Received: 13 December 2023 | Revised: 10 April 2024 | Accepted: 29 April 2024

DOI: 10.1111/1755-0998.13969

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

The great tit HapMap project: A continental-scale analysis of genomic variation in a songbird

Lewis G. Spurgin[1,2] | Mirte Bosse[3,4,5] | Frank Adriaensen[6] | Tamer Albayrak[7,8] | Christos Barboutis[9] | Eduardo Belda[10] | Andrey Bushuev[11] | Jacopo G. Cecere[12] | Anne Charmantier[13] | Mariusz Cichon[14] | Niels J. Dingemanse[15] | Blandine Doligez[16,17] | Tapio Eeva[18] | Kjell Einar Erikstad[19] | Vyacheslav Fedorov[11] | Matteo Griggio[20] | Dieter Heylen[2,6,21] | Sabine Hille[22] | Camilla A. Hinde[23] | Elena Ivankina[24] | Bart Kempenaers[25] | Anvar Kerimov[11] | Milos Krist[26] | Laura Kvist[27] | Veronika N. Laine[3,28] | Raivo Mänd[29] | Erik Matthysen[6] | Ruedi Nager[30] | Boris P. Nikolov[31] | Ana Claudia Norte[32] | Markku Orell[27] | Jenny Ouyang[33] | Gergana Petrova-Dinkova[31] | Heinz Richner[34] | Diego Rubolini[35] | Tore Slagsvold[36] | Vallo Tilgar[29] | János Török[37] | Barbara Tschirren[38] | Csongor I. Vágási[39] | Teru Yuta[40] | Martien A. M. Groenen[4] | Marcel E. Visser[3,4,41] | Kees van Oers[3] | Ben C. Sheldon[2] | Jon Slate[42]

Correspondence
Jon Slate, School of Biosciences, University of Sheffield, Sheffield, UK.
Email: j.slate@sheffield.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: NE/J012599/1; European Research Council, Grant/Award Number: 202487 and 339092; Biotechnology and Biological Sciences Research Council, Grant/Award Number: BB/N011759/1

Handling Editor: David Coltman

Abstract

A major aim of evolutionary biology is to understand why patterns of genomic diversity vary within taxa and space. Large-scale genomic studies of widespread species are useful for studying how environment and demography shape patterns of genomic divergence. Here, we describe one of the most geographically comprehensive surveys of genomic variation in a wild vertebrate to date; the great tit (*Parus major*) HapMap project. We screened *ca* 500,000 SNP markers across 647 individuals from 29 populations, spanning ~30 degrees of latitude and 40 degrees of longitude – almost the entire geographical range of the European subspecies. Genome-wide variation was consistent with a recent colonisation across Europe from a South-East European refugium, with bottlenecks and reduced genetic diversity in island populations. Differentiation across the genome was highly heterogeneous, with clear 'islands of differentiation', even among populations with very low levels of genome-wide differentiation. Low local recombination rates were a strong predictor of high local genomic differentiation (F_{ST}), especially in island and peripheral mainland populations, suggesting that the interplay between genetic drift and recombination causes highly

Deceased: Vyacheslav Fedorov and Matteo Griggio

For affiliations refer to page 10.

heterogeneous differentiation landscapes. We also detected genomic outlier regions that were confined to one or more peripheral great tit populations, probably as a result of recent directional selection at the species' range edges. Haplotype-based measures of selection were related to recombination rate, albeit less strongly, and highlighted population-specific sweeps that likely resulted from positive selection. Our study highlights how comprehensive screens of genomic variation in wild organisms can provide unique insights into spatio-temporal evolutionary dynamics.

KEYWORDS
adaptation, birds, ecological genetics, genomics/proteomics, molecular evolution, population genetics – empirical

1 | INTRODUCTION

Since the first studies of allozyme variation in humans (Harris, 1966) and *Drosophila* (Hubby & Lewontin, 1966; Lewontin & Hubby, 1966), there has been great interest in explaining how evolutionary and ecological processes shape the patterns of genetic variation observed within and among natural populations. One focus of research and debate in this area has been on quantifying the roles of adaptive and neutral processes in explaining observed levels of genetic variation (Nei, 2005). However, adaptation does not just occur in isolation, but acts on genetic variation that is also shaped by mutation, recombination, gene flow and genetic drift (Hartl & Clark, 1997; Hedrick, 2005). More recently there has been increased effort in understanding how these fundamental evolutionary forces operate in concert to generate and maintain the levels of genetic diversity commonly observed in natural populations (Chen et al., 2017; Feder et al., 2012; Lanfear et al., 2014).

The increasing feasibility of high-throughput sequencing and genotyping, alongside subsequent characterisation of genome-wide variation across large numbers of individuals, has revealed that at the genomic level, patterns of variation and divergence among natural populations and species are highly heterogeneous (Turner et al., 2005). A key feature of these 'genomic landscapes' of divergence that has received particular attention is the presence of so-called 'islands of differentiation': outlier regions of the genome with high levels of divergence estimated from statistics such as F_{ST} or d_{xy} that consistently emerge at the same loci in different comparisons between populations or related species (Nadeau et al., 2012; Poelstra et al., 2014; Renaut et al., 2013; Turner et al., 2005). Initially these regions were termed 'islands of speciation', and they were thought to arise as a result of reduced gene flow in genomic regions associated with reproductive isolation (Turner et al., 2005; Wolf & Ellegren, 2017). Subsequent research has revealed that highly heterogeneous patterns of genomic divergence can occur even in the complete absence of gene flow, as a result of recombination rate variation and linked selection (Booker et al., 2022; Cruickshank & Hahn, 2014; Noor & Bennett, 2009). In genomic regions of low recombination, selection for beneficial mutations (positive selection), or against deleterious mutations (background selection), will impact relatively large genomic regions as a result of elevated levels of linkage disequilibrium (LD) among sites. Selection within these regions reduces diversity within populations, and increases levels of differentiation among them, resulting in 'islands' of increased differentiation that persist over evolutionary time (Cruickshank & Hahn, 2014; Johri et al., 2020; Turner & Hahn, 2010). Another, less well-explored reason by which islands of divergence can arise is due to the differential effects of genetic drift in response to variation in effective population size across different genomic regions; something that may be particularly important in recently colonised populations (Campagna et al., 2015; Ma et al., 2018). These circumstances promote fixation of haplotypes and therefore result in either reduced or inflated local differentiation.

Comparing patterns of genomic differentiation among sets of populations or species at different stages of the divergence/speciation continuum is a powerful way of disentangling the forces that shape variation among populations. For example, across multiple *Heliconius* butterfly populations and species, patterns of genomic variation are shaped by a combination of gene flow and selection, particularly in genomic regions harbouring genes involved in wing patterning (Martin et al., 2013). In contrast, in *Helianthus* sunflowers, genomic architecture is the main driver of genomic differentiation across sets of populations (Renaut et al., 2013). Similarly, recent research in birds has revealed that differentiation landscapes are conserved across populations, species and even across avian families, with the same islands of differentiation arising among populations of distantly related species (Burri et al., 2015; Chase et al., 2021; Van Doren et al., 2017; Vijay et al., 2017). This latter pattern appears to have arisen, at least in part, as a result of a highly conserved synteny and recombination landscape in birds (Bravo et al., 2021; Zhang et al., 2014), with background selection in regions of low recombination producing recurrent islands of differentiation (Booker et al., 2020; Burri, 2017).

It is now clear that the recombination landscape and linked selection are key drivers of genomic variation within and among populations. However, we are only just beginning to understand how this linked selection interacts with other evolutionary forces to shape patterns of differentiation across natural populations and

species (Burri, 2017; Ellegren & Wolf, 2017; Jiang et al., 2023; Jiggins & Martin, 2017; Lohse, 2017; Perrier & Charmantier, 2019; Ravinet et al., 2017). A large-scale analysis of three-spined sticklebacks (*Gasterosteus aculeatus*) showed that islands of differentiation were more likely to arise in low recombination regions when gene flow occurred between populations (Samuk et al., 2017). There is also a significant impact of divergence time; in recently separated populations the differentiation landscape is most likely to reflect selective sweeps. Then, as divergence accumulates, genomic architecture is expected to play an increasingly important role in generating these genomic islands (Burri, 2017).

Because of their large effective population sizes and ecologically varied ranges, widespread continental species are excellent models for studying how demography and the environment shape genetic and phenotypic variation among populations. Insight into the evolutionary history of such species can be gained if genetic variation is characterised across much of its geographical range. Cross-population comparisons of genetic variation can then be utilised to make inferences about phylogeography, levels of gene flow between populations and how adaptation to different environmental and ecological conditions occurs (Perrier et al., 2020). The first large-scale studies were performed in humans – that is, the HapMap Projects (International HapMap, 2005; International HapMap et al., 2007, 2010) which characterised human genetic variation on different continents, with a view to determining the feasibility of association mapping studies. Similar studies have been conducted in domesticated species and their wild ancestors (Bovine HapMap et al., 2009; Chia et al., 2012; Kijas et al., 2012; Parejo et al., 2023), and in model organisms (Kirby et al., 2010; Lindblad-Toh et al., 2005). More recently, there is a growing appreciation that HapMap-type studies are useful for studying signatures of selection and adaptation in natural populations of species with large effective population sizes and high levels of gene flow (Alonso-Blanco et al., 2016; Horton et al., 2012; Mueller et al., 2018; Williamson et al., 2014).

The European great tit (*Parus major major*) is an excellent model for ecological and evolutionary studies (Gosler, 1993). As is the case with several avian species which are amenable to long-term study (Culina et al., 2021), a wealth of ecological data exists across multiple great tit populations (Charmantier et al., 2008; Dingemanse et al., 2012; Visser et al., 1998), enabling informed hypotheses about selection to be tested in this system. Phylogeographic research using mitochondrial DNA suggests that this species has experienced post-glacial range expansion through Central and Northern Europe, possibly from a single refugium in South-East Europe (Kvist et al., 1999). Most contemporary populations are characterised by large effective population sizes and low levels of genetic differentiation (Kvist et al., 2003; Lemoine et al., 2016). However, these previous cross-population molecular studies have relied on a modest number of microsatellite loci and mitochondrial DNA, making the detection of genomic regions under selection impossible. The genome of the great tit has been sequenced (Laine et al., 2016), and a high density panel of *ca* 500,000 SNP markers has been developed (Kim et al., 2018). A study of two European populations using this marker panel suggests that rapid adaptation has occurred at the genomic and phenotypic levels, with pronounced selection on morphology (Bosse et al., 2017).

Here, we perform a HapMap study of 647 unrelated individuals across 29 populations (Figure 1), to examine how genomic architecture, natural selection and population history have shaped patterns

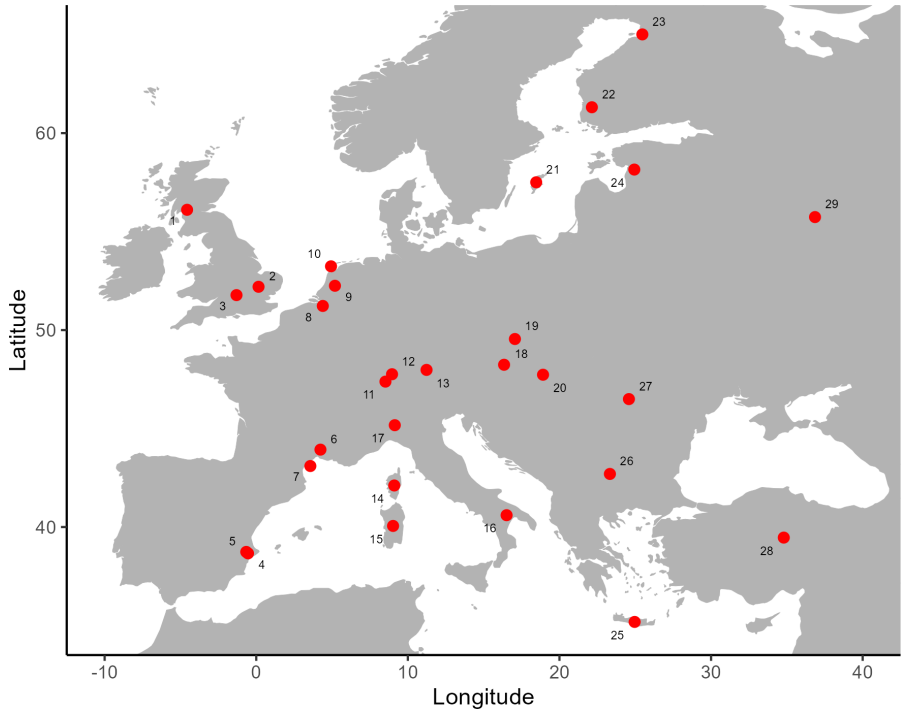

FIGURE 1 Sampling locations of great tit populations. Population names and sample sizes are given in Table S1, and numbers on the map correspond to the 'code' column in Table S1.

of genomic variation across recently colonised European great tit populations. Using a large SNP panel typed across all individuals, we first characterise genome-wide patterns of variation within and among populations, in order to infer population phylogeography. We then examine how variation is partitioned across the genome and test the hypothesis that highly divergent genomic regions have arisen in genomic regions of low recombination (Cruickshank & Hahn, 2014; Noor & Bennett, 2009; Perrier & Charmantier, 2019). Finally, we examine how genomic divergence accumulates along the colonisation route of this species, with the aim of inferring how recent natural selection and demography drive variation across the genome.

2 | MATERIALS AND METHODS

2.1 | Sampling and molecular methods

Samples were collected from 29 populations from 22 regions across Europe (Figure 1; Table S1). Samples were pooled into regions either based on geographical proximity (e.g. Cambridge and Wytham Woods) or on sample size (e.g. Romania and Bulgaria). An exploratory analysis considering all sampled populations separately yielded virtually identical results to those shown here, and in no cases did we observe substructure within pooled populations in our Admixture analyses (Figure S3).

Birds were trapped during the breeding season (April–May) from nest boxes, or using mist nets, and ringed with a uniquely numbered aluminium ring. Blood was taken via brachial or tarsal venipuncture and stored in either 1 mL Cell Lysis Solution (Gentra Puregene Kit, Qiagen, USA), Queen's buffer, or absolute ethanol. All samples were genotyped using a custom-made Affymetrix® great tit 650 K SNP chip at Edinburgh Genomics (Edinburgh, UK), following the approaches described elsewhere (Kim et al., 2018).

2.2 | Analyses

Unless stated otherwise, all population genetic statistics were calculated in PLINK version 1.9 (Purcell et al., 2007), and downstream analysis and plotting was carried out in R version 3.3 (R Core Team, 2013). For population genetic analyses, we used the filtering approaches outlined previously (Bosse et al., 2017). Briefly, we randomly removed individuals from pairs with relatedness values >0.4, and for demographic analyses we used a LD-pruned dataset (based on VIF > 0.2), with SNPs associated with an inversion on chromosome 1A (da Silva et al., 2019) removed. After filtering, a total of 647 samples (mean = 22.3 birds per population, range = 3–50) typed at 483,888 SNPs were retained for analysis.

In each population, we estimated LD (R^2) for each pair of markers within 50 kb on the same chromosome and compared this to physical distance between marker pairs. We calculated observed heterozygosity for each SNP and population using a reduced SNP dataset, which was pruned based on LD to remove all markers with $R^2 > 0.1$, then thinned with a probability of retaining each variant of 0.25. We calculated genome-wide (mean) F_{ST} between each pair of populations using the pruned and thinned dataset described above. Pairwise F_{ST} (linearised as $F_{ST}/(1-F_{ST})$) was compared to (natural logarithm) geographical distance between populations using Mantel tests, implemented in the Ecodist package in R (Goslee & Urban, 2007). We tested whether genetic structure was related to distance from candidate refugial populations (in Romania/Bulgaria, Turkey, Spain and Italy), using Pearson correlations. We also estimated population structure using Admixture version 1.3, with default settings (Alexander et al., 2009). We varied values of K from 1 to 10; by which point increasing values of K provided no informative information about population structure (see Section 3). Model support for each value of K was estimated by calculating five-fold cross-validation error. Finally, we visualised the evolutionary history among European great tit populations by generating a maximum likelihood tree in TreeMix version 1.13 (Pickrell & Pritchard, 2012). We rooted the tree using a sample of *P. minor* individuals sampled from Amur, Russia (Kim et al., 2018). We fitted models allowing for range of migration events (0–10), and used a window size of 500 SNPs (Pickrell & Pritchard, 2012). To assess model fit, we calculated the proportion of variance in relatedness between populations explained by each model (Pickrell & Pritchard, 2012).

Recombination rates at each locus were estimated by comparing the location of SNPs on the genome assembly (v1.1) with their location on the great tit linkage map (van Oers et al., 2014). Previous linkage mapping, using a lower density SNP chip, was independently carried out in UK and Netherlands great tit populations and the two maps were almost identical (van Oers et al., 2014). For the purposes of this analysis, we used SNPs and marker intervals from the UK comprehensive map. A total of 2706 SNPs were located on both the genome assembly and the linkage map. Thus, a position in Mb and cM of each of these SNPs is known. All other SNPs on the HD chip have a physical (Mb) position but no known linkage map position. The great tit genome v1.1 is 1.02 Gb long, so the average physical interval between mapped SNPs is ~376 Kb. The linkage map position of each unmapped SNP was estimated by interpolation; by taking the closest mapped SNP in either direction, and, assuming a constant recombination rate in the interval between those SNPs. For example, an unmapped SNP with physical position 1.4 Mb, flanked by mapped SNPs at 1.0 Mb/0.5 cM and 2.0 Mb/1.0 cM, would be estimated to be located at 0.5 + (1.4–1.0)/(2.0–1.0) × (1.0–0.5) = 0.7 cM. Having interpolated cM position of every SNP, the local recombination rate was calculated as the cM interval spanned by the nearest neighbouring SNPs, divided by the physical distance (bp) spanned by those same neighbouring SNPs. In other words, for the ith SNP, the recombination rate is estimated as the linkage distance between the $i-1$th and $i+1$th SNP, divided by the physical distance between the $i-1$th and $i+1$th SNP. For downstream analyses, local recombination rates were estimated by averaging across all SNPs in each 500 kb window. We calculated gene density in 10- and 500-kb windows using custom R scripts and the annotated great tit genome (v 1.1) (Laine et al., 2016).

We examined the genomic landscape of differentiation across European great tit populations by calculating F_{ST} in 10- and 500-kb bins, using python scripts obtained from Github (https://github.com/simonhmartin/genomics_general). We did not estimate d_{xy}, as this parameter is difficult to estimate accurately from single SNP loci (Cruickshank & Hahn, 2014). We also calculated standardised F_{ST} (zF_{ST}) by mean-centring windowed values and dividing them by the standard deviation among windows. We defined outlier regions as 500-kb bins with zF_{ST} values greater than 10. Finally, *Rsb* (Tang et al., 2007) was calculated for three pairwise comparisons (in Spain, Finland and the Netherlands), using the R package Rehh (Gautier & Vitalis, 2012), and this averaged into 500-kb windows. Unlike F_{ST}, the *Rsb* statistic gives an indication of which population an adaptively important haplotype is under positive selection in, and is thought to be less sensitive to local recombination rates (Tang et al., 2007). *Rsb* estimation requires phased genotype data, so phasing was performed using shapeIT2 v 2.r837 (O'Connell et al., 2014). The –duohmm argument was used to ensure that family information, where available, was used to improve the accuracy of phasing. The –effective-size parameter was set at 500,000 reflecting the large effective population size of European great tit populations (Laine et al., 2016). Local recombination rates (measured in cM/Mbp) were used in the map files. Comparisons between F_{ST}, *Rsb*, recombination rate and gene density were carried out using Spearman rank correlation (as the distributions of these statistics were highly skewed).

3 | RESULTS AND DISCUSSION

3.1 | Genetic diversity and population history

Sampling locations and sample sizes for each population are given in Figure 1 and Table S1. Levels of genetic diversity (π_{SNP}) were generally high, but we observed substantial variability among populations (Figure 2a). Similarly, LD declined rapidly with genomic distance in all populations, reaching baseline levels within ~5 kb in all populations, but also varied among populations (Figure S1). Highest levels of LD (and lowest levels of genetic diversity) were observed in the Mediterranean island populations of Crete (Greece) and Sardinia (Italy), with lowest levels of LD in Central and Western Europe (Figure S1). This is consistent with reduced effective population size in these island populations, either as result of the colonisation process or more recent bottlenecks, along with low levels of subsequent gene flow from the continent to the islands (James et al., 2016; Postma & van Noordwijk, 2005).

Mean genome-wide F_{ST} between all pairs of European great tit populations was 0.015, with a significant but weak pattern of isolation-by-distance (Mantel test; $t = 2.36$, df = 404, $r = .117$ (95% CI 0.077–0.173), $p = .011$; Figure S2). Instead, the highest levels of F_{ST} were found in comparisons involving the Mediterranean island populations of Corsica (France), Sardinia and Crete (Figure S2). Admixture analysis was consistent with this pattern (Figure 2, Figure S3); the $K = 2$ analysis assigned individuals in Sardinia and Corsica to one genetic cluster, and the remaining populations to the second. Thus, it is likely that much of the genetic structure between European great tit populations is a result of genetic drift in these relatively isolated island populations. Admixture analysis also revealed some structure between (mainly peripheral) mainland and larger island populations. At $K = 3$ (the model that best fitted the genetic data; Figure 2d, Figure S4), Spain was separated from the rest of mainland Europe. Increasing values of K (see Figure S3) resulted in the separation of populations in Scotland ($K = 4$), Sardinia (from Corsica; $K = 5$), southern France ($K = 6$), Crete ($K = 7$) and England ($K = 8$). The Admixture output at $K = 8$ (Figure 2e) gives the most detailed picture of genetic structure among European great tit populations. Further increases in K did not generate patterns of structure that corresponded to geographical variation (Figure S3) and were increasingly less well supported (Figure S4). Thus, even with hundreds of thousands of markers the Admixture analysis was unable to separate many of the European populations, confirming that levels of divergence are extremely low (Laine et al., 2016). PCA largely corroborated the Admixture results, with PC1 separating Corsica and Sardinia from the remaining populations, PC2 separating Spain, while PC3 and PC4 separated Scotland, England, Corsica, Sardinia and Crete (Figure S5).

Maximum likelihood analyses implemented in TreeMix showed that a model with no migration explained 97.8% of variance in relatedness between populations (Pickrell & Pritchard, 2012); increasing the number of migration events improved the percentage of relatedness explained, up to 99.7% when 10 migration events were fitted (Figure S6). In Figure 3, we display the maximum likelihood trees with zero to three migration events, after which the variance in relatedness explained plateaued when more migration events were added (Figure S6). The tree was generally characterised by short branch lengths, with the exception of the island populations of Sardinia and Crete, which were grouped with the population from mainland Italy (Figure 3). Thus, the TreeMix analysis is consistent with large populations and low overall genomic divergence, with the exception of the Mediterranean island populations, where genetic drift appears to have been stronger However, much (though not all) of the grouping that did occur among continental populations made geographical sense, with populations from Finland and Estonia grouped together, as were some populations from South-East Europe, and populations from England and Scotland (Figure 3). Interestingly, with three migration events TreeMix grouped the Spanish and Corsican populations, which is consistent with previous subspecies descriptions of European great tits (Clements, 2007). However, the plots with zero, one and two migration events do not cluster the Spanish populations with the Sardinian or Corsican ones, and we also note that the admixture and PCA analyses did not support a close relationship between them.

We next tested the hypothesis that great tits colonised Europe after the last ice age from a single refugium in South-East Europe. This scenario has been suggested before (Kvist et al., 1999), but due to the low number of genetic markers available the power to test this hypothesis was limited. Using our genome-wide panel of SNP

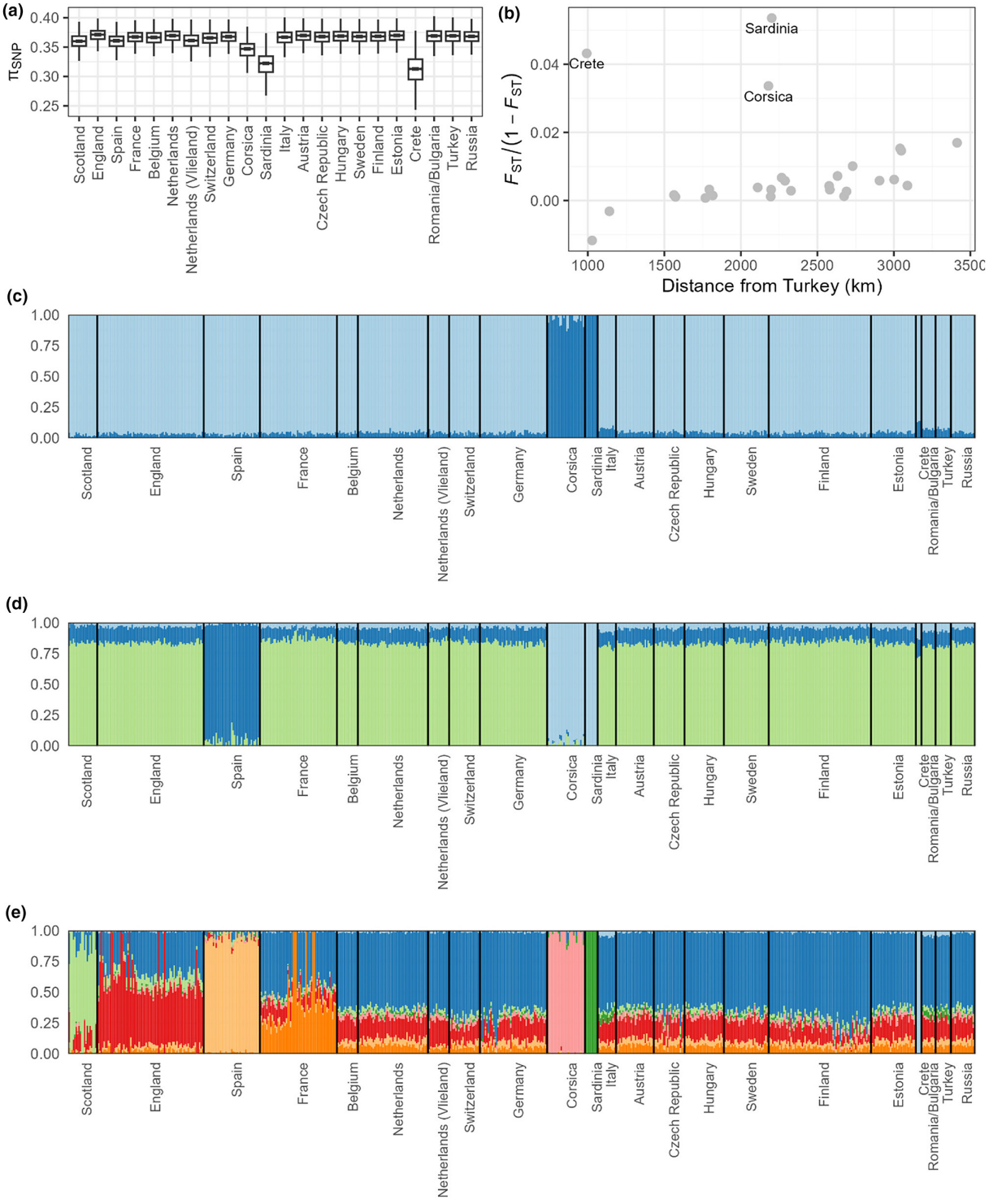

FIGURE 2 Genetic diversity and structure in European Great tit populations. (a) Nucleotide diversity within each population. (b) Pairwise F_{ST} in relation to geographic distance from Turkey, only including comparisons involving Turkey. (c-e) Output from Admixture analysis at $K=2$, $K=3$ and $K=8$. Population details can be found in Table S1.

markers, we compared genetic and geographic distance between each population and the proposed refugial populations. Because of the elevated structure in Corsica, Sardinia and Crete (Figure S2), we excluded comparisons involving these populations. We found that F_{ST} between Turkey and the remaining populations was significantly related to (natural logarithm) distance from Turkey ($r=.81$, $t=6.59$, df$=23$, $p<.001$; Figure 2b). The same relationship was not found for alternative potential refugial populations (Hewitt, 1999) in Spain ($r=.24$, $t=1.69$, df$=47$, $p=.10$) or southern Italy ($r=.35$, $t=1.81$, df$=23$, $p=.08$). Our results therefore lend empirical support to the hypothesis (Kvist et al., 1999) that after the last glacial maximum, great tits progressively colonised Central and Northern

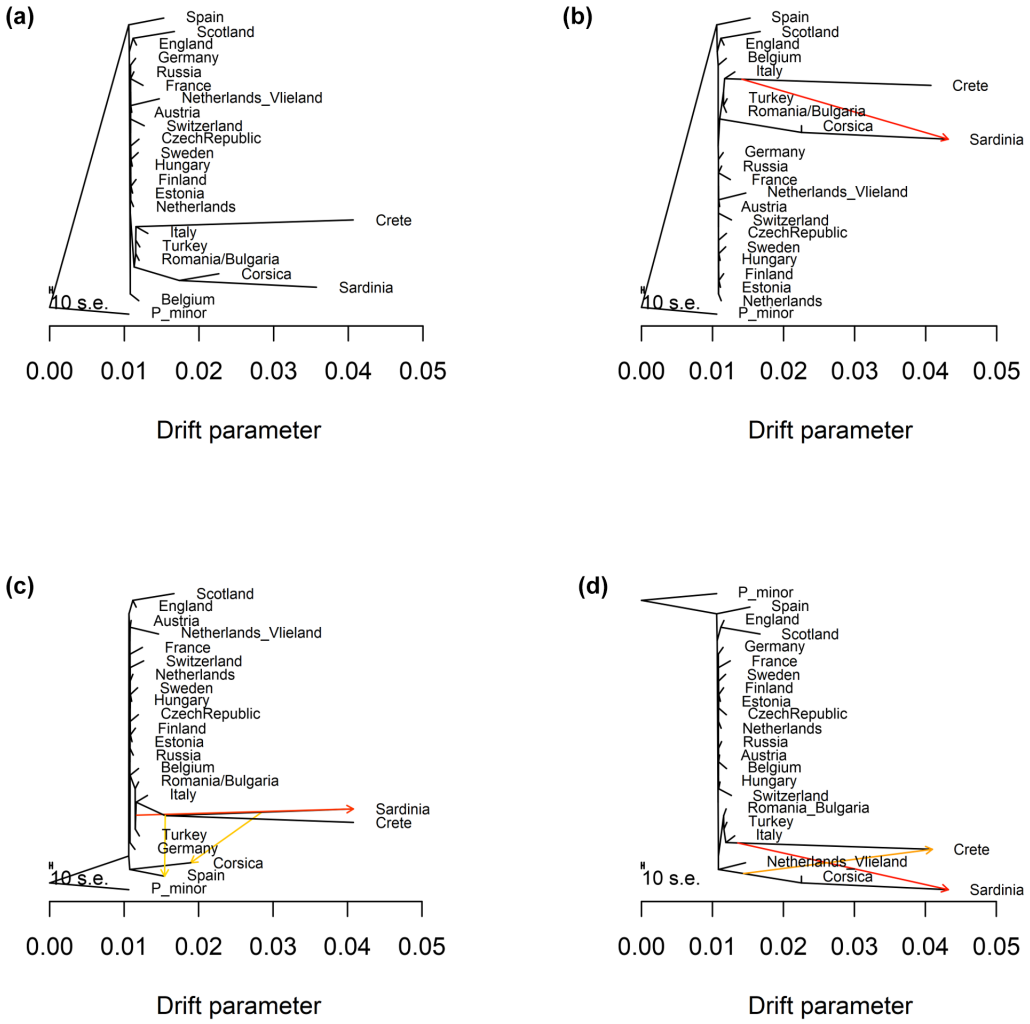

FIGURE 3 Maximum likelihood trees inferred by TreeMix, allowing (a) zero, (b) one, (c) two and (d) three migration events (analysis based on the populations pooled into 22 regions). The migration events (arrows) are coloured according to their weight (red = higher migration), and horizontal branch lengths are proportional to the amount of genetic drift that has occurred along the branch. A population of the great tit's sister species, *Parus minor*, was used as an outgroup. Population details are given in Table S1.

Europe primarily from a single refugium in the south-east. We cannot completely rule out a Spanish refugial population, given that it is the mainland population most divergent from the other populations. However, the admixture analysis and isolation-by-distance patterns find no evidence that the Spanish population has contributed to other populations. Clearly, although our sampling was extensive, it is not exhaustive, and more fine-scaled sampling in Eastern Europe and beyond would be required to determine a more precise location and extent of refugial great tit populations. Sampling in North Africa and West Asia would also be useful to determine whether further refugia exist, and to quantify the extent of admixture between European, African and Asian great tit populations.

3.2 | Genomic landscapes of differentiation

It is likely that many, and perhaps the majority, of wild populations are characterised by highly heterogeneous patterns of differentiation across the genome (Ravinet et al., 2017). This is the case in European great tits – despite the extremely low average F_{ST}, we found genomic regions with very high levels of genomic structure (maximum F_{ST} for 10- and 500-kb windows was 0.98 and 0.07 respectively). To examine how landscapes of genomic divergence have formed along the post-glacial colonisation route of European great tits, we calculated windowed F_{ST} in 500-kb bins between each population and the proposed refugial population in Turkey. We found that F_{ST} varied markedly across the genome in all comparisons (Figure 4; Figure S7). Outlier regions (windows with standardised F_{ST}, hereafter zF_{ST}, >10) were found in all comparisons apart from Crete and Sardinia, in which overall levels of divergence were highest, with some outlier regions found across multiple comparisons (Figure 4). Our results suggest, therefore, that genomic islands of differentiation can and do arise even among recently separated populations.

Considering all populations, F_{ST} calculated in 500-kb windows was strongly negatively correlated with local variation in recombination rate (Spearman correlation, $r = -.50$, $p < .001$). Although

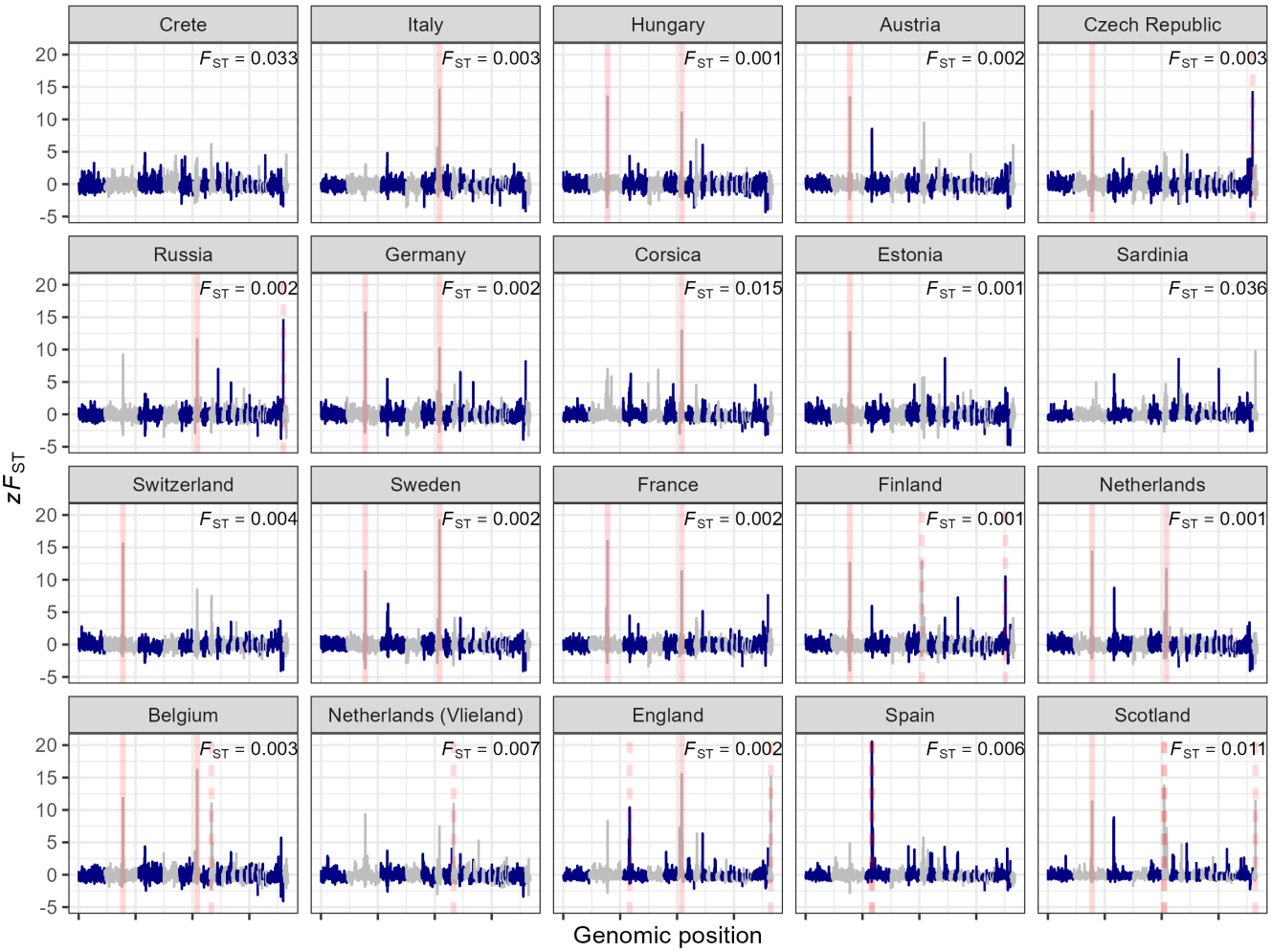

FIGURE 4 Landscapes of relative genomic differentiation in European great tit populations. zF_{ST} across the genome is averaged in 500-kb windows, with each panel displaying a pairwise comparison with the proposed refugial population in Turkey. Red lines represent F_{ST} outliers (windows with mean F_{ST} values at least 10 standard deviations greater than the global mean for that comparison) shared across more than two comparisons (solid red lines), or specific to one or two comparisons (dashed red lines). Mean, untransformed F_{ST} values are given in the top-right of each panel, and are fully displayed in Figure S7.

recombination rate and gene density were positively correlated, ($r = .16$, $p < .001$), F_{ST} was only weakly correlated with gene density in 500-kb windows ($r = -.09$, $p < .001$), and this correlation became almost zero when calculated in 10-kb windows ($r = -.01$, $p < .001$). If recombination rate and gene density are fitted in a linear model, recombination rate ($t = -8.46$, $p < .001$) but not gene density ($t = -0.35$, $p = .72$) explain significant variation in F_{st} measure in 500-kb windows. Thus, it appears that the negative relationship between F_{ST} and recombination rate is not driven by gene density. Examining how the relationship between genomic differentiation and recombination varied among populations revealed that F_{ST} was negatively related to recombination rate in almost all comparisons with Turkey (Figure 5). The relationship between F_{ST} and recombination rate was generally weak, but in a handful of populations this relationship was substantially stronger – most notably in the island populations of Corsica, Sardinia and Crete, and in the peripheral mainland populations of England, Scotland and Spain, which are the populations that are likely most susceptible to drift (Figure 5).

As a further examination of how allele frequency variation and haplotype structure may be shaped by neutral and adaptive processes, we compared F_{ST} and recombination rate with *Rsb* – a measure which aims to detect regions of the genome under positive selection by comparing extended haplotype homozygosity profiles between populations, and is expected to be less sensitive to variation in recombination rate because extended haplotypes should be present in both populations when recombination frequency is lower (Tang et al., 2007). Specifically, we compared the distributions of these statistics in 500 kb windows using three populations in the Netherlands, Finland and Spain, with the aim of exploring how *Rsb* varies with recombination rate and with F_{ST}. We found that, across the genome, correlations between absolute *Rsb* and recombination rate were only slightly weaker (Spearman rank *r* between −.25 and −.20) compared to those between F_{ST} and recombination for the same pairwise comparisons (*r* between −.36 and −.20), and that the regions of high F_{ST} in these comparisons tended to be in these same regions of low recombination and high *Rsb* (Figure S8). Notably,

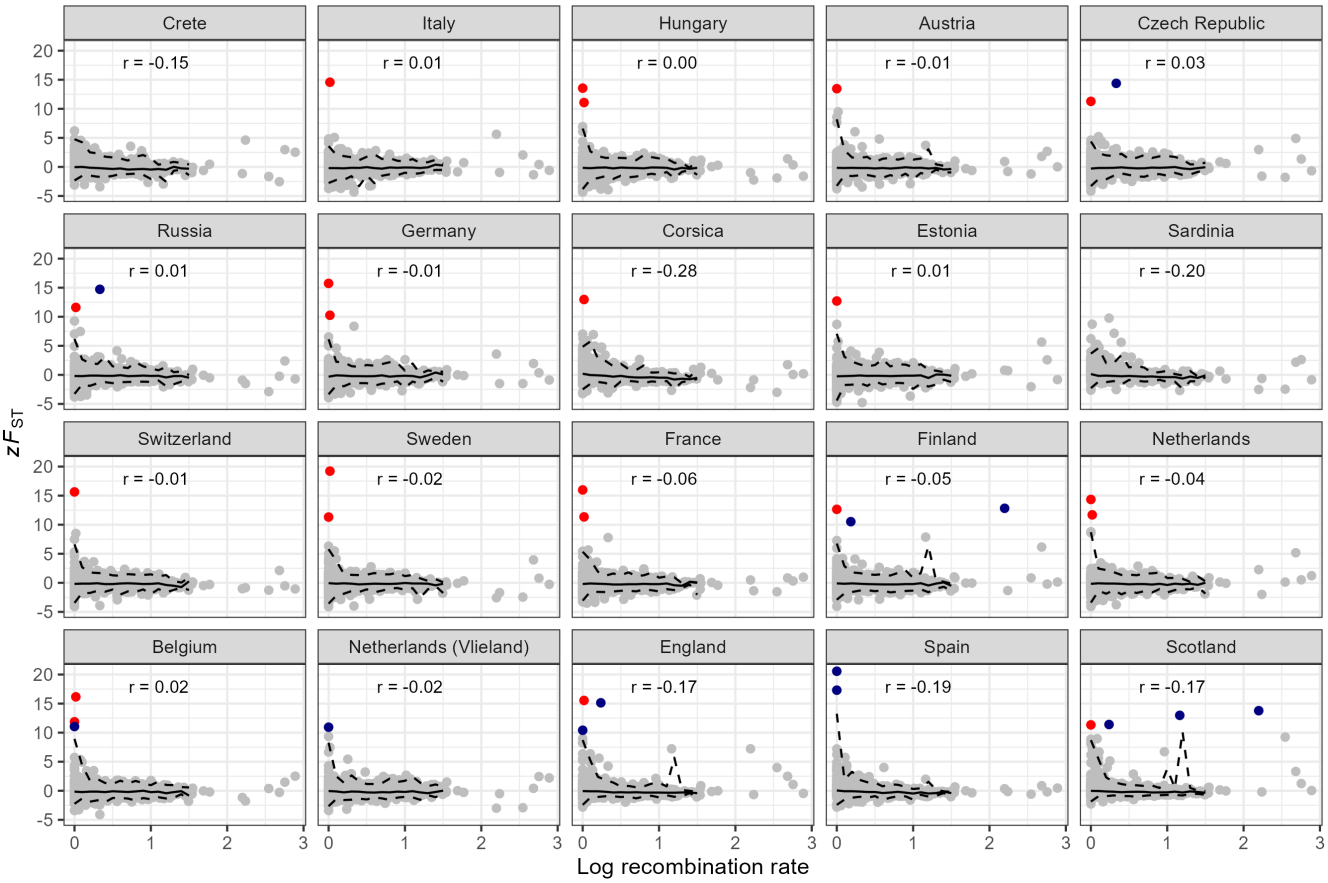

FIGURE 5 Genomic differentiation and recombination rate variation in European great tit populations. Each point is the mean of a 500-kb window, with each panel displaying a pairwise comparison with the proposed refugial population in Turkey. Coloured points represent F_{ST} outliers (mean standardised F_{ST} values of $zF_{ST} > 10$) shared across more than two comparisons (red), or specific to one or two comparisons (dark blue). Solid and dotted lines represent median and 99% quantiles of F_{ST} windows from bins of 0.1 log cM/Mbp. Spearman rank correlation coefficients are reported.

though, there were several cases where windows with extreme *Rsb* values occurred in regions of high recombination, and most high F_{ST} regions in high recombination regions were also *Rsb* outliers (Figure S8). Below we explore the possibility that such regions are strong candidates for recent adaptive evolution.

On the one hand, background selection is a key driver of genomic differentiation in birds (Delmore et al., 2018; Hejase et al., 2020; Van Doren et al., 2017; Vijay et al., 2017), and other organisms (Comeron, 2017; Shang et al., 2023), but is not expected to play a major role in driving islands of differentiation over the timescales (tens of thousands of years) relevant to this study (Burri, 2017; Samuk et al., 2017). On the other hand, differential effects of drift operating across the genome could contribute substantially to the highly heterogeneous patterns of differentiation observed in European great tits. Our finding that the relationship between recombination rate and F_{ST} is strongest in populations experiencing the greatest levels of drift (e.g. islands; Figure 5) is consistent with this – low recombining regions have reduced effective population size and increased rates of lineage sorting due to drift. Drift has been suggested to be a driver of heterogeneous genomic landscapes in other systems (Campagna et al., 2015; Ma et al., 2018; Sendell-Price et al., 2021), and our study suggests that it may play a key role in shaping genomic structure in widespread, continental species.

Outlier regions of very high differentiation ($zF_{ST} > 10$) often occurred in areas of the genome with low recombination rates (Figure 5, Table S2) and the recombination rate of outlier regions was marginally significantly lower than the genome-wide average (Wilcoxon test, $W = 6721.5$, $p = .054$). Of the 11 outlier regions, nine were found in only one or two comparisons, while the other two were found in 12 and 10 comparisons, respectively (Table S2). We hereafter refer to outlier regions found in one or two comparisons as 'population-specific' outlier regions, and to those found in more than two comparisons as 'shared' regions. Both shared outliers were in regions of very low recombination. However, some outliers were in regions of moderate or high recombination rate (Table S2).

Regions of high differentiation that are not shared among populations could potentially arise as a result of recent drift, but are also candidate regions for recent positive selection (Burri, 2017). Several lines of evidence suggest that this may be the case in European great tits. Firstly, population-specific regions tended to occur in regions of higher recombination than shared outliers, although low sample size prevents a formal statistical analysis of this. Second,

population-specific outliers tended to be found in the most peripheral European great tit populations, with three found in Scotland, and two in England, Spain and Finland; the remaining outlier regions were found in comparisons involving the Czech Republic, Russia, Vlieland (Netherlands) and Belgium (Table S2). Notably, these populations are not the lowest in genetic diversity and therefore drift effects are not necessarily stronger in these peripheral populations. Third, the *Rsb* tests, which indicate which population has undergone a selective sweep, also tend to find evidence of positive selection in the more peripheral population, as expected under an adaptive evolution scenario (Figure S8). This directionality cannot be assessed with F_{ST}, because it is a combined measure of allele frequency differences. Of course, across the genome, both positive selection and linked background selection are likely to be operating. Observational and experimental research shows that adaptation at range edges is a key feature shaping divergence among recently colonised and expanding populations (Hill et al., 2011; Sexton et al., 2009; Weiss-Lehman et al., 2017). Population-specific regions appeared less likely to be in regions of low recombination than shared regions (Figure 5), although the small number of regions precluded testing this hypothesis formally. Thus, it is likely that genomic architecture plays a key role in determining how both selection and drift have shaped genomic variation across the recent evolutionary history of European great tits.

To further explore how selection may have shaped variation in F_{ST} outlier regions, we estimated levels of nucleotide and haplotype diversity within these regions. Nucleotide diversity (π_{SNP}) in outlier regions varied from 0.21 to 0.48, and diversity in these regions was significantly lower than the genome-wide average (Wilcoxon test, $W = 5287.5$, $p = .001$; Figure S9A). However, there appeared to be little difference in nucleotide diversity between shared and population-specific regions (Figure S9A). Haplotype diversity varied substantially among regions, with haplotype richness ranging from 72 to 1033. Both haplotype richness and marker density in shared regions tended to be lower than those in population-specific regions (Figure S9B), consistent with lower recombination frequency. A detailed examination of haplotype structure in one shared and one population-specific region is displayed in Figure S10. The population-specific outlier region (to Finland, situated on chromosome 1A) was characterised by a complex structure, with a single haplotype at high frequency in Finland compared to other populations, indicating a population-specific selective sweep and high background diversity (Figure S10). It is unclear whether this haplotype is related to a previously described large inversion polymorphism on Chromosome 1A (da Silva et al., 2019). In contrast, the shared region on chromosome 2 was much less complex, demonstrating high frequency haplotypes that are found across a range of populations. Our data therefore suggest that examining patterns of haplotype diversity in outlier regions may help to separate recent episodes of positive selection from drift and background selection (Figures S9 and S10).

Potential candidate genes found within shared and population-specific outlier regions are displayed in Table S2. Perhaps most notable among these is *COL4A5*, a gene found to be associated with bill length, and under selection between populations in England and the Netherlands, in a previous great tit study (Bosse et al., 2017). Here we found that the *COL4A5* region is an F_{ST} outlier in England and Scotland, but not in any other European populations (Table S2). UK great tits have previously been described as a separate subspecies based on beak shape (Gosler, 1999), and our results here, combined with previous results, suggest that this divergence is the result of recent natural selection (Bosse et al., 2017). Another notable candidate gene potentially involved in beak morphology, and previously found to be under selection in UK great tits is *BMPR1A*, which plays a key role in palate development (Baek et al., 2011) and in this study was found in an outlier region in Scotland. Other candidate morphology and obesity genes in the population-specific outlier regions in the UK included *PPP1CB*, which may play a role in adipogenesis (Cho et al., 2015) and *GHITM*, which appears to have been subject to natural selection in human pygmy populations (Migliano et al., 2013). Thus, morphological and physiological traits may be involved in adaptation in great tits.

In addition to morphological candidates in the UK, we found distinct outlier regions in cold populations such as Scotland, Finland and Russia (Table S2), the outlier locus in the Russian population contained a candidate gene for thermal stress (*CDKN1B*) (Logan & Somero, 2011). Other genomic outlier regions contained potential candidate genes associated with malaria infection (*MRPL33*) (Videvall et al., 2015) and colour variation (*SOX10*) (Gunnarsson et al., 2011). This is thus far an exploratory analysis, and we are therefore reluctant to speculate whether these candidate genes are genuine targets for natural selection, and more reluctant still to speculate as to how selection might be driving variation at these regions. Regardless, these candidates will provide useful starting points for future genomic and ecological investigation.

HapMap style projects have been hugely informative in shaping our understanding of how natural selection operates in humans and other model species (International HapMap, 2005; Kirby et al., 2010). This study is one of the largest to date of genomic variation in a wild vertebrate, which has helped to reveal the evolutionary history of great tits, and to identify candidate genes and traits that may have been involved in adaptation during and/or after postglacial range expansion. Further, this work will form the foundation of many future analyses. Clearly, we have only touched on haplotype-based methods to infer adaptation here, and this will be the subject of future work. Environmental association approaches are also highly suited to detecting adaptation in widespread continental species (Coop et al., 2010; Frichot et al., 2013; Rellstab et al., 2015), and further work will test how variation in the environment has shaped patterns of genomic variation in great tits (Salmón et al., 2021). This combination of environmental and genomic data in species such as great tits, in which a wealth of ecological and genomic resources is available, is likely to generate interesting insights into the genetic and phenotypic basis of natural selection.

AFFILIATIONS

[1]School of Biological Sciences, Norwich Research Park, University of East Anglia, Norwich, UK
[2]Department of Biology, Edward Grey Institute, University of Oxford, Oxford, UK

[3]Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, The Netherlands
[4]Animal Breeding and Genomics, Wageningen University and Research, Wageningen, The Netherlands
[5]Department of Ecological Science, Animal Ecology Group, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
[6]Evolutionary Ecology Group, Department of Biology, University of Antwerp, Antwerp, Belgium
[7]Department of Biology, Science and art Faculty, Mehmet Akif Ersoy University, Istiklal Yerleskesi, Burdur, Turkey
[8]Biology Education, Buca Faculty of Education, Mathematics and Science Education, Dokuz Eylül University, İzmir, Turkey
[9]Hellenic Ornithological Society/BirdLife Greece, Athens, Greece
[10]Institut d'Investigació per a la Gestió Integrada de Zones Costaneres, Campus de Gandia, Universitat Politècnica de València, València, Spain
[11]Faculty of Biology, Lomonosov Moscow State University, Moscow, Russia
[12]Area Avifauna Migratrice, Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), Ozzano Emilia, Italy
[13]CEFE, Univ. Montpellier, CNRS, EPHE, IRD, Montpellier, France
[14]Institute of Environmental Sciences, Jagiellonian University, Kraków, Poland
[15]Behavioural Ecology, Faculty of Biology, LMU München, Planegg-Martinsried, Germany
[16]UMR CNRS 5558—LBBE, Biométrie et Biologie Évolutive, Villeurbanne, France
[17]Department of Ecology and Evolution, Animal Ecology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden
[18]Department of Biology, University of Turku, Turku, Finland
[19]Norwegian Institute for Nature Research, FRAM-High North Research Centre for Climate and the Environment, Tromsø, Norway
[20]Department of Biology, University of Padova, Padova, Italy
[21]Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium
[22]Institute of Wildlife Biology and Game Management, University of Natural Resources and Life Science, Vienna, Austria
[23]Behavioural Ecology Group, Department of Life Sciences, Anglia Ruskin University, Cambridgeshire, UK
[24]Faculty of Biology, Zvenigorod Biological Station, Lomonosov Moscow State University, Moscow, Russia
[25]Department of Ornithology, Max Planck Institute for Biological Intelligence, Seewiesen, Germany
[26]Department of Zoology, Faculty of Science, Palacký University, Olomouc, Czech Republic
[27]Department of Ecology and Genetics, University of Oulu, Oulu, Finland
[28]Zoology Unit, Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland
[29]Department of Zoology, University of Tartu, Tartu, Estonia
[30]School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow, Glasgow, UK
[31]Bulgarian Ornithological Centre, Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, Sofia, Bulgaria
[32]MARE – Marine and Environmental Sciences Centre, Department of Life Sciences, Faculty of Sciences and Technology, University of Coimbra, Coimbra, Portugal
[33]University of Nevada, Reno, Nevada, USA
[34]Evolutionary Ecology Lab, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland
[35]Dipartimento di Scienze e Politiche Ambientali, Università Degli Studi di Milano, Milan, Italy
[36]Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, Oslo, Norway
[37]Behavioural Ecology Group, Department of Systematic Zoology and Ecology, ELTE Eötvös Loránd University, Budapest, Hungary
[38]Centre for Ecology and Conservation, University of Exeter, Penryn, UK
[39]Evolutionary Ecology Group, Hungarian Department of Biology and Ecology, Babeș-Bolyai University, Cluj-Napoca, Romania
[40]Yamashina Institute for Ornithology, Abiko, Japan
[41]Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, Groningen, the Netherlands
[42]School of Biosciences, University of Sheffield, Sheffield, UK

ACKNOWLEDGEMENTS

This work was supported by grants from the Natural Environment Research Council (grant NE/J012599/1 to J.S. and B.C.S) and the European Research Council (grant 202487 to J.S. and grant 339092 – E-Response to MEV) L.G.S. was supported by fellowships from the Edward Grey Institute for Ornithology and the BBSRC (BB/N011759/1). We thank Claire Bloor, Geoff Scopes and Alessandro Davassi of Affymetrix for their help during the chip design and genotyping calling processes. Richard Talbot and Alison Downing of Edinburgh Genomics provided the genotyping service.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The code to reproduce the results is available on Github: https://github.com/lgs85/SpurginBosse_Hapmap. The data, including the Plink-formatted genotype files from all populations, and the downstream outputs are on Dryad: https://doi.org/10.5061/dryad.w3r2280z5.

BENEFIT-SHARING STATEMENT

Benefits arising from this research are the sharing on public databases of the data and the code used to generate the results and plots.

ORCID

Matteo Griggio https://orcid.org/0000-0002-3298-2905
Bart Kempenaers https://orcid.org/0000-0002-7505-5458
Milos Krist https://orcid.org/0000-0002-6183-686X
Veronika N. Laine https://orcid.org/0000-0002-4516-7002
Barbara Tschirren https://orcid.org/0000-0003-4806-4102
Martien A. M. Groenen https://orcid.org/0000-0003-0484-4545
Kees van Oers https://orcid.org/0000-0001-6984-906X
Jon Slate https://orcid.org/0000-0003-3356-5123

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Alonso-Blanco, C., Andrade, J., Becker, C., Bemm, F., Bergelson, J., Borgwardt, K. M., Cao, J., Chae, E., Dezwaan, T. M., Ding, W., Ecker, J. R., Exposito-Alonso, M., Farlow, A., Fitz, J., Gan, X., Grimm, D. G., Hancock, A. M., Henz, S. R., Holm, S., … 1001 Genomes Consortium. (2016). 1,135 Genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell*, *166*(2), 481–491. https://doi.org/10.1016/j.cell.2016.05.063

Baek, J. A., Lan, Y., Liu, H., Maltby, K. M., Mishina, Y., & Jiang, R. (2011). Bmpr1a signaling plays critical roles in palatal shelf growth and palatal bone formation. *Developmental Biology*, *350*(2), 520–531. https://doi.org/10.1016/j.ydbio.2010.12.028

Booker, T. R., Payseur, B. A., & Tigano, A. (2022). Background selection under evolving recombination rates. *Proceedings of the Royal Society B: Biological Sciences*, *289*, 20220782. https://doi.org/10.1098/rspb.2022.0782

Booker, T. R., Yeaman, S., & Whitlock, M. C. (2020). Variation in recombination rate affects detection of outliers in genome scans under

neutrality. *Molecular Ecology*, *29*(22), 4274–4279. https://doi.org/10.1111/mec.15501

Bosse, M., Spurgin, L. G., Laine, V. N., Cole, E. F., Firth, J. A., Gienapp, P., Gosler, A. G., McMahon, K., Poissant, J., Verhagen, I., Groenen, M. A. M., van Oers, K., Sheldon, B. C., Visser, M. E., & Slate, J. (2017). Recent natural selection causes adaptive evolution of an avian polygenic trait. *Science*, *358*(6361), 365–368. https://doi.org/10.1126/science.aal3298

Bovine HapMap Consortium, Gibbs, R. A., Taylor, J. F., Van Tassell, C. P., Barendse, W., Eversole, K. A., Gill, C. A., Green, R. D., Hamernik, D. L., Kappes, S. M., Lien, S., Matukumalli, L. K., McEwan, J. C., Nazareth, L. V., Schnabel, R. D., Weinstock, G. M., Wheeler, D. A., Ajmone-Marsan, P., Boettcher, P. J., … Dodds, K. G. (2009). Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*, *324*(5926), 528–532. https://doi.org/10.1126/science.1167936

Bravo, G. A., Schmitt, C. J., & Edwards, S. V. (2021). What have we learned from the first 500 avian genomes? *Annual Review of Ecology, Evolution, and Systematics*, *52*, 611–639. https://doi.org/10.1146/annurev-ecolsys-012121-085928

Burri, R. (2017). Dissecting differentiation landscapes: A linked selection's perspective. *Journal of Evolutionary Biology*, *30*(8), 1501–1505. https://doi.org/10.1111/jeb.13108

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

Campagna, L., Gronau, I., Silveira, L. F., Siepel, A., & Lovette, I. J. (2015). Distinguishing noise from signal in patterns of genomic divergence in a highly polymorphic avian radiation. *Molecular Ecology*, *24*(16), 4238–4251. https://doi.org/10.1111/mec.13314

Charmantier, A., McCleery, R. H., Cole, L. R., Perrins, C., Kruuk, L. E. B., & Sheldon, B. C. (2008). Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, *320*(5877), 800–803. https://doi.org/10.1126/science.1157174

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

Chen, J., Glémin, S., & Lascoux, M. (2017). Genetic diversity and the efficacy of purifying selection across plant and animal species. *Molecular Biology and Evolution*, *34*(6), 1417–1428.

Chia, J. M., Song, C., Bradbury, P. J., Costich, D., de Leon, N., Doebley, J., Elshire, R. J., Gaut, B., Geller, L., Glaubitz, J. C., Gore, M., Guill, K. E., Holland, J., Hufford, M. B., Lai, J., Li, M., Liu, X., Lu, Y., McCombie, R., … Ware, D. (2012). Maize HapMap2 identifies extant variation from a genome in flux. *Nature Genetics*, *44*(7), 803–807. https://doi.org/10.1038/ng.2313

Cho, Y. L., Min, J. K., Roh, K. M., Kim, W. K., Han, B. S., Bae, K. H., Lee, S. C., Chung, S. J., & Kang, H. J. (2015). Phosphoprotein phosphatase 1CB (PPP1CB), a novel adipogenic activator, promotes 3T3-L1 adipogenesis. *Biochemical and Biophysical Research Communications*, *467*(2), 211–217. https://doi.org/10.1016/j.bbrc.2015.10.004

Clements, J. F. (2007). *Clements checklist of birds of the world*. Comstock Pub. Associates/Cornell University Press.

Comeron, J. M. (2017). Background selection as null hypothesis in population genomics: Insights and challenges from drosophila studies. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *372*(1736), 20160471. https://doi.org/10.1098/rstb.2016.0471

Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, *185*(4), 1411–1423. https://doi.org/10.1534/genetics.110.114819

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

Culina, A., Adriaensen, F., Bailey, L. D., Burgess, M. D., Charmantier, A., Cole, E. F., Eeva, T., Matthysen, E., Nater, C. R., Sheldon, B. C., Saether, B.-E., Vriend, S. J. G., Zajkova, Z., Adamík, P., Aplin, L. M., Angulo, E., Artemyev, A., Barba, E., Barišić, S., … Visser, M. E. (2021). Connecting the data landscape of long-term ecological studies: The SPI-birds data hub. *The Journal of Animal Ecology*, *90*(9), 2147–2160. https://doi.org/10.1111/1365-2656.13388

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

Delmore, K. E., Lugo Ramos, J. S., Van Doren, B. M., Lundberg, M., Bensch, S., Irwin, D. E., & Liedvogel, M. (2018). Comparative analysis examining patterns of genomic differentiation across multiple episodes of population divergence in birds. *Evolution Letters*, *2*(2), 76–87. https://doi.org/10.1002/evl3.46

Dingemanse, N. J., Bouwman, K. M., van de Pol, M., van Overveld, T., Patrick, S. C., Matthysen, E., & Quinn, J. L. (2012). Variation in personality and behavioural plasticity across four populations of the great tit *Parus major*. *The Journal of Animal Ecology*, *81*(1), 116–126. https://doi.org/10.1111/j.1365-2656.2011.01877.x

Ellegren, H., & Wolf, J. B. W. (2017). Parallelism in genomic landscapes of differentiation, conserved genomic features and the role of linked selection. *Journal of Evolutionary Biology*, *30*(8), 1516–1518. https://doi.org/10.1111/jeb.13113

Feder, J. L., Egan, S. P., & Nosil, P. (2012). The genomics of speciation-with-gene-flow. *Trends in Genetics*, *28*(7), 342–350. https://doi.org/10.1016/j.tig.2012.03.009

Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, *30*(7), 1687–1699. https://doi.org/10.1093/molbev/mst063

Gautier, M., & Vitalis, R. (2012). Rehh: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics*, *28*(8), 1176–1177. https://doi.org/10.1093/bioinformatics/bts115

Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*, 1–19. https://doi.org/10.18637/jss.v022.i07

Gosler, A. (1993). *The great tit*. Hamlyn.

Gosler, A. (1999). A comment on the validity of the British great tit *Parus major newtoni*. *Bulletin of the British Ornithologists Club*, *119*, 47–55.

Gunnarsson, U., Kerje, S., Bed'hom, B., Sahlqvist, A. S., Ekwall, O., Tixier-Boichard, M., Kämpe, O., & Andersson, L. (2011). The dark brown plumage color in chickens is caused by an 8.3-kb deletion upstream of SOX10. *Pigment Cell & Melanoma Research*, *24*(2), 268–274. https://doi.org/10.1111/j.1755-148X.2011.00825.x

Harris, H. (1966). Enzyme polymorphisms in man. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *164*(995), 298–310. https://doi.org/10.1098/rspb.1966.0032

Hartl, D. L., & Clark, A. G. (1997). *Principles of population genetics* (3rd ed.). Sinauer.

Hedrick, P. W. (2005). *Genetics of populations* (3rd ed.). Jones and Bartlett.

Hejase, H. A., Salman-Minkov, A., Campagna, L., Hubisz, M. J., Lovette, I. J., Gronau, I., & Siepel, A. (2020). Genomic islands of differentiation in a rapid avian radiation have been driven by recent selective sweeps. *Proceedings of the National Academy of Sciences of the*

United States of America, *117*(48), 30554–30565. https://doi.org/10.1073/pnas.2015987117

Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, *68*(1–2), 87–112. https://doi.org/10.1111/j.1095-8312.1999.tb01160.x

Hill, J. K., Griffiths, H. M., & Thomas, C. D. (2011). Climate change and evolutionary adaptations at species' range margins. *Annual Review of Entomology*, *56*, 143–159. https://doi.org/10.1146/annurev-ento-120709-144746

Horton, M. W., Hancock, A. M., Huang, Y. S., Toomajian, C., Atwell, S., Auton, A., Muliyati, N. W., Platt, A., Sperone, F. G., Vilhjálmsson, B. J., Nordborg, M., Borevitz, J. O., & Bergelson, J. (2012). Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature Genetics*, *44*(2), 212–216. https://doi.org/10.1038/ng.1042

Hubby, J. L., & Lewontin, R. C. (1966). A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. *Genetics*, *54*(2), 577–594. https://doi.org/10.1093/genetics/54.2.577

International HapMap Consortium. (2005). A haplotype map of the human genome. *Nature*, *437*(7063), 1299–1320. https://doi.org/10.1038/nature04226

International HapMap Consortium, Altshuler, D. M., Gibbs, R. A., Peltonen, L., Altshuler, D. M., Gibbs, R. A., Peltonen, L., Dermitzakis, E., Schaffner, S. F., Yu, F., Peltonen, L., Dermitzakis, E., Bonnen, P. E., Altshuler, D. M., Gibbs, R. A., de Bakker, P. I. W., Deloukas, P., Gabriel, S. B., Gwilliam, R., … McEwen, J. E. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature*, *467*(7311), 52–58. https://doi.org/10.1038/nature09298

International HapMap Consortium, Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., Belmont, J. W., Boudreau, A., Hardenbol, P., Leal, S. M., Pasternak, S., Wheeler, D. A., Willis, T. D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., … Stewart, J. (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature*, *449*(7164), 851–861. https://doi.org/10.1038/nature06258

James, J. E., Lanfear, R., & Eyre-Walker, A. (2016). Molecular evolutionary consequences of Island colonization. *Genome Biology and Evolution*, *8*(6), 1876–1888. https://doi.org/10.1093/gbe/evw120

Jiang, Z., Song, G., Luo, X., Zhang, D., Lei, F., & Qu, Y. (2023). Recurrent selection and reduction in recombination shape the genomic landscape of divergence across multiple population pairs of green-backed tit. *Evolution Letters*, *7*(2), 99–111. https://doi.org/10.1093/evlett/qrad005

Jiggins, C. D., & Martin, S. H. (2017). Glittering gold and the quest for Isla de Muerta. *Journal of Evolutionary Biology*, *30*(8), 1509–1511. https://doi.org/10.1111/jeb.13110

Johri, P., Charlesworth, B., & Jensen, J. D. (2020). Toward an evolutionarily appropriate null model: Jointly inferring demography and purifying selection. *Genetics*, *215*(1), 173–192.

Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Porto Neto, L. R., San Cristobal, M., Servin, B., McCulloch, R., Whan, V., Gietzen, K., Paiva, S., Barendse, W., Ciani, E., Raadsma, H., McEwan, J., Dalrymple, B., & International Sheep Genomics Consortium Members. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology*, *10*(2), e1001258. https://doi.org/10.1371/journal.pbio.1001258

Kim, J. M., Santure, A. W., Barton, H. J., Quinn, J. L., Cole, E. F., Great Tit HapMap Consortium, Visser, M. E., Sheldon, B. C., Groenen, M. A. M., van Oers, K., & Slate, J. (2018). A high-density SNP chip for genotyping great tit (*Parus major*) populations and its application to studying the genetic architecture of exploration behaviour. *Molecular Ecology Resources*, *18*(4), 877–891. https://doi.org/10.1111/1755-0998.12778

Kirby, A., Kang, H. M., Wade, C. M., Cotsapas, C., Kostem, E., Han, B., Furlotte, N., Kang, E. Y., Rivas, M., Bogue, M. A., Frazer, K. A., Johnson, F. M., Beilharz, E. J., Cox, D. R., Eskin, E., & Daly, M. J. (2010). Fine mapping in 94 inbred mouse strains using a high-density haplotype resource. *Genetics*, *185*(3), 1081–1095. https://doi.org/10.1534/genetics.110.115014

Kvist, L., Martens, J., Higuchi, H., Nazarenko, A. A., Valchuk, O. P., & Orell, M. (2003). Evolution and genetic structure of the great tit (*Parus major*) complex. *Proceedings of the Royal Society B: Biological Sciences*, *270*(1523), 1447–1454. https://doi.org/10.1098/rspb.2002.2321

Kvist, L., Ruokonen, M., Lumme, J., & Orell, M. (1999). The colonization history and present-day population structure of the European great tit (*Parus major major*). *Heredity*, *82*, 495–502. https://doi.org/10.1038/sj.hdy.6885130

Laine, V. N., Gossmann, T. I., Schachtschneider, K. M., Garroway, C. J., Madsen, O., Verhoeven, K. J. F., de Jager, V., Megens, H.-J., Warren, W. C., Minx, P., Crooijmans, R. P. M. A., Corcoran, P., Great Tit HapMap Consortium, Sheldon, B. C., Slate, J., Zeng, K., van Oers, K., Visser, M. E., & Groenen, M. A. M. (2016). Evolutionary signals of selection on cognition from the great tit genome and methylome. *Nature Communications*, *7*, 10474. https://doi.org/10.1038/ncomms10474

Lanfear, R., Kokko, H., & Eyre-Walker, A. (2014). Population size and the rate of evolution. *Trends in Ecology & Evolution*, *29*(1), 33–41. https://doi.org/10.1016/j.tree.2013.09.009

Lemoine, M., Lucek, K., Perrier, C., Saladin, V., Adriaensen, F., Barba, E., Belda, E. J., Charmantier, A., Cichon, M., Eeva, T., Grégoire, A., Hinde, C., Johnsen, A., Komdeur, J., Mänd, R., Matthysen, E., Norte, A. C., Pitala, N., Sheldon, B. C., … Richner, H. (2016). Low but contrasting neutral genetic differentiation shaped by winter temperature in European great tits. *Biological Journal of the Linnean Society*, *118*(3), 668–685. https://doi.org/10.1111/bij.12745

Lewontin, R. C., & Hubby, J. L. (1966). A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, *54*(2), 595–609. https://doi.org/10.1093/genetics/54.2.595

Lindblad-Toh, K., Wade, C. M., Mikkelsen, T. S., Karlsson, E. K., Jaffe, D. B., Kamal, M., Clamp, M., Chang, J. L., Kulbokas, E. J., 3rd, Zody, M. C., Mauceli, E., Xie, X., Breen, M., Wayne, R. K., Ostrander, E. A., Ponting, C. P., Galibert, F., Smith, D. R., DeJong, P. J., … Lander, E. S. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, *438*(7069), 803–819. https://doi.org/10.1038/nature04338

Logan, C. A., & Somero, G. N. (2011). Effects of thermal acclimation on transcriptional responses to acute heat stress in the eurythermal fish *Gillichthys mirabilis* (Cooper). *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, *300*(6), R1373–R1383. https://doi.org/10.1152/ajpregu.00689.2010

Lohse, K. (2017). Come on feel the noise – From metaphors to null models. *Journal of Evolutionary Biology*, *30*(8), 1506–1508. https://doi.org/10.1111/jeb.13109

Ma, T., Wang, K., Hu, Q., Xi, Z., Wan, D., Wang, Q., Feng, J., Jiang, D., Ahani, H., Abbott, R. J., Lascoux, M., Nevo, E., & Liu, J. (2018). Ancient polymorphisms and divergence hitchhiking contribute to genomic islands of divergence within a poplar species complex. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(2), E236–E243. https://doi.org/10.1073/pnas.1713288114

Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., & Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in Heliconius butterflies. *Genome Research*, *23*(11), 1817–1828. https://doi.org/10.1101/gr.159426.113

Migliano, A. B., Romero, I. G., Metspalu, M., Leavesley, M., Pagani, L., Antao, T., Huang, D.-W., Sherman, B. T., Siddle, K., Scholes, C., Hudjashov, G., Kaitokai, E., Babalu, A., Belatti, M., Cagan, A., Hopkinshaw, B., Shaw, C., Nelis, M., Metspalu, E., … Kivisild, T.

(2013). Evolution of the pygmy phenotype: Evidence of positive selection fro genome-wide scans in African, Asian, and Melanesian pygmies. *Human Biology*, *85*(1–3), 251–284. https://doi.org/10.3378/027.085.0313

Mueller, J. C., Kuhl, H., Boerno, S., Tella, J. L., Carrete, M., & Kempenaers, B. (2018). Evolution of genomic variation in the burrowing owl in response to recent colonization of urban areas. *Proceedings of the Biological Sciences*, *285*(1878), 20180206. https://doi.org/10.1098/rspb.2018.0206

Nadeau, N. J., Whibley, A., Jones, R. T., Davey, J. W., Dasmahapatra, K. K., Baxter, S. W., Quail, M. A., Joron, M., ffrench-Constant, R. H., Blaxter, M. L., Mallet, J., & Jiggins, C. D. (2012). Genomic islands of divergence in hybridizing Heliconius butterflies identified by large-scale targeted sequencing. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*(1587), 343–353. https://doi.org/10.1098/rstb.2011.0198

Nei, M. (2005). Selectionism and neutralism in molecular evolution. *Molecular Biology and Evolution*, *22*(12), 2318–2342. https://doi.org/10.1093/molbev/msi242

Noor, M. A., & Bennett, S. M. (2009). Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity (Edinb)*, *103*(6), 439–444. https://doi.org/10.1038/hdy.2009.151

O'Connell, J., Gurdasani, D., Delaneau, O., Pirastu, N., Ulivi, S., Cocca, M., Traglia, M., Huang, J., Huffman, J. E., Rudan, I., & Marchini, J. (2014). A general approach for haplotype phasing across the full spectrum of relatedness. *PLoS Genetics*, *10*(4), e1004234. https://doi.org/10.1371/journal.pgen.1004234

Parejo, M., Talenti, A., Richardson, M., Vignal, A., Barnett, M., & Wragg, D. (2023). AmelHap: Leveraging drone whole-genome sequence data to create a honey bee HapMap. *Scientific Data*, *10*(1), 198.

Perrier, C., & Charmantier, A. (2019). On the importance of time scales when studying adaptive evolution. *Evolution Letters*, *3*(3), 240–247. https://doi.org/10.1002/evl3.86

Perrier, C., Rougemont, Q., & Charmantier, A. (2020). Demographic history and genomics of local adaptation in blue tit populations. *Evolutionary Applications*, *13*(6), 1145–1165. https://doi.org/10.1111/eva.13035

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8*(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Poelstra, J. W., Vijay, N., Bossu, C. M., Lantz, H., Ryll, B., Muller, I., Baglione, V., Unneberg, P., Wikelski, M., Grabherr, M. G., & Wolf, J. B. (2014). The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, *344*(6190), 1410–1414. https://doi.org/10.1126/science.1253226

Postma, E., & van Noordwijk, A. J. (2005). Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature*, *433*(7021), 65–68. https://doi.org/10.1038/nature03083

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

R Core Team. (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlovic, 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

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370. https://doi.org/10.1111/mec.13322

Renaut, S., Grassa, C. J., Yeaman, S., Moyers, B. T., Lai, Z., Kane, N. C., Bowers, J. E., Burke, J. M., & Rieseberg, L. H. (2013). Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nature Communications*, *4*, 1827. https://doi.org/10.1038/ncomms2833

Salmón, P., Jacobs, A., Ahrén, D., Biard, C., Dingemanse, N. J., Dominoni, D. M., Helm, B., Lundberg, M., Senar, J. C., Sprau, P., & Sprau, P. (2021). Continent-wide genomic signatures of adaptation to urbanisation in a songbird across Europe. *Nature Communications*, *12*(1), 2983. https://doi.org/10.1038/s41467-021-23027-w

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

Sendell-Price, A. T., Ruegg, K. C., Robertson, B. C., & Clegg, S. M. (2021). An Island-hopping bird reveals how founder events shape genome-wide divergence. *Molecular Ecology*, *30*(11), 2495–2510. https://doi.org/10.1111/mec.15898

Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 415–436. https://doi.org/10.1146/annurev.ecolsys.110308.120317

Shang, H., Field, D. L., Paun, O., Rendón-Anaya, M., Hess, J., Vogl, C., Liu, J., Ingvarsson, P. K., Lexer, C., & Leroy, T. (2023). Drivers of genomic landscapes of differentiation across a Populus divergence gradient. *Molecular Ecology*, *32*(15), 4348–4361. https://doi.org/10.1111/mec.17034

Tang, K., Thornton, K. R., & Stoneking, M. (2007). A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biology*, *5*(7), 1587–1602. https://doi.org/10.1371/journal.pbio.0050171

Turner, T. L., & Hahn, M. W. (2010). Genomic islands of speciation or genomic islands and speciation? *Molecular Ecology*, *19*(5), 848–850. https://doi.org/10.1111/j.1365-294X.2010.04532.x

Turner, T. L., Hahn, M. W., & Nuzhdin, S. V. (2005). Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biology*, *3*(9), e285. https://doi.org/10.1371/journal.pbio.0030285

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. M., Crooijmans, R., Sheldon, B. C., Visser, M. E., Slate, J., & Groenen, M. A. M. (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

Videvall, E., Cornwallis, C. K., Palinauskas, V., Valkiunas, G., & Hellgren, O. (2015). The avian transcriptome response to malaria infection. *Molecular Biology and Evolution*, *32*(5), 1255–1267. https://doi.org/10.1093/molbev/msv016

Vijay, N., Weissensteiner, M., Burri, R., Kawakami, T., Ellegren, H., & Wolf, J. B. W. (2017). Genomewide patterns of variation in genetic diversity are shared among populations, species and higher-order taxa. *Molecular Ecology*, *26*(16), 4284–4295. https://doi.org/10.1111/mec.14195

Visser, M. E., van Noordwijk, A. J., Tinbergen, J. M., & Lessells, C. M. (1998). Warmer springs lead to mistimed reproduction in great tits (*Parus major*). *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *265*(1408), 1867–1870. https://doi.org/10.1098/rspb.1998.0514

Weiss-Lehman, C., Hufbauer, R. A., & Melbourne, B. A. (2017). Rapid trait evolution drives increased speed and variance in experimental range expansions. *Nature Communications*, *8*, 14303. https://doi.org/10.1038/ncomms14303

Williamson, R. J., Josephs, E. B., Platts, A. E., Hazzouri, K. M., Haudry, A., Blanchette, M., & Wright, S. I. (2014). Evidence for widespread positive and negative selection in coding and conserved noncoding

regions of *Capsella grandiflora*. *PLoS Genetics*, *10*(9), e1004622. https://doi.org/10.1371/journal.pgen.1004622

Wolf, J. B., & Ellegren, H. (2017). Making sense of genomic islands of differentiation in light of speciation. *Nature Reviews. Genetics*, *18*(2), 87–100. https://doi.org/10.1038/nrg.2016.133

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*, *346*(6215), 1311–1320. https://doi.org/10.1126/science.1251385

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Spurgin, L. G., Bosse, M., Adriaensen, F., Albayrak, T., Barboutis, C., Belda, E., Bushuev, A., Cecere, J. G., Charmantier, A., Cichon, M., Dingemanse, N. J., Doligez, B., Eeva, T., Erikstad, K. E., Fedorov, V., Griggio, M., Heylen, D., Hille, S., Hinde, C. A., … Slate, J. (2024). The great tit HapMap project: A continental-scale analysis of genomic variation in a songbird. *Molecular Ecology Resources*, *00*, e13969. https://doi.org/10.1111/1755-0998.13969