Received: 29 May 2019 | Revised: 7 November 2019 | Accepted: 10 December 2019

DOI: 10.1111/1755-0998.13134



RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

# A genome-wide linkage map for the house sparrow (*Passer domesticus*) provides insights into the evolutionary history of the avian genome

Ingerid J. Hagen[1,2] | Sigbjørn Lien[3] | Anna M. Billing[1] | Tore O. Elgvin[4] | Cassandra Trier[4] | Alina K. Niskanen[1,5] | Maja Tarka[1,6] | Jon Slate[7] | Glenn-Peter Sætre[4] | Henrik Jensen[1]

[1]Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway

[2]Norwegian Institute for Nature Research (NINA), Trondheim, Norway

[3]Centre for Integrative Genetics, Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, Ås, Norway

[4]Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, Oslo, Norway

[5]Ecology and Genetics Research Unit, University of Oulu, Oulu, Finland

[6]Department of Biology, Lund University, Lund, Sweden

[7]Department of Animal and Plant Sciences, University of Sheffield, Western Bank, Sheffield, UK

**Correspondence**
Ingerid J. Hagen, Norwegian Institute for Nature Research (NINA), Trondheim, Norway.
Email: ingerid.arnesen@nina.no

**Funding information**
NTNU; The Norwegian Research Council, Grant/Award Number: 191847, 221956 and 223257; The Norwegian Directorate for Nature Management; The EU-commission

## Abstract

The house sparrow is an important model species for studying physiological, ecological and evolutionary processes in wild populations. Here, we present a medium density, genome wide linkage map for house sparrow (*Passer domesticus*) that has aided the assembly of the house sparrow reference genome, and that will provide an important resource for ongoing mapping of genes controlling important traits in the ecology and evolution of this species. Using a custom house sparrow 10 K iSelect Illumina SNP chip we have assigned 6,498 SNPs to 29 autosomal linkage groups, based on a mean of 430 informative meioses per SNP. The map was constructed by combining the information from linkage with that of the physical position of SNPs within scaffold sequences in an iterative process. Averaged between the sexes; the linkage map had a total length of 2,004 cM, with a longer map for females (2,240 cM) than males (1,801 cM). Additionally, recombination rates also varied along the chromosomes. Comparison of the linkage map to the reference genomes of zebra finch, collared flycatcher and chicken, showed a chromosome fusion of the two avian chromosomes 8 and 4A in house sparrow. Lastly, information from the linkage map was utilized to conduct analysis of linkage disequilibrium (LD) in eight populations with different effective population sizes ($N_e$) in order to quantify the background level LD. Together, these results aid the design of future association studies, facilitate the development of new genomic tools and support the body of research that describes the evolution of the avian genome.

**KEYWORDS**

linkage disequilibrium, linkage map, population genomics, reference genome

## 1 | INTRODUCTION

Recent technological advances have decreased the cost of next generation sequencing and made the technology broadly available for non-model species. Whole genome sequencing and the construction of draft genome assemblies have become routine and are currently applied to non-commercial species for the purpose of studies in ecology and evolution (Ellegren, 2014) . While a draft genome assembly accurately describes the physical position of genes along scaffolds, it does not provide information about the order and orientation of scaffolds and high-level chromosome organization. As such, evolutionary information cannot be found in highly fragmented assemblies. In contrast, linkage maps are inaccurate at a fine scale due to difficulties of resolving recombination between adjacent markers, but highly valuable to describe genome structure at the chromosome level. Combining the two resources can be used to order and orientate scaffolds along linkage groups to generate chromosome-scale assemblies (Fierst, 2015).

The comprehensive information in a reference genome cannot be fully utilized without further knowledge about the genetic distance between loci and the pattern of inheritance. Linkage maps represent an invaluable tool in this respect, as they describe the genetic distance along the chromosomes; that is the likelihood that two loci are inherited together and the frequency at which their association is broken by crossing over during meiosis. Recombination is a fundamentally important force in evolution, as it creates new haplotypes by breaking up associations between linked loci, thus increasing haplotype variation and improving the efficiency of natural selection. Whilst advantageous gene combinations may be disrupted, recombination may also break associations between disadvantageous alleles and favourable mutations at linked loci (Kawakami et al., 2017). The rate of recombination is therefore expected to correlate with rate of adaptive evolution.

The rate of recombination is not uniform, but elevated in genomic hotspots, also called 'fragile sites' characterized by specific repeats (see Aguilera & Gomez-Gonzalez, 2008 for review). Genome-wide and site-specific recombination rates are species specific and information in linkage maps cannot be extrapolated between taxa, although some general trends exist. Due to at least one obligate crossover for each chromosome pair prior to segregation, the rate of recombination per physical unit of distance generally increases with decreasing chromosome size (Ellegren, 2010; Kaback, Guacci, Barber, & Mahon, 1992). In the zebra finch (*Taeniopygia guttata*), hotspots were detected on average every 215 kb, with a total of nearly 4,000 hotspots (Singhal et al., 2015). These were in the proximity of GC rich regions, gene promotor regions and transcription start/stop sites. In zebra finch and collared flycatcher (*Ficedula albicollis*), these hotspots are often associated with chromosome ends, whilst areas near centromeric regions are characterized as recombination deserts (Backström et al., 2010; Bolívar, Mugal, Nater, & Ellegren, 2016; Kawakami et al., 2017). Different species of finches show greater conservation in the location of recombination hotspots than do mammals (Singhal et al., 2015), perhaps in concordance with the higher degree of synteny in bird genomes compared to mammals (Backström et al., 2010; Ellegren et al., 2012). Moreover, several studies spanning a wide variety of taxa have reported sex specific recombination rates, also termed heterochiasmy (Hansson, Åkesson, Slate, & Pemberton, 2005; Johnston, Huisman, Ellis, & Pemberton, 2017; Lien et al., 2011), although no clear pattern exists with respect to the sex that displays the elevated recombination rate.

Linkage maps were previously limited to humans (*Homo sapiens*) (Weissenbach et al., 1992) and model or life stock species such as mice (*Mus musculus*) (Copeland et al., 1993), sheep (*Ovis aries*) (Crawford et al., 1995), zebra finch (Backström et al., 2010) or Atlantic salmon (*Salmo salar*) (Lien et al., 2011). With the advancement of genotyping technology, linkage maps from many wild populations have become available, and include great reed warbler (*Acrocephalus arundinaceus*) (Åkesson, Hansson, Hasselquist, & Bensch, 2007), red deer (*Cervus elaphus*) (Johnston et al., 2017), collard flycatcher (*Ficedula albicollis*) (Kawakami et al., 2014), great tit (*Parus major*) (van Oers et al., 2014) and Soay sheep (Slate et al., 2010). These linkage maps greatly facilitate association studies in wild populations, often with the aim of mapping genes for ecologically important phenotypes. In this context, the extent of linkage disequilibrium (LD) is important: LD describes the nonrandom association between alleles at different loci. A long LD range suggests that fewer markers are needed to detect an association, but it also implies that the underlying causative variants will be more difficult to identify (Lindblad-Toh et al., 2005). Higher LD is the result of genetic drift, natural selection, and population structure/admixture, whilst the breakdown of LD is mainly caused by recombination (De La Vega et al., 2005). Because the effective size of a population reflects the rate of genetic drift, LD is expected to increase with decreasing population size. Information about LD in different focal populations is therefore important when carrying out association studies.

Importantly, linkage maps provide insight into the evolution of the avian genome, which is characterized by a high chromosome number and a karyotype consisting of a mixture of microchromosomes and a few larger macrochromosomes that may differ by as much as an order of magnitude in size (Rodionov, 1996). Despite long evolutionary divergence times, the avian karyotype has proved to be remarkably well conserved compared to other amniotic taxa such as mammals (Backström et al., 2010, 2008; Ellegren et al., 2012; Kawakami et al., 2014). Passerine birds diverged from Galliformes approximately 100 million years ago (MYA) (Tuinen, Sibley, & Hedges, 2000), and while Passeriformes is the largest and most diverse bird clade, the synteny is similar in all species investigated (Åkesson et al., 2007; Backström et al., 2010; Kawakami et al., 2014; van Oers et al., 2014).

The house sparrow (*Passer domesticus*) is a small passerine bird that has served as a model species of a number of evolutionary, physiological and ecological studies (Anderson, 2006). The house sparrow is native to Europe and Asia, and due to introductions aided by humans, the species is now established on all major habitable land masses across the globe (Anderson, 2006). House sparrows are nonmigratory with a low dispersal rate from their

natal area (Pärn, Ringsby, Jensen, & Sæther, 2012) and populations and individuals can be monitored over multiple years. The house sparrow is commensal with humans, individuals are easily captured using mist nets, and the species will readily make use of nest boxes, which again facilitates the access to eggs and chicks. Together, these characteristics make populations of house sparrows suitable for pedigree studies. Until recently, genomic resources for the house sparrow were lacking and the genetic tools were limited to microsatellites (Dawson et al., 2012). A custom made 10 K Illumina SNP array represented the first genomic resource (Hagen et al., 2013) and has greatly facilitated further development of genetic tools. A major recent achievement has been the house sparrow reference genome (Elgvin et al., 2017), which was developed in concert with the house sparrow linkage map presented herein. The house sparrow linkage map is constructed based on 6,498 genome wide SNPs and a pedigree of 862 individuals included in 105 families, and represents an important resource for further population studies, association studies and for comparative evolutionary analysis. We have estimated recombination rate across the house sparrow genome and investigated how it varies with sex, chromosome size and distance from chromosome ends. Moreover, we have identified chromosomal synteny between house sparrow and zebra finch, flycatcher and chicken, respectively, and estimated the global genomic LD in eight geographically separated house sparrow populations of varying effective sizes.

## 2 | MATERIALS AND METHODS

### 2.1 | Blood sampling and pedigree

The samples used in the linkage analysis were collected from 1910 individuals from four different island populations: Aldra (*N* = 348), Hestmannøy (*N* = 448), Leka (*N* = 511) and Vega (*N* = 603) (see Table 1) during the years 1998–2011. Due to the isolated nature of the populations (Hagen et al., 2013), these samples constitute four different pedigrees. The pedigrees were constructed using individual microsatellite genotypes and knowledge of which individuals were potential mothers and fathers of offspring produced in different years in the program CERVUS 3.0 (Kalinowski, Taper, & Marshall, 2007); for details see Billing et al. (2012) and Kvalnes et al. (2017) . Additionally, we used samples collected from the islands of Gjærøy (*N* = 89), Indre Kvarøy (*N* = 94), Myken (*N* = 52) and Nesøy (*N* = 36). These individuals were not included in the pedigrees used in the linkage analysis but were used for estimating population-specific levels of linkage disequilibrium. The estimated average annual effective population size ($N_e$) for each of the island populations are from Stubberud et al., (2017) and are presented in Table 1. Extraction of the blood samples and subsequent genotyping on the 10 K Illumina iSelect HD BeadChip is described in detail in Hagen et al. (2013). The results were scored using GenomeStudio (Illumina) following the manufacturer's guidelines (Illumina, Inc. 2010). We used PLINK version 1.07 (Purcell et al., 2007) for further quality control. Markers with a minor allele frequency (MAF) < 0.01 were removed. First, we removed markers that deviated significantly from Hardy-Weinberg equilibrium frequencies within the four populations used in linkage mapping (*n* = 163). Next, markers with a minor allele frequency (MAF) < 0.005 (*n* = 1,462), and individuals with >10% missing genotypes were removed. Finally, we removed SNPs with >20 Mendelian errors based on parent-offspring links in the microsatellite pedigrees (*n* = 73). After quality control our data set consisted of 6,764 (autosomal) SNPs typed in 1,898 individuals.

### 2.2 | Construction of the linkage map

The linkage map was constructed in concert with the house sparrow reference genome (Elgvin et al., 2017). A modified version of the CRIMAP 2.4 software (Green, Falls, & Crooks, 1990), including added utilities that prepare text files to facilitate downstream analysis of large marker sets provided by Xuelu Liu and Michael Grosz (Monsanto), were used for the map construction. Preceding the mapping procedure, the microsatellite pedigrees were corrected using ca. 3,700 of the SNPs, and split into 105 separate and smaller families with eight individuals on average (range: 4–37), in which both parents were known for on average seven individuals in the family. Among families, 64% included only individuals for which both parents were known, whereas 39% of the families included at least one individual with only one known parent. No individuals had two missing parents. Initially, SNPs were assigned to linkage groups (LGs) based on pairwise linkages and the grouping algorithm implemented in the AUTOGROUP option of the program. Of the initial 8,957 SNPs, the linkage analysis assigned 6,498 SNPs to 29 autosomal LGs which comprise the linkage map presented in this study. Additionally, 456 SNPs were assigned to an unmapped linkage group corresponding to the Z chromosome. After the initial grouping of SNPs, markers on the 29 autosomal LGs were ordered using the BUILD and FLIPSN options in CRIMAP. Following this, 120 bp flanking each SNP was positioned in the house sparrow genome assembly using BLAST and used to assign scaffolds to LGs, order and orientate scaffolds within LGs, and build sequences for 29 autosomal chromosomes in *P. domesticus* (Elgvin et al., 2017). If a scaffold contained only one marker in the linkage map, the scaffold was assigned the same orientation as in the zebra finch genome. Following the construction of chromosome sequences, the SNP-order within each LG was fine-tuned using physical positions of the SNPs in the chromosome. The CHROMPIC option in CRIMAP was then used to phase genotypes within linkage groups. CHROMPIC output were utilized to identify double recombination events caused by single SNPs. Such events are considered highly dubious, most probably caused by genotyping errors, and the genotypes were zeroed out before constructing the final version of linkage maps. Finally, multipoint linkage maps for the 29 autosomal linkage groups in the linkage map were constructed using the FIXED option of CRIMAP and

**TABLE 1** Number of samples (after quality control) and average annual effective population sizes ($N_e$) for each of the island populations

| Island population | Sampling years | Number of samples | Population $N_e$ |
|---|---|---|---|
| Nesøy | 1995, 2003, 2011 | 36 | 8.30 |
| Gjærøy | 1995, 2003, 2011 | 89 | 15.90 |
| Indre Kvarøy | 1995, 2003, 2011 | 93 | 18.10 |
| Myken | 1995, 2003, 2011 | 52 | 7.32 |
| Hestmannøy* | 1998–2011 | (448) 490 | 53.95 |
| Aldra* | 1998–2011 | (348) 348 | 8.60 |
| Leka* | 2002–2010 | (511) 511 | 28.38 |
| Vega* | 2002–2010 | (603) 603 | 37.00 |

*Note:* For the islands with three sampling years, the sample consisted of adults present on the island in one or more of those years. For Aldra, the sample consisted of all individuals (adults, fledglings and fledged juveniles) present on the island in the period 1998–2011. For Hestmannøy the sample consisted of a large proportion of adults present on the island during the time period 1998–2011. For Leka and Vega, the samples consisted of a large proportion of adults present during the time period 2002–2010. Populations donated with * were used in the assembly of the linkage map. Numbers in parenthesis represent the number of individuals from each population that were used in the linkage map assembly, all other numbers represent individuals used in LD analysis.

the Kosambi map function (Kosambi, 1943) . Chromosome nomenclature in house sparrow was determined by alignments against the zebra finch, flycatcher and chicken genomes and named Pdom 1–15, Pdom 17–28, Pdom 1A and LGE22, where Pdom is derived from the scientific name.

## 2.3 | Analysis of recombination rate and genome structure

We investigated the relationship between physical length and genetic length of all mapped chromosomes using linear regression in R version 3.5.3 (R Core Team, 2018). The mean recombination rate of macro- and microchromosomes (larger and smaller than 35 Mbp), according to the karyotype described in Bulatova, Radjabli, and Panov (1972) was compared using t-statistics in R. To characterize how recombination rate varies with distance from chromosome ends, recombination rates were characterized in 1 Mbp bins along the chromosomes and the physical length of each chromosome adjusted to a relative size of 0 (start) to 1 (end). Due to size dependent recombination patterns, we visualised the recombination landscape of macrochromosomes (Pdom 1A and 1–8) and microchromosomes < (Pdom 9–28) separately.

## 2.4 | Chromosomal rearrangements

For visualization of the alignments of the house sparrow genome with the genomes of zebra finch, chicken and collared flycatcher, we used CIRCOS version 0.67-7 (Krzywinski et al., 2009). To obtain input files for CIRCOS, we aligned the house sparrow genome (GenBank accession MBAE00000000) with each of the three target genomes (GenBank accession numbers AGTO02000000, ABQF00000000 and AADN00000000) using Satsuma SYNTENY version 2.0 (Grabherr et al., 2010). To get a measure of the percentage of interchromosomal rearrangements on the 29 mapped house sparrow chromosomes compared to zebra finch, chicken and collared flycatcher, we estimated the pairwise length of rearranged sequence for each chromosome. Chicken, flycatcher or zebra finch sequences that were in opposite direction than in the house sparrow genome were defined as "negative length", and the total length of "negative sequence" was summarised for each species. Due to chromosomal rearrangements between species, house sparrow chromosomes 1A and 1 were compared with chicken chromosome 1, and house sparrow chromosome 8 was compared with chicken chromosomes 4 and 8. House sparrow chromosome 8 was compared with zebra finch and collared flycatcher chromosomes 4A and 8.

## 2.5 | Validation of interchromosomal fusion

We utilized REAPR (v 0.1.9) (Hunt et al., 2013) and the 8 and 10 kb mate-pair sequencing libraries from the house sparrow genome assembly to further validate the fusion of chromosomes 4A and 8 identified in the linkage map. REAPR uses short reads to score bases and long insert libraries to identify assembly errors. To identify potential assembly errors, both mate-pair libraries were separately mapped to the house sparrow assembly using the default settings of BWA-MEM (v 0.7.17) (Li, 2013) with the -M parameter added for Picard compatibility. The resulting bam files were sorted with Picard (v 1.9.6) SortSam and duplicates were marked with Picard MarkDuplicates with validation stringency set to "lenient". The full REAPR pipeline of assembly evaluation was then run using the REAPR pipeline

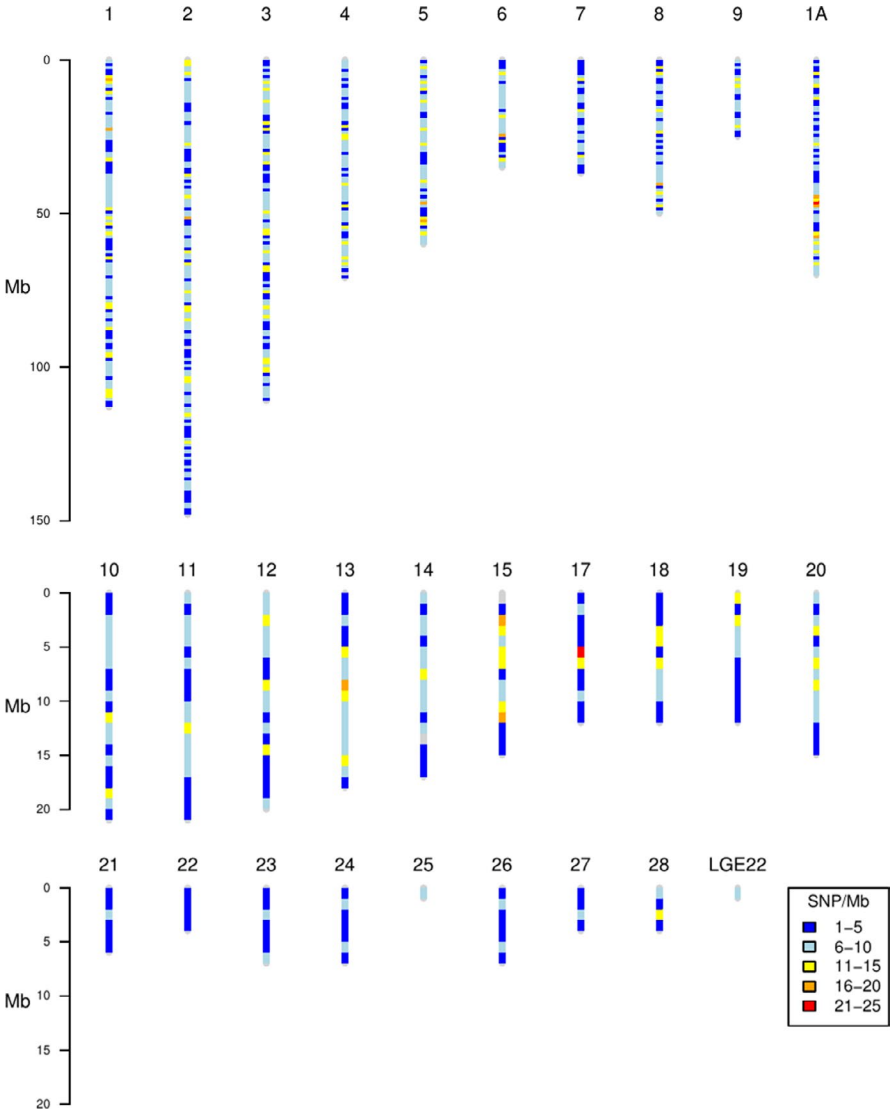


**FIGURE 1** Marker density for all mapped autosomal house sparrow (*Passer domesticus*) chromosomes indicated in 1 Mbp bins. The vertical length of each chromosome represents the physical size (Mb) [Colour figure can be viewed at wileyonlinelibrary.com]

command and default settings. Lastly, the location of the fusion between chromosomes 8 and 4A was compared with results from the REAPR analysis to determine if REAPR agreed with the fusion.

## 2.6 | Analysis of linkage disequilibrium

For the analysis of LD, we used 6,490 SNPs for 28 autosomal LGs (excluding the eight SNPs on LGE22), and genotypes from individuals that originate from the following eight populations: Aldra, Gjærøy, Hestmannøy, Indre Kvarøy, Leka, Myken, Nesøy and Vega (see Table 1). The genotypes were based on the data set described in Hagen et al. (2013) with additional 271 samples genotyped on the house sparrow 10K Illumina SNP array. The additional samples were stored, extracted and genotyped as previously described in Hagen et al. (2013). Analysis of LD were carried out in PLINK version 1.07 (Purcell et al., 2007) and R version 3.5.3 (R Core Team, 2018). For measures of LD, we used the squared correlation coefficient $r^2$, which is considered the most robust parameter of LD (Kruglyak, 1999) . We quantified LD in six bins of 5 Kbp spanning 0–30 Kbp, seven bins of 10 Kbp spanning 30–100 Kbp as well as additional bins spanning 100–200 Kbp, 200–500 Kbp, 500–1,000 Kbp, 1,000–2,000 Kbp and 2,000–3,000 Kbp. Linkage disequilibrium was estimated separately in each population. As a measure of background LD, we used the LD estimate obtained for the 2,000–3,000 Kbp bin. The relationship between background LD and average effective population sizes for each of the eight populations was tested using a linear model in R version 3.5.3 (R Core Team, 2018). As estimates of $N_e$ for seven of the islands (Aldra, Gjerøy, Hestmannøy, Indre Kvarøy, Nesøy, Leka and Vega) we used the estimates of local demographic variance $N_e$ in Stubberud et al. (2017) (Table 1). Importantly, these $N_e$ estimates account for effects of overlapping generations and environmental stochasticity on the rate of genetic drift, in addition to variation in population size, and among-individual variation in survival and reproduction within the local population (Engen, Lande, Sæther, & Gienapp, 2010). For the eighth island, Myken, only adult population size N was available. Hence, to obtain $N_e$ we multiplied the mean observed adult population size at Myken during the years 1993–2009 ($N$ = 20; Baalsrud et al., 2014) with the average $N_e/N$ ratio (using generation time T = 1.97) for the other seven islands ($N_e/N$ = 0.366).

**TABLE 2** Number of SNP markers, physical size (bp), average distance between markers (bp), and male, female and average recombination rate (cM) for each autosomal chromosome in the house sparrow linkage map

| Chromosome | Number of markers | Total physical length (bp) | Average distance between markers (bp) | cM | | |
|---|---|---|---|---|---|---|
| | | | | Male | Female | Average |
| 1 | 813 | 112,674,304 | 138,591 | 127.3 | 176.6 | 150.5 |
| 1A | 517 | 69,873,230 | 135,151 | 96.9 | 122.3 | 108.4 |
| 2 | 1,007 | 147,836,772 | 146,809 | 133.2 | 202.9 | 165.9 |
| 3 | 837 | 110,987,800 | 132,602 | 115.1 | 174.9 | 143.4 |
| 4 | 522 | 70,348,008 | 134,766 | 100.6 | 139.3 | 118.7 |
| 5 | 491 | 61,081,188 | 124,402 | 103.6 | 139.4 | 120.6 |
| 6 | 249 | 35,007,241 | 140,591 | 61.3 | 87.8 | 73.9 |
| 7 | 236 | 36,521,674 | 154,753 | 69.2 | 94.5 | 81.4 |
| 8 | 368 | 49,693,984 | 185,425 | 111.5 | 138.2 | 123.7 |
| 9 | 169 | 25,224,909 | 149,260 | 59.2 | 75.7 | 66.9 |
| 10 | 138 | 21,077,231 | 152,734 | 54.5 | 69.0 | 61.7 |
| 11 | 132 | 20,431,416 | 154,783 | 55.4 | 69.9 | 61.9 |
| 12 | 137 | 19,787,094 | 144,431 | 56.0 | 62.5 | 59.0 |
| 13 | 142 | 18,021,457 | 126,912 | 52.4 | 59.0 | 55.5 |
| 14 | 99 | 16,467,355 | 166,337 | 48.3 | 61.1 | 54.3 |
| 15 | 121 | 14,042,894 | 116,057 | 50.0 | 57.1 | 53.2 |
| 17 | 80 | 11,241,730 | 140,522 | 55.6 | 59.1 | 56.1 |
| 18 | 79 | 11,529,515 | 145,943 | 55.1 | 56.0 | 55.2 |
| 19 | 70 | 11,116,859 | 158,812 | 54.1 | 53.2 | 53.2 |
| 20 | 108 | 14,776,391 | 136,818 | 45.0 | 45.2 | 45.0 |
| 21 | 27 | 5,712,519 | 211,575 | 23.2 | 20.1 | 21.9 |
| 22 | 14 | 3,666,047 | 261,861 | 36.4 | 47.1 | 41.8 |
| 23 | 27 | 7,033,470 | 260,499 | 50.5 | 57.2 | 53.3 |
| 24 | 31 | 7,077,728 | 228,314 | 52.3 | 59.0 | 55.1 |
| 25 | 6 | 476,507 | 79,418 | 16.0 | 11.6 | 13.9 |
| 26 | 25 | 6,896,631 | 275,865 | 66.3 | 65.2 | 65.7 |
| 27 | 15 | 3,737,950 | 249,197 | 16.2 | 13.7 | 15.1 |
| 28 | 30 | 3,534,834 | 117,828 | 26.7 | 19.6 | 23.2 |
| LGE22 | 8 | 467,554 | 5,8,444 | 9.4 | 2.5 | 6.0 |
| Total | 6,498 | 916,344,292 | 159,610 | 1,801.3 | 2,239.7 | 2,004.5 |

*Note:* For accurate marker position and marker names see Table S1.

## 3 | RESULTS

### 3.1 | Linkage analysis

The linkage analysis mapped 6,498 SNPs to 29 autosomal chromosomes (1–15, 17–28, 1A and LGE22) (Figure 1, Table 2). For accurate physical and genetic marker positions and information on number of informative meioses for each marker, see Table S1. The sex averaged linkage map had a length of 2,004.5 cM, with a longer map for females (2,239.7 cM) than for males (1,801.3 cM), a difference corresponding to a 1.24 times higher recombination rate in females than males (Table 2).

### 3.2 | Recombination rate and genome structure

It is apparent that house sparrow chromosomes lack the strong s-shaped or sigmoid recombination rate pattern that is seen in zebra finch (Figure 2). Instead, recombination rates along the chromosomes more closely resembles the linear pattern that is seen in chicken. However, the recombination rate varied both within and between chromosomes in the house sparrow genome: recombination rates were greater for microchromosomes (6.41 cM/Mbp) than for macrochromosomes (1.78 cM/Mbp) (one-sided *t* test, $p = .001$) (Figure 3). Additionally, the recombination pattern along the chromosomes differs between micro- and macrochromosomes. Macrochromosomes

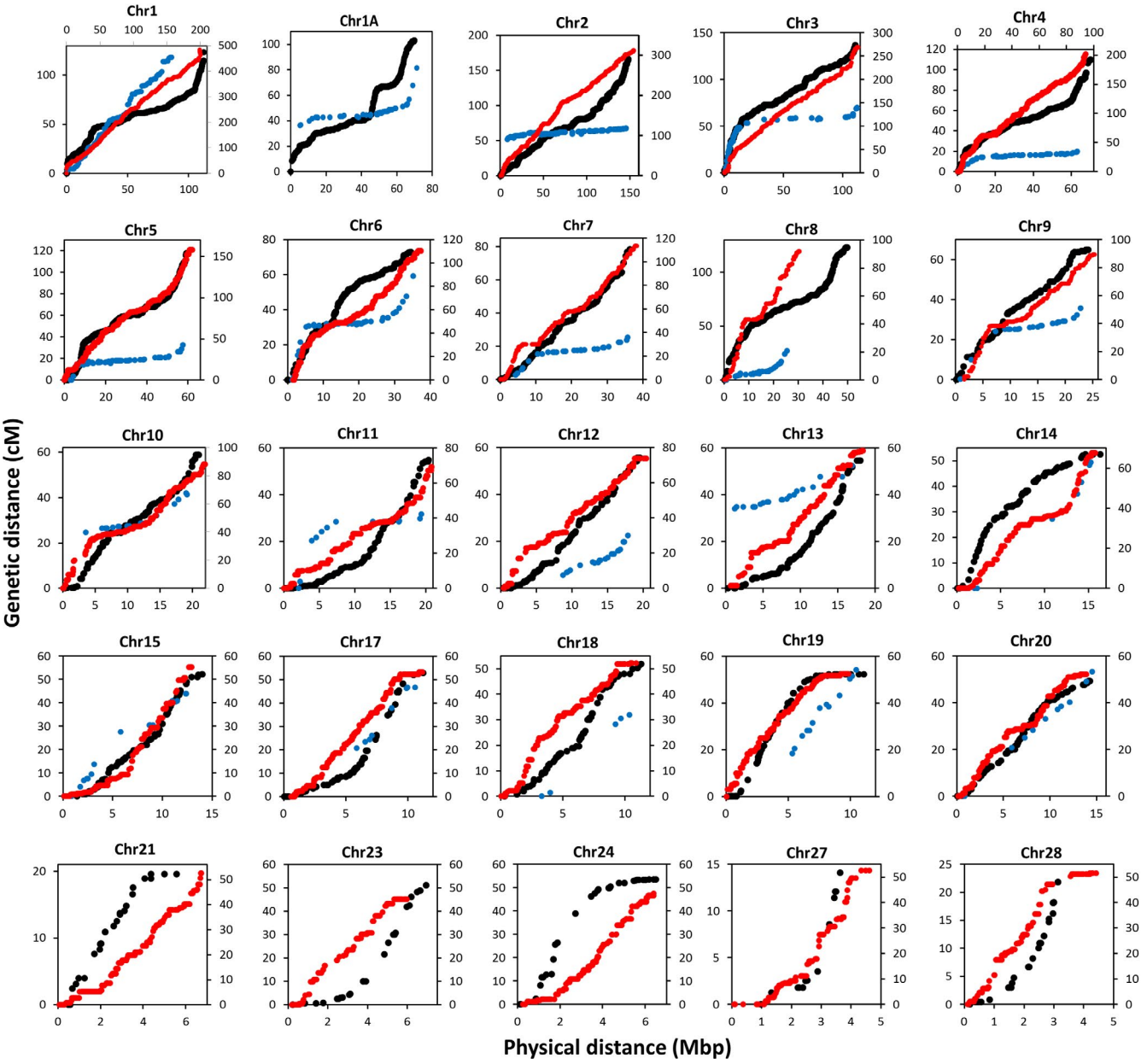


**FIGURE 2** The relationship between the physical and genetic position for all markers on each mapped autosomal chromosome for three avian species. Black dots are house sparrow (this study), red dots are chicken (Groenen et al., 2009) and blue dots are zebra finch (Backström et al., 2010). The *x*-axis represents the physical size of the chromosomes, based on the house sparrow genome assembly (Elgvin et al., 2017). Due to the chicken linkage map being significantly longer than the two passerine species, the chicken genome is always presented on a secondary *y*-axis (right), and for chromosome 1 and 4, a secondary x-axis (top). House sparrow lacks chromosome 4A, which for house sparrow is fused with chromosome 8. Only chromosomes with 15 or more markers are included. For zebra finch, data was not available for chromosomes 21, 22, 24, 26, 27 and 28 [Colour figure can be viewed at wileyonlinelibrary.com]

appear to have increased recombination rates towards the ends: the average recombination rate in the first and last 20% of the large chromosomes was 2.57 cM/Mbp, whilst the average recombination rate in these chromosomes' middle part (from 21% to 79% of relative length) was 1.09 cM/Mbp (Figure 4a). Microchromosomes lack this pattern and instead show large variation in recombination rate along the entire chromosome length (Figure 4b). Still, physical and genetic size for the different house sparrow chromosomes correspond closely ($r^2$ = .87, $p$ < .000, slope = 0.95, intercept = 36.8 ± 3.72) (Figure 5). The mean genome-wide recombination rate was 4.82 cM/Mbp.

## 3.3 | Chromosomal rearrangements

In construction of the linkage map, we detected a novel interchromosomal rearrangement in the house sparrow genome: the fusion of the Passerine chromosome 8 and the Passerine chromosome

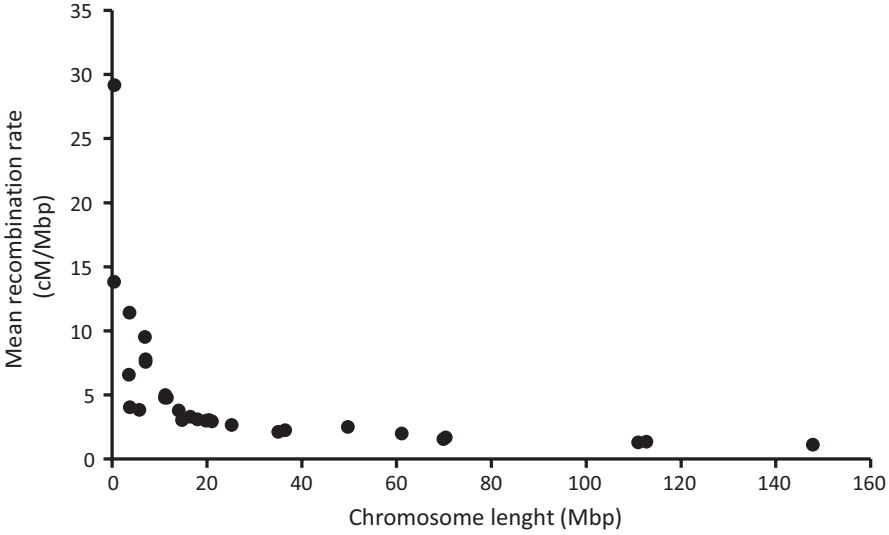


**FIGURE 3** Relationship between mean recombination rate (cM/Mbp) and physical chromosome size (Mbp) for 29 autosomal house sparrow chromosomes

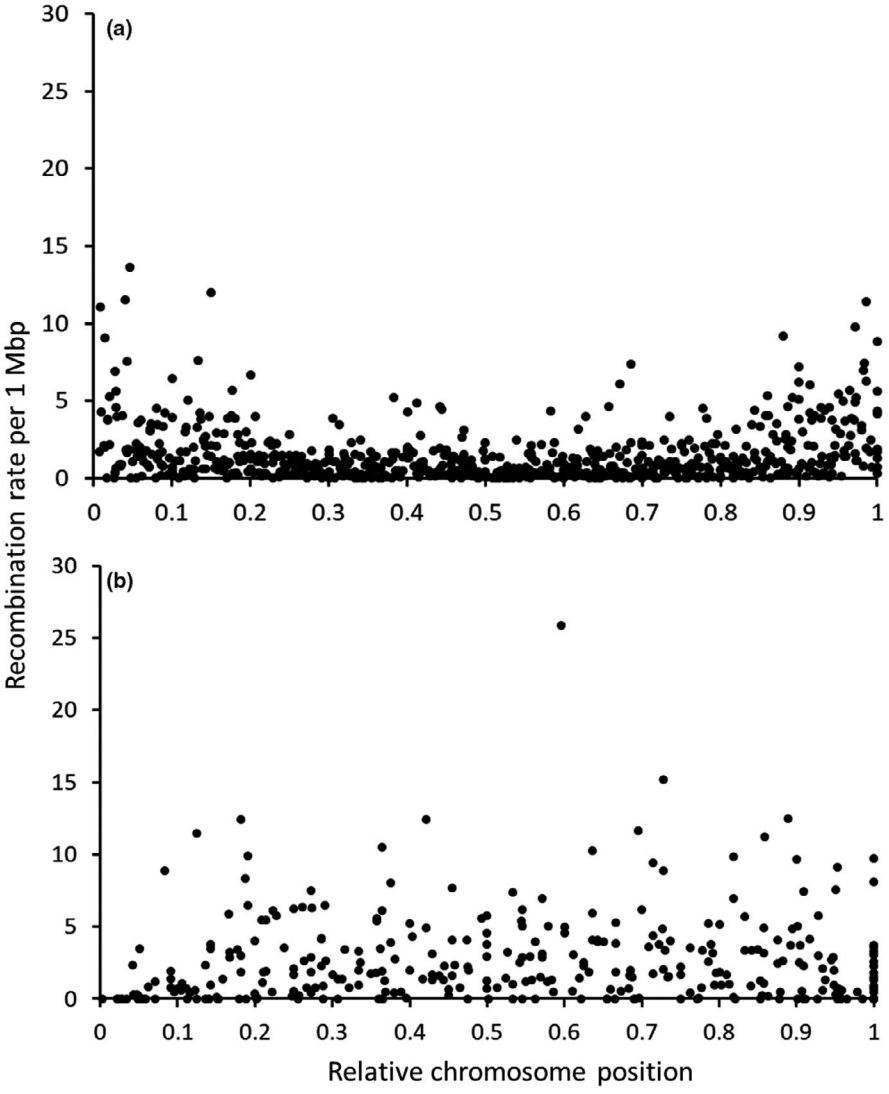


**FIGURE 4** Recombination rate per 1 Mbp windows along house sparrow macrochromosomes (a) and microchromosomes (b)

4A. Analysis of Pdom 8 using REAPR verified the assembly of the chromosome as a whole, thus supporting the fusion and the detected inversions in Pdom 8 (see Supporting information Results and Table S2 for details). The part added to chromosome 8 in house sparrow is the part of chicken chromosome 4 which constitutes the separate chromosome 4A in zebra finch, great tit

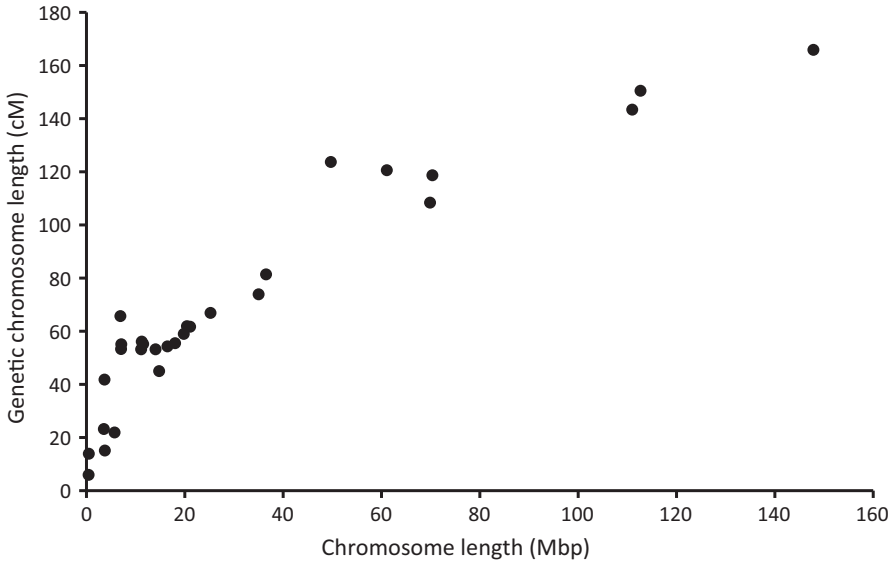


**FIGURE 5** Relationship between genetic size (cM) and physical size (Mbp) for the 29 autosomal chromosomes in the house sparrow linkage map

and collared flycatcher (Figure 6). As in zebra finch and collared flycatcher, chromosome 1 in house sparrow is split into 1 and 1A, whilst chicken has retained the ancestral chromosome 1. Beyond this, the house sparrow chromosome karyotype largely resembles that of zebra finch and collared flycatcher (Figure 6). The separate linkage group LGE22, which is found in zebra finch, great tit and flycatcher was also detected in house sparrow and is homologous to regions in the chicken linkage groups LG2 and LGE22C19W28_E50C23.

Despite the high degree of conserved synteny seen between house sparrow, zebra finch, collared flycatcher and chicken, several inversions were detected when intrachromosomal marker orders were compared (Table S3). When comparing the percentage of inverted sequence between house sparrow and chicken, collared flycatcher and zebra finch respectively, we found that only a few chromosomes had near complete conserved synteny of marker order between the different species. Chromosomes 6, 9 and 17 showed a high degree of conserved synteny between all three species. The total percentage of "negative sequence" between house sparrow and chicken was 47%, for collared flycatcher 51% and for zebra finch 50% (see Table S3).

### 3.4 | Linkage disequilibrium

Our results show that LD ($r^2$) decreased from values of 0.15–0.20 for markers up to 5 Kbp apart to 0.11–0.14 in the 10–15 kbp bin (Figure 7). Thereafter, LD decreased more slowly until reaching background level at around 1,000–2,000 Kbp. The populations with the lowest effective population size ($N_e$) (Aldra, Myken and Nesøy) showed both higher initial LD as well as higher background LD than did the populations of larger $N_e$ (Leka, Vega, Hestmannøy, Indre Kvarøy and Gjærøy). A linear regression shows some evidence of a positive relationship ($r^2$ = .44; $p$ = .07, intercept = −2.88 ± 0.20, estimate for $N_e$ = −0.02 ± 0.01) between background LD and effective population size for the eight populations.

## 4 | DISCUSSION

We have mapped 29 autosomal chromosomes for the house sparrow (Figure 1, Table 2). Additionally, 456 markers have been assigned to an unmapped reference sequence that corresponds to the zebra finch Z chromosome (Elgvin et al., 2017). As the house sparrow has an expected haploid chromosome number of 38 (Bulatova et al., 1972), it is apparent that eight chromosomes (linkage groups) have not been mapped. It is likely that these represent a corresponding number of microchromosomes, which by nature are particularly challenging to map and assemble: they are often rich in tandem repeats, have a high GC content and often contain gene families that are difficult to sequence such as the MHC gene complex on Gga16, which has proved difficult to assemble (Warren et al., 2017). Despite a high-density linkage map and excellent reference genome for collared flycatcher, for which the karyotype is unknown, Kawakami et al. (2014) estimated that approximately five or six microchromosomes remain unmapped. Similarly, for zebra finch, which has a haploid chromosome number of 40 (Itoh & Arnold, 2005), eight microchromosomes remain unmapped (Backström et al., 2010).

The mean recombination rate reported in this study is comparable to that found in chicken (6.02 cM/Mbp) (Groenen et al., 2000), zebra finch (3.18 cM/Mbp) (Stapley, Birkhead, Burke, & Slate, 2008) and collared flycatcher (3.1 cM/Mbp) (Kawakami et al., 2014). Recombination rates in the zebra finch have been shown to be strikingly heterogeneous along chromosomes, with most recombination events being concentrated towards the chromosome ends (Backström et al., 2010), i.e., plots of recombination rate against physical distance have a sigmoidal shape. Even though there is a tendency for increased recombination rate towards telomeric ends on the large chromosomes (Figure 4a), the house sparrow does not seem to have the pattern of very low recombination rate in middle parts of the chromosomes that is seen in zebra finch (Figure 2). Instead, the house sparrow displays a recombination pattern that

resembles the uniform distribution seen in collared flycatcher and chicken (Kawakami et al., 2014) (Figure 2). As expected (Groenen et al., 2000; Stapley et al., 2008), there was an inverse relationship between chromosome size and recombination rate as well as a near linear relationship between genetic distance and physical size of chromosomes (Figure 3) similar to what has been reported for other Passerine species (Backström et al., 2010; Kawakami et al., 2014).

Of the 10 largest house sparrow chromosomes, eight are either submetacentric or subtelocentric, whilst two are telocentric (Bulatova et al., 1972). Apart from chromosome 2, which is the largest, the identity of the different chromosomes cannot be discerned based on karyotype. Also, the position of the centromere for the 28 smaller microchromosomes is unknown (Bulatova et al., 1972). As recombination rate increases with distance from centromere (Akhunov et al., 2003), we expect the different arms on acrocentric chromosomes to have different recombination rates. It is therefore possible that recombination rate as a function of distance from both the centromere and chromosome ends would provide the

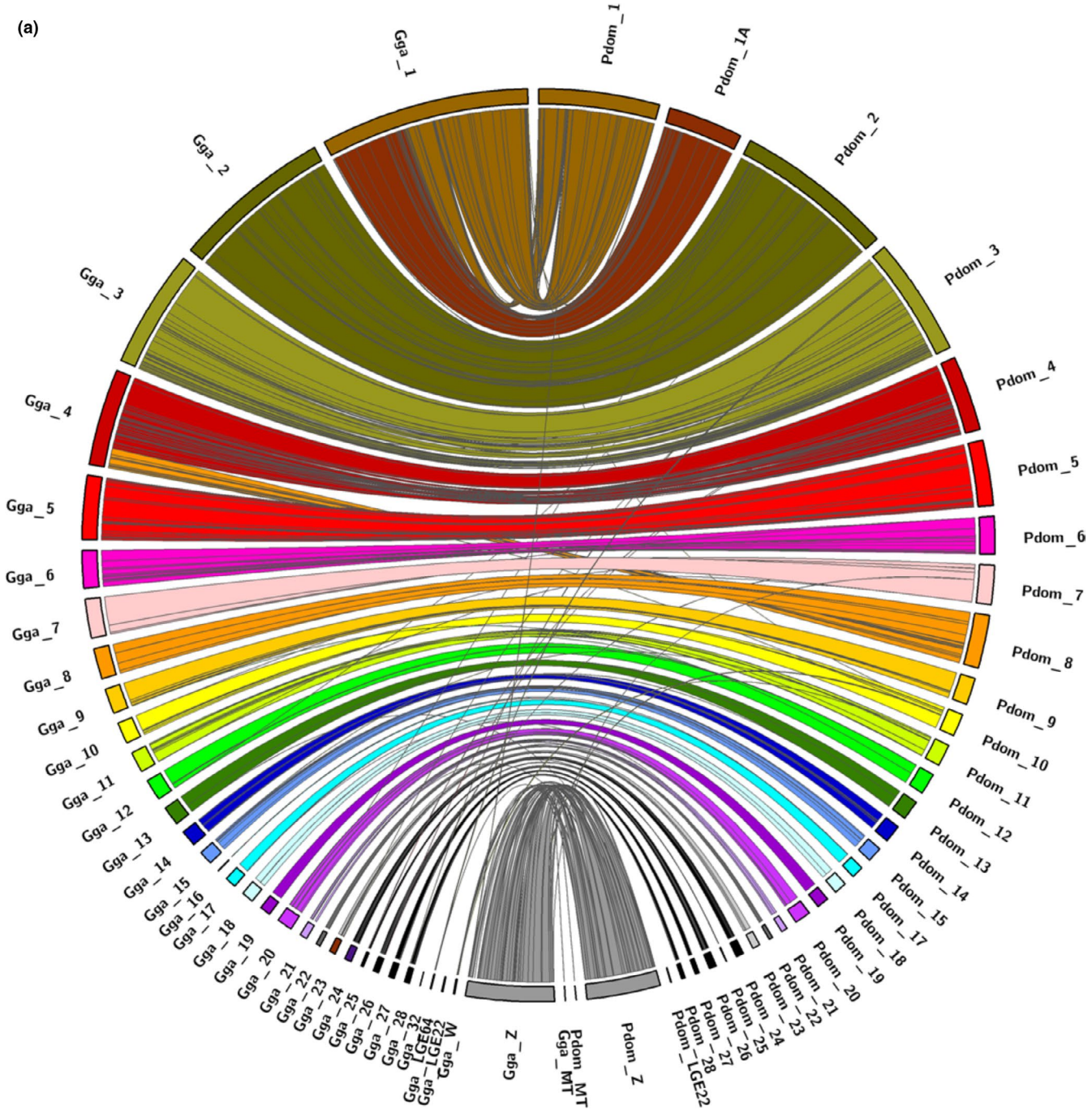


**FIGURE 6** Comparative circular visualization of the house sparrow genome with the genomes of (a) chicken, (b) collared flycatcher and (c) and zebra finch. The plots show the full house sparrow reference genome with the 29 mapped autosomal chromosomes as well as the unmapped Z chromosome, for which scaffold assembly was based on synteny with the zebra finch genome [Colour figure can be viewed at wileyonlinelibrary.com]

most accurate model for describing the heterogeneous recombination rates along house sparrow macrochromosomes. Therein lies the challenge of achieving an accurate estimate for the centromere position for the different chromosomes; data which cannot be discerned from the karyotype described in Bulatova et al. (1972).

Although the avian karyotype is highly conserved across species (Derjusheva, Kurganova, Habermann, & Gaginskaya, 2004), several rearrangements have been reported, especially involving chromosome 4. Chicken chromosome 4 is likely to be a result of a fusion between the ancestral chromosome 4 and another chromosome (Griffin, Robertson, Tempest, & Skinner, 2007). In zebra finch, great tit and collared flycatcher, chromosome 4 is split into 4 and 4A, respectively (Kawakami et al., 2014; Stapley et al., 2008; van Oers et al., 2014). Microsatellite-based linkage maps suggest that chromosome 4 is split into two or more linkage groups in blue tit as well (Hansson et al., 2009; Nietlisbach et al., 2015). Moreover, evidence of sex linkage on parts of chromosome 4A has been found in great reed warbler (*Acrocephalus arundinaceus*), skylark (*Alauda arvensis*) and common whitethroat (*Sylvia communis*), indicating a fusion of parts of 4A to the Z chromosome

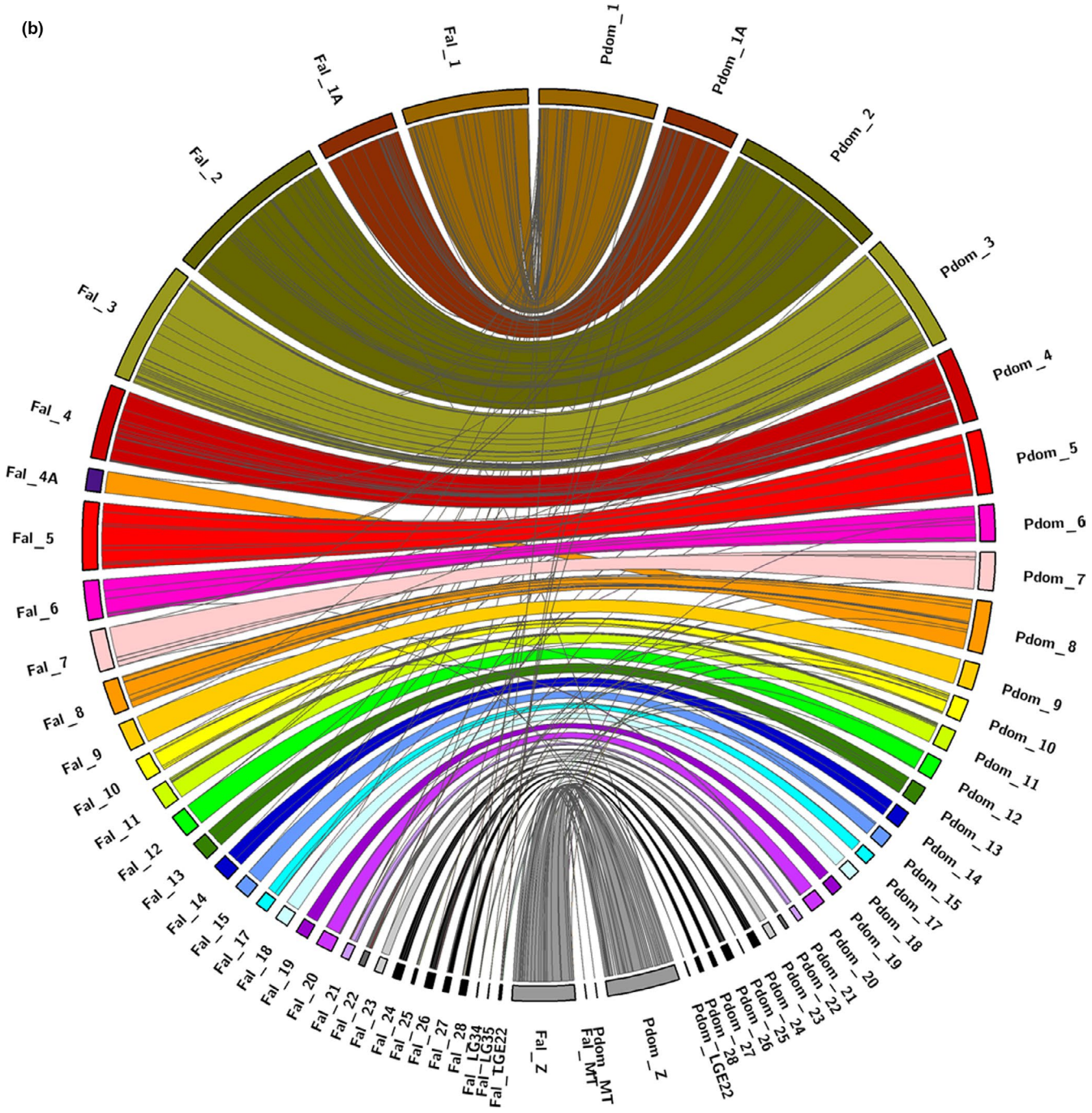


**FIGURE 6** (Continued)

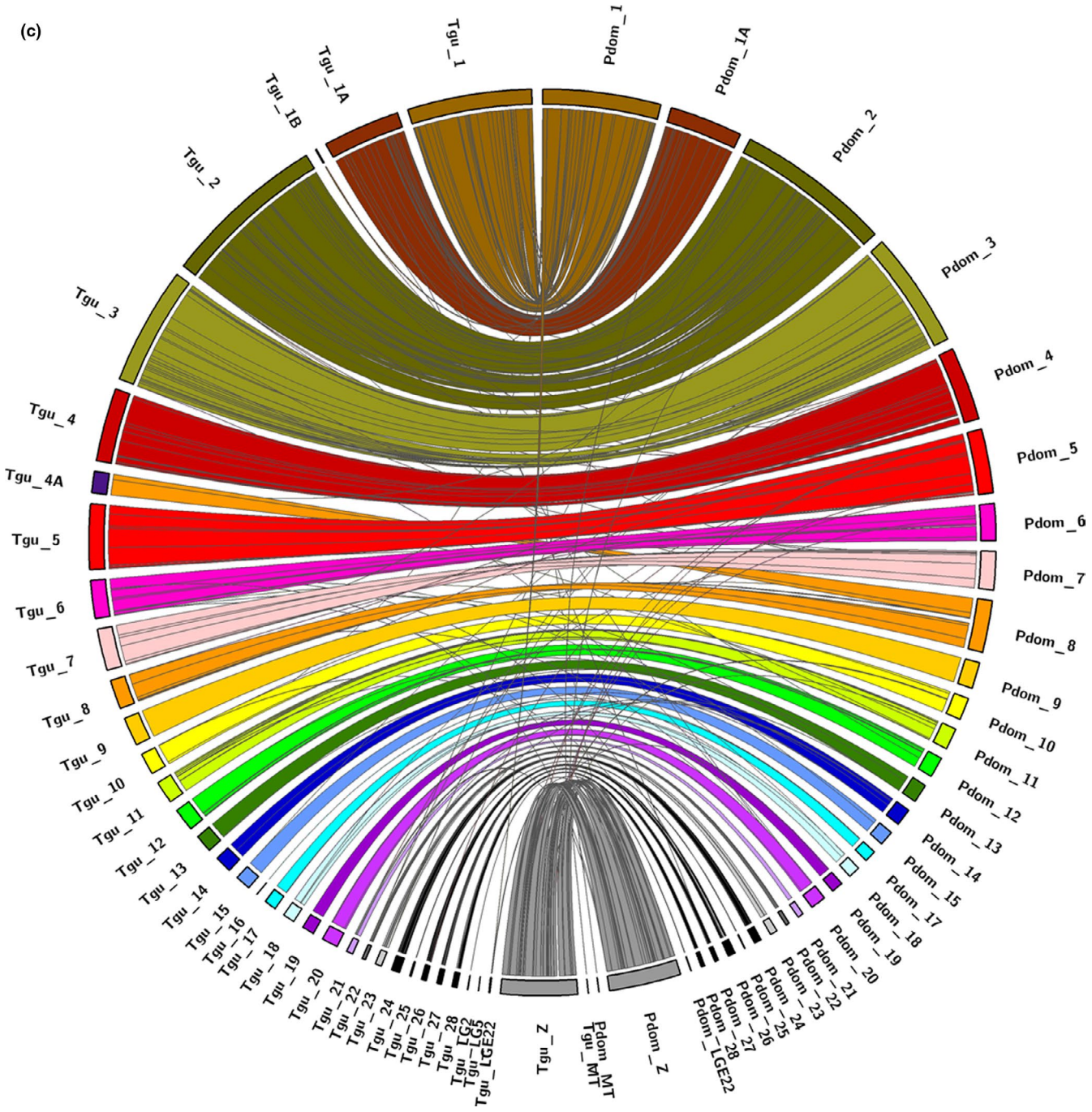


**FIGURE 6** (Continued)

in these species (Pala et al., 2012). In the current study, we have shown that house sparrow lacks the corresponding chromosome 4A, due to a fusion of 4A with chromosome 8 (Figure 6). These results enlighten the interesting evolutionary history of the avian chromosome 4.

As expected, when the house sparrow linkage map and reference genome are compared to that of other species, several intrachromosomal inversions are found (see Table S3). Only three chromosomes (6, 9 and 17) have a similar and conserved marker order along the chromosomes across species. This lends support to the findings in Kawakami et al. (2014) where the same chromosomes in collared flycatcher were involved with few inversions relative to the rest of the genome when compared to chicken and zebra finch. Divergence times between house sparrow and chicken, collared flycatcher and zebra finch are approximately 100, 40 and 35 MYA, respectively (Griffin et al., 2007; Tuinen et al., 2000). When the house sparrow genome is compared to the three species, we find that approximately 50% of the genome is involved with intrachromosomal inversions in all three cases, despite differing divergence times.

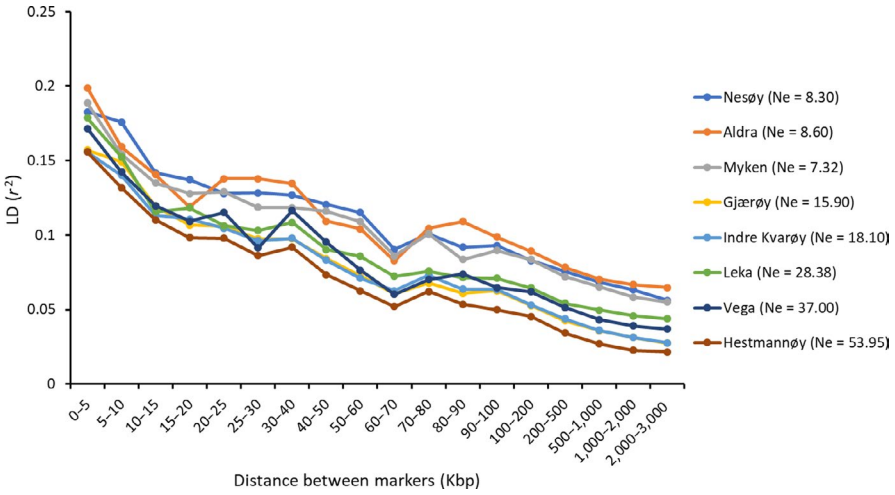


**FIGURE 7** Linkage disequilibrium ($r^2$) in eight insular house sparrow populations of different effective population sizes ($N_e$). LD was estimated within 5 Kbp windows from 0–30 Kbp, 10 Kbp window from 30–100 Kbp, from 100–200 Kbp, 200–500 Kbp and 500–1,000 Kbp and in 1,000 Kbp windows from 1,000–3,000 Kbp [Colour figure can be viewed at wileyonlinelibrary.com]

Within Passeriformes, most research suggests higher recombination rates in females. The most extreme heterochiasmy has been detected in the great reed warbler, with female recombination rate being 2.15 times higher than in males (Hansson et al., 2005). Female recombination rate also exceeds that of males in blue tits with a ratio of 1.18 (Hansson et al., 2009), and in zebra finches with a ratio of 1.02 (Backström et al., 2010). In the current study, recombination rate was found to be 1.24 times higher in female house sparrows compared to males. Interestingly, Kawakami et al. (2014) found the opposite in collared flycatchers, where males recombine more frequently than females, with a ratio of 1.1, whilst little genome wide heterochiasmy was detected in great tit (van Oers et al., 2014). Heterochiasmy is found in a wide variety of taxa, with no clear pattern emerging and it is intriguing to see that also within Passeriformes, opposite trends appear when more species are investigated. The evolutionary benefit of heterochiasmy remains unresolved and several hypotheses have been proposed (for review, see Hedrick, 2007 and references therein). For instance, it has been suggested that pleiotropy (Haldane, 1922) and haploid selection (Lenormand, 2003) should lead to decreased recombination rates in the heterogameic sex. Our results, where female are the heterogameic sex, do not support these hypotheses. Moreover, Trivers (1988) suggested that in species where one sex experiences more variation in reproductive success than the other, the sex with higher variation should recombine less to maintain successful haplotypes. In house sparrows, the difference in sex specific variance in reproductive success is small, but slightly male biased (Jensen et al., 2004). Thus, based on this scenario, males are expected to experience a relatively small evolutionary advantage by recombining less. An alternative hypothesis predicts that there should be higher recombination rate in the sex with the most variation in reproductive success, because an elevated recombination rate tends to coevolve with any trait under selection (Korol & Iliadi, 1994). The overall lower recombination rate seen in house sparrow males lends little support to this hypothesis, unless specific chromosomes are considered. The most pronounced heterochiasmy was seen on the macrochromosomes, and on chromosomes Pdom 2–8 in particular, where females had a recombination rate 1.4 times higher than males. Interestingly, the reverse pattern was seen for the smallest chromosomes Pdom 25 and Pdom 28, respectively (Table 2). Directional selection on a trait can increase recombination (Otto & Lenormand, 2002) and it is possible that the aforementioned microchromosomes represent candidate regions that are particularly affected by sexually antagonistic selection (van Oers et al., 2014).

Based on already established research, our results on linkage disequilibrium are in concordance with expectations: LD drops relatively quickly, the populations with comparably low $N_e$ display higher LD for loci in close proximity of each other (0–5 Kbp), and populations with low $N_e$ also have higher levels of LD at longer intermarker distances (Figure 7). The LD estimates for Leka and Vega are overall higher compared to the other populations considering their effective population sizes. It is possible that this is a result of selection experiments for body size that were carried out on Leka and Vega during the years 2002–2005 (Kvalnes et al., 2017). Strong selection is expected to increase LD (De La Vega et al., 2005). Based on these LD analyses, it is apparent that in order to detect associations between a phenotypic trait and a locus in these house sparrow populations, the SNP markers should be spaced evenly along the genome with intervals of less than approximately 5–10 Kbp (Figure 7). If possible, the markers should be spaced closer in regions of higher recombination, such as towards the ends of the macrochromosomes. With the current 10K Illumina SNP array, markers are spaced at an average interval of 18 Kbp, at which point the LD has dropped to 0.12 (averaged across all eight populations) and is therefore likely to miss associations with traits where the causative locus is not in proximity of a marker. The house sparrow populations on which the linkage map is based has also been used as focal populations for two different GWA studies to map genes for morphology in house sparrows. These GWA studies made use of the same 6,498 SNPs as in the linkage map (Silva et al., 2017) and a denser panel of ca. 183,000 SNPs (Lundregan et al., 2018), respectively. More significant hits were found when the denser panel of SNPs and an average intermarker distance ca. 6 Kbp was used (Lundregan et al., 2018), thus illustrating our conclusion that an average intermarker distance of 18 Kbp is not dense enough to detect associations in GWA studies given the level of LD in the house sparrow populations used in these studies.

We have presented a genome wide, medium-density linkage map for an ecological and genetic model species, the house sparrow. These data will greatly facilitate ongoing association studies in wild populations and aid our understanding of evolutionary adaptations in a Passerine species. Moreover, we have characterized a novel interchromosomal translocation, and thus added to the body of research that elucidates the evolution of the avian genome.

## ACKNOWLEDGEMENTS

We thank fieldworkers and students for their contributions to fieldwork, and Randi Røsbak for her assistance with laboratory analyses. Thanks also to Arild Husby, Takeshi Kawakami, Anna Santure, and Helle T. Baalsrud for fruitful discussions. Finally, we extend our gratitude to the residents of the study system, whose cooperation and friendly attitude has been invaluable. This study was supported by grants from NTNU, The Norwegian Research Council (programmes STORFORSK, Strategic University Program in Conservation Biology, projects 191847 and 221956) The Norwegian Directorate for Nature Management, and The EU-commission (project METABIRD). This work was also partly supported by the Research Council of Norway through its Centres of Excellence funding scheme, project 223257). The fieldwork was carried out in accordance with permits from the Ringing Centre at Stavanger Museum, Norway. Whole genome resequencing of house sparrows was carried out at Genome Quebec, McGill University, Canada. Genotyping on the custom house sparrow Illumina iSelect 10K SNP array was carried out at the Genomics Core Facility, NTNU.

## AUTHOR CONTRIBUTIONS

I.J.H. prepared the samples for genotyping, conducted statistical analysis and wrote the paper with input from S.L., A.M.B,, A.K.N., M.T., J.S., and H.J., S.L. conducted the linkage map analysis. I.J.H., H.J., A.M.B., T.O.E., C.T., and G.P.-S. contributed to the development of the linkage map. A.K.N., and M.T. conducted statistical analysis. H.J. designed the study and is the senior scientist.

## DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available in the Dryad Digital Repository with the identifier https://doi.org/10.5061/dryad.00000000n.

## ORCID

*Ingerid J. Hagen* https://orcid.org/0000-0003-1028-3940
*Sigbjørn Lien* https://orcid.org/0000-0002-5159-1950
*Alina K. Niskanen* https://orcid.org/0000-0003-2017-2718
*Maja Tarka* https://orcid.org/0000-0001-9651-1101
*Jon Slate* https://orcid.org/0000-0003-3356-5123
*Henrik Jensen* https://orcid.org/0000-0001-7804-1564

## REFERENCES

Aguilera, A., & Gomez-Gonzalez, B. (2008). Genome instability: A mechanistic view of its causes and consequences. *Nature Reviews Genetics*, *9*(3), 204–217.

Åkesson, M., Hansson, B., Hasselquist, D., & Bensch, S. (2007). Linkage mapping of AFLP markers in a wild population of great reed warblers: Importance of heterozygosity and number of genotyped individuals. *Molecular Ecology*, *16*(11), 2189–2202. https://doi.org/10.1111/j.1365-294X.2007.03290.x

Akhunov, E. D., Goodyear, A. W., Geng, S., Qi, L.-L., Echalier, B., Gill, B. S., … Dvorak, J. (2003). The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. *Genome Research*, *13*(5), 753–763. https://doi.org/10.1101/gr.808603

Anderson, T. R. (Ed.) (2006). *Biology of the ubiquitous house sparrow: From genes to populations*. New York, NY: Oxford University Press.

Baalsrud, H. T., Sæther, B.-E., Hagen, I. J., Myhre, A. M., Ringsby, T. H., Pärn, H., & Jensen, H. (2014). Effects of population characteristics and structure on estimates of effective population size in a house sparrow metapopulation. *Molecular Ecology*, *23*(11), 2653–2668. https://doi.org/10.1111/mec.12770

Backstrom, N., Forstmeier, W., Schielzeth, H., Mellenius, H., Nam, K., Bolund, E., … 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

Backström, N., Karaiskou, N., Leder, E. H., Gustafsson, L., Primmer, C. R., Qvarnström, A., & Ellegren, H. (2008). A gene-based genetic linkage map of the collared flycatcher (*Ficedula albicollis*) reveals extensive Synteny and gene-order conservation during 100 million years of avian evolution. *Genetics*, *179*(3), 1479–1495. https://doi.org/10.1534/genetics.108.088195

Billing, A. M., Lee, A. M., Skjelseth, S., Borg, Å. A., Hale, M. C., Slate, J., … Jensen, H. (2012). Evidence of inbreeding depression but not inbreeding avoidance in a natural house sparrow population. *Molecular Ecology*, *21*(6), 1487–1499. https://doi.org/10.1111/j.1365-294X.2012.05490.x

Bolívar, P., Mugal, C. F., Nater, A., & Ellegren, H. (2016). Recombination rate variation modulates gene sequence evolution mainly via GC-biased gene conversion, not Hill-Robertson interference, in an avian system. *Molecular Biology and Evolution*, *33*(1), 216–227. https://doi.org/10.1093/molbev/msv214

Bulatova, N. S., Radjabli, S. I., & Panov, E. N. (1972). Karyological description of three species of the genus Passer. *Experientia*, *28*(11), 1369–1371. https://doi.org/10.1007/bf01965351

Copeland, N., Jenkins, N., Gilbert, D., Eppig, J., Maltais, L., Miller, J., … et, A. L. (1993). A genetic linkage map of the mouse: Current applications and future prospects. *Science*, *262*(5130), 57. https://doi.org/10.1126/science.8211130

Crawford, A. M., Dodds, K. G., Ede, A. J., Pierson, C. A., Montgomery, G. W., Garmonsway, H. G., … Kappes, S. W. (1995). An autosomal genetic linkage map of the sheep genome. *Genetics*, *140*(2), 703–724.

Dawson, D. A., Horsburgh, G. J., Krupa, A. P., Stewart, I. R. K., Skjelseth, S., Jensen, H., … Burke, T. (2012). Microsatellite resources for Passeridae species: A predicted microsatellite map of the house sparrow *Passer domesticus*. *Molecular Ecology Resources*, *12*(3), 501–523. https://doi.org/10.1111/j.1755-0998.2012.03115.x

De La Vega, F. M., Isaac, H., Collins, A., Scafe, C. R., Halldórsson, B. V., Su, X., … Gilbert, D. A. (2005). The linkage disequilibrium maps of three human chromosomes across four populations reflect their demographic history and a common underlying recombination pattern. *Genome Research*, *15*(4), 454–462. https://doi.org/10.1101/gr.3241705

Derjusheva, S., Kurganova, A., Habermann, F., & Gaginskaya, E. (2004). High chromosome conservation detected by comparative chromosome painting in chicken, pigeon and passerine birds. *Chromosome Research*, *12*(7), 715. https://doi.org/10.1023/B:CHRO.0000045779.50641.00

Elgvin, T. O., Trier, C. N., Tørresen, O. K., Hagen, I. J., Lien, S., Nederbragt, A. J., … Sætre, G.-P. (2017). The genomic mosaicism

of hybrid speciation. *Science Advances*, *3*(6), e1602996–https://doi.org/10.1126/sciadv.1602996

Ellegren, H. (2010). Evolutionary stasis: The stable chromosomes of birds. *Trends in Ecology & Evolution*, *25*(5), 283–291. https://doi.org/10.1016/j.tree.2009.12.004

Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, *29*(1), 51–63. https://doi.org/10.1016/j.tree.2013.09.008

Ellegren, H., Smeds, L., Burri, R., Olason, P. I., Backstrom, N., Kawakami, T., … Wolf, J. B. W. (2012). The genomic landscape of species divergence in Ficedula flycatchers. *Nature*, *491*(7426), 756–760.

Engen, S., Lande, R., Sæther, B. E., & Gienapp, P. (2010). Estimating the ratio of effective to actual size of an age-structured population from individual demographic data. *Journal of Evolutionary Biology*, *23*(6), 1148–1158. https://doi.org/10.1111/j.1420-9101.2010.01979.x

Fierst, J. L. (2015). Using linkage maps to correct and scaffold de novo genome assemblies: Methods, challenges, and computational tools. *Frontiers in Genetics*, *6*, 220. https://doi.org/10.3389/fgene.2015.00220

Grabherr, M. G., Russell, P., Meyer, M., Mauceli, E., Alföldi, J., Di Palma, F., & Lindblad-Toh, K. (2010). Genome-wide synteny through highly sensitive sequence alignment: Satsuma. *Bioinformatics*, *26*(9), 1145–1151. https://doi.org/10.1093/bioinformatics/btq102

Green, P., Falls, K., & Crooks, S. (1990). *Documentation for CRI-MAP*. St. Louis, MO: Washington University. Version 2.4.

Griffin, D. K., Robertson, L. B. W., Tempest, H. G., & Skinner, B. M. (2007). The evolution of the avian genome as revealed by comparative molecular cytogenetics. *Cytogenetic and Genome Research*, *117*(1–4), 64–77. https://doi.org/10.1159/000103166

Groenen, M. A. M., Cheng, H. H., Bumstead, N., Benkel, B. F., Briles, W. E., Burke, T., … Vignal, A. (2000). A consensus linkage map of the chicken genome. *Genome Research*, *10*(1), 137–147. https://doi.org/10.1101/gr.10.1.137

Groenen, M. A., Wahlberg, P., Foglio, M., Cheng, H. H., Megens, H.-J., Crooijmans, R. P., … Wong, G. K.-S. (2009). A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Research*, *19*(3), 510–519.

Hagen, I. J., Billing, A. M., Rønning, B., Pedersen, S. A., Pärn, H., Slate, J., & Jensen, H. (2013). The easy road to genome-wide medium density SNP screening in a non-model species: Development and application of a 10 K SNP-chip for the house sparrow (*Passer domesticus*). *Molecular Ecology Resources*, *13*(3), 429–439. https://doi.org/10.1111/1755-0998.12088

Haldane, J. B. S. (1922). Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, *12*(2), 101–109. https://doi.org/10.1007/BF02983075

Hansson, B., Åkesson, M., Slate, J., & Pemberton, J. M. (2005). Linkage mapping reveals sex-dimorphic map distances in a passerine bird. *Proceedings of the Royal Society of London B: Biological Sciences*, *272*(1578), 2289–2298. https://doi.org/10.1098/rspb.2005.3228

Hansson, B., Ljungqvist, M., Dawson, D. A., Mueller, J. C., Olano-Marin, J., Ellegren, H., & Nilsson, J. A. (2009). Avian genome evolution: Insights from a linkage map of the blue tit (*Cyanistes caeruleus*). *Heredity*, *104*(1), 67–78.

Hedrick, P. W. (2007). Sex: Differences in mutation, recombination, selection, gene flow, and genetic drift. *Evolution*, *61*(12), 2750–2771. https://doi.org/10.1111/j.1558-5646.2007.00250.x

Hunt, M., Kikuchi, T., Sanders, M., Newbold, C., Berriman, M., & Otto, T. D. (2013). REAPR: A universal tool for genome assembly evaluation. *Genome Biology*, *14*(5), R47. https://doi.org/10.1186/gb-2013-14-5-r47

Itoh, Y., & Arnold, A. (2005). Chromosomal polymorphism and comparative painting analysis in the zebra finch. *Chromosome Research*, *13*(1), 47–56. https://doi.org/10.1007/s10577-005-6602-x

Jensen, H., Sæther, B.-E., Ringsby, T. H., Tufto, J., Griffith, S. C., & Ellegren, H. (2004). Lifetime reproductive success in relation to morphology in the house sparrow *Passer domesticus*. *Journal of Animal Ecology*, *73*(4), 599–611. https://doi.org/10.1111/j.0021-8790.2004.00837.x

Johnston, S. E., Huisman, J., Ellis, P. A., & Pemberton, J. M. (2017). A high density linkage map reveals sexual dimorphism in recombination landscapes in red deer (*Cervus elaphus*). *G3: Genes|genomes|genetics*, *7*, 2859–2870.

Kaback, D. B., Guacci, V., Barber, D., & Mahon, J. W. (1992). Chromosome size-dependent control of meiotic recombination. *Science*, *256*(5054), 228. https://doi.org/10.1126/science.1566070

Kalinowski, S. T., Taper, M. L., & Marshall, T. C. (2007). Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, *16*(5), 1099–1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x

Kawakami, T., Mugal Carina, 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., … 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

Korol, A. B., & Iliadi, K. G. (1994). Increased recombination frequencies resulting from directional selection for geotaxis in Drosophila. *Heredity*, *72*(1), 64–68. https://doi.org/10.1038/hdy.1994.7

Kosambi, D. D. (1943). The estimation of map distances from recombination values. *Annals of Eugenics*, *12*(1), 172–175. https://doi.org/10.1111/j.1469-1809.1943.tb02321.x

Kruglyak, L. (1999). Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genetics*, *22*(2), 139–144. https://doi.org/10.1038/9642

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., … Marra, M. A. (2009). Circos: An information aesthetic for comparative genomics. *Genome Research*, *19*(9), 1639–1645. https://doi.org/10.1101/gr.092759.109

Kvalnes, T., Ringsby, T. H., Jensen, H., Hagen, I. J., Rønning, B., Pärn, H., … Saether, B.-E. (2017). Reversal of response to artificial selection on body size in a wild passerine. *Evolution*, *71*(8), 2062–2079. https://doi.org/10.1111/evo.13277

Lenormand, T. (2003). The evolution of sex dimorphism in recombination. *Genetics*, *163*(2), 811–822.

Li, H. (2013). *Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM*. arXiv.

Lien, S., Gidskehaug, L., Moen, T., Hayes, B. J., Berg, P. R., Davidson, W. S., … Kent, M. P. (2011). A dense SNP-based linkage map for Atlantic salmon (*Salmo salar*) reveals extended chromosome homeologies and striking differences in sex-specific recombination patterns. *BMC Genomics*, *12*(1), 615. https://doi.org/10.1186/1471-2164-12-615

Lindblad-Toh, K., Wade, C. M., Mikkelsen, T. S., Karlsson, E. K., Jaffe, D. B., Kamal, M., … Lander, E. S. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, *438*, 803. https://doi.org/10.1038/nature04338

Lundregan, S. L., Hagen, I. J., Gohli, J., Niskanen, A. K., Kemppainen, P., Ringsby, T. H., … Jensen, H. (2018). Inferences of genetic architecture of bill morphology in house sparrow using a high-density SNP array point to a polygenic basis. *Molecular Ecology*, *27*(17), 3498–3514. https://doi.org/10.1111/mec.14811

Nietlisbach, P., Camenisch, G., Bucher, T., Slate, J., Keller, L. F., & Postma, E. (2015). A microsatellite-based linkage map for song sparrows (*Melospiza melodia*). *Molecular Ecology Resources*, *15*(6), 1486–1496. https://doi.org/10.1111/1755-0998.12414

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

Pala, I., Naurin, S., Stervander, M., Hasselquist, D., Bensch, S., & Hansson, B. (2012). Evidence of a neo-sex chromosome in birds. *Heredity*, *108*(3), 264–272. http://www.nature.com/hdy/journal/v108/n3/suppinfo/hdy201170s1.html

Pärn, H., Ringsby, T. H., Jensen, H., & Sæther, B.-E. (2012). Spatial heterogeneity in the effects of climate and density-dependence on dispersal in a house sparrow metapopulation. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1726), 144–152. https://doi.org/10.1098/rspb.2011.0673

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*, 559–575. https://doi.org/10.1086/519795

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rodionov, A. V. (1996). Micro vs. macro: Structural-functional organization of avian micro- and macrochromosomes. *Genetika*, *32*(5), 597–608.

Silva, C. N. S., McFarlane, S. E., Hagen, I. J., Rönnegård, L., Billing, A. M., Kvalnes, T., … Husby, A. (2017). Insights into the genetic architecture of morphological traits in two passerine bird species. *Heredity*, *119*, 197. https://doi.org/10.1038/hdy.2017.29

Singhal, S., Leffler, E. M., Sannareddy, K., Turner, I., Venn, O., Hooper, D. M., … Przeworski, M. (2015). Stable recombination hotspots in birds. *Science*, *350*(6263), 928–932. https://doi.org/10.1126/science.aad0843

Slate, J., Santure, A. W., Feulner, P. G. D., Brown, E. A., Ball, A. D., Johnston, S. E., & Gratten, J. (2010). Genome mapping in intensively studied wild vertebrate populations. *Trends in Genetics*, *26*(6), 275–284. https://doi.org/10.1016/j.tig.2010.03.005

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

Stubberud, M. W., Myhre, A. M., Holand, H., Kvalnes, T., Ringsby, T. H., Sæther, B.-E., & Jensen, H. (2017). Sensitivity analysis of effective population size to demographic parameters in house sparrow populations. *Molecular Ecology*, *26*(9), 2449–2465. https://doi.org/10.1111/mec.14057

Trivers, R. (1988). Sex differences in rates of recombination and sexual selection. In R. L. Michod & B. R. Levin (Eds.). *The evolution of sex* (pp. 270–286). Sunderland, MA: Sinauer Press Sunderland.

Tuinen, M. V., Sibley, C. G., & Hedges, S. B. (2000). The early history of modern birds inferred from DNA sequences of nuclear and mitochondrial ribosomal genes. *Molecular Biology and Evolution*, *17*(3), 451–457. https://doi.org/10.1093/oxfordjournals.molbev.a026324

van Oers, K., Santure, A. W., De Cauwer, I., van Bers, N. E., Crooijmans, R. P., Sheldon, B. C., … 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*, 307. https://doi.org/10.1038/hdy.2013.107

Warren, W. C., Hillier, L. D. W., Tomlinson, C., Minx, P., Kremitzki, M., Graves, T., … Cheng, H. H. (2017). A new chicken genome assembly provides insight into avian genome structure. *G3: Genes|genomes|genetics*, *7*(1), 109. https://doi.org/10.1534/g3.116.035923

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., … Lathrop, M. (1992). A second-generation linkage map of the human genome. *Nature*, *359*, 794. https://doi.org/10.1038/359794a0

## SUPPORTING INFORMATION

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

**How to cite this article:** Hagen IJ, Lien S, Billing AM, et al. A genome-wide linkage map for the house sparrow (*Passer domesticus*) provides insights into the evolutionary history of the avian genome. *Mol Ecol Resour*. 2020;20:544–559. https://doi.org/10.1111/1755-0998.13134