chromosomal distribution of recombination rates and methylation sites revealed CpG sites with high methylation patterns and reduced

recombination at local scales. This pattern was more evident in micro-chromosomes (See chromosomes 18, 23, 24,28,29**,** Fig. 2A, Fig. S4).

Recombination rates are highly associated with CpGi density; however, in chapter II, we found discrepant patterns at the local level, with high CpGi density in low recombining regions or the opposite. Thus, we wanted to see whether high or low CpG methylation patterns could be involved. For that, we analysed specific chromosomes, and the intrachromosomal distribution revealed hypermethylated CpGs (with methylation percentage >80%) associated with low recombining regions, which contain a considerable amount of CpGi (Fig. 2A see chromosomes 6, 18, 22,29). These results support the idea that not only the presence of CpGi but also their methylation status can contribute to the variation of recombination rates within chromosomes.

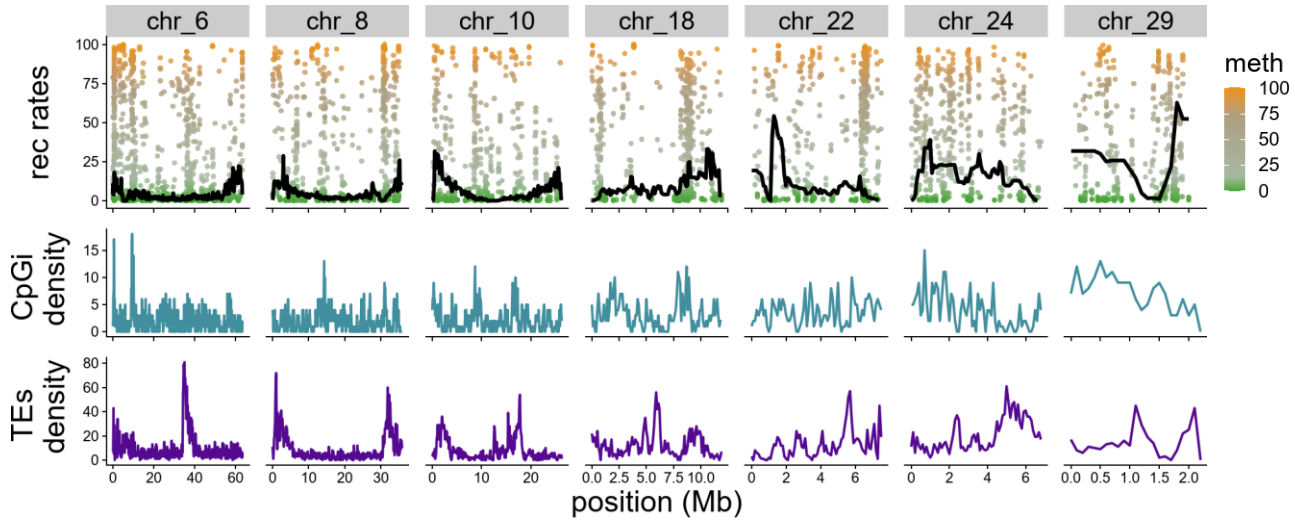


Figure 2. Recombination rates and CpG methylation association at fine-scale. Intrachromosomal distribution of recombination rates and CpG methylation sites in the top panel. The colour scale represents the methylation percentage, from orange denoting high methylated sites to green representing the low methylated sites across selected chromosomes. The recombination rates are from medium-distance migrants SE population and were calculated in 100 kb windows. CpG islands density (in the medium panel) and TE coverage (in the bottom panel) were calculated in 100 kb windows across selected chromosomes.

Methylation decreases at regulatory regions and increases at transposable elements

We calculated CpG methylation as a function of distance from the closest TSS of annotated genes spanning 200 kb upstream and downstream regions. We found a significant drop in CpGs methylation percentage in the TSS and promoter (2 kb upstream from TSS) region (Fig. 3). The methylation percentage in the surrounding upstream and downstream areas fluctuate between 35-45% methylation.

The pairwise genome-wide comparison of TE coverage and CpG methylation in 100 kb windows revealed a slightly positive association (P=0.11, R<0.001). At the local level, we identified regions where hypermethylated sites coincide with high density and coverage of transposable elements (see chromosomes 6,8,10 Fig. 2B).

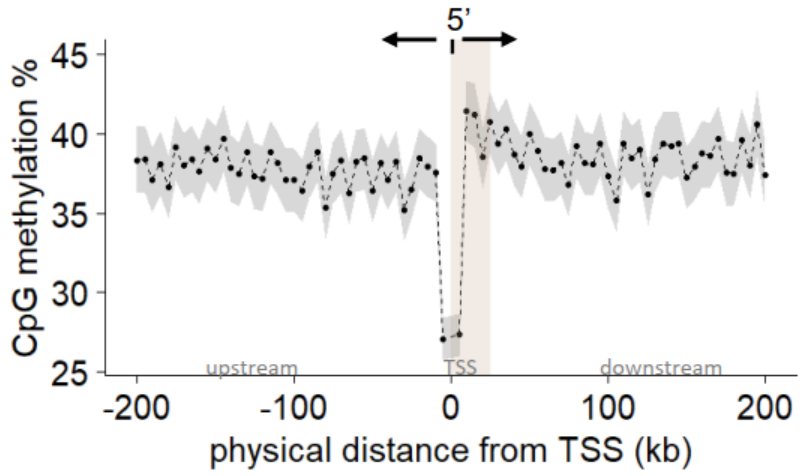


Figure 3. CpG methylation drops at regulatory regions. Average CpG methylation as a function of distance to the nearest Transcription Start Site (TSS), the orange shadow denotes mRNA of annotated genes. Dots represent the average CpG methylation among all sites for each 5kb window, and the grey shadow shows the 95% confidence interval.

3.5 Discussion

We found that recombination is more prevalent at intermediate methylation levels (between 30-50 % and 70-80 %). Similar results were reported in the flycatcher, where the recombination rate increased between 20-40 % methylation. This was further supported by the greater amount of CpGs identified in HRR at intermediate stages (30-70 % methylation percentage) compared to LRR.

In line with our expectations, CpG sites showed higher methylation patterns at LRR compared to HRR; however, this was only evident in micro-chromosomes. In macro-chromosomes, we saw the opposite, which may be attributed to the high amount of CpGs within LRR skewed towards the lowest methylation levels (0-10 %). Peri-centromeres could yield this pattern as it was found in plants (Yelina et al., 2015). Moreover, we cannot discard that some of the LRR are regions where recombination events weren't captured and perhaps don't truly represent regions with absence of recombination. Generally, the high amount of CpGs in low (0-10 %) and very high methylation ranges (80-100%) genome-wide that we observe in our data is a typical distribution in bisulfite methylation data. This pattern may also result from population-specific polymorphic sites (C to T), misinterpreted as differential DNA methylation sites between individuals. To distinguish between actual methylation at a CpG site versus the presence of an SNP, masking the outlier values along with C-T SNPs could be one possibility for controlling for that potential bias and preserving the methylation data for known SNP sites with rare variants (reviewed in LaBarre et al., 2019).

It was also evident the presence of hypermethylated CpGs in high recombining regions. This pattern was typically seen at the ends of macro-chromosomes and may also contribute to the positive genome-wide pairwise correlation between recombination rates and methylation calculated in 100 kb windows. This result goes in contrast to our initial hypothesis or expectation. A similar association was found in the flycatcher in promoter regions (Westerberg, 2019). The correlation analysis in our study encompass all shared sites between different tissues and genome-wide patterns and may not reflect the association of methylation and recombination in specific genomic features. Contrasting or controlling for covarying factors such as tissue and sex-specific methylation sites could also yield a different correlation with recombination rates. Additionally, calculating methylation information in windows may not reflect the methylation fluctuation at specific regions entirely, and this will need a further test with shorter window sizes.

The relationship between recombination rates and CpG methylation will also depend on the location and architecture of chromosomes. For instance, it has been reported that the hypomethylation of the nonrecombining chromosome (ZAL2m) in white-throated sparrows is potentially associated with a molecular mechanism where double-strand breaks can recruit DNA methyltransferases and increase DNA methylation in specific regions (Sun et al., 2021). Hence, the scale and the dynamics of specific genomic regions might have an impact on this association.

At the intra-chromosomal distribution, we could identify regions enriched with hypermethylated CpG sites and reduced recombination, likely reflecting functional elements. In line with that, we evidenced a decrease in methylation at TSS and 5kb upstream regions corresponding to promoters. The surrounding areas fluctuate between 20-50 % CpG methylation. Therefore, genomic features like regulatory elements are probably between this methylation range and not at very high or very low methylation levels. This also highlights the importance of characterising at fine or local scales the location and regulation of specific genomic features and recombination rates. For example, we found a region with a considerable amount of CpGi coinciding with low recombination rates. Interestingly, CpGs with high methylation (>70 % methylation) were enriched in these regions. This goes with the hypothesis that not only the presence of CpGi will increase recombination rates and their methylation patterns.

We confirmed the positive association of TEs density with DNA methylation as reported previously (Wyler et al., 2020). The accumulation of TEs may trigger DNA methylation to repress their expression and potential deleterious effect on the host (Deniz et al., 2019). The intrachromosomal distribution generally showed this pattern, however, with some exceptions. Here, we included all TEs, regardless of the type. Different subfamilies of TEs and species-specific or non-species-specific TEs could yield other associations, as was reported previously in the correlation with recombination rates (see chapter II, Bascón-Cardozo et al., 2022a, Bours et al., 2022). Interestingly, in chapters II, we identified specific regions, such as within chromosomes 6, 10, and 18, with highly differential recombination patterns among populations and species (in our comparison with the Garden warbler). Here, we found the same regions enriched with high amount of TEs and hypermethylated CpGs.

This study opens the possibility of further studying the association of CpG methylation with genomic features and open chromatin regions. Our results support that epigenetic regulation,

particularly methylation of CpGs, might play an important role in regulating and changing recombination rates across the genomic landscape.

Acknowledgments

We are very grateful to Leonida Fusani for offering us the possibility of using additional tissue from captured birds from an earlier study focused on the retina (Fusani et al., 2014). We thank Juan Sebastian Lugo Ramos who dissected the focal four tissues of all birds, which were used as part of a bigger cross species comparative project by Johanna Klughammer, who generously allowed us to use the RRBS data for our independent analysis. All treatments were carried out under the permission of The Nature Conservation Agency of Latvia (permit no. D3.6/37 of 26/08/2011), the IATA Guidelines, the University of Ferrara, and the Italian Ministry of Health. This study was supported by the German Academic Exchange Service (DAAD scholarship grant to KBC), the Max Planck Society (MPRG grant MFFALIMN0001 to ML), and the DFG (project Z02 within SFB 1372 – Magnetoreception and Navigation in Vertebrates to ML).

3.6 References

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., & Mason, C. E. (2012). methylKit: A comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, *13*(10), R87. https://doi.org/10.1186/gb-2012-13-10-r87

Allis, C. D., & Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nature Reviews Genetics*, *17*(8), 8. https://doi.org/10.1038/nrg.2016.59

Bascón-Cardozo, K., Bours, A., Manthey, G., Pruisscher, P., Durieux, G., Dutheil, J., Odenthal-Hesse, L., & Liedvogel, M. (2022a). *Fine-scale map reveals highly variable recombination rates associated with genomic features in the European blackcap* [Preprint]. Preprints. https://doi.org/10.22541/au.165423614.49331155/v1

Bascón-Cardozo, K., Bours, A., Ishigohoka, J., Odenthal-Hesse, L., & Liedvogel, M. (2022b). Historical recombination maps diverge between Eurasian blackcap populations with distinct migratory strategies. https://doi.org/10.22541/au.166790167.72861799/v1

Berthold, P. (1973). Relationships Between Migratory Restlessness and Migration Distance in Six Sylvia Species*. *Ibis*, *115*(4), 594–599. https://doi.org/10.1111/j.1474-919X.1973.tb01998.x

Bours, A., Pruisscher, P., Bascón-Cardozo, K., Odenthal-Hesse, L., & Liedvogel, M. (2022). *The blackcap (Sylvia atricapilla) genome reveals a species-specific accumulation of LTR retrotransposons* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-2166848/v1

Buard, J., Barthès, P., Grey, C., & de Massy, B. (2009). Distinct histone modifications define initiation and repair of meiotic recombination in the mouse. *The EMBO Journal*, *28*(17), 2616–2624. https://doi.org/10.1038/emboj.2009.207

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & Development*, *25*(10), 1010–1022. https://doi.org/10.1101/gad.2037511

Deniz, Ö., Frost, J. M., & Branco, M. R. (2019). Regulation of transposable elements by DNA modifications. *Nature Reviews Genetics*, *20*(7), 7. https://doi.org/10.1038/s41576-019-0106-6

Derks, M. F. L., Schachtschneider, K. M., Madsen, O., Schijlen, E., Verhoeven, K. J. F., & van Oers, K. (2016). Gene and transposable element methylation in great tit (Parus major) brain and blood. *BMC Genomics*, *17*(1), 332. https://doi.org/10.1186/s12864-016-2653-y

Ellermeier, C., Higuchi, E. C., Phadnis, N., Holm, L., Geelhood, J. L., Thon, G., & Smith, G. R. (2010). RNAi and heterochromatin repress centromeric meiotic recombination. *Proceedings of the National Academy of Sciences*, *107*(19), 8701–8705. https://doi.org/10.1073/pnas.0914160107

Fusani, L., Bertolucci, C., Frigato, E., & Foà, A. (2014). Cryptochrome expression in the eye of migratory birds depends on their migratory status. *Journal of Experimental Biology*, *217*(6), 918–923. https://doi.org/10.1242/jeb.096479

Jones, P. A. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, *13*(7), 7. https://doi.org/10.1038/nrg3230

Kawakami, T., Mugal, C. F., 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. *Molecular Ecology*, *26*(16), 4158–4172. https://doi.org/10.1111/mec.14197

Kent, T. V., Uzunović, J., & Wright, S. I. (2017). Coevolution between transposable elements and recombination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160458. https://doi.org/10.1098/rstb.2016.0458

Klughammer, J., Romanovskaia, D., Nemc, A., Posautz, A., Seid, C., Schuster, L. C., Keinath, M. C., Lugo Ramos, J. S., Kosack, L., Evankow, A., Prinz, D., Kirchberger, S., Ergüner, B., Datlinger, P., Fortelny, N., Schmidl, C., Farlik, M., Skjærven, K., Bergthaler, A., … Bock, C. (2022). *Comparative analysis of genome-scale, base-resolution DNA methylation profiles across 580 animal species* [Preprint]. Genomics. https://doi.org/10.1101/2022.06.18.496602

Krueger, F., & Andrews, S. R. (2011). Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics (Oxford, England)*, *27*(11), 1571–1572. https://doi.org/10.1093/bioinformatics/btr167

LaBarre, B. A., Goncearenco, A., Petrykowska, H. M., Jaratlerdsiri, W., Bornman, M. S. R., Hayes, V. M., & Elnitski, L. (2019). MethylToSNP: Identifying SNPs in Illumina DNA methylation array data. *Epigenetics & Chromatin*, *12*, 79. https://doi.org/10.1186/s13072-019-0321-6

Lam, I., & Keeney, S. (2015). Non-paradoxical evolutionary stability of the recombination initiation landscape in yeast. *Science (New York, N.Y.)*, *350*(6263), 932–937. https://doi.org/10.1126/science.aad0814

Lim, W.-J., Kim, K. H., Kim, J.-Y., Jeong, S., & Kim, N. (2019). Identification of DNA-Methylated CpG Islands Associated With Gene Silencing in the Adult Body Tissues of the Ogye Chicken Using RNA-Seq and Reduced Representation Bisulfite Sequencing. *Frontiers in Genetics*, *10*. https://www.frontiersin.org/articles/10.3389/fgene.2019.00346

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria. URL https://www.R-project.org/.

Singhal, S., Leffler, E. M., Sannareddy, K., Turner, I., Venn, O., Hooper, D. M., Strand, A. I., Li, Q., Raney, B., Balakrishnan, C. N., Griffith, S. C., McVean, G., & Przeworski, M. (2015). Stable recombination hotspots in birds. *Science (New York, N.Y.)*, *350*(6263), 928–932. https://doi.org/10.1126/science.aad0843

Sun, D., Layman, T. S., Jeong, H., Chatterjee, P., Grogan, K., Merritt, J. R., Maney, D. L., & Yi, S. V. (2021). Genome-wide variation in DNA methylation linked to developmental stage and chromosomal suppression of recombination in white-throated sparrows. *Molecular Ecology*, *30*(14), 3453–3467. https://doi.org/10.1111/mec.15793

Westerberg, I. (2019). CpG islands, but not their methylation level, are key in the regulation of meiotic recombination in chicken (Gallus gallus). https://www.diva-portal.org/smash/get/diva2:1285224/FULLTEXT01.pdf

Wyler, M., Stritt, C., Walser, J.-C., Baroux, C., & Roulin, A. C. (2020). Impact of Transposable Elements on Methylation and Gene Expression across Natural Accessions of Brachypodium distachyon. *Genome Biology and Evolution*, *12*(11), 1994–2001. https://doi.org/10.1093/gbe/evaa180

Yelina, N., Diaz, P., Lambing, C., & Henderson, I. R. (2015). Epigenetic control of meiotic recombination in plants. *Science China Life Sciences*, *58*(3), 223–231. https://doi.org/10.1007/s11427-015-4811-x

Yelina, N. E., Lambing, C., Hardcastle, T. J., Zhao, X., Santos, B., & Henderson, I. R. (2015). DNA methylation epigenetically silences crossover hot spots and controls chromosomal domains of meiotic recombination in Arabidopsis. *Genes & Development*, *29*(20), 2183–2202. https://doi.org/10.1101/gad.270876.115

Zamudio, N., Barau, J., Teissandier, A., Walter, M., Borsos, M., Servant, N., & Bourc'his, D. (2015). DNA methylation restrains transposons from adopting a chromatin signature permissive for meiotic recombination. *Genes & Development*, *29*(12), 1256–1270. https://doi.org/10.1101/gad.257840.114

Zhong, Z., Feng, S., Duttke, S. H., Potok, M. E., Zhang, Y., Gallego-Bartolomé, J., Liu, W., & Jacobsen, S. E. (2021). DNA methylation-linked chromatin accessibility affects genomic architecture in Arabidopsis. *Proceedings of the National Academy of Sciences*, *118*(5), e2023347118. https://doi.org/10.1073/pnas.2023347118

3.7 Supplementary materials

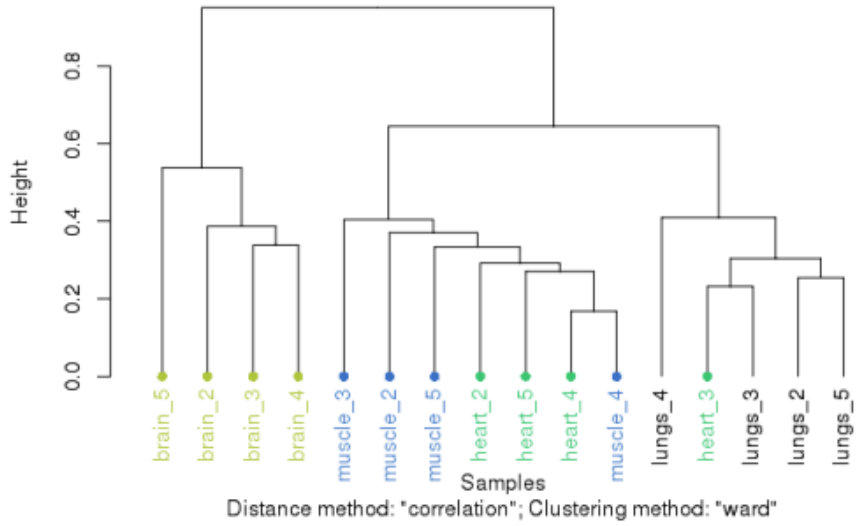


Figure S1. CpG methylation clustering among samples from different tissues

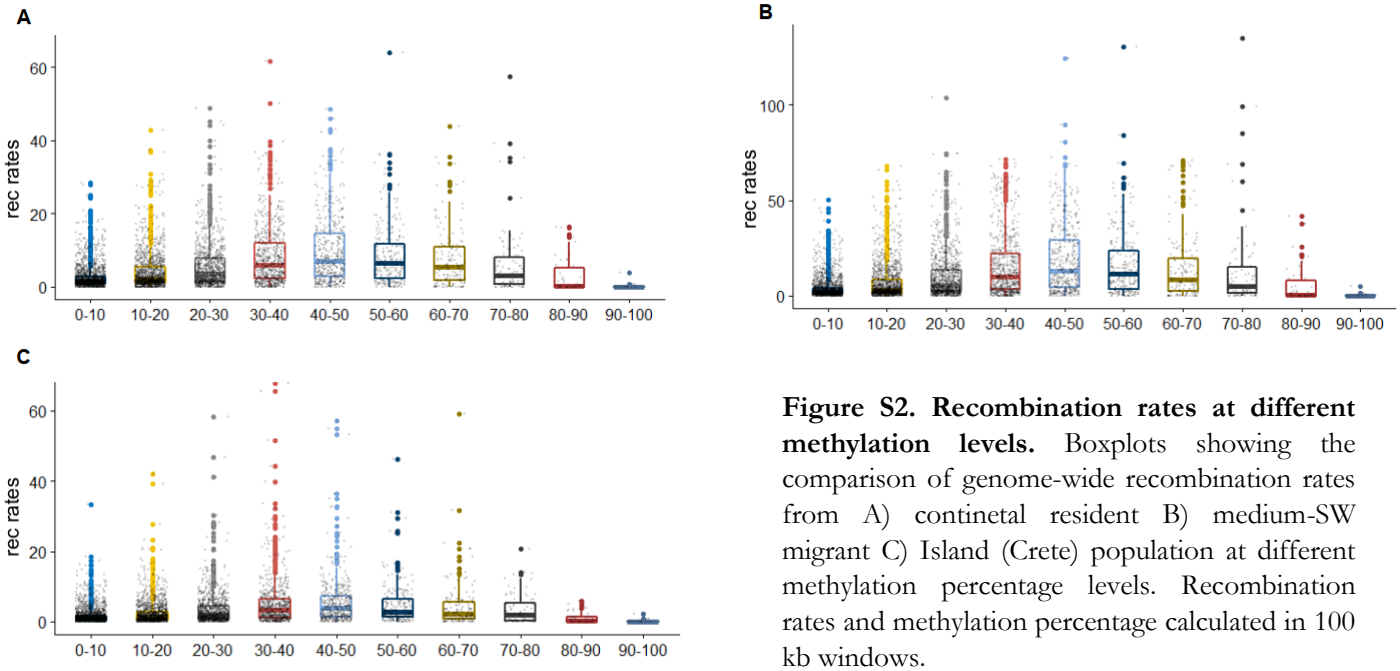


Figure S2. Recombination rates at different methylation levels. Boxplots showing the comparison of genome-wide recombination rates from A) continetal resident B) medium-SW migrant C) Island (Crete) population at different methylation percentage levels. Recombination rates and methylation percentage calculated in 100 kb windows.

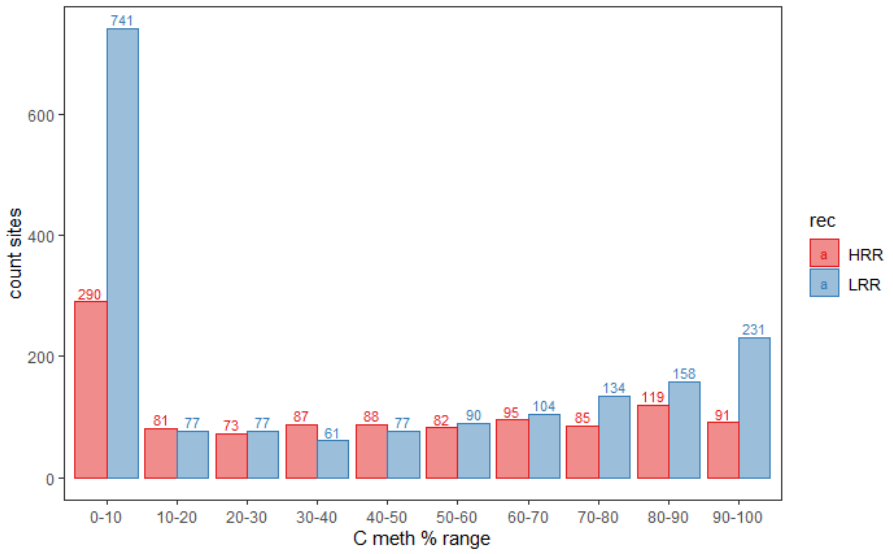


Figure S3. Distribution of CpGs overlapping HRR (in red) and LRR (in blue) across the methylation range.

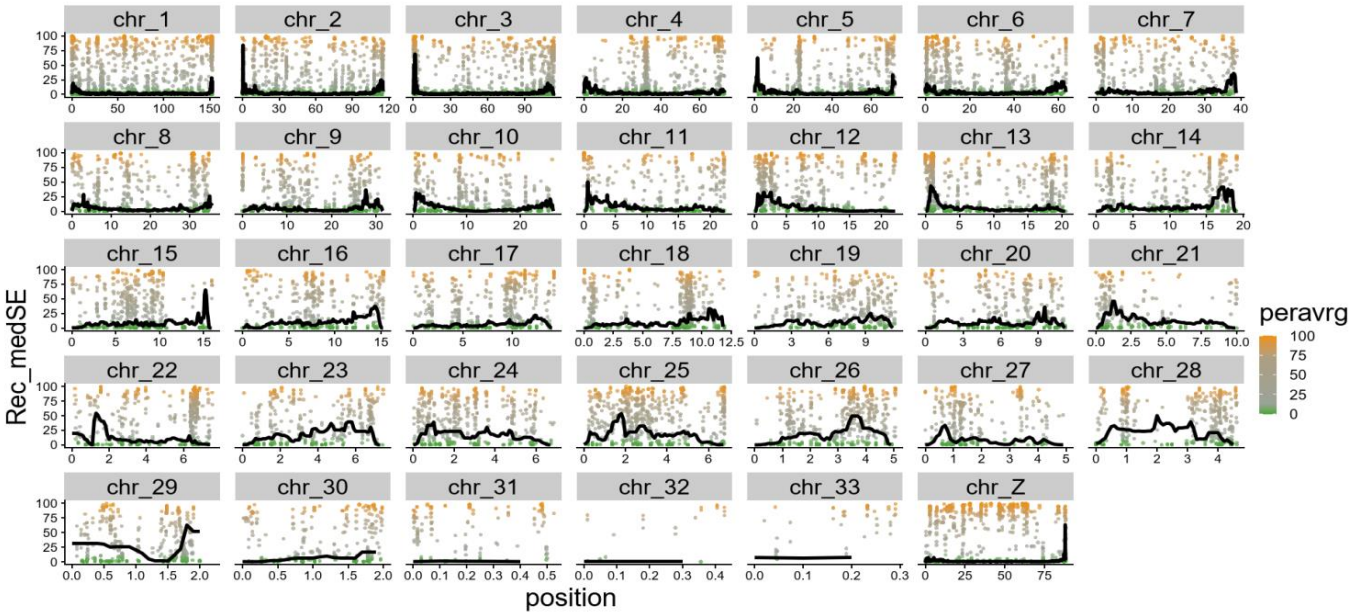


Figure S4. CpGs methylation and recombination rates. Intrachromosomal distribution of CpGs methylation and recombination rates for all the chromosomes. The recombination map belongs to the med-SE migratory population and the rates were calculated in 100 kb windows. Colour scale represents the methylation percentage, from orange denoting high methylated sites to green representing the low methylated site.

CHAPTER IV

Historical recombination maps diverge between Eurasian blackcap populations with distinct migratory strategies

Karen Bascón-Cardozo[1#], Andrea Bours[1], Jun Ishigohoka[1], Linda Odenthal-Hesse*[1#] and Miriam Liedvogel*[1,2#]

Affiliations:

[1] Max Planck Institute for Evolutionary Biology, MPRG Behavioural Genomics, Plön, Germany
[2] Institute of Avian Research, Wilhelmshaven, Germany

*equal contribution

Submitted manuscript: This manuscript was submitted to Molecular Ecology in November 7, 2022

Published preprint: https://www.authorea.com/users/486808/articles/594051-historical-recombination-maps-diverge-between-eurasian-blackcap-populations-with-distinct-migratory-strategies
DOI:10.22541/au.166790167.72861799/v1

4.1 Abstract

Recombination generates new combination of alleles, whereby it maintains haplotype diversity and enhances the efficacy of selection. Despite the apparent stasis in positioning recombination events in birds, recombination rates differ widely across the genome and within species. The causes of recombination rate variation and its evolutionary impact on natural populations remain poorly understood. We used whole-genome resequencing data of 167 individuals of the Eurasian blackcap (*Sylvia atricapilla*) to characterise the historical recombination landscape variation at broad and fine scales among populations with distinct migratory phenotypes. We additionally evaluated the interplay between recombination rates with patterns of genetic diversity, population divergence (based on Fst and dxy), and potential signs of selection. Our comparative analyses revealed: i) Lower divergence of recombination maps at the broad scale and higher variability at fine scales. Resident island populations showed higher variability in recombination patterns among them and with continental populations. Recombination rates were more conserved in continental populations regardless of the migratory phenotype. ii) The degree of divergence between recombination maps correlated with population differentiation. It could also recapitulate population-specific demographic history and genetic structure. iii) Recombination rates correlated negatively with Fst and positively with nucleotide diversity and dxy, suggesting that recombination may reduce the effect of linked selection over the loss of neutral diversity. We identified chromosomal regions with potential signs of linked selection. This study evidences that recombination is a variable trait that shapes the diversity and evolution of population differentiation in the blackcap.

Keywords

historical recombination maps, nucleotide diversity, population divergence, linked selection, population genomics, bird migration

4.2 Introduction

Meiotic recombination facilitates novel combinations of alleles, contributes to generating and maintaining genetic diversity along the genome, and potentially leads populations to adapt and evolve over prolonged periods of time (Felsenstein, 1974; Keightley & Otto, 2006). When selection acts over physically linked loci (i.e., Hill-Robertson interference, Hill & Robertson, 1966), recombination breaks up this linkage. It reshuffles haplotypes, fundamentally influencing the efficacy of selection in two ways: First, modulating the speed of purging deleterious mutations and spreading advantageous alleles in the population (Betancourt et al., 2009; Crow & Kimura, 1965; Hickey & Golding, 2018). Secondly, affecting the strength of linked selection and maintaining neutral diversity physically linked to the loci that are under either positive (selective sweeps) or negative selection (background selection) (Castellano et al., 2020; Charlesworth et al., 1993; Chase & Mugal, 2022; Cutter & Payseur, 2013). Recombination further plays a critical role in influencing the degree of gene flow and introgression between populations (Martin et al., 2019; Schumer et al., 2018). High recombination rate is associated with greater population divergence (dxy) (Burri et al., 2015; Nelson et al., 2021), whereas low recombining regions are more related to loss of nucleotide diversity due to linked selection, population differentiation, and speciation (Burri et al., 2015; Cutter & Payseur, 2013; Henderson & Brelsford, 2020; Samuk et al., 2017). Given the impact of recombination on the evolution and adaptation of populations, variation in recombination rates emerges as an exciting phenomenon to be investigated in light of population genomics.

Variability of recombination rates at different scales has been widely described over the last few years, spanning from variation within the genome, between sexes and individuals, to populations, species as well as different taxa (Peñalba & Wolf, 2020; Peterson & Payseur, 2021; Smukowski & Noor, 2011; Stapley et al., 2017). Birds lack a functional PRDM9 protein responsible for positioning and initiating recombination events in many mammals. The majority of recombination events ("hotpots") cluster in promoter-like regions or regions where open chromatin stretches allow access to the transcription machinery, presumably conferring more conservation to the recombination landscape (Singhal et al., 2015). Recombination landscapes in birds appear conserved at broad scales (genome-wide); however, remarkable differences in recombination rates are evident at the fine scale (Kawakami et al., 2017; Singhal et al., 2015; Bascon Cardozo et al., 2022). Recombination variation within the avian genome is associated with the presence and

distribution of specific genomic features. High recombination rates are enriched in GC-rich content sequences, including CpG islands, regulatory regions such as promoters, transcription start sites, 5'UTR regions, and specific families of transposable elements (TEs) (Bascón-Cardozo et al., 2022; Kawakami et al., 2017; Peñalba et al., 2020; Singhal et al., 2015; Smeds et al., 2016). Low recombining regions associate with intergenic regions, chromosomal rearrangements (i.e. inversions) and long tandem repeats (Bascón-Cardozo et al., 2022; da Silva et al., 2019; Ishigohoka et al., 2021; Kawakami et al., 2017). Moreover, the recombination rate is negatively associated with chromosome size (Backström et al., 2010; Bascón-Cardozo et al., 2022; Kawakami et al., 2017; Peñalba et al., 2020; Smeds, et al., 2016). Despite these striking correlative patterns, broad and fine scale recombination rate variation and its implications in natural populations remain poorly studied. To disentangle recombination variation and its impact throughout evolution in natural populations, it is crucial to use consistent approaches to estimate and systematically compare recombination maps at different scales, across and within species.

The Eurasian blackcap, *Sylvia atricapilla* (referred to as blackcap in the following) is a powerful study system for investigating genetic variation in behavioural traits such as migration in an evolutionary framework. Blackcaps are night-migratory songbirds encompassing populations that exhibit the entire repertoire of migratory phenotypes, including variation in the propensity to migrate, migratory distance, and orientation (e.g., Berthold et al., 1992; Helbig, 1991; Merlin & Liedvogel, 2019). Populations of blackcaps throughout the distribution range started to diverge as recent as ~30,000 years ago (ya), resulting in an overall low genetic differentiation, as well as limited gene flow between migratory populations, as suggested by the contemporary genetic structure between a subset of the populations studied previously (Delmore et al., 2020). At the same time, these populations differ in their demographic history, indicating that they may evolve differently and have potentially experienced different strengths of selective pressures. Hence, blackcaps provide an ideal empirical scenario to study recombination variation and evaluate their association with population dynamics.

In a previous study, we characterised the fine-scale historical recombination map in the blackcap system, revealing variable recombination rates across the genome and within chromosomes associated with specific genomic features (see chapter II, Bascón-Cardozo et al., 2022). In this context, we were able to identify divergent chromosomal regions between the blackcap and one of its closest sister species, the garden warbler (*Sylvia borin*). Comparative studies using genetic linkage

maps were also carried out in more distantly related species (Backström et al., 2010; Peñalba et al., 2020), however little is known about the variation between different populations within the same species. This raised our interest to focus our evaluation of recombination rate variation on the population scale, specifically contrasting populations with divergent migratory phenotypes and evaluating the association with patterns of nucleotide diversity, population differentiation and potential signatures of selection. For that, we first inferred population-scaled recombination rates of twelve populations using an approach based on Linkage Disequilibrium (LD) taking population-specific demography into account. With this, we assessed the following questions: (i) Do historical recombination maps vary between populations with different migratory strategies at broad and fine scales? ii) What is the association of recombination rates with specific population-genetic parameters such as nucleotide diversity and Tajima's D within each population, and (iii) is the variation of recombination rate associated with relative and absolute population divergence (Fst, dxy)?

4.3 Materials and Methods

Genomic Data

We analysed whole genome resequencing data of 167 blackcap individuals across the species' range, including all migratory phenotypes published in Delmore et al. (2020) and Ishigohoka et al. (2021). The resident populations include continental residents (cont-res) (see Table S1), as well as island populations, which include Azores (AZO), Cape Verde (CAVE), Canary Islands (CAN), Madeira (MAD), Mallorca (MAL), and Crete (CRE). Continental migrants are separated by migratory distance: short-, medium-, and long-distance (short, medium, long) and further subdivided by migratory orientation: south-west (SW), north-west (NW), and south-east (SE) (see **table S1**).

Sequence reads obtained by Illumina NextSeq 500, HiSeq 4000 or NovaSeq 5000 were mapped to a chromosomal-level reference genome (available under NCBI BioProject PRJNA558064, accession number GCA_009819655.1) at a median coverage of 15.1 ± 11.1X. The sample collection and genome assembly details are further described in Ishigohoka et al., 2021. A VCF containing polymorphic sites (SNPs) (published in Ishigohoka et al., 2021) was filtered using GATK version 4.1.6.0 (McKenna et al., 2010) and VCFtools (Danecek et al., 2011) with the following: minimum Genotype Quality of 20, a minimum depth of 10, a minor allele count (mac)

of 1 and depth per site calculated as three times higher than the average (=< 10443), while ensuring only variant sites remain. After extracting each population separately, we used VCFtools to filter out sites with missingness greater than or equal to 0.7 to yield a collection of high-quality SNPs for downstream analysis (the number of SNPs retained for each population is summarised in Table S1. We additionally removed singletons for the estimation and analysis of recombination rates.

Inference and comparison of recombination maps between different populations

We estimated population recombination rates for each population separately using Pyrho (Kamm et al., 2016; Spence & Song, 2019), an LD-based approach that estimates and scales the per-generation per-base recombination rate (r) using historical population-specific effective population sizes (Ne) and mutation rate. Demography for each population was inferred as described below (2.3 demography inference), and we used the mutation rate of a related species, the collared flycatcher ($4.6x10^{-09}$ site/generation) (Smeds et al., 2016) for all the estimations. We generated lookup tables with pyrho "make table". Then, we input unphased genotypes in VCF format and assign a block penalty of 20 and a window size of 50 for the recombination rates estimation with pyrho "optimise". Recombination rates were inferred for each population separately. The estimates of Pyrho (scaled-recombination per site per generation rate) were converted to cM/Mb, and the linkage genetic map (genetic distance in cM) was calculated for each population with a custom script (Bascón-Cardozo et al., 2022).

To compare recombination maps between all populations, we calculated average recombination rates in 100 kb non-overlapping windows (as in Bascón-Cardozo et al., 2022) and performed genome-wide pairwise comparisons between all populations considering only autosomes. We measured the correlation coefficient for all the comparisons using Kendall's rank correlation test in R version 4.0.5 (R Core Team 2021) and visualised them as a heatmap with ggplot (Wickham, 2016). To illustrate local recombination patterns, we plotted the intra-chromosomal distribution of recombination rates for all populations. Finally, we calculated the recombination rate variance per window, including all the populations, to identify regions across the genome with higher variability in recombination patterns.

Additionally, we calculated the average recombination rate per chromosome in cM/Mb, weighted by the physical distance between each pair of sites where recombination was estimated, for each population. We calculated and compared average recombination rates among all populations using

nonparametric pairwise Wilcoxon rank sum tests, Bonferroni corrected statistics with no assumption of equal variances (Mann-Whitney test) with the function function **'**pairwise.wilcox.test**'** in R.

Demography inference

For demography inference, we used whole-genome resequencing data, including 167 blackcaps used in Ishigohoka et al., 2021. The SNP data in VCF (variant call format) were statistically phased using SHAPEIT2 (Delaneau et al., 2013), using the blackcap genetic map inferred in (Bascón-Cardozo et al., 2022) with Pyrho. We further polarised the phased SNPs according to allele frequencies in outgroup samples (five garden warblers and three African hill babblers). We used Relate (Speidel et al., 2019) to infer genome-wide genealogy of the 167 individuals. Coalescent patterns of samples within and across focal populations were extracted from the inferred genealogy. Effective historical population sizes and relative cross-coalescent rates were computed based on the distribution of coalescent times.

Simulation analysis

We performed a simulation to investigate whether different demographic histories affect the inference of recombination rate using samples of diverse populations. For that, we used msprime (Baumdicker et al., 2022) to simulate an ancestral population of 1,000,000 diploids which splits at 10,000 generations ago into three populations that undergo three different demographic trajectories: constant size, ten times expansion, and 1/10 contraction over the time until the present. We simulated a chromosome of 10 Mb with a mutation rate of 4.6×10^{-9} (/base-pair/generation) and constant recombination rates: 1, 4 and 10 times the mutation rates. We sampled 50 individuals per population and saved phased genotypes in VCF format. Subsequently, we estimated recombination rates for the three populations separately using Pyrho with block penalty of 25 and window size of 50. Then recombination rates (r) were converted to cM/Mb and compared among the three populations.

Evaluating recombination association with nucleotide diversity and population divergence

To characterise the relationship between recombination rates and patterns of diversity and population divergence, we calculated nucleotide diversity (π) for each population and F_{st} and dxy for all population pairs *using* a VCF file containing all callable sites (meaning polymorphic and non-

polymorphic sites that had mapping support) and varying tools from the "genomics general" toolkit (Martin et al., 2020). *We used parseVCFs.py* (https://github.com/simonhmartin/genomics_general/VCF_processing release 0.4) to filter the described VCF with min-depth >5 and max <60 and ran popgenWindows.py (https://github.com/simonhmartin/genomics_general release 0.4) with a window size of 100 kb, ensuring that at least 10 kb were called in each window. Missingness was set to 0.7 (minData), taking the proportion of individuals covering at least 10kb per window. Genome-wide pairwise comparisons between recombination rates with nucleotide diversity *(π)*, $F_{st,}$ and dxy were carried out in 100 kb windows for all populations. We used the "ggcorrplot' package (https://github.com/kassambara/ggcorrplot) in R to measure Kendall's rank correlation coefficients with a significance threshold of 5 %.

Studies in flycatchers and mice suggested that Fst outliers tend to occur in regions where recombination is low or suppressed (Burri et al., 2015; Cruickshank & Hahn, 2014; Nachman & Payseur, 2012; reviewed in Ravinet et al., 2017), which we also assess for the blackcap genome here. We first identified populations with elevated genome-wide Fst values. Subsequently, we identified windows with Fst outlier taking the 95th and 99th percentiles, and we plotted them together with the recombination maps of the respective populations.

Since we used recombination rate estimations for each population and evaluated the correlation coefficients among all pairs of populations, we additionally assessed the relationship of dissimilarity in recombination maps with population differentiation (Fst). We calculated recombination dissimilarity by subtracting the genome-wide pairwise correlation measurements among all pairs of populations from the maximum value (1). We estimate Spearman and Kendall's correlation between recombination dissimilarity and Fst values for all population pairs. To support our statistical analysis, we performed Mantel tests implemented in the R-package 'vegan' (Oksanen et al. 2020), which statistically compares distance matrices. For that, we converted our data sets of recombination dissimilarity to a matrix and the pairwise comparisons of Fst values to a genetic distance matrix (Fst matrix), including all populations. We compared both matrices using the Mantel test with the Kendall-rank method and randomly permuting 9,999 times the rows and columns of one of the matrices. We performed all statistical analyses encompassing all populations. We also ran the same analysis on a subset of the data contrasting all continental with all island populations.

Recent studies suggest that elevated levels of divergence (dxy) occur in regions with elevated recombination rates (Burri et al., 2015; Nelson et al., 2021). Analogous to these studies, we defined Highly Recombining Regions (HRR) as windows with an LD-based recombination rate exceeding four times the background recombination rate (genome-wide-average recombination rate) for each population. We then calculated the average dxy within HRR for each population weighted by the number of sites within each window. We compared these with the genome-wide weighted average dxy separately for each population pair using a one-sample Wilcoxon signed-rank test.

Recombination and Tajima's D

To allow for an association between recombination rate and possible patterns of selection, we calculated Tajima's D to distinguish between genomic regions neutrally evolving and regions evolving via non-random processes. Tajima's D calculates the difference between two measures of genetic diversity: the mean number of pairwise differences and the number of segregating sites. We calculated Tajima's D for each population in 100 kb windows using VCFtools. Then, we characterised their association with recombination rates by performing partial correlations using "ppcor" package (Kim, 2015) and treating nucleotide diversity as a confounding variable.

To complement our analysis, we also included the annotated gene density described in Bascón-Cardozo et al., 2022 and calculated gene count in 100 kb windows along the genome. Using Kendall's rank coefficient, we measured the genome-wide association of gene density with recombination rates for each population. Finally, we visualized the distribution of gene density together with recombination rates, nucleotide diversity, and Tajima's D using the R package ggplot.

4.4 Results

Our analyses revealed that at the broad scale, genome-wide recombination maps of different blackcap populations across the species' distribution range appear highly similar (Fig. 1A). Higher correlation coefficients were observed among migratory populations (Kendall's tau (r_τ)>0.82, $p<0.0001$), except for the long-distance migrants, which are mainly characterized by low SNP density and the lowest correlation coefficients of all comparisons (r_τ=0.4-0.5, $p<0.0001$) (Fig. 1A). The continental resident population was more similar to continental migratory populations compared to resident island populations. Island populations showed lower levels of inter-

population correlation with variable patterns among islands and also with resident and migratory continental populations (Fig. 1A). Correlation coefficients between the island and continental populations were highest in Madeira and Canary Islands (Fig. 1A). When we calculated and compared genome-wide recombination rate averages between populations, we found that they ranged between 0.8 to 6.2 cM/Mb (Table S2, Fig.1B). Island populations showed lower genome-wide average recombination rates compared to continental populations (Wilcoxon rank sum test $p<0.0001$), with the exception of long-distance migrants that showed overall the lowest average recombination rate (0.8 cM/Mb, Fig. 1B). Generally, recombination rates were higher in micro-chromosomes than macro-chromosomes in all populations, and chromosome Z showed one of the lowest averages of recombination rates (Fig. 1B).

In contrast to the nearly consistent broad-scale patterns across blackcap populations, recombination rates were highly variable at a finer scale. Intra-chromosomal comparisons revealed divergent patterns of recombination between populations across different chromosomes (Fig. 1C**,** Fig. S1, S2): Island populations showed different patterns among populations, e.g., chromosome 18 in the Crete population (Fig. 1C, Fig. S2), whereas continental populations showed more variability in the smoothness of recombination maps (Fig. 1C, Fig. S1**)**. The variation was more prominent in micro-chromosomes (e.g., chromosomes 18, 23, 26, 28, 29 in Fig. 1C**,** Fig. S4) than in macro-chromosomes. In the largest macro-chromosomes, the recombination landscapes were more consistent among populations (Fig. S4), with high recombination rates towards the end of chromosomes and lower recombination rates in the center (e.g., chromosomes 1, 2, 4, Fig. 1C, Fig. S1, S2). Recombination rates in chromosome Z were conserved across populations at both broad and fine scales. All populations consistently show suppressed recombination along the entire Z chromosome and a peak of historical recombination at one chromosomal end which may represent the pseudoautosomal region (PAR) (Fig. S1, S2).

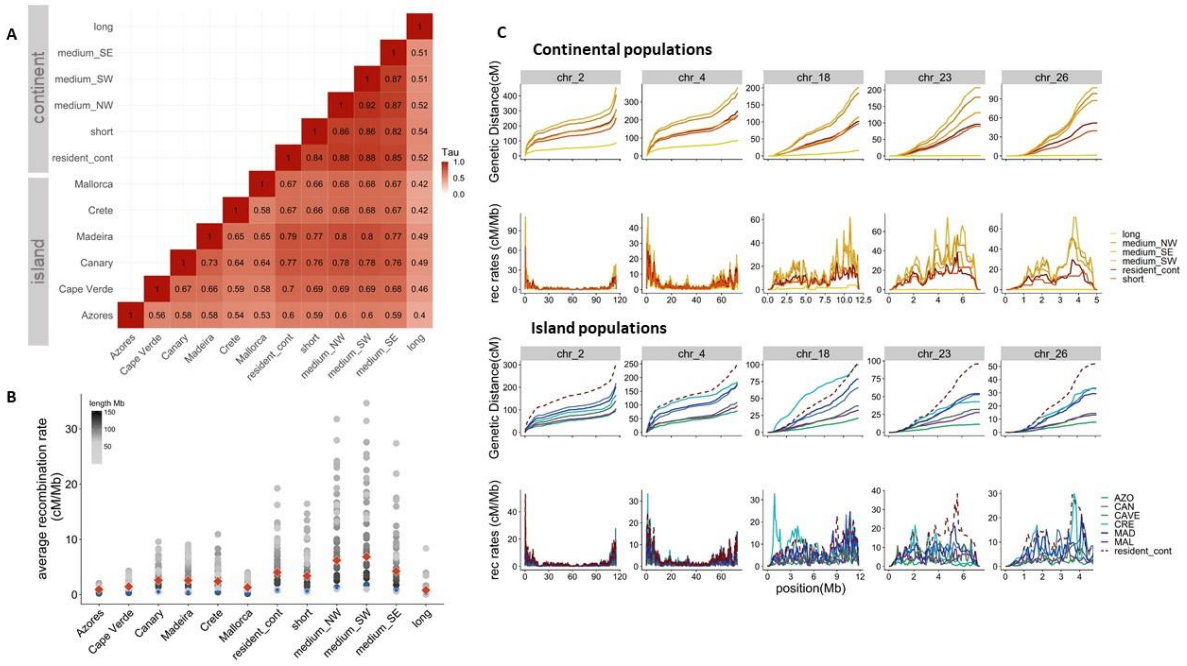


Figure 1. Recombination variation at different scales A) Correlation heatmap of recombination maps between all pairs of populations differing in migratory phenotype. Populations are classified between the island (resident) and continental (migratory and resident) populations. Heatmaps show Kendal rank test correlations of pairwise comparisons between recombination rates calculated in 100 kb non-overlapping windows. **B)** Average recombination rates are given for each chromosome and all populations. The coloration gradient of dots reflects the genome-wide average recombination rate for larger (black) to shorter (light-grey) chromosomes. Chromosome Z is shown in blue dots. Red diamonds represent the genome-wide recombination rate average for each population. Mean values of recombination rates are significantly different between populations (Wilcoxon rank sum test, p<0,001), except for two populations from the Canary (CAN) and Madeira (MAD) islands (Wilcoxon rank sum test, p=0.25). **C)** Distribution of intra-chromosomal recombination rates (in cM/Mb) and genetic distance maps (in cM) across selected macro- (chromosomes 2 and 4) and micro-chromosomes (chromosomes 18, 23, 26) from migratory and resident populations subdivided into continental populations (top panel) and island populations (bottom panel) including continental residents in dashed lines.

The strong relationship between recombination and nucleotide diversity

All populations show a strong positive association between recombination rate and nucleotide diversity (π) (r_τ>0.55, p<0.0001, Table S2). The lowest coefficient of positive correlation was observed in the long-distance migrants (r_τ=0.44, p<0.001). Genomic regions where decreased nucleotide diversity coincided with low or suppressed recombination were shared across all

populations (Fig. 2). Interestingly, in some chromosomes, potential targets of selection represented as genes were not entirely depleted within these regions (see chromosomes 16, 22,26 Fig. 2). In chromosomes 20 and 29, a substantial number of genes was more evident (Fig. 2).

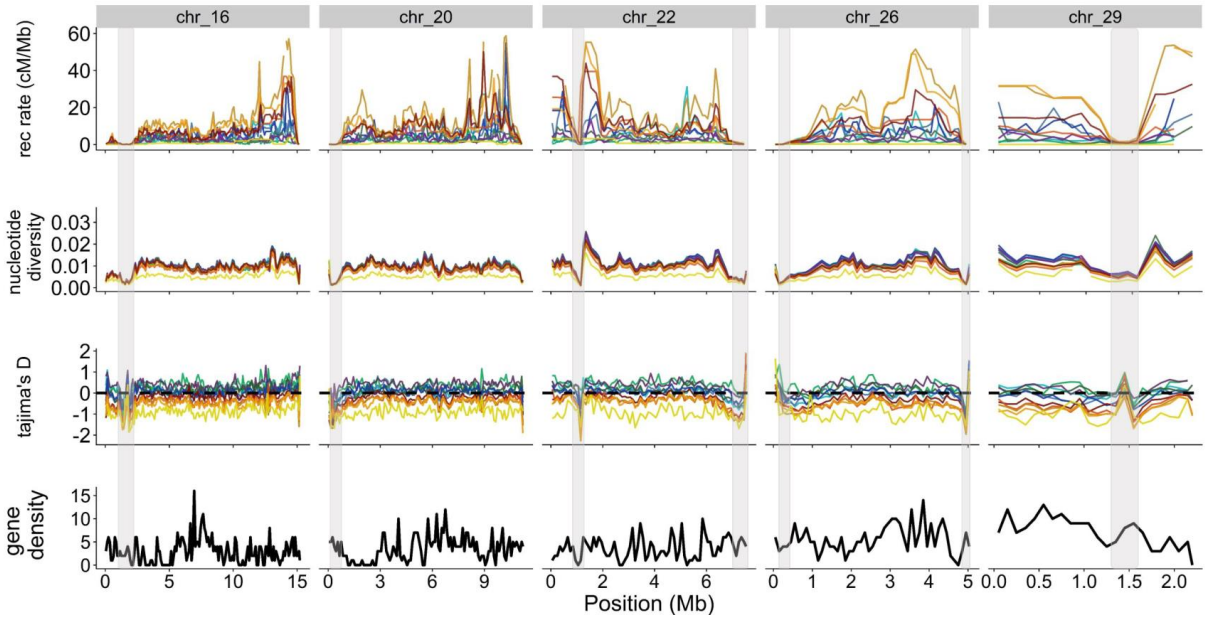


Figure 2. Chromosomal regions with possible signs of linked selection. Plots showing nucleotide diversity (π), recombination rates (rec rate), Tajima's D, and gene density distributions across selected chromosomes calculated in 100 kb non-overlapping windows for all populations (color coding as in Figure 1). The positive association between recombination rate and nucleotide diversity is reflected by highly similar patterns in all the populations. Grey bars highlight regions of low nucleotide diversity, reduced recombination rates and distortion in the site-frequency spectrum indicated by high or low patterns of Tajima's D and variable gene density.

Variation in recombination rates and distortion of the Site-frequency spectrum (SFS) as potential means of selection

We found recombination rates weakly correlated with Tajima's D in all blackcap populations when we performed partial correlations controlling for nucleotide diversity for each population (Table S2). The correlation was slightly positive in continental populations and insignificant for the short-distance population. In contrast, island populations show a negatively weak correlation Table S2). Genome-wide average Tajima's D in island populations was greater (≥ 0) compared to continental populations, which showed values below 0 (Fig. S5). Interestingly, when we looked into the intra-

chromosomal distribution of Tajima's D, we identified a distortion, either a peak or drop in the site-spectrum frequency (SFS), coinciding with some of the chromosomal regions characterised by decreased recombination rates, low nucleotide diversity and intermediate gene density described in the previous section (Fig. 2 see chromosome 16,20,26 and 29).

The interplay of variable recombination rates with relative and absolute population divergence

The pairwise comparisons among all populations revealed different associations between relative divergence or differentiation (Fst) and recombination rates between populations (Fig. 3, Fig. S6). Fst calculated for each population pair with their respective recombination maps was negatively correlated in some, but not all populations (Fig. 3, Fig. S6). Fst outliers (95^{th} and 99^{th} percentile) in these populations occurred predominantly in regions of low or suppressed recombination. This was consistent in both recombination maps from the populations where Fst was calculated (see CRE-MAL Fig. 3A, also see cont_res-CRE, CAVE-CRE, short-MAD, AZO-CRE in Fig. S6). Among continental populations, short, medium-distance migrants and resident populations show negative correlation coefficients between recombination rate and Fst when compared with the long-distance population (for instance, Fst resident_cont-long: r_τ =-0.31 for resident_cont and r_τ =-0.068 for long). In the comparison between medium-distance populations with different migratory orientations and short-distance populations, a weak positive association of recombination and Fst was seen (Fst medium-SE_short, r_τ =0.06 for medium_SE, r_τ =0.08 for short). However, not all populations showed a consistent association. For instance, the Fst calculated between Azores and Mallorca (Fst AZO-MAL) was significantly associated with recombination rates of Mallorca but not Azores (Fig. 3B). Additionally, in the comparison between the Azores and medium_SW migrants, Fst AZO-medium_SW correlated positively with recombination rates of medium_SW (r_τ=0.13) but negatively with recombination rates of AZO (r_τ =-0.12, see also: Fst AZO-medium_SE, Fst AZO_medium_NW, and Fst short-MAL Fig. S6). Thus, the recombination rate in one population was differentially associated with Fst compared to the other population.

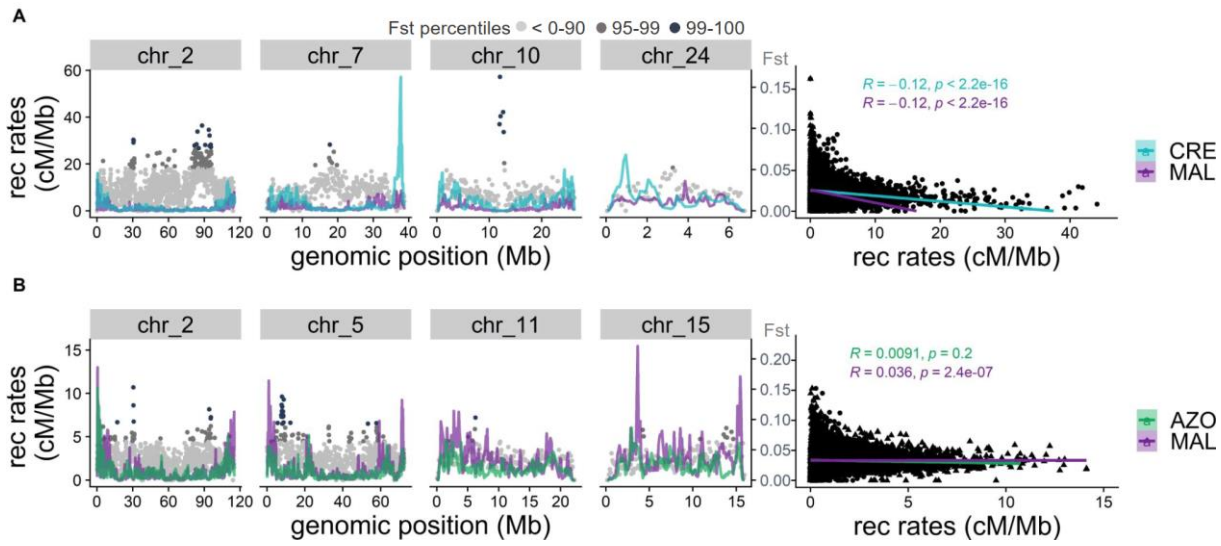


Figure 3. Population differentiation and recombination rates variation. Recombination rates (rec rate) and Fst outliers in gray (with 95th percentile in black and 99th percentile in dark blue) across selected chromosomes for two island population comparisons: (A) between Crete (CRE)-Mallorca (MAL) and (B) between Azores (AZO) -Mallorca (MAL) on the left panels. Kendall correlations between Fst and recombination rates calculated in 100kb windows for the respective pair of populations are shown on the right panels. Colour coding for populations as in Figure 1.

In populations where recombination maps are less conserved such as island populations, recombination patterns showed higher variability in regions of elevated Fst in some of them (See AZO-MAL Fig. 3B). Therefore, we evaluated the relation between genome-wide dissimilarity of recombination maps between populations and population differentiation (Fst). This revealed a positive association (r_{τ}=0.29, p<0.001, Fig. 4), which was further supported by comparing recombination dissimilarity matrices and Fst matrices for all population pairs (Mantel's test r=0.13, p<0.001, 9999 permutations). This relationship was driven by island populations (Fig. 4), and when we separated island populations from continental populations, the association in island populations became even stronger (r_{τ}=0.45, p<0.05, Mantel's r=0.23 p<0.001, Fig. 4). Whereas, the correlation pattern within continental populations disappeared (in Mantel's r= 0.1, p>0.05) or turned negative (r_{τ} =-0.61, p<0.01, Fig. 4).

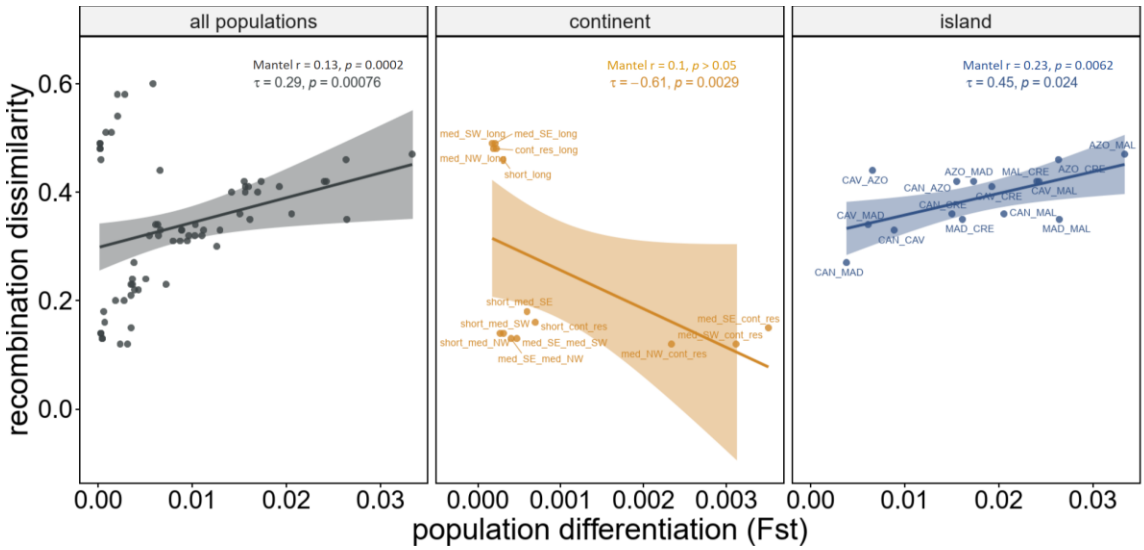


Figure 4. Dissimilarity in recombination maps is associated with population differentiation in island populations. Correlation between genome-wide recombination dissimilarity and population differentiation (Fst) in all the populations, continental populations, and island populations. Correlations were measured with Mantel statistics and Kendall's rank tests.

Our results revealed a positive association between recombination rates and absolute divergence (dxy) between all populations (r_τ>0.55 -r_τ>0.36, p<0.0001). Again, the long-distance migratory population showed the weakest correlation coefficient, albeit positive. When we focused only on dxy within HRR, we found a significantly greater dxy average compared to the genome-wide dxy in all population pairs (Wilcoxon test p<2.2e-16, Fig. 5), suggesting an increase in population divergence within HRR.

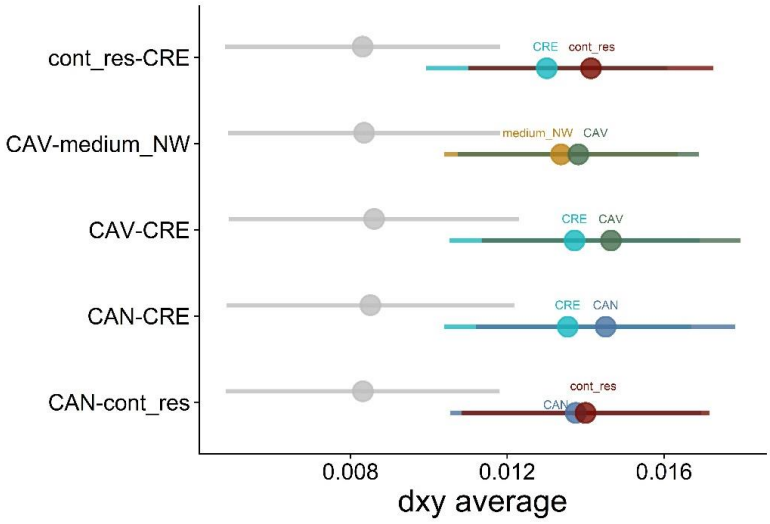


Figure 5. **Greater population divergence (Dxy) in High Recombining Regions (HRR).** Dot plots showing the comparison of genome-wide average dxy calculated for the focal population pairs across the whole genome in grey and within HRR corresponding to each population (denoted in color codes as in Figure 1). The plot shows a subset of selected populations with higher genome-wide dxy for visualization. The dots represent the means and error bars at the 95% confidence level. All the comparisons for each population pair are significantly different (Wilcoxon. test, p<0,001).

4.5 Discussion

Recombination rate variation between different populations at broad and fine scales

We compared historical recombination maps at genome-wide (broad-scale) and intra-chromosomal (fine-scale) levels in 100 kb windows among blackcap populations with different migratory phenotypes. Our data shows that recombination maps of distinct populations have diverged at fine scales, while patterns at broader scales are more conserved. This pattern of similarity in broad-scale recombination rates has previously been found between and within other bird species (Kawakami et al., 2017; Peñalba et al., 2020; Singhal et al., 2015) as well as in mammals (Betancourt et al., 2009; Campbell et al., 2016; Myers et al., 2005), fish (Shanfelter et al., 2019), and reptiles (Schield et al., 2020). The variation of recombination rates at the genome-wide scale can be attributed to genetic factors, epigenetic regulation (e.g., DNA methylation), and mutations in genes involved in meiosis and/or double-strand break repair pathways (Brand et al., 2018; Charlesworth, 2018). Hence, the conservation of recombination rates at broad-scales could reflect

molecular mechanisms shaping recombination dynamics. For example, it is often observed the enrichment of recombination events in open chromatin regions and functional elements in organism lacking the protein PRDM9 such as birds and yeast, where recombining patterns are often stable (Lam & Keeney, 2015; Singhal et al., 2015). This similarity in genome-wide recombination rates could also be a common pattern in organisms containing recombination hotspots since this is not observed in organisms that lack of them, such as *Drosophila* (Samuk et al., 2020).

Our observed variation in the level of correlation among recombination landscapes of distinct blackcap populations may recapitulate population structure and demography (Delmore et al., 2020, Bours et al. in preparation). Blackcap population stratification suggests the majority of island populations cluster separately, while continental migrant populations form one genetic cluster. Even continental resident populations are more similar to the continental migratory population than to the resident populations on islands (Delmore et al., 2020). This mirrors the higher pairwise correlation of recombination maps among continental populations regardless of the migratory phenotype, suggesting that recombination patterns are predominantly formed by genetic divergence and the evolutionary history of the populations. Contact zones across a migratory divide separating populations with distinct migratory behaviour exist in central Europe between medium-distance SW and SE migrants, in line with that, recombination rates between both populations share the highest similarities.

Long-distance migrants were the exception, showing lower correlation coefficients with all the rest of the populations, probably due to their overall low SNP density (despite good coverage), diversity, and sample size. In contrast to other populations, samples for the long-distance migrant phenotype were not exclusively sampled at the breeding locations but included samples taken during migration, thus, the sample pool likely covers a broader population range. Adding further individuals with known breeding origins to complement the samples from this phenotype in the future would be ideal for clarifying whether recombination rate in this population is indeed exceptional due to molecular, biological, and evolutionary reasons or because the resolution was too low in our current study to allow for accurate characterisation of this population.

Reflecting what is known for the genetic structure, island populations show less similarity among themselves and with continental populations, except for Crete, which clusters closer to the continental populations in comparison to the rest of the island populations and didn't show an

exceptionally high correlation with continental populations. The Mallorca population showed higher recombination rate divergence with islands and continental populations, mirroring the separation reported in PC2. Macaronesian islands form two genetic clusters, one between the Canary Islands and Madeira which are closer to the migrant continental populations and this was reflected in higher correlation in recombination maps between both populations and with the continent. A second cluster was formed by Cape Verde and Azores that did not show increased similarity levels compared to the rest of the populations. Thus, additional factors, for example, population sub-structure likely also contribute to observed patterns of similarity and variation of recombination maps between populations.

Population demography is another important factor that affects the variation in recombination rates. Our simulations with bird-specific parameters corroborated that the approach used in this study to infer recombination rates is not biased by population-specific demography (Fig. S3). Therefore, the divergent recombination maps may reflect biological factors and different evolutionary histories in the populations. Blackcap populations split into different migratory phenotypes about 30,000 ya. After the split, island populations experienced a reduction of effective population size (Ne), whereas Ne in continental populations increased (Delmore et al. 2020). Consistent with that, the genome-wide average of Tajima's D was significantly lower in continental populations compared to island populations, indicative of expansion and contraction events in these populations (Tajima, 1989). Perhaps, due to the low Ne in island populations, selection (directional or purifying selection) is less efficient (selection relaxation), and genetic drift has larger effects (Charlesworth, 2009; Gravel, 2016; Ohta, 2013; Wright, 1931), resulting in divergent recombination landscapes. Continental populations with higher Ne may have more substantial selective pressures over migratory traits, maintaining "optimal" recombination rates and lowering the effect of genetic drift, which may contribute to the conservation of recombination maps. Probably that also relates to the difference in genome-wide recombination rates among continental populations showing higher recombination rates than island populations. However, as populations differ in population size and SNP density and additional population-specific factors may influence recombination rates, one should be cautious when comparing absolute values. The pairwise correlation between recombination rates and Tajima's D was positive for continental populations and negative for island populations, probably reflecting the role of recombination in the efficacy

of selection in populations with different histories. However, correlations were only weakly significant.

Our results reveal variation at fine scales between different populations of the Eurasian blackcap. We identified chromosomal regions with divergent recombination patterns between populations across several chromosomes, predominantly in micro-chromosomes. The level of variation differed between populations that exhibit different migratory behaviour. More specifically, resident island populations showed higher variation in their recombination patterns, whereas continental populations mostly varied in the intensity (smoothness) of recombination at specific genomic regions, possibly influenced by variation in sample sizes and SNP densities between populations. Incidences of local recombination variation have also been described in other bird species (Backström et al., 2010; Kawakami et al., 2017; Stapley et al., 2010; van Oers et al., 2014), and may therefore represent a common phenomenon in birds. This fine-scale variation could be attributed to structural variation such as inversions and large deletions suppressing recombination, which have been described in several studies in the blackcap and other species (Hooper & Price, 2017; Ishigohoka et al., 2021; Morgan et al., 2017; Völker et al., 2010). In addition, the presence of other specific genomic features, such as retrotransposons, the genetic and epigenetic regulation of CpG islands, and particular genes may also affect local recombination variation (Bascón-Cardozo et al., 2022; Kawakami et al., 2017; Peñalba et al., 2020; reviewed in Stapley et al., 2017). The areas identified with more significant variability in recombination rates between populations could be associated with structural variants or differential gene regulation providing exciting targets for further exploration in genotype-phenotype association studies. Perhaps some of these regions are associated in the evolution and regulation of the migratory behaviour. However, this is purely speculative at present, as migratory and resident blackcap populations also experience different selective environmental pressures that could influence recombination variation among populations.

Previous studies reported both an increase and decrease in recombination rates with changes in temperature, solar radiation, and precipitation (Dreissig et al., 2019; Lloyd et al., 2018; Zhang et al., 2017). Populations studied here are also subjected to different environmental pressures, particularly between the island and continental populations, as well as broad latitudinal changes with varying temperatures and precipitation levels (Cropper 2013). Furthermore, migrant populations encounter different biotic and abiotic conditions throughout their migratory journey and may have different

adaptation strategies to adjust to different habitats. Potentially, many adaptations to changing environments might also be influencing recombination rates and driving recombination differentiation between continental migrants and resident populations.

Putative regions under linked selection: Conserved regions in low recombination associated with reduced nucleotide diversity and potential signs of selection

Our results corroborate that a strong relationship exists between nucleotide diversity and recombination rate across the genome, as has been shown for other bird species (Kawakami et al., 2017; Mugal et al., 2013; Singhal et al., 2015), as well as a broad range of other taxa (reviewed in Cutter & Payseur, 2013; McGaugh et al., 2012; Smukowski & Noor, 2011). This result hints at linked selection, indicating that the effects of selection on linked neutral (or weakly selected) diversity are strong in the blackcap. This is particularly evident in conserved regions across all populations, where low recombination rates coincide with low diversity. As recombination events could be undetectable due to the lack of suitable markers (Hudson & Kaplan, 1985; Stephens, 1986), we cannot exclude that these regions may represent regions where historical recombination events are masked rather than genuinely absent. However, given the high quality of our dataset and the remarkably consistent pattern across all populations, they may be indicative of regions with strong linked selection. Some of the regions of low recombination could also represent regions under background selection (Burri et al., 2015; Chase & Mugal, 2022; McGaugh et al., 2012, reviewed in Cutter & Payseur, 2013). Signatures of background selection would show a pattern of reduced recombination and reduced diversity associated with a considerable density of genes (Cutter & Payseur, 2013; Talla et al., 2019; Vijay et al., 2016). Several identified regions match this consistent pattern, e.g., chromosomes 20 and 29. It could be the case that recombination in regions with a high gene density is disfavored. Thus, a majority of recombination events reshuffling these alleles may be deleterious and thus rapidly purged from the population (reviewed in Dapper & Payseur, 2017; Otto et al., 1994). Distortion in the site-frequency spectrum (SFS), measured here with Tajima's D, can be identified at these regions where recombination and nucleotide density are low, patterns that are indicative of background selection and selective sweeps (Campos et al., 2014; Cvijović et al., 2018; Kim, 2006). However, it is essential to point out that Tajima's D can be influenced by recombination and linkage disequilibrium (Thornton, 2005), and thus interpretations should be treated with caution.

Moreover, recurrent hard sweeps and putative chromosomal inversions would yield a similar pattern of low recombination rates coinciding with reduced diversity (Andolfatto, 2001; Sanchez-Donoso et al., 2022). Here, we would expect a drop in Tajima's D (skewing the SFS towards rare variants over neutral expectation) to signal selective sweeps (Kim, 2006), a pattern that we observe, for example, in chromosomes 20 and 22 and 26. Thus, even though we cannot fully disentangle whether low nucleotide diversity is a side-effect, a cause, or a consequence of low recombination, these loci are relevant for further investigation. After all, these regions may indicate long-term background selection, selective sweeps, and harbours of structural variants.

Recombination rates association with genetic differentiation (Fst) and absolute divergence (dxy)

Although genetic differentiation between blackcap populations is generally low, we identified regions where Fst outliers coincide with low recombining regions. This pattern has previously been reported for other bird species , as well as other taxa (Burri, 2017 Burri et al., 2015; Henderson & Brelsford, 2020; Samuk et al., 2017). Interestingly not all of the populations showed the same pattern: we also found examples of populations where Fst outliers were enriched in regions with high recombination rates, and these regions show variable recombination maps between populations. We found the divergence of recombination maps among populations (recombination dissimilarity) in association with population differentiation (Fst), which indicates that not only the suppression of recombination but also the divergence in recombination rates among populations could contribute to population differentiation. Similar results were also reported in plants such as wheat (Danguy des Déserts et al., 2021) and cocoa tree, where population differentiation decreased the percentage of hotspots overlapping among populations (Schwarzkopf et al., 2020). Additionally, chromosomal rearrangements could also cause these patterns, as has been described in Alves et al., 2014, where genetic differentiation given by a chromosomal inversion was positively correlated with dissimilarity in recombination maps.

The relation between recombination dissimilarity and Fst was most evident in island populations with generally more variable recombination landscapes and greater Fst. In continental populations, the correlation flipped to negative or not significant, mostly driven by the long-distance migrants, which again behaves differently from all other populations showing high recombination dissimilarity and very low Fst. Nevertheless, it is also important to consider that the relative measure of divergence (Fst) can be inflated in regions where absolute genetic diversity is reduced

and linked selection is strong (Burri et al., 2015; Cruickshank & Hahn, 2014; Martin et al., 2019; Noor & Bennett, 2009).

We also included absolute divergence (dxy) in our analysis and found a higher divergence between populations in HRR. This was a consistent pattern across all the populations. Together with our findings between recombination rates, nucleotide diversity, and Fst, this suggests that our blackcap populations may exemplify the model where genetic variation is affected by linked selection and recombination variation proposed by Nachman and Payseur (2012). Even though gene flow is a variable that was not considered in this study, previous work reported limited (or not detectable) gene flow among blackcap populations (Delmore et al., 2020). The linked selection model (Nachman & Payseur, 2012) was also reported in *Ficedula* flycatchers, another passerine species (Burri et al. 2015) where the heterogeneous landscape of differentiation evolves mainly as the result of background selection and selective sweeps in genomic regions of low recombination. Similarly, our results may lead to the identification of important putative genomic regions under background selection, selective sweeps, and regions with divergent recombination rates that may contribute to blackcap populations' differentiation and evolution.

In this study, we deciphered recombination rate variation from genome-wide to local chromosomal scale between blackcap populations across a wide geographical distribution, including island and continental populations and comprising a wide range of migratory phenotypes. Despite the generally assumed conserved recombination rates in birds and low differentiation between populations, which we also recovered on the broader scale, we detect striking differences in recombination rates at the fine scale between different populations of the same species. The interplay between recombination rate variation, nucleotide diversity, and population divergence suggests that linked selection may potentially contribute to the evolution of blackcap populations differentiation.

The correlation of population differentiation with reduced recombination rates was confirmed and recombination landscapes in blackcap populations were shown to vary with genetic divergence between populations. Our findings support the idea of recombination maps as a variable phenotype and highlight the importance of including them in population genomic analyses.

Acknowledgments

We would like to thank Julien Dutheil for helpful discussions and suggestions and all members of the MPRG Behavioural Genomics, the MPRG Biological Clocks, and the Department of Evolutionary Genetics at the MPI for Evolutionary Biology for constructive discussions during seminar presentations.

This work was supported by the German A*cademic* Exchange Service (DAAD scholarship to KBC, grant 57381412), the Max Planck Society (MPRG grant MFFALIMN0001 to ML), and the DFG (project Z02 within SFB 1372 – Magnetoreception and Navigation in Vertebrates to ML).

Data accessibility

The primary and alternate haplotype assemblies for the Eurasian blackcap can be found under NCBI BioProject PRJNA558064, accession numbers GCA_009819655.1 and GCA_009819715.1 (Ishigohoka et al. 2021). The final dataset and all scripts used for the analyses will be uploaded to Dryad upon acceptance of the manuscript.

Author Contributions

ML, LOH and KBC designed the study. KBC carried out population-specific recombination rate estimates, respective downstream analyses with population features and simulated data with support from LOH and ML. JI inferred population-specific demography and performed simulation analysis. The VCF file was generated and summary statistics estimated by AB. KBC wrote the manuscript with input from LOH and ML. All authors provided comments and feedback on the manuscript.

4.6 References

Alves, J. M., Chikhi, L., Amorim, A., & Lopes, A. M. (2014). The 8p23 Inversion Polymorphism Determines Local Recombination Heterogeneity across Human Populations. *Genome Biology and Evolution*, *6*(4), 921–930. https://doi.org/10.1093/gbe/evu064

Andolfatto, P. (2001). Adaptive hitchhiking effects on genome variability. *Current Opinion in Genetics & Development*, *11*(6), 635–641. https://doi.org/10.1016/S0959-437X(00)00246-X

Backström, N., Forstmeier, W., Schielzeth, H., Mellenius, H., Nam, K., Bolund, E., Webster, M. T., Öst, T., Schneider, M., Kempenaers, B., & Ellegren, H. (2010). The recombination landscape of the zebra finch Taeniopygia guttata genome. *Genome Research*, *20*(4), 485–495. https://doi.org/10.1101/gr.101410.109

Bascón-Cardozo, K., Bours, A., Manthey, G., Pruisscher, P., Durieux, G., Dutheil, J., Odenthal-Hesse, L., & Liedvogel, M. (2022). *Fine-scale map reveals highly variable recombination rates associated with genomic features in the European blackcap* [Preprint]. Preprints. https://doi.org/10.22541/au.165423614.49331155/v1

Baumdicker, F., Bisschop, G., Goldstein, D., Gower, G., Ragsdale, A. P., Tsambos, G., Zhu, S., Eldon, B., Ellerman, E. C., Galloway, J. G., Gladstein, A. L., Gorjanc, G., Guo, B., Jeffery, B., Kretzschumar, W. W., Lohse, K., Matschiner, M., Nelson, D., Pope, N. S., … Kelleher, J. (2022). Efficient ancestry and mutation simulation with msprime 1.0. *Genetics*, *220*(3), iyab229. https://doi.org/10.1093/genetics/iyab229

Berthold, P., Helbig, A. J., Mohr, G., & Querner, U. (1992). Rapid microevolution of migratory behaviour in a wild bird species. *Nature*, *360*(6405), 6405. https://doi.org/10.1038/360668a0

Betancourt, A. J., Welch, J. J., & Charlesworth, B. (2009). Reduced Effectiveness of Selection Caused by a Lack of Recombination. *Current Biology*, *19*(8), 655–660. https://doi.org/10.1016/j.cub.2009.02.039

Brand, C. L., Cattani, M. V., Kingan, S. B., Landeen, E. L., & Presgraves, D. C. (2018). Molecular Evolution at a Meiosis Gene Mediates Species Differences in the Rate and Patterning of Recombination. *Current Biology*, *28*(8), 1289-1295.e4. Scopus. https://doi.org/10.1016/j.cub.2018.02.056

Burri, R. (2017). Interpreting differentiation landscapes in the light of long-term linked selection. *Evolution Letters*, *1*(3), 118–131. https://doi.org/10.1002/evl3.14

Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S.-A., Sætre, G.-P., Török, J., & Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Research*, *25*(11), 1656–1665. https://doi.org/10.1101/gr.196485.115

Campbell, C. L., Bhérer, C., Morrow, B. E., Boyko, A. R., & Auton, A. (2016). A Pedigree-Based Map of Recombination in the Domestic Dog Genome. *G3 Genes|Genomes|Genetics*, *6*(11), 3517–3524. https://doi.org/10.1534/g3.116.034678

Campos, J. L., Halligan, D. L., Haddrill, P. R., & Charlesworth, B. (2014). The Relation between Recombination Rate and Patterns of Molecular Evolution and Variation in Drosophila melanogaster. *Molecular Biology and Evolution*, *31*(4), 1010–1028. https://doi.org/10.1093/molbev/msu056

Castellano, D., Eyre-Walker, A., & Munch, K. (2020). Impact of Mutation Rate and Selection at Linked Sites on DNA Variation across the Genomes of Humans and Other Homininae. *Genome Biology and Evolution*, *12*(1), 3550–3561. https://doi.org/10.1093/gbe/evz215

Charlesworth, B. (2009). Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*(3), 195–205. https://doi.org/10.1038/nrg2526

Charlesworth, B. (2018). Evolution: Increased Recombination Caused by a Single Gene. *Current Biology*, *28*(8), R342–R344. https://doi.org/10.1016/j.cub.2018.02.072

Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, *134*(4), 1289–1303. https://doi.org/10.1093/genetics/134.4.1289

Charlesworth, D. (2017). Evolution of recombination rates between sex chromosomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160456. https://doi.org/10.1098/rstb.2016.0456

Chase, M. A., & Mugal, C. F. (2022). *The role of recombination dynamics in shaping signatures of direct and indirect selection across the* Ficedula *flycatcher genome* [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2022.08.11.503468

Cropper, T. (2013). The weather and climate of Macaronesia: Past, present and future. *Weather*, *68*(11), 300–307. https://doi.org/10.1002/wea.2155

Crow, J. F., & Kimura, M. (1965). Evolution in Sexual and Asexual Populations. *The American Naturalist*. https://doi.org/10.1086/282389

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*(13), 3133–3157. https://doi.org/10.1111/mec.12796

Cutter, A. D., & Payseur, B. A. (2013). Genomic signatures of selection at linked sites: Unifying the disparity among species. *Nature Reviews Genetics*, *14*(4), 4. https://doi.org/10.1038/nrg3425

Cvijović, I., Good, B. H., & Desai, M. M. (2018). The Effect of Strong Purifying Selection on Genetic Diversity. *Genetics*, *209*(4), 1235–1278. https://doi.org/10.1534/genetics.118.301058

da Silva, V. H., Laine, V. N., Bosse, M., Spurgin, L. G., Derks, M. F. L., van Oers, K., Dibbits, B., Slate, J., Crooijmans, R. P. M. A., Visser, M. E., & Groenen, M. A. M. (2019). The Genomic Complexity of a Large Inversion in Great Tits. *Genome Biology and Evolution*, *11*(7), 1870–1881. https://doi.org/10.1093/gbe/evz106

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danguy des Déserts, A., Bouchet, S., Sourdille, P., & Servin, B. (2021). Evolution of recombination landscapes in diverging populations of bread wheat. *Genome Biology and Evolution*, evab152. https://doi.org/10.1093/gbe/evab152

Dapper, A. L., & Payseur, B. A. (2017). Connecting theory and data to understand recombination rate evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160469. https://doi.org/10.1098/rstb.2016.0469

Delaneau, O., Zagury, J.-F., & Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nature Methods*, *10*(1), 1. https://doi.org/10.1038/nmeth.2307

Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Lugo Ramos, J. S., Durieux, G., Ishigohoka, J., & Liedvogel, M. (2020). The evolutionary history and genomics of European blackcap migration. *ELife*, *9*, e54462. https://doi.org/10.7554/eLife.54462

Dreissig, S., Mascher, M., & Heckmann, S. (2019). Variation in Recombination Rate Is Shaped by Domestication and Environmental Conditions in Barley. *Molecular Biology and Evolution*, *36*(9), 2029–2039. https://doi.org/10.1093/molbev/msz141

Felsenstein, J. (1974). THE EVOLUTIONARY ADVANTAGE OF RECOMBINATION. *Genetics*, *78*(2), 737–756. https://doi.org/10.1093/genetics/78.2.737

Gravel, S. (2016). When Is Selection Effective? *Genetics*, *203*(1), 451–462. https://doi.org/10.1534/genetics.115.184630

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*(1), 9–12. https://doi.org/10.1007/BF00172133

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

Hickey, D. A., & Golding, G. B. (2018). The advantage of recombination when selection is acting at many genetic Loci. *Journal of Theoretical Biology*, *442*, 123–128. https://doi.org/10.1016/j.jtbi.2018.01.018

Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetics Research*, *8*(3), 269–294. https://doi.org/10.1017/S0016672300010156

Hooper, D. M., & Price, T. D. (2017). Chromosomal inversion differences correlate with range overlap in passerine birds. *Nature Ecology & Evolution*, *1*(10), 1526–1534. https://doi.org/10.1038/s41559-017-0284-6

Hudson, R. R., & Kaplan, N. L. (1985). STATISTICAL PROPERTIES OF THE NUMBER OF RECOMBINATION EVENTS IN THE HISTORY OF A SAMPLE OF DNA SEQUENCES. *Genetics*, *111*(1), 147–164. https://doi.org/10.1093/genetics/111.1.147

Ishigohoka, J., Bascón-Cardozo, K., Bours, A., Fuß, J., Rhie, A., Mountcastle, J., Haase, B., Chow, W., Collins, J., Howe, K., Uliano-Silva, M., Fedrigo, O., Jarvis, E. D., Pérez-Tris, J., Illera, J. C., & Liedvogel, M. (2021). *Recombination suppression and selection affect local ancestries in genomes of a migratory songbird* [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2021.12.22.473882

Kamm, J. A., Spence, J. P., Chan, J., & Song, Y. S. (2016). Two-Locus Likelihoods Under Variable Population Size and Fine-Scale Recombination Rate Estimation. *Genetics*, *203*(3), 1381–1399. https://doi.org/10.1534/genetics.115.184820

Kawakami, T., Mugal, C. F., 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. *Molecular Ecology*, *26*(16), 4158–4172. https://doi.org/10.1111/mec.14197

Keightley, P. D., & Otto, S. P. (2006). Interference among deleterious mutations favours sex and recombination in finite populations. *Nature*, *443*(7107), 89–92. https://doi.org/10.1038/nature05049

Kim, S. (2015). ppcor: An R Package for a Fast Calculation to Semi-partial Correlation Coefficients. *Communications for Statistical Applications and Methods*, *22*(6), 665–674. https://doi.org/10.5351/CSAM.2015.22.6.665

Kim, Y. (2006). Allele Frequency Distribution Under Recurrent Selective Sweeps. *Genetics*, *172*(3), 1967–1978. https://doi.org/10.1534/genetics.105.048447

Lam, I., & Keeney, S. (2015). Non-paradoxical evolutionary stability of the recombination initiation landscape in yeast. *Science (New York, N.Y.)*, *350*(6263), 932–937. https://doi.org/10.1126/science.aad0814

Lloyd, A., Morgan, C., H. Franklin, F. C., & Bomblies, K. (2018). Plasticity of Meiotic Recombination Rates in Response to Temperature in Arabidopsis. *Genetics*, *208*(4), 1409–1420. https://doi.org/10.1534/genetics.117.300588

Martin, S. H., Davey, J. W., Salazar, C., & Jiggins, C. D. (2019). Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLOS Biology*, *17*(2), e2006288. https://doi.org/10.1371/journal.pbio.2006288

Martin, S. H., Singh, K. S., Gordon, I. J., Omufwoko, K. S., Collins, S., Warren, I. A., Munby, H., Brattström, O., Traut, W., Martins, D. J., Smith, D. A. S., Jiggins, C. D., Bass, C., & ffrench-Constant, R. H. (2020). Whole-chromosome hitchhiking driven by a male-killing endosymbiont. *PLOS Biology*, *18*(2), e3000610. https://doi.org/10.1371/journal.pbio.3000610

McGaugh, S., Smukowski, C., Manzano-Winkler, B., Himmel, T., & Noor, M. (2012). Recombination modulates how selection affects linked sites in Drosophila. *Nature Precedings*, 1–1. https://doi.org/10.1038/npre.2012.7005.1

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Merlin, C., & Liedvogel, M. (2019). The genetics and epigenetics of animal migration and orientation: Birds, butterflies and beyond. *Journal of Experimental Biology*, *222*(Suppl_1), jeb191890. https://doi.org/10.1242/jeb.191890

Morgan, A. P., Gatti, D. M., Najarian, M. L., Keane, T. M., Galante, R. J., Pack, A. I., Mott, R., Churchill, G. A., & de Villena, F. P.-M. (2017). Structural Variation Shapes the Landscape of Recombination in Mouse. *Genetics*, *206*(2), 603–619. https://doi.org/10.1534/genetics.116.197988

Mugal, C. F., Nabholz, B., & Ellegren, H. (2013). Genome-wide analysis in chicken reveals that local levels of genetic diversity are mainly governed by the rate of recombination. *BMC Genomics*, *14*(1), 86. https://doi.org/10.1186/1471-2164-14-86

Myers, S., Bottolo, L., Freeman, C., McVean, G., & Donnelly, P. (2005). A Fine-Scale Map of Recombination Rates and Hotspots Across the Human Genome. *Science*, *310*(5746), 321–324. https://doi.org/10.1126/science.1117196

Nachman, M. W., & Payseur, B. A. (2012). Recombination rate variation and speciation: Theoretical predictions and empirical results from rabbits and mice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1587), 409–421. https://doi.org/10.1098/rstb.2011.0249

Nelson, J. T., Cornejo, O. E., & Ag1000G Consortium. (2021). *Evolutionary implications of recombination differences across diverging populations of* Anopheles [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2021.02.04.429659

Noor, M. a. F., & Bennett, S. M. (2009). Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity*, *103*(6), 6. https://doi.org/10.1038/hdy.2009.151

Ohta, T. (2013). Molecular Evolution: Nearly Neutral Theory. In *ELS*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0001801.pub4

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., ... & Maintainer, H. W. (2019). Vegan: community ecology package. Ordination methods, diversity analysis and other functions for community and vegetation ecologists, 05–26. Version 2.5-1.

Otto, S. P., Feldman, M. W., & Christiansen, F. B. (1994). Some Advantages and Disadvantages of Recombination. In S. A. Levin (Ed.), *Frontiers in Mathematical Biology* (pp. 198–211). Springer. https://doi.org/10.1007/978-3-642-50124-1_12

Peñalba, J. V., Deng, Y., Fang, Q., Joseph, L., Moritz, C., & Cockburn, A. (2020). Genome of an iconic Australian bird: High-quality assembly and linkage map of the superb fairy-wren (Malurus cyaneus). *Molecular Ecology Resources*, *20*(2), 560–578. https://doi.org/10.1111/1755-0998.13124

Peñalba, J. V., & Wolf, J. B. W. (2020). From molecules to populations: Appreciating and estimating recombination rate variation. *Nature Reviews Genetics*, *21*(8), 476–492. https://doi.org/10.1038/s41576-020-0240-1

Peterson, A. L., & Payseur, B. A. (2021). Sex-specific variation in the genome-wide recombination rate. *Genetics*, *217*(1), iyaa019. https://doi.org/10.1093/genetics/iyaa019

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, *30*(8), 1450–1477. https://doi.org/10.1111/jeb.13047

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria. URL https://www.R-project.org/

Samuk, K., Manzano-Winkler, B., Ritz, K. R., & Noor, M. A. F. (2020). Natural Selection Shapes Variation in Genome-wide Recombination Rate in Drosophila pseudoobscura. *Current Biology*, *30*(8), 1517-1528.e6. https://doi.org/10.1016/j.cub.2020.03.053

Samuk, K., Owens, G. L., Delmore, K. E., Miller, S. E., Rennison, D. J., & Schluter, D. (2017). Gene flow and selection interact to promote adaptive divergence in regions of low recombination. *Molecular Ecology*, *26*(17), 4378–4390. https://doi.org/10.1111/mec.14226

Sanchez-Donoso, I., Ravagni, S., Rodríguez-Teijeiro, J. D., Christmas, M. J., Huang, Y., Maldonado-Linares, A., Puigcerver, M., Jiménez-Blasco, I., Andrade, P., Gonçalves, D., Friis, G., Roig, I., Webster, M. T., Leonard, J. A., & Vilà, C. (2022). Massive genome inversion drives coexistence of divergent morphs in common quails. *Current Biology*, *32*(2), 462-469.e6. https://doi.org/10.1016/j.cub.2021.11.019

Schield, D. R., Pasquesi, G. I. M., Perry, B. W., Adams, R. H., Nikolakis, Z. L., Westfall, A. K., Orton, R. W., Meik, J. M., Mackessy, S. P., & Castoe, T. A. (2020). Snake Recombination Landscapes Are Concentrated in Functional Regions despite PRDM9. *Molecular Biology and Evolution*, *37*(5), 1272–1294. https://doi.org/10.1093/molbev/msaa003

Schumer, M., Xu, C., Powell, D. L., Durvasula, A., Skov, L., Holland, C., Blazier, J. C., Sankararaman, S., Andolfatto, P., Rosenthal, G. G., & Przeworski, M. (2018). Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science*, *360*(6389), 656–660. https://doi.org/10.1126/science.aar3684

Schwarzkopf, E. J., Motamayor, J. C., & Cornejo, O. E. (2020). Genetic differentiation and intrinsic genomic features explain variation in recombination hotspots among cocoa tree populations. *BMC Genomics*, *21*(1), 332. https://doi.org/10.1186/s12864-020-6746-2

Shanfelter, A. F., Archambeault, S. L., & White, M. A. (2019). Divergent Fine-Scale Recombination Landscapes between a Freshwater and Marine Population of Threespine Stickleback Fish. *Genome Biology and Evolution*, *11*(6), 1552–1572. https://doi.org/10.1093/gbe/evz090

Singhal, S., Leffler, E. M., Sannareddy, K., Turner, I., Venn, O., Hooper, D. M., Strand, A. I., Li, Q., Raney, B., Balakrishnan, C. N., Griffith, S. C., McVean, G., & Przeworski, M. (2015). Stable recombination hotspots in birds. *Science (New York, N.Y.)*, *350*(6263), 928–932. https://doi.org/10.1126/science.aad0843

Smeds, L., Mugal, C. F., Qvarnström, A., & Ellegren, H. (2016). High-Resolution Mapping of Crossover and Non-crossover Recombination Events by Whole-Genome Re-sequencing of an Avian Pedigree. *PLOS Genetics*, *12*(5), e1006044. https://doi.org/10.1371/journal.pgen.1006044

Smeds, L., Qvarnström, A., & Ellegren, H. (2016). Direct estimate of the rate of germline mutation in a bird. *Genome Research*, *26*(9), 1211–1218. https://doi.org/10.1101/gr.204669.116

Smukowski, C. S., & Noor, M. a. F. (2011). Recombination rate variation in closely related species. *Heredity*, *107*(6), 496–508. https://doi.org/10.1038/hdy.2011.44

Speidel, L., Forest, M., Shi, S., & Myers, S. R. (2019). A method for genome-wide genealogy estimation for thousands of samples. *Nature Genetics*, *51*(9), 9. https://doi.org/10.1038/s41588-019-0484-x

Spence, J. P., & Song, Y. S. (n.d.). Inference and analysis of population-specific fine-scale recombination maps across 26 diverse human populations. *Science Advances*, *5*(10), eaaw9206. https://doi.org/10.1126/sciadv.aaw9206

Stapley, J., Birkhead, T. R., Burke, T., & Slate, J. (2010). Pronounced inter- and intrachromosomal variation in linkage disequilibrium across the zebra finch genome. *Genome Research*, *20*(4), 496–502. https://doi.org/10.1101/gr.102095.109

Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation in recombination frequency and distribution across eukaryotes: Patterns and processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160455. https://doi.org/10.1098/rstb.2016.0455

Stephens, J. C. (1986). ON THE FREQUENCY OF UNDETECTABLE RECOMBINATION EVENTS. *Genetics*, *112*(4), 923–926. https://doi.org/10.1093/genetics/112.4.923

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585–595. https://doi.org/10.1093/genetics/123.3.585

Talla, V., Soler, L., Kawakami, T., Dincă, V., Vila, R., Friberg, M., Wiklund, C., & Backström, N. (2019). Dissecting the Effects of Selection and Mutation on Genetic Diversity in Three Wood White (Leptidea) Butterfly Species. *Genome Biology and Evolution*, *11*(10), 2875–2886. https://doi.org/10.1093/gbe/evz212

Thornton, K. (2005). Recombination and the Properties of Tajima's D in the Context of Approximate-Likelihood Calculation. *Genetics*, *171*(4), 2143–2148. https://doi.org/10.1534/genetics.105.043786

van Oers, K., Santure, A. W., De Cauwer, I., van Bers, N. E., Crooijmans, R. P., Sheldon, B. C., Visser, M. E., Slate, J., & Groenen, M. A. (2014). Replicated high-density genetic maps of two great tit populations reveal fine-scale genomic departures from sex-equal recombination rates. *Heredity*, *112*(3), 307–316. https://doi.org/10.1038/hdy.2013.107

Vijay, N., Bossu, C. M., Poelstra, J. W., Weissensteiner, M. H., Suh, A., Kryukov, A. P., & Wolf, J. B. W. (2016). Evolution of heterogeneous genome differentiation across multiple contact zones in a crow species complex. *Nature Communications*, *7*(1), 1. https://doi.org/10.1038/ncomms13195

Völker, M., Backström, N., Skinner, B. M., Langley, E. J., Bunzey, S. K., Ellegren, H., & Griffin, D. K. (2010). Copy number variation, chromosome rearrangement, and their association with recombination during avian evolution. *Genome Research*, *20*(4), 503–511. https://doi.org/10.1101/gr.103663.109

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer.

Wright, S. (1931). Evolution in Mendelian Populations. *Genetics*, *16*(2), 97–159.

Zhang, K., Wu, X.-C., Zheng, D.-Q., & Petes, T. D. (2017). Effects of Temperature on the Meiotic Recombination Landscape of the Yeast Saccharomyces cerevisiae. *MBio*, *8*(6), e02099-17. https://doi.org/10.1128/mBio.02099-17

4.7 Supplementary material

Table S1: Blackcap populations. The number of samples, polymorphic sites (SNPs), recombination average, and standard deviation for all the populations.

phenotype	population	abbreviation & color code	number samples	SNPs_VCF	Rec mean	Rec mean_sd
resident island	Azores	AZO	13	9068568	0,9	0,9
resident island	Canary Islands	CAN	10	17279079	2,6	3,4
resident island	Cape Verde	CAVE	11	14537013	1,5	1,7
resident island	Madeira	MAD	13	12752338	2,6	3,5
resident island	Mallorca	MAL	10	8488456	1,3	1,7
resident island	Crete	CRE	9	10109173	2,4	4,1
resident continent	continental resident	resident_cont	23	22393970	4,0	5,5
migrant	short distance migrant South-West	short	11	11439383	3,4	5,1
migrant	medium distance North-West	medium_NW	16	21389963	6,2	9,8
migrant	medium distance South-East	medium_SE	18	12522249	4,2	6,6
migrant	medium distance South-West	medium_SW	25	26283834	6,8	10,7
migrant	long distance South_East	long	8	3772197	0,8	1,2

Table S2. Kendall's rank correlation coefficient for the pair-wise comparison between recombination rates with nucleotide diversity (π), Tajima's D, and gene density was calculated in 100 kb windows for all the populations. The association between recombination rates and Tajima's D was measured with a partial correlation controlling for the effect of nucleotide diversity.

	Kendall's Tau		
pop	**π**	**Taj D**	**gene density**
AZO	0.59	-0.07	0.12
CAN	0.66	-0.06	0.14
CAVE	0.66	-0.08	0.12
CRE	0.63	-0.11	0.16
MAD	0.67	ns	0.15
MAL	0.61	-0.12	0.15
cont-res	0.63	0.07	0.14
short	0.61	ns	0.13
med_NW	0.58	0.08	0.15
med_SE	0.56	0.09	0.17
med_SW	0.57	0.11	0.16
long	0.44	0.06	ns

all correlations are significant $p<0.0001$,
ns= non-significant

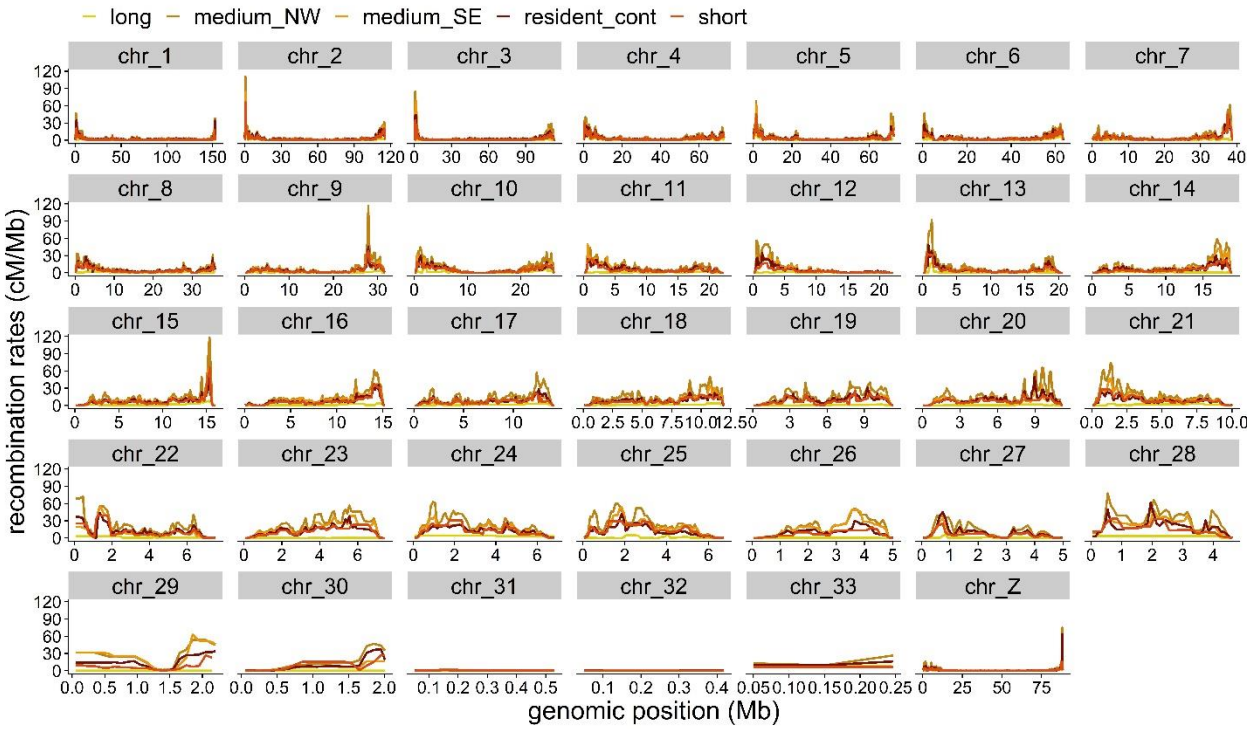


Figure S1 Recombination rates across all chromosomes for all continental populations, including migratory and resident populations. Recombination rates were calculated in 100 kb windows.

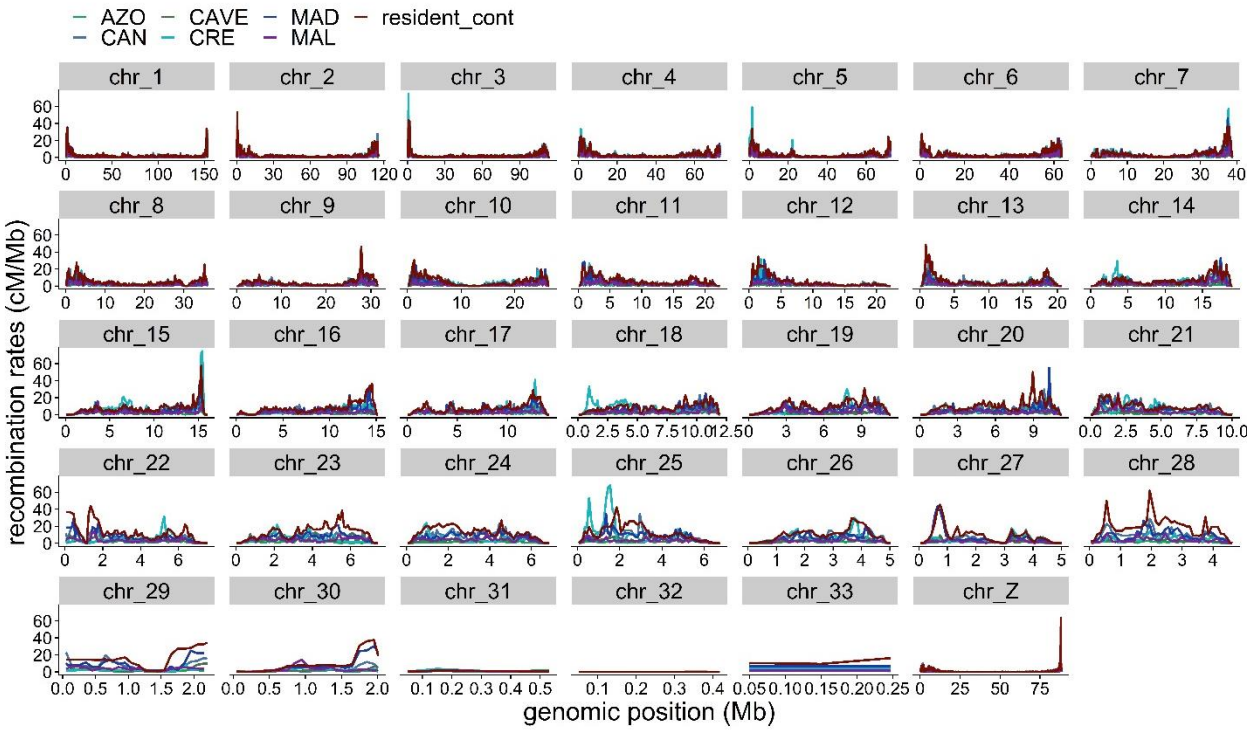
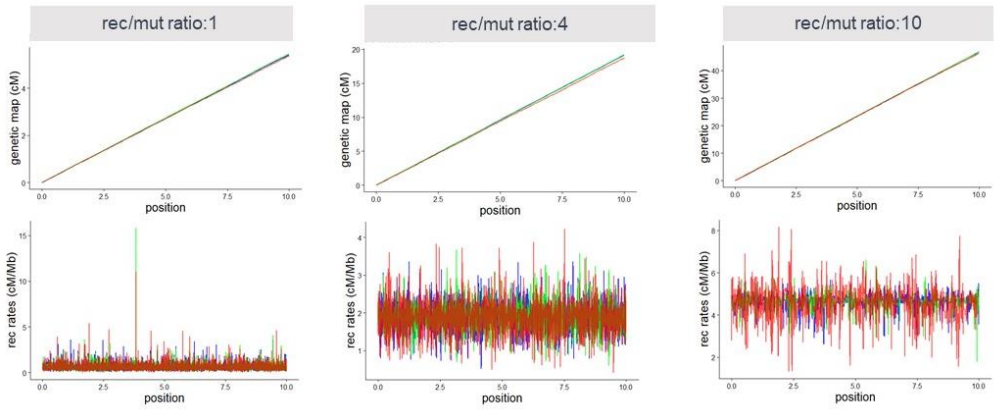


Figure S2 Recombination rates across chromosomes in resident populations, including continental residents and all island residents. Recombination rates were calculated in 100 kb windows.



Pop	gen map (cM)	rec mean (cM/Mb)
A	5.42	0.54
B	5.46	0.55
C	5.37	0.54

Pop	gen map (cM)	rec mean (cM/Mb)
A	19,15	1,92
B	19,09	1,91
C	18,71	1,89

Pop	gen map (cM)	rec mean (cM/Mb)
A	46.78	4.67
B	46.80	4.68
C	46.30	4.63

Figure S3. Genetic maps (upper panel) and recombination rates (lower panel) estimated from simulated data using msprime with recombination/mutation ratio of 1 (left), 4 (middle) and 10 (right) from three populations ABC that evolved under different demographic trajectories: Pop A (constant population size), pop B (expansion), and pop C (contraction). Tables show genetic map (in cM) and recombination rate average (in cM/Mb) for each population.

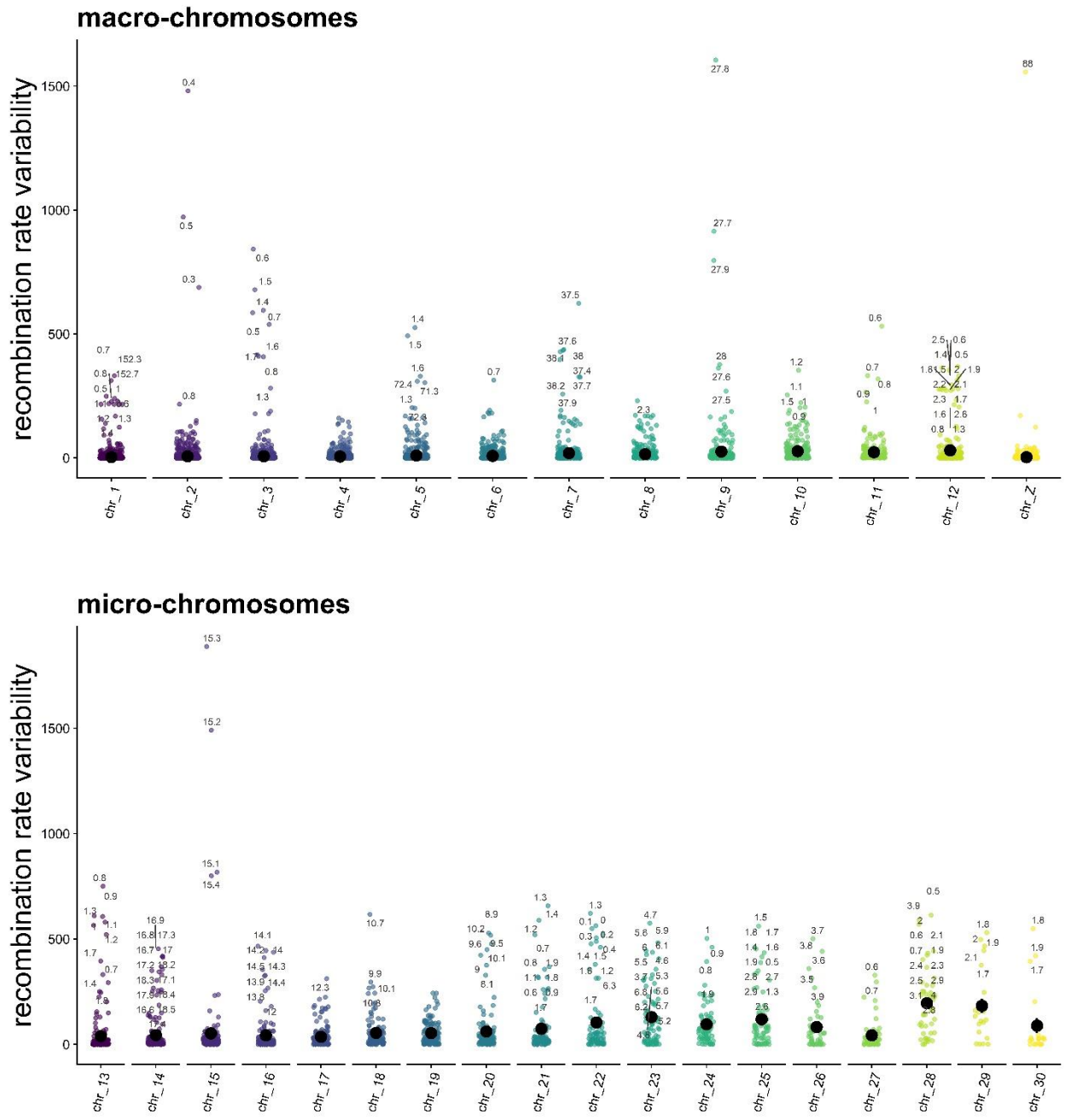


Figure S4. Dot plot and error plot comparing the variance calculated per window (100 kb), including all populations among all macro- (upper panel) and micro-chromosomes (lower panel). Each dot represents a window within the respective chromosome. The start of the windows with the highest variability (variance >250) within each chromosome is shown with the numbers in Mb.

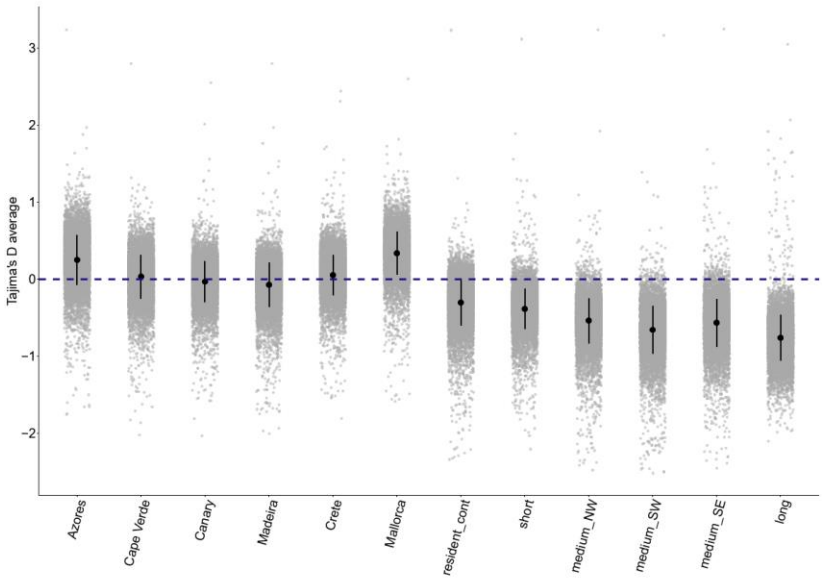


Figure S5. Dot plot and error plot comparing the genome-wide average of Tajima's'D among all populations. Tajima's D means between all populations are significantly different (Wilcoxon rank sum test, $p<0.001$) except for the comparison between medium_SE and long (both SE migrantory direction). Tajima's D was calculated in 100 kb windows, and the average was weighted by the number of sites per window.

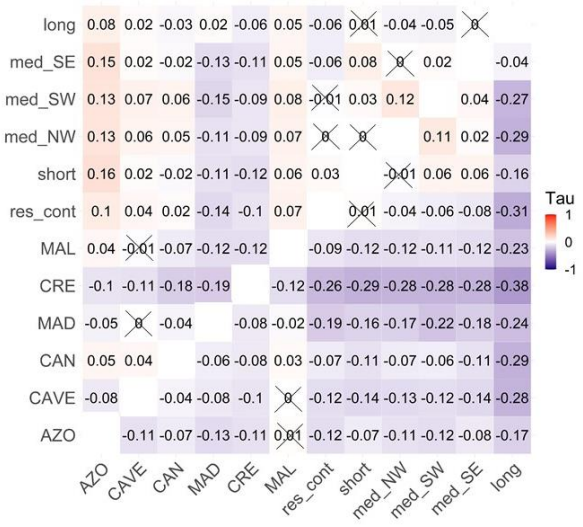


Figure S6. Heatmap showing the pairwise correlation of recombination rates and Fst calculated in 100 kb windows across all population pairs. Fst was calculated between populations of the respective row (y-axis) and column (x-axis) combination. The upper-diagonal shows the correlation coefficients for the comparison between Fst and recombination maps from populations in the y-axis. The lower-diagonal shows the correlation coefficients for the comparison between Fst and the recombination maps from populations in the x-axis. All correlations are significant $p<0.001$, except for those indicated with a cross for non-significant.

CHAPTER V

General discussion and Final Remarks

This thesis characterises the historical recombination landscape of a high-quality bird genome and the interplay of recombination variation with genomic features, including different scales. The main findings reveal that recombination rates vary across the genome, among populations, and between different lineages. Specific genomic features contribute to shape this variation and the implications of the divergent recombination landscape at broad and fine-scales may impact the evolution of genomes, populations and species.

5.1 The heterogeneous recombination rate across the blackcap genome: what is new?

This thesis initially reports a heterogeneous historical recombination landscape among and within chromosomes in the blackcap. The results confirmed that genome-wide recombination rate variation is associated with specific genomic features, such as GC content, including CpG islands and nucleotide diversity, as previous studies reported in other bird species (Kawakami et al., 2014; Peñalba et al., 2020; Singhal et al., 2015). Similarly, a negative correlation between chromosome size and recombination rates was evident. However, what is new and how does this advance our understanding on the relationship of genomic features with recombination rate variation at genome-level? The new findings that this thesis provides are the following:

1) The association of recombination rate with genes will depend not only on their density but also on their regulation. Most recombination events are prevalent in cis-regulatory regions (e.g., transcription start sites (TSS), promoters, 5'UTR regions) and not necessarily in whole gene bodies.

2) CpG islands are important determinants of recombination variation at genome-wide and local scales. The location and epigenetic regulation of CpGi may have an influence on recombination rates at local levels. In support of this, chapter II shows that increased recombination rates are prevalent in regulatory elements containing CpGi compared to the same elements without CpGi. Additionally, some of the intrachromosomal regions identified with contrasting patterns of high

CpGi densities in low recombining regions showed hypermethylated CpGi patterns (see chapter III).

3) The interplay between recombination and TEs involves different factors. Very few studies report this association in bird species (Kawakami et al., 2017). I found in chapter II that this association will differ according to the insertion sites and TEs family. LINEs appeared associated with high recombination rates, whereas SINEs and LTRs showed the contrary. In Bours et al., 2022, we found that the variation is also associated with species-specific TEs or non-species-specific TEs, which can give us a proxy for TEs age. Species-specific TEs (considered as "young" TEs), represented mainly by LTRs, are negatively correlated, whereas non-species-specific TEs ("old" TEs) did not show a significant correlation with recombination rates. These results contrast the rationale that "young" TEs could evidence a more recent epigenetic status where the accessibility to the transcription machinery is greater and thus the correlation with recombination. Whereas "old" TEs might get expanded and rich fixation, potentially repressing chromatin structure and therefore counteracting with recombination rates or contributing to recombination suppression, as reported in flycatchers (Kawakami et al., 2017). However, species-specific TEs encompass a wider age range of TEs. Further investigation is needed to disentangle whether species-specific TEs are recently active, such as identifying *de-novo* TEs with phylogenetic trees or TE insertion site sequences in combination with transcriptomic data (Goerner-Potvin & Bourque, 2018). These data could then be integrated with recombination maps, including the identification of recombination hotspots within active TEs.

The genomic characterisation at local scales provides evidence of the well-described recombination suppression in chromosomal rearrangements such as inversions (Farré et al., 2013; Fransz et al., 2016; Hooper & Price, 2017; Völker et al., 2010; Xu et al., 2019). In Ishigohoka et al., 2021, putative polymorphic inversions were identified in some chromosomes and we found that recombination was suppressed in heterozygotes. One of the polymorphic inversions was potentially associated with balancing selection (e.g., chromosome 12), exemplifying the scenario when recombination suppression is favourable to maintain DNA stretches with loci in positive LD and enriched with advantageous combinations of alleles (Otto et al., 1994). The integration of findings at the genome level allowed the identification of additional intrachromosomal regions of interest for further study. Some of these regions appear as candidates for putative centromeric regions characterised by

suppressed recombination, enrichment of long tandem repeats and TEs, and high CpG methylation patterns (see for example, chromosomes 6 and 10 in chapters II and III).

Usually, non-prdm9 organisms which follow the "default" recombination mechanism show high recombining regions enriched in genomic elements associated with open chromatin regions and/or low methylation CpGs (Baker et al., 2017; Schield et al., 2020; Singhal et al. 2015, see chapter II). In chapter III, I evidenced a high recombination rate at the intermediate stages of CpG methylation. In addition, low recombining regions were significantly associated with high CpG methylation only in micro-chromosomes, where recombination rate is usually much higher than in macro-chromosomes. Nevertheless, TSS show a deep decline in methylation in these regions. Overall, this thesis provides strong evidence that regulatory regions, with usually low CpG methylation patterns and low repetitive sequences, gather the majority of recombination events across the genome (chapter II, III, Appendix Supplementary). As a future outlook, assessing the interaction between chromatin status and recombination variation across the genome would be ideal to complement our understanding on the impact of the epigenome over the historical recombination maps in the blackcap. For example, including approaches that provide information on chromatin-accessibility and the location of open-chromatin regions across the genome (e.g. ATAC-seq, DNA-seq). These data then can be combined with the identified high or low recombining regions across the genome.

5.2 Possible implications of recombination divergence among distinct populations

Blackcaps encompass a variety of migratory phenotypes from long-distance migrants to entirely resident populations providing an excellent system to study genetics and evolution of the migratory trait. This thesis evaluated the similarity and divergence of recombination landscapes among blackcap populations and the interplay with patterns of nucleotide diversity and population differentiation, presented in chapter IV. This chapter provides strong evidence that correlation levels between recombination maps can recapitulate known demographic history and genetic structure of populations (see chapter IV). Blackcap populations split as recently as ~30.000 ya resulting in very low genetic differentiation, mostly detected between migrant and resident populations. In line with this, the greater differentiation of recombination maps was evident between the island compared to continental populations. In contrast, medium-distance migrants with distinct migratory orientations showed the highest similarity in recombination landscapes

from all population comparisons reflecting the low, and almost no detectable, genetic differentiation between them(Delmore et al., 2020; Mettler et al., 2013). Inbreeding between these populations due to limited gene flow detected is likely (Delmore et al., 2020) and may also contribute to the conservation of recombination rates.

Additionally, here we report that population differentiation is associated with the divergence of recombination landscapes. Resident island populations show more divergence recombination patterns at broad and fine scales. This may result from the smaller Ne in these populations owing to bottleneck events after the split and, thus, higher effects of genetic drift. Smaller Ne can also be related to the relaxation of selection, which can generate the loss of traits, for example, migration in the case of resident populations, since so far migration is known as the ancestral status in the blackcaps (Perez-Tris et al., 2004; Voelker & Light, 2011). Perhaps selection is stronger and more efficient in migratory populations, which require the maintenance of "optimal" recombination rates and control of specific genomic regions involved in migration, thus yielding more conserved recombination maps genome-wide. But all of this is pure speculation, and particular tests are needed such as detecting whether the differences in recombination maps between populations are driven by natural selection. For example, applying Q_{ST}-F_{ST} approach to test if the observed differences between populations in a quantitative trait (recombination rate) is greater than expected on the basis of drift alone , as applied with direct estimates of recombination in Samuk et al. (2020). Additionally, simulation to evaluate how much the recombination rate inference is affected by selection will be required to discard potential artifacts.

Whether the conservation or divergence of recombination maps in continental or island populations is adaptative is still unknown and perhaps not easy to disentangle since the time from the split is still very recent, and further analysis will be required. In addition, it is unclear whether the recombination rate has modulated the rate of adaptive evolution across the genome in blackcaps during the evolution of migration to residency or whether the evolution of migration (or residency) entailed increased or reduced recombination rates in specific genes that control this behaviour. To investigate that, an integrative approach would be ideal A starting point could be investigating deeper into the genomic windows with higher variability of recombination rates among migratory and resident populations reported in this thesis. Look if they coincide with specific transcription factors, candidate genes controlling migration, and regions with signs of positive selection identified in previous studies (Delmore et al., 2020). Another approach would be

to examine the correlation between a measure of the rate of protein evolution, such as the ratio between nonsynonymous and synonymous mutations (dN/dS), with the recombination rate specific for identified candidate regulatory regions and genes involved in migration (Delmore et al., 2020), combined with gene expression and methylation profiles.

5.3 The role of recombination in the blackcap system

After many years of empirical and theoretical research, recombination is known as an important determinant of the strength of selection and rates of adaptive evolution in populations and also within genomes in some taxa (Campos et al., 2014; Castellano et al., 2016; Grandaubert et al., 2019; Tiley & Burleigh, 2015; Webster & Hurst, 2011). Different mechanisms can lead to different evolutionary paths. The fact that recombination in non-PRDM9 organisms such as birds is prevalent in regulatory regions may suggest that recombination plays a specific role in increasing selection efficacy, maintaining diversity, and eventually contributing to adaptation. For example, recombination in yeast, a non-PRDM9 organism, was a determinant of adaptation to standing variation, allowing different strains to adapt to new environments (Kosheleva & Desai, 2018). This thesis shows that the recombination landscape is heterogeneous, and this variation, particularly at local scales, may significantly influence the evolution of genes, genomes, sex, and populations.

Recombination is linked to the efficacy of selection in two main ways: by reducing HRI, and shaping the strength of linked selection. The latter is when recombination minimizes the effect of linked selection on the reduction of genetic diversity at neutral sites linked to selection targets. In this thesis, we report two primary pieces of evidence that point to the pervasive role of recombination rate shaping signatures of linked selection (background or selective sweeps) described previously in flycatchers (Burri et al., 2015; Chase & Mugal, 2022). One is the positive and negative genome-wide association between recombination rates with nucleotide diversity and Fst, respectively (see chapters II and IV). Second, the identification of genomic islands of differentiation characterised by elevated Fst relative to the background and recombination rate reduction (see chapters IV and Ishigohoka et al., 2021). We additionally report genomic regions with potential signs of long-term linked selection in chapters IV. To know if background selection, selective sweep, or perhaps chromosomal rearrangements yield those patterns, further analysis will be required. Altogether, our results, particularly in chapters IV, point to linked selection that may significantly contribute to shaping the heterogeneous genomic landscape of diversity and

differentiation in blackcaps populations; thus, recombination is a key determinant in these processes.

Whether recombination is influencing directly the efficacy of selection and adaptation in blackcaps is still an open question, which will require further studies, including measurements as a proxy for the efficacy of selection. For example, to assess the presence of Hill-Robertson interference (HRi) by measuring nonsynonymous over synonymous ratio of diversity $\pi N/\pi S$. To quantify the mode and strength of selection by measuring the rate of protein adaptation and comparing the nonsynonymous over synonymous ratio of divergence dN/dS. Additionally, applying codon usage bias as a measure of the proportion of 'preferred' codons in a gene as predicted to be greater in regions where selection is more efficient (high recombining regions) (reviewed in Webster & Hurst, 2011).

5.4 Recombination variation between lineages

The blackcap and its closes sister species, the garden warbler, diverged ~14-16 mya (Voelker & Light, 2011). We found that the recombination landscapes from these species are conserved at genome-wide scales. Garden warblers are long-distance migrants (Bairlein, 1987; Voelker & Light, 2011). The comparison of this species with blackcap continental populations revealed a higher correlation between recombination maps, whereas the correlation with island resident populations was lower. Recombination maps between species can be difficult to compare since many factors may contribute to variations, such as the density of markers, variable levels of nucleotide diversity, demographic history, population structure (and substructure), environmental conditions, selective pressures, mutation rates, rearing conditions (reviewed in Peñalba & Wolf, 2020; Smukowski & Noor, 2011; Stapley et al., 2017). For example, finches exhibit lower recombination rates than chicken (Backström et al., 2010), probably due to domestication in chicken.

The divergence in recombination maps between garden warblers and blackcaps became more evident at the fine scale, particularly in micro-chromosomes. Some intrachromosomal regions with variable patterns were associated with specific genomic features, and perhaps these regions potentially contribute to lineage divergence. For example, a genomic region in chromosome 10 enriched with TEs showed suppressed recombination in the blackcap but not in the garden warbler, suggesting that recombination suppression is probably specific for the blackcap. Species-specific chromosomal rearrangements can also generate divergence in recombination maps among lineages (reviewed in Stapley et al., 2017), which will be worth to further investigate including the

putative inversion identified in the blackcap genome. On the other hand, conserved regions of high and reduced recombination were identified in chapter II. Perhaps some of the conserved regions between these species, which are characterised by reduced recombination and high differentiation, represent homologous differentiation islands shared among lineages. Hence, these regions could contribute to the conservation of parallel patterns of differentiation as described in other passerines (Chase et al., 2021; Kawakami et al., 2017; Van Doren et al., 2017).

5.5 Strengths and limitations of the methodological approach

The inference of recombination rates using LD-based approaches is useful to characterise the recombination landscape of populations or species, including many generations. Our methodological approach overcomes the potential biases that different demographic history between populations generates in the estimates (Dapper & Payseur, 2018; Kamm et al., 2016). Additionally, simulation analysis confirmed that our estimates are robust against population-specific demography (see chapter IV). Pyrho initially estimates rho and scales it to r using the mutation rate and Ne. However, this scaling can also generate some biases; for example, Watterson theta is assumed constant, but that is not always the case and can fluctuate across the genome. One way to alleviate that bias is to include all callable sites or rates of heterozygosity in downstream analysis. Using species-specific mutation rates will also be ideal. Moreover, it is suggested that performing haplotype phasing and genotype imputation or using pedigree information might be optimal for overcoming specific biases in recombination inference (Hassan et al., 2021). However, getting large pedigrees becomes more complex when natural populations are studied. Captivity or controlled reared conditions could be applied, but these could lead to biases due to inbreeding. Nevertheless, the inference of recombination rates using pedigree data and LD-based approaches appears significantly correlated (Kawakami et al., 2017; Shanfelter et al., 2019; Singhal et al., 2015), even when unphased genotype data is used (Spence & Song, 2019).

Another point is that LD-based approaches are marker-density dependent. Our estimations and analysis benefit from the high-quality species-specific reference genome, allowing us to call SNPs with good coverage. However, we don't discard that some differences in recombination rates between populations could be artifacts or whether low recombining regions represent the absence of recombination or undetected markers. Combining bioinformatic tools with direct estimates (e.g., sperm typing, ChIP-SSeq) can help corroborate the presence of high or low recombining regions. Direct methods distinguish between gene conversion and crossing-over, which may be beneficial

to consider for specific analysis at local scales. For example, GC-biased generated by gene conversion (gBGC) interferes with natural selection and has consequences on genome evolution (Wallberg et al., 2015). It can affect levels of diversity, the site spectrum frequency, and the rate of nucleotide substitutions which can be mistaken as signatures of selection and influence some statistics (Berglund et al., 2009; Galtier et al., 2009; Webster & Smith, 2004).

Finally, this thesis suggests some factors to consider when comparing recombination maps: i) Using similar approaches with the same parameters is relevant, ii) The scales of comparison are essential. Including broad and fine scales will generate a more comprehensive view of the similarity and divergence of recombination maps. iii) Absolute values of recombination rate should be interpreted with caution since they may be influenced by the SNP density, diversity, and the unequal number of samples of the populations.

5.6 General conclusion

The present work integrates genomic and epigenetic components to decipher the heterogeneous recombination landscape of a migratory songbird at the genome, populations, and inter-specific levels. The main findings reveal that recombination events are prevalent in functional elements where CpG methylation is typically low. Recombination rate is a variable phenotype and may play a fundamental role in adaptation, shaping genomic diversity, and population differentiation of blackcaps. Additionally, the pieces of evidence presented here hint at recombination as an essential factor influencing the efficacy of linked selection. This thesis contributes to broadening our knowledge on the role of recombination in population genetics of a bird species. This is one of the starting steps to eventually unrevealing the underlying genetic variation behind complex traits such as migration.

5.7 References

Backström, N., Forstmeier, W., Schielzeth, H., Mellenius, H., Nam, K., Bolund, E., Webster, M. T., Öst, T., Schneider, M., Kempenaers, B., & Ellegren, H. (2010). The recombination landscape of the zebra finch Taeniopygia guttata genome. *Genome Research*, *20*(4), 485–495. https://doi.org/10.1101/gr.101410.109

Bairlein, F. (1987). The migratory strategy of the Garden Warbler: A survey of field and laboratory data. *Ringing & Migration*, *8*(2), 59–72. https://doi.org/10.1080/03078698.1987.9673903

Baker, Z., Schumer, M., Haba, Y., Bashkirova, L., Holland, C., Rosenthal, G. G., & Przeworski, M. (2017). Repeated losses of PRDM9-directed recombination despite the conservation of PRDM9 across vertebrates. *ELife*, *6*, e24133. https://doi.org/10.7554/eLife.24133

Berglund, J., Pollard, K. S., & Webster, M. T. (2009). Hotspots of Biased Nucleotide Substitutions in Human Genes. *PLoS Biology*, *7*(1), e1000026. https://doi.org/10.1371/journal.pbio.1000026

Bours, A., Pruisscher, P., Bascón-Cardozo, K., Odenthal-Hesse, L., & Liedvogel, M. (2022). *The blackcap (Sylvia atricapilla) genome reveals a species-specific accumulation of LTR retrotransposons* [Preprint]. In Review. https://doi.org/10.21203/rs.3.rs-2166848/v1

Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., Suh, A., Dutoit, L., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S.-A., Sætre, G.-P., Török, J., & Ellegren, H. (2015). Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Research*, *25*(11), 1656–1665. https://doi.org/10.1101/gr.196485.115

Campos, J. L., Halligan, D. L., Haddrill, P. R., & Charlesworth, B. (2014). The Relation between Recombination Rate and Patterns of Molecular Evolution and Variation in Drosophila melanogaster. *Molecular Biology and Evolution*, *31*(4), 1010–1028. https://doi.org/10.1093/molbev/msu056

Castellano, D., Coronado-Zamora, M., Campos, J. L., Barbadilla, A., & Eyre-Walker, A. (2016). Adaptive Evolution Is Substantially Impeded by Hill–Robertson Interference in Drosophila. *Molecular Biology and Evolution*, *33*(2), 442–455. https://doi.org/10.1093/molbev/msv236

Chase, M. A., Ellegren, H., & Mugal, C. F. (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*(9), 2179–2196. https://doi.org/10.1111/evo.14234

Chase, M. A., & Mugal, C. F. (2022). *The role of recombination dynamics in shaping signatures of direct and indirect selection across the* Ficedula *flycatcher genome* [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2022.08.11.503468

Dapper, A. L., & Payseur, B. A. (2018). Effects of Demographic History on the Detection of Recombination Hotspots from Linkage Disequilibrium. *Molecular Biology and Evolution*, *35*(2), 335–353. https://doi.org/10.1093/molbev/msx272

Delmore, K., Illera, J. C., Pérez-Tris, J., Segelbacher, G., Lugo Ramos, J. S., Durieux, G., Ishigohoka, J., & Liedvogel, M. (2020). The evolutionary history and genomics of European blackcap migration. *ELife*, *9*, e54462. https://doi.org/10.7554/eLife.54462

Farré, M., Micheletti, D., & Ruiz-Herrera, A. (2013). Recombination Rates and Genomic Shuffling in Human and Chimpanzee—A New Twist in the Chromosomal Speciation Theory. *Molecular Biology and Evolution*, *30*(4), 853–864. https://doi.org/10.1093/molbev/mss272

Fransz, P., Linc, G., Lee, C.-R., Aflitos, S. A., Lasky, J. R., Toomajian, C., Ali, H., Peters, J., van Dam, P., Ji, X., Kuzak, M., Gerats, T., Schubert, I., Schneeberger, K., Colot, V., Martienssen, R., Koornneef,

M., Nordborg, M., Juenger, T. E., … Schranz, M. E. (2016). Molecular, genetic and evolutionary analysis of a paracentric inversion in Arabidopsis thaliana. *The Plant Journal*, *88*(2), 159–178. https://doi.org/10.1111/tpj.13262

Galtier, N., Duret, L., Glémin, S., & Ranwez, V. (2009). GC-biased gene conversion promotes the fixation of deleterious amino acid changes in primates. *Trends in Genetics*, *25*(1), 1–5. https://doi.org/10.1016/j.tig.2008.10.011

Goerner-Potvin, P., & Bourque, G. (2018). Computational tools to unmask transposable elements. *Nature Reviews Genetics*, *19*(11), 11. https://doi.org/10.1038/s41576-018-0050-x

Grandaubert, J., Dutheil, J. Y., & Stukenbrock, E. H. (2019). The genomic determinants of adaptive evolution in a fungal pathogen. *Evolution Letters*, *3*(3), 299–312. https://doi.org/10.1002/evl3.117

Hassan, S., Surakka, I., Taskinen, M.-R., Salomaa, V., Palotie, A., Wessman, M., Tukiainen, T., Pirinen, M., Palta, P., & Ripatti, S. (2021). High-resolution population-specific recombination rates and their effect on phasing and genotype imputation. *European Journal of Human Genetics*, *29*(4), 4. https://doi.org/10.1038/s41431-020-00768-8

Hooper, D. M., & Price, T. D. (2017). Chromosomal inversion differences correlate with range overlap in passerine birds. *Nature Ecology & Evolution*, *1*(10), 1526–1534. https://doi.org/10.1038/s41559-017-0284-6

Ishigohoka, J., Bascón-Cardozo, K., Bours, A., Fuß, J., Rhie, A., Mountcastle, J., Haase, B., Chow, W., Collins, J., Howe, K., Uliano-Silva, M., Fedrigo, O., Jarvis, E. D., Pérez-Tris, J., Illera, J. C., & Liedvogel, M. (2021). *Recombination suppression and selection affect local ancestries in genomes of a migratory songbird* [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2021.12.22.473882

Kamm, J. A., Spence, J. P., Chan, J., & Song, Y. S. (2016). Two-Locus Likelihoods Under Variable Population Size and Fine-Scale Recombination Rate Estimation. *Genetics*, *203*(3), 1381–1399. https://doi.org/10.1534/genetics.115.184820

Kawakami, T., Mugal, C. F., 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. *Molecular Ecology*, *26*(16), 4158–4172. https://doi.org/10.1111/mec.14197

Kawakami, T., Smeds, L., Backström, N., Husby, A., Qvarnström, A., Mugal, C. F., 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. *Molecular Ecology*, *23*(16), 4035–4058. https://doi.org/10.1111/mec.12810

Kosheleva, K., & Desai, M. M. (2018). Recombination Alters the Dynamics of Adaptation on Standing Variation in Laboratory Yeast Populations. *Molecular Biology and Evolution*, *35*(1), 180–201. https://doi.org/10.1093/molbev/msx278

Mettler, R., Schaefer, H. M., Chernetsov, N., Fiedler, W., Hobson, K. A., Ilieva, M., Imhof, E., Johnsen, A., Renner, S. C., Rolshausen, G., Serrano, D., Wesołowski, T., & Segelbacher, G. (2013). Contrasting Patterns of Genetic Differentiation among Blackcaps (Sylvia atricapilla) with Divergent Migratory Orientations in Europe. *PLOS ONE*, *8*(11), e81365. https://doi.org/10.1371/journal.pone.0081365

Otto, S. P., Feldman, M. W., & Christiansen, F. B. (1994). Some Advantages and Disadvantages of Recombination. In S. A. Levin (Ed.), *Frontiers in Mathematical Biology* (pp. 198–211). Springer. https://doi.org/10.1007/978-3-642-50124-1_12

Peñalba, J. V., Deng, Y., Fang, Q., Joseph, L., Moritz, C., & Cockburn, A. (2020). Genome of an iconic Australian bird: High-quality assembly and linkage map of the superb fairy-wren (Malurus cyaneus). *Molecular Ecology Resources*, *20*(2), 560–578. https://doi.org/10.1111/1755-0998.13124

Peñalba, J. V., & Wolf, J. B. W. (2020). From molecules to populations: Appreciating and estimating recombination rate variation. *Nature Reviews Genetics*, *21*(8), 8. https://doi.org/10.1038/s41576-020-0240-1

Perez-Tris, J., Bensch, S., Carbonell, R., Helbig, A., & Tellería, J. L. (2004). Historical Diversification of Migration Patterns in a Passerine Bird. *Evolution*, *58*(8), 1819–1832. https://doi.org/10.1111/j.0014-3820.2004.tb00464.x

Samuk, K., Manzano-Winkler, B., Ritz, K. R., & Noor, M. A. F. (2020). Natural Selection Shapes Variation in Genome-wide Recombination Rate in Drosophila pseudoobscura. *Current Biology*, *30*(8), 1517-1528.e6. https://doi.org/10.1016/j.cub.2020.03.053

Schield, D. R., Pasquesi, G. I. M., Perry, B. W., Adams, R. H., Nikolakis, Z. L., Westfall, A. K., Orton, R. W., Meik, J. M., Mackessy, S. P., & Castoe, T. A. (2020). Snake Recombination Landscapes Are Concentrated in Functional Regions despite PRDM9. *Molecular Biology and Evolution*, *37*(5), 1272–1294. https://doi.org/10.1093/molbev/msaa003

Shanfelter, A. F., Archambeault, S. L., & White, M. A. (2019). Divergent Fine-Scale Recombination Landscapes between a Freshwater and Marine Population of Threespine Stickleback Fish. *Genome Biology and Evolution*, *11*(6), 1552–1572. https://doi.org/10.1093/gbe/evz090

Singhal, S., Leffler, E. M., Sannareddy, K., Turner, I., Venn, O., Hooper, D. M., Strand, A. I., Li, Q., Raney, B., Balakrishnan, C. N., Griffith, S. C., McVean, G., & Przeworski, M. (2015). Stable recombination hotspots in birds. *Science (New York, N.Y.)*, *350*(6263), 928–932. https://doi.org/10.1126/science.aad0843

Smukowski, C. S., & Noor, M. a. F. (2011). Recombination rate variation in closely related species. *Heredity*, *107*(6), 496–508. https://doi.org/10.1038/hdy.2011.44

Spence, J. P., & Song, Y. S. (2019.). Inference and analysis of population-specific fine-scale recombination maps across 26 diverse human populations. *Science Advances*, *5*(10), eaaw9206. https://doi.org/10.1126/sciadv.aaw9206

Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation in recombination frequency and distribution across eukaryotes: Patterns and processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160455. https://doi.org/10.1098/rstb.2016.0455

Tiley, G. P., & Burleigh, J. G. (2015). The relationship of recombination rate, genome structure, and patterns of molecular evolution across angiosperms. *BMC Evolutionary Biology*, *15*(1), 194. https://doi.org/10.1186/s12862-015-0473-3

Van Doren, B. M., Campagna, L., Helm, B., Illera, J. C., Lovette, I. J., & Liedvogel, M. (2017). Correlated patterns of genetic diversity and differentiation across an avian family. *Molecular Ecology*, *26*(15), 3982–3997. https://doi.org/10.1111/mec.14083

Voelker, G., & Light, J. E. (2011). Palaeoclimatic events, dispersal and migratory losses along the Afro-European axis as drivers of biogeographic distribution in Sylvia warblers. *BMC Evolutionary Biology*, *11*(1), 163. https://doi.org/10.1186/1471-2148-11-163

Völker, M., Backström, N., Skinner, B. M., Langley, E. J., Bunzey, S. K., Ellegren, H., & Griffin, D. K. (2010). Copy number variation, chromosome rearrangement, and their association with recombination during avian evolution. *Genome Research*, *20*(4), 503–511. https://doi.org/10.1101/gr.103663.109

Wallberg, A., Glémin, S., & Webster, M. T. (2015). Extreme Recombination Frequencies Shape Genome Variation and Evolution in the Honeybee, Apis mellifera. *PLOS Genetics*, *11*(4), e1005189. https://doi.org/10.1371/journal.pgen.1005189

Webster, M., & Hurst, L. (2011). Direct and indirect consequences of meiotic recombination: Implications for genome evolution. *Trends in Genetics : TIG*, *28*, 101–109. https://doi.org/10.1016/j.tig.2011.11.002

Webster, M. T., & Smith, N. G. C. (2004). Fixation biases affecting human SNPs. *Trends in Genetics*, *20*(3), 122–126. https://doi.org/10.1016/j.tig.2004.01.005

Xu, L., Auer, G., Peona, V., Suh, A., Deng, Y., Feng, S., Zhang, G., Blom, M. P. K., Christidis, L., Prost, S., Irestedt, M., & Zhou, Q. (2019). Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds. *Nature Ecology & Evolution*, *3*(5), 834–844. https://doi.org/10.1038/s41559-019-0850-1

Acknowledgements

First, I would like to thank Miriam and Linda for giving me the chance to work on this project, letting me explore new topics, for all the consistent support, encouragement, and patience. I learnt so much from you both and I am grateful for that. Thank you for always respecting my cultural background and for your empathy. Miriam, your guidance, expertise, and all your valuable advice were essential for this project and my scientific development. Linda, your support, fairness and enthusiasm were fundamental to me, thank you for sharing many of your experiences. Our meetings and conversations inspired me in so many ways.

I would also like to thank my Thesis Advisory Committee: Julien Dutheil and Oscar Puebla for the valuable input and fruitful discussions for this project. To Bernhard Haubold, Elena Damm, Kristian Ullrich, Loukas Theodosiou who also contributed to my academic growth. A special thanks to the IT members: Derk, Kristian and Martin, for the great technical support and patience. I am also grateful for all collaborators who provided samples and enabled me to analyse this dataset for my PhD project. I acknowledge my scholarship DAAD under the grant: 57381412f for funding my PhD and also to IMPRS. Angela Donner, I appreciate all your help and the nice vibe. Thank you to Eva Stukenbrock and Marc Bramkamp for accepting being part of my dissertation committee.

I am thankful with the current and past members of my research group. The allopatric phase was not easy but we learnt so much form each other. The discussions about birds, migration, and genomics were encouraging. Thanks to the Biological Clocks group, it was always nice to learn about moon cycles and the evolution of clunios, and to the Evolutionary Genetics department for the interesting discussions and presentations.

I am very grateful to my friends at the MPI, for making the PhD an easier journey. I want to thank my friends who made the first years so enjoyable. Tania gracias por tu apoyo incondicional y por siempre estar ahí cuando lo necesito, darme ánimos, consejos; por inspirarme y guiarme, incluso a pesar de la distancia. Ana Teles, eres un ser de luz! Thank you for all the positive vibe you transmit, for always be there, for encouraging me in difficult times, for all the memories collected, the laughs and conversations about life: "La vida siempre nos enseña". I am so grateful for having the chance to meet kind and genuine friends like you Zahra, we had great times together and your support,

despite of the distance, was huge to me. Ezgi and Fede all the moments together were so inspiring. Filipa, sharing office was so much fun, I appreciate all your support and all the help with this thesis. Maria, being neighbours during the pandemia and supporting each other was so fulfilling, all the moments together always inspire me. A pesar de la distancia te he sentido cerquita este tiempo y lo agradezco mucho. Mallavi, your jokes and nice vibe always brought me joy. Gillian sharing our passion for animals was exceptional. Andrea, girl we passed through so many things together during these years, I feel fortunate to share the whole journey with you. Thank you for always being so kind, supportive and empathic.

A PhD in Plön during a pandemia is much smoother with friends around. A big thanks to Amor, for the great energy that you transmit, to Elio, Christina, Pauline, Fernanda, Demetris, Nikhil, Roman for all the nice moments together. Lavisha, sharing Indian traditions was awesome and all the delicious food made me feel like home. Wiola, mermaiding with you was always amazing. Carolina, Artemis and Natasha, I enjoyed so much the dinners, conversations, and parties with you girls. Karem and Stella having the chance to meet you girls was one of the best things in these last years. Many more people contributed to the great times in Plön, here it goes a big thanks!

It was also gratifying getting to know people outside of the science bubble during this time, Vera your friendship is very valuable, the diving adventures was much more exciting with you. Wendy and Nicole, reviving culture was so fulfilling.

To my people in Bolivia, Yumi, gracias por siempre apoyarme, por la energía que transmites y por estar ahí cuando lo necesito. Indy, las charlas de los domingos y reencuentros siempre me han motivado. Ceci y Mari Yapu compartir la experiencia de vivir en otro continente ha sido una de las mejores cosas. Gracias por todo el apoyo, los viajes y las visitas que siempre me han recargado. Lau, Mari, Caro, Pao, Vale, Monich y Pame gracias por siempre estar ahí, esperarme con ansias en mis visitas y hacerme sentir en casa. A mis "Wayruritos", nuestros reencuentros siempre me han recargado e inspirado.

A huge thanks to Sven Frehn for bringing the balance that my life needed in the last years of the PhD. Thank you for all the emotional support, the patience, for bearing with me all the stressful moments (specially in the end). I appreciate the help reading my thesis and understanding about bird migration and genomics. Gracias por darme calma y alegrar hasta los días mas dificiles.

This PhD is dedicated to my family, without them this wouldn't have been possible. Abuelitos, Basi and Toffy perderlos ha sido una de las partes más difíciles. Soy afortunada de aun poder sentirlos. Gracias por toda la enseñanza y valores que me han dejado. Especialmente, quiero agradecer a mi madre Ana María y hermana Daniela por ser mi luz, inspiración y fortaleza. No ha sido fácil estar lejos, pero doy gracias por siempre apoyar mis decisiones, por su amor incondicional y paciencia. El ejemplo de mi madre y valores inculcados han hecho quien soy y todo lo que estoy logrando. Un gracias especial a mi prima, Katy, por siempre poder contar contigo, por tus consejos y el soporte. También agradecer a mi padre Jose Luis por el apoyo durante estos años y al resto de mi familia.

Author Contributions

Chapter II: Fine-scale map reveals highly variable recombination rates associated with genomic features in the European blackcap

Personal contributions: I optimised the methodological approach to estimate historical recombination rates. I performed recombination rate estimates and characterised a variety of genomic features such as GC content, CpG islands, nucleotide density and genome complexity. I additionally performed the analysis and interpretation of results. I wrote the manuscript with input from LOH and ML and all the co-authors provided comments on earlier versions of the manuscript.

Author contributions: ML and LOH designed and jointly supervised the study. JYD provided input for the rrecombination rate estimates and analysis. Characterisation of genomic features, genome complexity and their respective associations with recombination rate were carried out by KBC with support from LOH and ML. Mapping and variant calling was done by AB, genome annotation was generated by GM with input from GD, annotation of retrotransposons was done by PP.

Chapter III: The relationship of CpG methylation shared among tissues with recombination rates across the blackcap genome

Personal contributions: LOH, ML and I designed the project. I analysed the bisulfite sequencing data and characterised the association of CpG methylation with recombination rates. I interpreted the results, performed the visualization and wrote the chapter.

Author contributions: LOH and ML provided input for the interpretation of results and gave feedback for the writing and reviewing of the chapter. BM provided support with the bioinformatic analyses.

Chapter IV: Historical recombination maps diverge between blackcap populations with distinct migratory phenotype

Personal contributions: LOH, ML and I designed the project. I estimated population-specific recombination rates and performed analysis with simulated data. By integrating the population

features estimates, I analysed and interpret the association with recombination maps among populations. I performed the analysis and interpretation of the results with input from LOH, ML, AB and JI. I wrote the manuscript with input from LOH and ML, all the co-authors provided comments on earlier versions of the manuscript.

Author contributions: JI inferred population-specific historical demography and performed simulation analysis. AB generated the VCF file and estimated summary statistics. JI and AB contribute with the interpretation of the results related to demography and population structure. All authors provided comments and feedback on the manuscript.

Additional contributions to papers not included in this thesis that I co-authored during my PhD research:

The blackcap (*Sylvia atricapilla*) genome reveals a species-specific accumulation of LTR retrotransposons

Personal contributions: As co-author of this study, I provided recombination maps and GC content and CpG islands density estimates. I contribute with bioinformatic analysis and scripts for the partial correlations. I additionally, provide input for the interpretation of TEs and recombination correlations results. Together with the rest of the co-authors, I provided feedback for the writing and editing of the manuscript.

Author contributions: PP, LOH an ML designed research; AB performed data preparation and data analyses. PP generated TE landscapes and annotation. AB and PP wrote the manuscript together with ML and LOH, all co-authors commented on the manuscript.

Recombination suppression and selection affect local ancestries in genomes of a migratory songbird

Personal contributions: As co-authors of this study, I contributed with the population-specific recombination analysis by estimating recombination rates for intrachromosomal regions corresponding to chromosomal inversion and genomic differentiated regions. I provided the recombination map of the blackcap and garden warbler and performed synteny analysis. I participated in the reviewing and editing rounds for different manuscript versions.

Author contributions: JI conceptualised the study with support from ML. The reference assembly was generated by JF, AR, JM, BH, WC, JC, KH, MU, OF, and EDJ within the framework of the

Vertebrate Genomes Project, sample material was provided byJP and JCI provided blood samples of new blackcap individuals and WGS resequencing was carried out through a cooperation agreement with JF. VCF was generated and variant calling performed by AB. All population genomic analyses (except recombination map inference). simulations, and phylogenetics were carried out by JI. JI visualised data and wrote the original manuscript with input by KBC, AB, OF, EDJ and ML. All authors reviewed and edited the final version of the manuscript.

Authors given in alphabetic order (after KBC):

KBC: Karen Bascon-Cardozo, **AB:** Andrea Bours, **AR:** Arang Rhie, **BH:** Bettina Haase, **BM:** Britta Meyer, **ED:** Erich D. Jarvis, **GD:** Gillian Durieux, **GM**: Georg Manthey**, JM:** Jacquelyn Mountcastle, **JF:** JaninaFuß, **JP:** Javier Pérez-Tris, **JC:** Joanna Collins, **JI:** Juan CarlosIllera, **JYD:** Julien Yann Dutheil , **JI:** Jun Ishigohoka, **KH:** Kerstin Howe, **LOH:** Linda Odenthal-Hesse; **MU:** Marcela Uliano-Silva, **ML:** Miriam Liedvogel, **OF:** Olivier Fedrigo, **PP:** Peter Pruisscher, **WC:** William Chow

Declaration/ AFFIDAVIT

I hereby declare that:

i. Apart from my supervisor's guidance, the content and design of this thesis is the product of my own work. The co-authors' contributions are listed in the dedicated section;

ii. This thesis has not been already submitted either partially or wholly as part of a doctoral degree to another examination body, and no other materials are published or submitted for publication than indicated in the thesis;

iii. The preparation of the thesis has been subjected to the Rules of Good Scientific Practice of the German Research Foundation;

iv. Prior to this thesis, I have not attempted and failed to obtain a doctoral degree.

Plön, November 2022:

Karen Bascón Cardozo