Check for updates

Received: 4 December 2019 | Revised: 11 May 2020 | Accepted: 18 May 2020

DOI: 10.1111/eva.13035

SPECIAL ISSUE ORIGINAL ARTICLE

Evolutionary Applications Open Access WILEY

Demographic history and genomics of local adaptation in blue tit populations

Charles Perrier[1,2] | Quentin Rougemont[3] | Anne Charmantier[1]

[1]Centre d'Ecologie Fonctionnelle et Evolutive UMR 5175 , CNRS , Univ Montpellier, CNRS , EPHE , IRD , Univ Paul Valéry Montpellier 3, Montpellier, France

[2]Centre de Biologie pour la Gestion des Populations, UMR CBGP, INRAE, CIRAD, IRD, Montpellier SupAgro, Univ Montpellier, Montpellier, France

[3]Département de Biologie, Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, Québec, Canada

Correspondence
Centre de Biologie pour la Gestion des Populations|UMR CBGP, INRAE|CIRAD|IRD|Montpellier SupAgro|Univ Montpellier.
Email: charles5perrier@icloud.com

Funding information
European Research Council, Grant/Award Number: ERC-2013-StG-337365-SHE

Abstract

Understanding the genomic processes underlying local adaptation is a central aim of modern evolutionary biology. This task requires identifying footprints of local selection but also estimating spatio-temporal variations in population demography and variations in recombination rate and in diversity along the genome. Here, we investigated these parameters in blue tit populations inhabiting deciduous *versus* evergreen forests, and insular *versus* mainland areas, in the context of a previously described strong phenotypic differentiation. Neighboring population pairs of deciduous and evergreen habitats were weakly genetically differentiated (F_{ST} = 0.003 on average), nevertheless with a statistically significant effect of habitat type on the overall genetic structure. This low differentiation was consistent with the strong and long-lasting gene flow between populations inferred by demographic modeling. In turn, insular and mainland populations were moderately differentiated (F_{ST} = 0.08 on average), in line with the inference of moderate ancestral migration, followed by isolation since the end of the last glaciation. Effective population sizes were large, yet smaller on the island than on the mainland. Weak and nonparallel footprints of divergent selection between deciduous and evergreen populations were consistent with their high connectivity and the probable polygenic nature of local adaptation in these habitats. In turn, stronger footprints of divergent selection were identified between long isolated insular versus mainland birds and were more often found in regions of low recombination, as expected from theory. Lastly, we identified a genomic inversion on the mainland, spanning 2.8 Mb. These results provide insights into the demographic history and genetic architecture of local adaptation in blue tit populations at multiple geographic scales.

KEYWORDS

adaptation, ecological genetics, population genetics | genomics

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2020 The Authors. *Evolutionary Applications* published by John Wiley & Sons Ltd

1 | INTRODUCTION

Local adaptation in heterogeneous environments imposing divergent selection on counterparts of a same species has fascinated scientists for decades (Blanquart, Kaltz, Nuismer, & Gandon, 2013; Kawecki & Ebert, 2004). However, it is only with the recent advance in sequencing technologies that numerous recent empirical studies started discovering the genomic features and processes underlying local adaptation in heterogeneous environments (Hoban et al., 2016; Manel et al., 2016; Savolainen, Lascoux, & Merilä, 2013; Tiffin & Ross-Ibarra, 2014). Especially, many studies used so-called F_{ST} genome scans between populations (Excoffier, Hofer, & Foll, 2009; Narum & Hess, 2011; de Villemereuil, Frichot, Bazin, François, & Gaggiotti, 2014) to detect genes under selection, and potentially implicated in local adaptations (Hohenlohe et al., 2010; Lamichhaney et al., 2016; Perrier, Ferchaud, Sirois, Thibault, & Bernatchez, 2017; Reid et al., 2016). Nonetheless, genome scans need improvements, for example, for studying polygenic variation (Rockman, 2012; Wellenreuther & Hansson, 2016). Genome scans also need to be interpreted in the light of recombination variation along the genome (Booker, Yeaman, & Whitlock, 2020; Stapley, Feulner, Johnston, Santure, & Smadja, 2017) since it shapes the potential extent of diversity and of differentiation in the different regions of the genome (Burri et al., 2015; Cutter & Payseur, 2013; Rougemont et al., 2019; Tine et al., 2014). Genome scans results need to be interpreted in the light of the demographic history of populations since (a) spatio-temporal variation in gene flow and of effective population size affect the possibility of local adaptations (Lenormand, 2002) and the level of false positive due to genetic drift and (b) regions implicated in reproductive isolation accumulated during allopatry showing increased differentiation upon secondary contacts (SC) can be misinterpreted as genuine footprints of recent divergent selection (Bierne, Roze, & Welch, 2013; Bierne, Welch, Loire, Bonhomme, & David, 2011). Lastly, structural variants such as inversions also require special attention, since they have the potential to protect adaptive gene sets from recombination (Kirkpatrick, 2006, 2010; Stapley et al., 2017; Wellenreuther, Merot, Berdan, & Bernatchez, 2019) and hence enable their persistence and rapid redeployment in heterogeneous environments despite high gene flow (Lowry & Willis, 2010; Sinclair-Waters et al., 2017; Todesco et al., 2019).

Here, we aimed at investigating various demographic and genomic aspects of the adaptive divergence of a well-studied passerine bird, the Blue tit (*Cyanistes caeruleus*). Populations of small insectivorous passerines have long been used to study local adaptation, in both quantitative genetics and population genetics frameworks (Broggi, Hohtola, Orell, & Nilsson, 2005; Carbonell, Perez-Tris, & Telleria, 2003; Laaksonen, Sirkiä, & Calhim, 2015). In particular, several blue tit populations breeding in heterogeneous habitats in Southern France (Figure 1a) offer an ideal context to study local adaptation. Four of them (two deciduous and two evergreen, circled in black in Figure 1a) have been subject to a long-term project spanning more than 40 years (Blondel et al., 2006; Charmantier, Doutrelant, Dubuc-Messier, Fargevieille, & Szulkin, 2016). These populations show marked quantitative phenotypic differences (Figure 1b,c), notably in morphological (e.g., tarsus length and body mass), life-history (lay date and clutch size), and behavioral traits (e.g., song characteristics and handling aggression) (Charmantier et al., 2016). These phenotypic differences were found at two spatial scales. First,

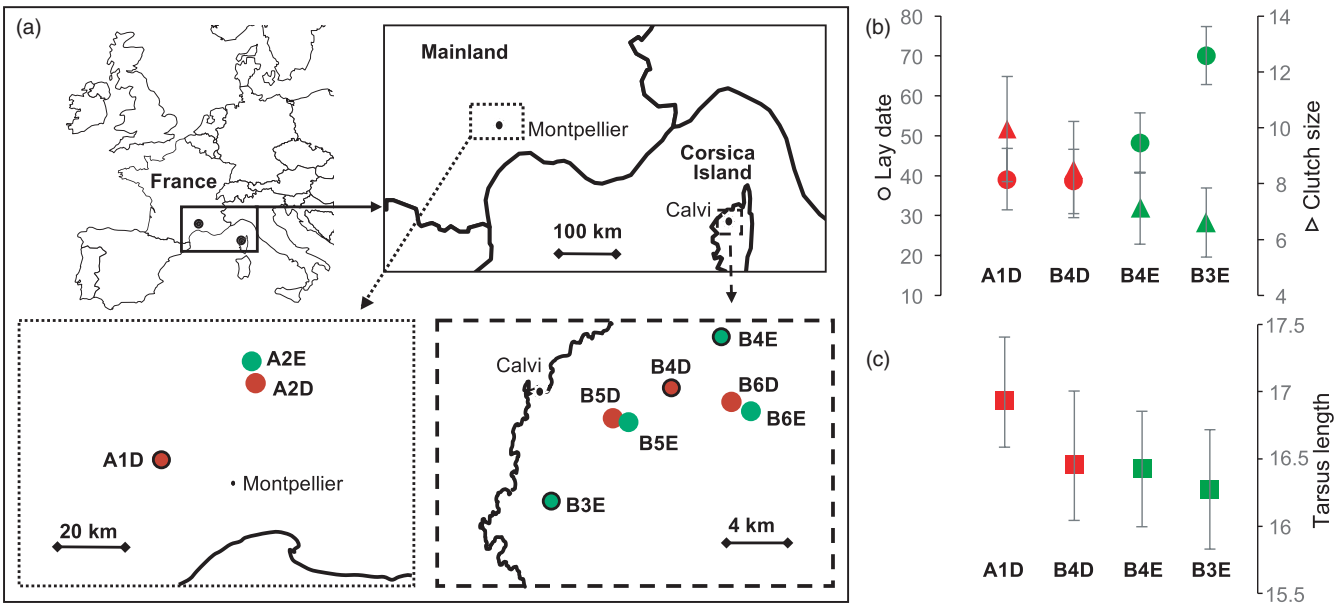


FIGURE 1 Map of the sampling locations of blue tit populations on the mainland and in Corsica (a) and phenotypic differences (b). In red, habitats dominated by deciduous oaks and in green, habitats dominated by evergreen oaks. Sites with long-term monitoring are circled in black on the map (A1D corresponds to D-Rouvière in Charmantier et al. 2016, B3E to D-Pirio, B4D to D-Muro and B4E to E-Muro). On figures b & c are presented illustrative phenotypic differences (with standard deviations) between the four main populations with long-term monitoring. Are shown laying date (days), clutch size and male tarsus length (mm) (data from table 1 in Charmantier et al., 2016 Evol Appl, see this reference for detailed information for these measures and for differences in many other traits)

1752457, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

birds breeding in deciduous forest habitats are taller, more aggressive, and lay larger and earlier broods than birds in evergreen forests (see table 1 in Charmantier et al., 2016). Strikingly, neighboring populations in deciduous and evergreen habitats are weakly genetically differentiated (Dubuc-Messier et al., 2018; Porlier, Garant, Perret, & Charmantier, 2012; Szulkin, Gagnaire, Bierne, & Charmantier, 2016) despite the short spatial scale, which questioned the mechanisms of persistence of the observed phenotypic differentiation against presumably large gene flow. Second, insular blue tits from Corsica, that might have been isolated since the sea level rise after the last glacial maximum (the sea level raised of 120 m from 17,000 to 5,000 years ago [Jouet et al., 2006; Lambeck & Bard, 2000]), are smaller and more colorful than their mainland counterparts (again, see table 1 in Charmantier et al., 2016) and are listed as a separate subspecies (*Cyanistes caeruleus caeruleus* on mainland Europe and *Cyanistes caeruleus ogliastrae* mainly in Corsica and Sardinia). Overall, traits displaying these strong phenotypic differences had heritabilities ranging from 0.20–0.43 (e.g., for lay date, [Caro et al., 2009]) to 0.42–0.60 (e.g., for tarsus length, [Delahaie et al., 2017; Perrier, Delahaie, & Charmantier, 2018; Teplitsky et al., 2014]), are classically related with fitness, and hence could be involved in a local adaptation process. In particular, the heritability of lay date and the breeding time gap between populations could result in reproductive isolation by breeding time, limiting gene flow, and favoring local adaptation. The studied traits were typically quantitative (Charmantier et al., 2016), hence probably controlled by a polygenic architecture (Perrier et al., 2018) composed of many loci with small individual effects, as found in similar traits for other passerine birds (Bosse et al., 2017; Hansson et al., 2018; Lundregan et al., 2018; Santure et al., 2013). Overall, given their phenotypic, demographic, and genetic characteristics, these blue tit populations provide an ideal framework to study the genomic architecture of polygenic adaptation in heterogeneous environments.

We investigated genome-wide patterns of genetic diversity and differentiation and the demographic history between several populations of blue tits from Southern France, in heterogeneous forest habitats (deciduous vs. evergreen) and in insular (Corsica island) and mainland areas (Mainland France) in order to better understand the determinants of their local adaptation. The analysis was based on Resequencing of birds sampled in four sites studied in the context of a long-term project (Blondel et al., 2006; Charmantier et al., 2016) together with three additional pairs of deciduous and evergreen forests in order to test for parallel evolution (see Figure 1a). First, we investigated variation in genetic diversity and differentiation in order to verify that habitat type and geographic distance explained a significant proportion of the genetic structure between populations (Szulkin et al., 2016). Second, we investigated the historical and contemporary demography of each deciduous and evergreen population pair in order to better understand the origin of their differentiation, and the demography between the populations on mainland France and on Corsica Island in order to estimate their split time and their subsequent connectivity and effective population size. Third, we screened the genome for F_{ST} outlier loci and inversions

TABLE 1 Study populations' names, abbreviations, habitat types, locations, approximate forest area (ha), altitude (m), and sample sizes

Name	Abbreviation	Region	Site	Type of oaks	Latitude	Longitude	Forest area	Altitude	N-General Dataset	N-Filtered Dataset
D-Rouviere	A1D	Mainland	Rouviere	Deciduous	43.664000	3.668000	465	248	203	60
D-Quissac	A2D	Mainland	Quissac	Deciduous	43.891036	3.996931	23	104	35	27
E-Quissac	A2E	Mainland	Quissac	Evergreen	43.926673	3.989759	415	116	31	24
E-Pirio	B3E	Corsica	Pirio	Evergreen	42.376000	8.750000	400	145	180	60
D-Muro	B4D	Corsica	Muro	Deciduous	42.551000	8.923000	188	283	215	60
E-Muro	B4E	Corsica	Muro	Evergreen	42.589000	8.967000	80	99	116	60
D-Zilia	B5D	Corsica	Zilia	Deciduous	42.523685	8.873689	13	102	42	38
E-Zilia	B5E	Corsica	Zilia	Evergreen	42.520633	8.884057	17	267	38	38
D-OlmiCappella	B6D	Corsica	OlmiCappella	Deciduous	42.540078	8.993915	20	1,027	44	44
E-OlmiCappella	B6E	Corsica	OlmiCappella	Evergreen	42.529414	9.021652	25	881	43	43
Total/average									947	454

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

potentially implicated in adaptation in the two habitat types and between Corsica and Mainland France. Rather than identifying the genes under selection, we wanted to test whether the adaptation had a polygenic or oligogenic architecture and whether outlier loci would be more frequent in regions with low recombination rate. We discuss our results in the light of the recent literature in population genomics, to decipher the role of genomic and demographic variations on the genetic and phenotypic divergence between blue tits in deciduous versus evergreen habitats and in mainland versus insular contexts.

2 | METHODS

2.1 | Sites and sampling

Nine hundred forty seven blue tits (*Cyanistes caeruleus*) were captured in 10 locations (Figure 1a, Table 1). Three locations are located in the South of mainland France (codes staring with "A": A1D, A2D, A2E) and seven in Corsica (staring with "B": B3E, B4D, B4E, B5D, B5E, B6D, B6E). The numbers from 1 to 6 indicated the sampling area and hence the population pairs. Half of these sites were dominated by the deciduous downy oak *Quercus pubescens* (indicated by "D" in the code), and the other half are dominated by the evergreen holm oak *Q. ilex* (indicated by "E"). Four of these locations (A1D, B3E, B4D, B4E) are monitored as part of a long-term research program started in 1976 and have been described in previous studies (Blondel et al., 2006; Charmantier et al., 2016). The six other sites (A2D, A2E, B5D, B5E, B6D, and B6E) were sampled in order to obtain further replicates of neighboring deciduous and evergreen populations. Capture and handling of the birds was conducted under permits provided by the *Centre de Recherches sur la Biologie des Populations d'Oiseaux* (CRBPO) and by the *Direction Départementale des Services Vétérinaires* (DDSV). Birds were captured during the reproductive period, from early April to late June, on their nesting territory. Birds were captured either in nest boxes during the feeding of their young (in the four sites studied on the long term) or using nets in the vicinity of their nest (in the other sites). Birds were banded with a unique metal ring provided by the CRBPO. Five to 20 μl of blood was sampled from a neck or a wing vein from breeding blue tits. Blood was stored at 4°C in Queen's buffer (Seutin, White, & Boag, 1991).

2.2 | Molecular biology and sequencing

DNA extractions were achieved using Qiagen DNeasy Blood & Tissue kits and were randomized across sites. DNA was quantified using first a NanoDrop ND8000 spectrophotometer and then a Qubit 2.0 fluorometer with the DNA HS assay kit (Life Technologies). DNA quality was examined on agarose gels. Library preparation using RAD-seq (restriction-site-associated DNA sequencing; (Baird et al., 2008)) with the enzyme *SbfI* was done by Montpellier GenomiX (MGX) platform (CNRS, Montpellier). Each individual was identified using a unique six nucleotides tag, individuals were multiplexed in equimolar proportions by libraries of 36 individuals, and each library was sequenced on a lane of an Illumina HiSeq 2000. Single-end sequencing was used to produce 100 bp reads. This design was used to obtain an average read depth of 50x. The DNA of three individuals was extracted twice and processed in different libraries to test for reliability of the genotyping process.

2.3 | Bioinformatics and data filtering

Raw sequences were inspected with *FastQC* (Andrews, 2010) for quality controls. Potential fragments of Illumina adapters were trimmed with *Cutadapt* (Martin 2011), allowing for a 10% mismatch in the adapter sequence. Reads were filtered for overall quality, demultiplexed, and trimmed to 85bp using *process_radtags*, from the *Stacks software pipeline* 1.39 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013), allowing for one mismatch in the barcode sequence. Sequencing RAD-tags resulted in a median value of 5,449,564 reads per individual. *BWA-MEM* 0.7.13 (Li & Durbin, 2009) was used to map individual sequences against the reference genome of the great tit *Parus major* (Laine et al., 2016) and to produce *sam* files using default options. Indeed, although great tit and blue tit diverged about 7–14 Ma (Päckert et al., 2007), the use of a reference genome over a de-novo approach is very often highly recommended (see, e.g., Rochette & Catchen, 2017). It is also interesting to note that synteny is typically highly conserved in birds (Backström et al., 2008), suggesting that aligning reads on a divergent species would still allow analyses based on SNP position along the genome. On average, 93% of the raw reads were mapped against the genome (Table S1). *Samtools* 0.1.19 (Li et al., 2009) was used to build and sort *bam* files. We used *pstacks* to treat *bam* files, align the reads as assembled loci, and call SNPs in each locus. We used a minimum depth of coverage (m) of 5, the SNP model, and alpha = 0.05 (chi-square significance level required to call a heterozygote or homozygote). *cstacks* was used to build the catalog of loci, allowing three mismatches between sample loci when building the catalog. *sstacks* was used to match loci against the catalog. Lastly, *populations* program in *Stacks* was used to genotype individuals. In this program, relatively permissive filters were applied, to retain SNP genotyped in at least 50% of individuals (all individuals from all sites grouped) and with a minimum read depth of 4. 350,941 SNP and 947 individuals were obtained in this "GeneralDataset."

We hence applied additional filters to this GeneralDataset using the programs *VCFtools* (Danecek et al., 2011) and *Plink* (Purcell et al., 2007). We filtered for a minimum average read depth of 10 and a maximum average read depth of 100 (corresponding approximately to the 5%–95% distribution of read depth). SNP genotyped in <80% of the individuals were removed. We removed SNPs with observed heterozygosity ≥0.65 among individuals on Corsica or on mainland to reduce the potential occurrence of stacked paralogues. Individuals below 85% genotyping rate were removed. Identity-by-state (IBS) was estimated for three replicated individuals (from the

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

DNA extraction to SNP calling) in order to investigate reliability of the entire genotyping process. The IBS measured between replicates was high, ranging from 0.9989 to 0.9999, indicating very low genotyping error rate. These replicates were then removed from the dataset. Using the R packages *gdsfmt* and *SNPRelate* (Zheng et al., 2012), we measured the realized genomic relatedness (GRM) between individuals in order to remove highly related individuals (i.e., full-sibs and parent–offspring). For each pair of individuals with relatedness ≥0.35, we removed one individual. This procedure was applied in order to limit biases due to highly related individuals. Indeed, since we sampled breeding birds at their nests for the four sites studied on the long term and that they tend to disperse relatively close to their nests, we expected a higher percentage of related individuals than there actually is in the population. We then limited the number of individuals to 60, chosen at random, in populations (A1D, B3E, B4D, & B4E) in which a large number of individuals were genotyped. This limitation was achieved in order to limit differences in analyses precision between populations due to unequal sample sizes (all individuals will be used in another ongoing study of the genomic architecture of quantitative adaptive traits). We hence removed potential monomorphic SNPs and created the dataset "FilteredDataset." 454 individuals (Table 1) and 144,773 SNPs were kept in this final dataset used for all the genomic analyses excepted analyses of demographic history using ABC. The median genotyping rate across all SNPs for these individuals was 0.981. The median read depth across genotypes (SNPs × Individual) was 49.2x. The number of SNPs per chromosome ranged from 143 (LGE22) to 15,738 (chromosome 2). Further filtering (e.g., MAF) was often operated, depending on the analyses, and thereafter mentioned if it was the case. To perform analyses of demographic history using ABC analyses, we produced a haplotype VCF file for these 454 aforementioned individuals with the populations module of Stacks and we filtered it as explained in the Note S1.

2.4 | Analysis of genetic structure and effects of environmental variables

We used PCAs with the function *snpgdsPCA* from *snprelate* to depict genetic structure between individuals. PCAs were run for the entire dataset, for mainland and Corsican populations separately and for each pairs of deciduous and evergreen populations. We inferred admixture proportions for each individual using *Admixture* 1.23 (Alexander, Novembre, & Lange, 2009) with *K*-values ranging from 1 to 12 and 1,000 bootstraps. The different clustering solutions were inferred by plotting cross-validations errors and by plotting individual admixture proportion. The effect of environmental variables, forest phenology (E vs. D), geographic distance (latitude and longitude), and altitude, on genomic differentiation, was measured using a redundancy analysis (RDA) (Forester, Jones, Joost, Landguth, & Lasky, 2015; Legendre & Fortin, 2010) as implemented in the R package *Vegan* (Oksanen et al., 2007). We investigated the proportion of genetic variability explained by a constraining covariance matrix consisting of phenology, latitude, longitude, and altitude for each individual. We tested the global significance of the model using 1,000 permutations. We ran marginal effects permutations to address the significance of each variable. Then, we focused on the effect of phenology alone, using partial RDA to take into account the effect of latitude, longitude, and altitude. Significance was tested running 1,000 permutations. For the PCA, the admixture analysis and the RDA, we selected from the FilteredDataset the SNPs with more than 95% genotyping rate, MAF >0.05, retaining one SNP per locus and we remove SNP in linkage disequilibrium using the *Plink* command "indep 50 5 2." Genome-wide differentiation between each sampling location was measured with Weir and Cockerham's F_{ST} estimator (Weir & Cockerham, 1984) implemented in *StAMPP* (Pembleton, Cogan, & Forster, 2013). F_{ST} was estimated for all the SNPs, the ones on the autosomes and the ones on the sex chromosome Z separately. Significance was assessed using 1,000 bootstraps replicates.

2.5 | Analysis of demographic history

Alternative models of divergence history including the effects of selection at linked sites affecting *Ne* and of differential introgression (*m*) were compared using an ABC framework modified from Roux et al. (2016). Linked selection either under the form of hitchhiking of neutral loci linked to a selective sweep (Maynard Haigh & Smith, 1974) or under the form of background selection (Charlesworth, Morgan, & Charlesworth, 1993) has strong effects in regions of low recombination and have been shown to influence model choice and parameter estimates (Ewing & Jensen, 2016; Schrider, Shanku, & Kern, 2016). The same is true when populations accumulate reproductive incompatibilities during the divergence process: The resulting barrier to gene flow reduces the effective migration rate along the genome (Barton & Bengtsson, 1986) and not accounting for it can affect demographic model choice and parameter estimates (Roux, Tsagkogeorga, Bierne, & Galtier, 2013; Sousa, Carneiro, Ferrand, & Hey, 2013). Moreover, including selected loci in demographic inferences can reveal the deeper origins of population divergence (Bierne et al., 2013). Six scenarios were compared for the four pairs of deciduous and evergreen populations in order to test whether the divergence between different habitats was not due to a divergence in different historical refugia but to a contemporary ecological divergence. We included a model of panmixia (PAN), a model of equilibrium corresponding to the island model with two populations (EQ), a model of isolation with migration (IM), a model of SC, a model of divergence with migration during the first generations, *that is*, ancestral migration (AM), and a model of strict isolation (SI). The prior and details of the simulation pipeline are fully described in the Note S1. The PAN model assumes that the two focal populations descent from a single panmictic population characterized by its effective size (Ne). The EQ model (equivalent to the island model) assumes that the population is subdivided into two discrete populations of sizes N_{pop1} and N_{pop2} that are connected by continuous gene flow at a constant rate each generation. In this model, the divergence time is not a parameter. The IM, SI, SC, and AM models all assume that an ancestral

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

population of size N_{ANC} splits at T_{split} into two daughter populations of sizes N_{pop1} and $N_{pop2.}$ Under SI, no subsequent gene flow occurs. Under AM model, gene flow occurs from T_{split} to T_{am} and is followed by a period without gene flow. Under IM, gene flow is continuous after T_{split}. Under, the SC model, T_{split} is followed by a period of SI, after which a SC starts T_{sc} generations. The EQ, IM, AM, and SC models included migration as $M = 4\ N_0.m$, with $M_{1\leftarrow 2}$ being the number of migrants from population 2 to population 1 and $M_{2\leftarrow 1}$ being the reverse. The effect of linked selection and barriers to gene flow were accounted for, by modeling heterogeneous population size (Ne) and heterogeneous migration (m), respectively. Such heterogeneity was modeled using beta distributions as hyper-prior on each of these two parameters. These resulted in four alternative versions for models with gene flow EQ, AM, IM, and SC (NhomoMhomo, NhomoMhetero, NheteroMhomo, and NheteroMhetero) and two versions for PAN and SI (Nhomo and Nhetero). We used a modified ABC (Csilléry, Blum, Gaggiotti, & François, 2010) pipeline from Rougemont and Bernatchez (2018) to perform model selectionand to estimate posterior probabilities of parameters.

We also investigated the historical demography of the populations from Corsica as compared to the ones from the mainland. Gene flow has probably been impossible since the last deglaciation and sea level rise. Therefore, we compared models of ancient migration (AM) and SI. The pipeline described in the above section, integrating linked selection and barriers to gene flow, was run between the A1D samples (chosen on the mainland because it had the largest sample size) and B4D (chosen in Corsica because it had both a large sample size and was from the same habitat as A1D). We used the same ABC pipeline for the model selection procedure and parameter estimation as described in the above section. Finally, we attempted to convert demographic parameter into biological units assuming a mutation rate of 1e-8 mutations/bp/generations.

2.6 | Genomic diversity

Genome-wide genetic diversity was inferred for each sampling location in the dataset by measuring observed heterozygosity (Ho), proportion of polymorphic loci, and MAF spectrums. For each population and chromosome, and subsequently on average for the entire genome, linkage disequilibrium (LD) decay was measured with *Plink* and smoothed in R. To contrast with long-term Ne estimates from coalescent simulation in our ABC modeling (see "Analysis of demographic history"), we also inferred recent Ne for each population using *SNeP* V1.1 (Barbato, Orozco-terWengel, Tapio, & Bruford, 2015), which uses LD data, with a MAF ≥0.05 filter per population. We investigated the nature of SNP variation, *that is*, synonymous or nonsynonymous, blasting all RAD sequences on the reference genome and the transcriptome of the great tit (ftp://ftp.ncbi.nih.gov/genomes/Parus_major/; Santure, Gratten, Mossman, Sheldon, & Slate, 2011) using *blastx* (McGinnis & Madden, 2004). We kept hits with at least 90% similarity and a minimum amino acid sequence length alignment of 25. We kept only SNPs for which both the alternative and reference allele yields the same score. Finally, we tested for differences in the distribution of run of homozygosity (ROH) between the mainland and Corsica that may have resulted from smaller Ne and larger inbreeding in Corsica versus the mainland. We used plink 1.9 to estimate the length and number of ROH. We required a window of 500 kb to be homozyguous in order to be considered as a ROH, and with a maximum of 100 SNP.

2.7 | Identification of genomic footprints of selection

We used three methods to search for outlier SNPs potentially under divergent selection between blue tit populations. First, we used *Bayescan* V2.1 (Foll & Gaggiotti, 2008) to search for SNPs potentially under divergent selection at two geographic levels: (a) between the mainland and Corsican populations and (b) between each local pairs of evergreen (E) and deciduous (D) populations, *that is*, A2D versus A2E, B4D versus B4E, B5D versus B5E, B6D versus B6E. We filtered each of the five datasets for a minimum MAF of 0.05. We used default parameters except for prior odds that were set at 10,000 in order to limit false positives. We investigated the parallelism across pairs of D–E environments. Second, we estimated F_{ST} along the genome using either a 200 kb sliding average with *VCFtools* or the function "snpgdsSlidingWindow" from the package SNPrelate to estimate F_{ST} in blocks of 50 SNPs moving by 5 SNPs. The second window strategy was used in order to compensate for lower SNP density in regions of low recombination that tend to exaggerate the contribution of individual SNPs in these regions and to dilute the individual SNP contribution in regions of high recombination (see Perrier & Charmantier, 2019 for a broader comment on this). Third, we used an RDA as an alternative method to search for SNPs putatively implicated in multilocus adaptation (a) between populations in deciduous and evergreen habitats, and (b) between populations in Corsica and on the mainland. As suggested by Forester, Lasky, Wagner, and Urban (2018), such a multivariate method may be more suitable than univariate ones to detect weaker footprints of adaptation that are expected in polygenic adaptations in response to complex environmental heterogeneity. Using a similar procedure as described earlier in the methods, we used two RDAs constrained to investigate the effect of phenology (a) or of the geography (b). We then used a three standard deviation cutoff as suggested by Forester et al. (2018) to list loci with outlier loading scores on the first RDA axes. We compared the loci found using these different methods. We reported in which genes these outliers were found (the list of genes can be found together with the genome published by Laine et al., 2016 on NCBI).

2.8 | Variation of genomic differentiation with recombination rate

We investigated variation of F_{ST} with local recombination rate and whether SNP outliers were more often found in regions of low

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

recombination than elsewhere in the genome. We estimated local recombination rates using a coalescent method implemented in *Ldhat* (McVean, 2004) using linkage disequilibrium signal. Following these authors' recommendation, the dataset was split in blocks of 2,000 SNPs with 500 overlapping SNPs. A MAF of 0.05 and a maximum of 5% of missing data were allowed. The local recombination rate, $\rho = 4Ne\ r$, was then estimated in each block independently with the Bayesian reversible jump MCMC scheme implemented in interval. We used a block penalty of 5, with 30 million MCMC iterations sampled every 5,000 iterations. The first 250,000 iterations were discarded. To speed up computations, we used the precomputed two locus likelihood table for $n = 190$ and assuming theta = 0.001. We estimated ρ in a composite dataset with individuals from every population, habitats, and sex. We tested for correlations between SNP F_{ST} and recombination rate using linear models, for Corsica–mainland and for deciduous–evergreen differentiation, independently, and we represented the correlation using a LOESS fit. We tested whether outliers found using Bayescan and the RDA method for both Corsica–mainland and deciduous–evergreen differentiation were more often found in regions of low recombination than elsewhere in the genome using chi-square (χ^2) tests, and we represented the pattern using histograms.

2.9 | Detection of genomic inversions

We searched for potential genomic inversions using a variety of descriptive statistics. First, we searched for genomic regions having a particularly low recombination rate and large long-distance linkage disequilibrium, nevertheless associated with a high density of SNPs and therefore unlikely to correspond to peri-centromeric regions but rather to local suppression of recombination that may be due to inversions. Second, we implemented a PCA sliding window analysis in order to identify portions of the genome with individuals carrying an inversion at the homozygous or heterozygous state or individuals exempt from the inversion (Ma & Amos, 2012). We first used 10Mb windows sliding by 1 Mb and then 1 Mb windows sliding by 100 kb, in order to use enough SNPs to perform PCAs. We then used *Lostruct* (Li & Ralph, 2019), with $k = 2$, sliding by 100 SNPs, to identify particular blocks of linked SNPs explaining an abnormally high proportion of variance between two groups of individuals (e.g., inverted and noninverted).

We detected one putative inversion. We verified, using admixture with $K = 2$ for analyzing SNPs from the inversion, that putative heterozygous individuals had an admixture ratio close to 1:1 of both putative inverted and noninverted homozygous clusters. We then inspected variations of F_{ST} (per SNP) and π (per 10kb window, using *vcftools*) along the genome between individuals that were inverted homozygous, noninverted homozygous, and heterozygous for the inversion, looking for potentially reduced diversity and increased differentiation at the inversion. We looked for potential salient variations in read depth in and around the putative inversion. We also tested whether the detected inversion was at Hardy–Weinberg equilibrium (using a common χ^2 test) and whether the frequency of the inversion varied geographically and between evergreen and deciduous habitats.

To study the history of the putative inversion identified, we aimed at measuring intra- and interspecific genetic distance and absolute divergence, at the inversion and for the entire genome, for several blue tits and great tits. To do that, we first generated a new SNP dataset by running stacks with the same pipeline and parameters as explained earlier, with two mainland great tits, two insular great tits, four mainland blue tits homozygous for the inversion, four mainland blue tits not carrying the inversion, and four Corsican blue tits not carrying the inversion (the inversion was not found in Corsica). Second, we selected the SNPs from the region of the inversion. Third, Da was measured between the five aforementioned groups of individuals using *PopGenome* (Pfeifer, Wittelsbürger, Ramos-Onsins, & Lercher, 2014). Using this measure of Da, we estimated approximately the inversion apparition time using $T = Da/2\mu$ (Hudson, Kreitma, Aguadé, 1987), assuming a standard mutation rate of 1e-8 and with the simplifying assumption of no gene flow and no introgression. Lastly, we represented the divergence between these individuals using an UPGMA tree of bitwise distance using the R package *poppr* 1.1.1 (Kamvar, Tabima, & Grünwald, 2014).

2.10 | Gene ontology

We used the R package topGO (Alexa & Rahnenführer, 2009) to investigate the potential gene ontologies (GO) that were statistically enriched for the sets of genes identified among outliers and the inversion, compared to the entire list of genes in which all the SNPs from the entire dataset were found. For the outlier tests, we used GO analyses for each gene list obtained using the different outlier identification methods but also for aggregated lists of all of the gene identified for deciduous–evergreen tests or for Corsica–mainland tests. We used the GO referenced for the zebra finch, T. Guttata (tguttata_gene_ensembl). We report a Fisher enrichment test p-value and a p-value after applying a Benjamini and Hochberg correction for multiple testing to control for false discovery rate.

3 | RESULTS

3.1 | Genetic structure and effects of environmental variables

The admixture analysis suggested the existence of two main distinct genetic groups corresponding to mainland and island populations (Figure 2a). Increasing *K* values contributed to delineate the populations, showing the existence of a weak structure between the six main sampling areas. However, the coefficient of variation increased with *K*. There was very little evidence for genetic structure within each of the four pairs of D- and E- populations. PCAs revealed a clear structure between individuals from the mainland and the island

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

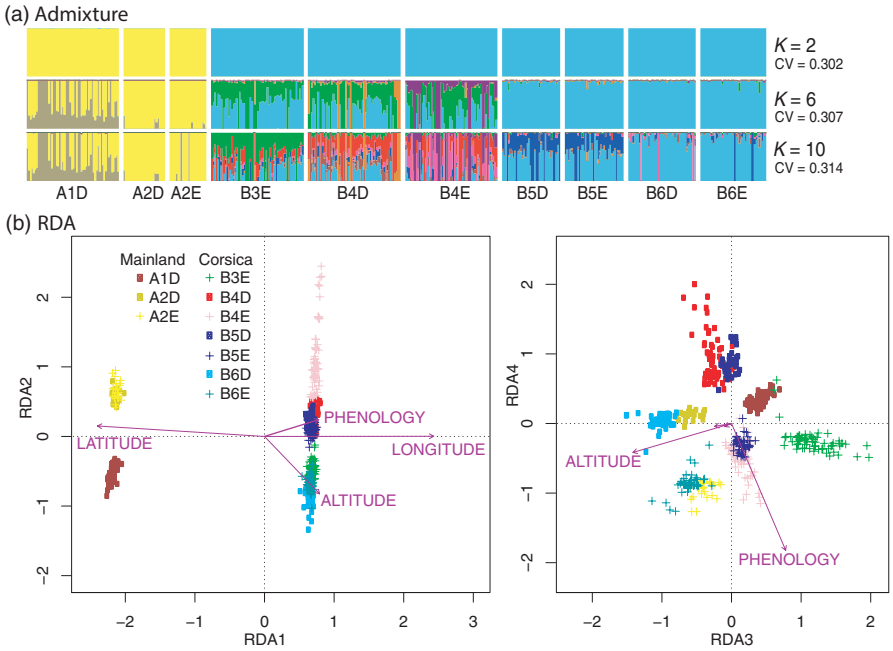


FIGURE 2 Population structure between populations. (a) Admixture for $K = 2$, 6 and 10 (Vertical lines correspond to individual admixture; colors correspond to distinct genetic groups). (b) RDA showing the influence of latitude, longitude, altitude, and phenology on genetic structure

but little structure within each of these two groups (Figure S1). The full RDA model was significant ($p < .001$, Table S2, Figure 2b), as well as the effects of each variable tested in the model: phenology (*i.e.*, E vs. D, $p = .006$), latitude ($p < .001$), longitude ($p < .001$), and altitude ($p < .001$). The first axis of the global RDA explained 4.54% of the variance and was correlated mainly with latitude (−0.987) and longitude (+0.999). The fourth axis of this global RDA displayed the strongest correlation with phenology (−0.856) and explained 0.27% of the variance. The partial RDA model conditioning for the effects of latitude, longitude, and altitude was globally significant ($p < .001$), as well as the effect of phenology ($p < .001$). The axis of the partial RDA explained 0.29% of the variance and was correlated with phenology (0.922).

All pairwise F_{ST} comparisons were significantly different from zero (Table 2). Average F_{ST} calculated between island and mainland populations was 0.081. It was on average 0.003 between close populations within the mainland and 0.005 within the island. F_{ST} was low, on average 0.003, for each of the four pairs of D- and E- populations, ranging from 0.0006 between B6D and B6E to 0.0079 between B4D and B4E. The F_{ST} estimated on the Z chromosome was on average 1.52 times higher than on the autosomes; 1.62 times higher when comparing populations from the mainland versus Corsica , and 1.44 times higher when comparing neighboring populations in the island or in the mainland.

3.2 | Demographic history

When deciphering the historical demography of deciduous and evergreen population pairs, the hierarchical model choice procedure strongly rejected models of SI, ancient migration, and of PAN, which were associated with a posterior probability of 0 (Table S3). Instead, models with gene flow were highly supported. In three of the four pairs, the equilibrium model (EQ) received the highest posterior probability with P(EQ) = 0.99 and 0.95, 0,94 in the B6D versus B6E, B5D versus B5E and B4D versus B4E comparisons, respectively. In the A2D versus A2E comparisons, the best supported model was the isolation with migration model, P(IM) = 0.74 with the second best model being SC, with P(SC) = 0.23. Comparing the two models against each other while excluding all remaining models yield unambiguous support for IM, with P(IM) = 0.99 (Table S3). Across all models, comparisons with heterogeneous gene flow and heterogeneous effective population size were not supported, indicating that, if genetic barriers or linked selection were at play they could not be detected. Demographic parameters were estimated for each pair of populations under the best model (Table S4). Posterior distributions were well differentiated from their prior indicating that estimated parameters were confidently estimated (Figure S2). Effective population size (N_e) was slightly higher in deciduous than in evergreen habitats (Ne-ABC in Table 2, Figure 3a, Table S4) and tended to be higher on the mainland than on the island. In a number of comparisons, our migration estimates reached the prior upper bounds. Therefore, we ran a new set of simulations with wider priors (Note S1). Migration rates were not different between deciduous and evergreen habitats: In two instances, the rate of migration was higher from the deciduous to the evergreen, and in two other instances, the reverse was true (Figure S3).

Deciphering the historical demography between Corsica and the mainland, the ancient migration models (AM) strongly outperformed the SI models, with p(AM) = 0.99 (Table S5). For the AM model, the simplest model without linked selection or barriers to gene flow was the best supported, with $p = .76$. Parameters estimates revealed posterior generally well differentiated from the prior providing increased confidence (Figure S4). Effective population size of the mainland was 397,685 [CI = 244,375–600,105], around 7.5 times greater than the one of the island, 52,990 [CI = 33,065–78,475] (Figure 3b,

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Diversity and differentiation across populations. Fixation index between populations (F_{ST}), effective population size (Ne) estimated with ABC (Ne-ABC) using models for deciduous–evergreen population pairs or for Corsica–mainland (A1D-B4D), Ne estimated with a LD method (Ne-LD), and observed heterozygosity (Ho). Color gradients indicate differentiation intensity (Darker red for higher differentiation and darker blue for lowest differentiation). Underlined values indicate deciduous and evergreen pairs

	F_{ST}										Ne-ABC			
	A1D	A2D	A2E	B3E	B4D	B4E	B5D	B5E	B6D	B6E	Deciduous–Evergreen	Corsica–Mainland	Ne-LD	Ho
A1D		0.0045	0.0043	0.0812	0.0830	0.0853	0.0796	0.0804	0.0785	0.0784	Na	397,685	355	0.171
A2D	0.0045		0.0029	0.0832	0.0849	0.0872	0.0799	0.0810	0.0793	0.0792	400,005	na	161	0.167
A2E	0.0043	0.0029		0.0828	0.0846	0.0871	0.0796	0.0807	0.0790	0.0787	100,055	na	142	0.168
B3E	0.0812	0.0832	0.0828		0.0058	0.0089	0.0067	0.0071	0.0056	0.0058	Na	Na	334	0.149
B4D	0.0830	0.0849	0.0846	0.0058		0.0079	0.0063	0.0067	0.0061	0.0062	43,070	52,990	310	0.149
B4E	0.0853	0.0872	0.0871	0.0089	0.0079		0.0093	0.0097	0.0083	0.0085	37,100	na	271	0.148
B5D	0.0796	0.0799	0.0796	0.0067	0.0063	0.0093		0.0025	0.0035	0.0032	46,520	na	210	0.145
B5E	0.0804	0.0810	0.0807	0.0071	0.0067	0.0097	0.0025		0.0036	0.0033	42,800	na	216	0.146
B6D	0.0785	0.0793	0.0790	0.0056	0.0061	0.0083	0.0035	0.0035		0.0006	46,885	na	268	0.146
B6E	0.0784	0.0792	0.0787	0.0058	0.0062	0.0085	0.0032	0.0033	0.0006		162,940	na	260	0.146

Ne-ABC in Table 2, Table S6). Our analysis also indicated a strong population size change during the process of population isolation, since the ancestral Ne was estimated at only 240 [CI = 65–685]. Split time was estimated at 1,417,480 [CI = 740,000–2,394,000] generations ago, hence around 3.2 Ma (assuming a 2.3 years generation time [Charmantier, Kruuk, Blondel, & Lambrechts, 2004]) and gene flow subsequently stopped 4,540 [CI = 2,500–6360] generations ago, hence around 10,000 years ago, corresponding closely with the end of the last glacial maximum. The AM rate between Corsica and the mainland, from the split in two populations to the end of the latest gene flow between these populations, was inferred asymmetric, with a five times larger migration rate from the island to the mainland (*m ~ 1.9e-4)* than the other way around *(m ~ 3.8e-5*). However, this asymmetry did not generate a strong difference in gene flow between the two groups, given the lower effective population size on the island compared to the mainland, (i.e., respective number of migrants of 10 and 15 from Corsica to the mainland and in the reverse direction, with overlapping credible intervals, see details in Table S6).

3.3 | Genetic diversity

The mainland populations displayed significantly higher patterns of observed heterozygosity (on average 0.169, Table 2) than Island populations (on average 0.147), all *t* tests' *p*-values <.001. On the contrary, there was no significant difference of observed heterozygosity at smaller spatial scale among populations within the island (either same or different tree phenology, E- and D-) or within the mainland (all *t* tests' *p*-value >.05). More SNPs were polymorphic in mainland populations (77,306–80,616 SNPs for a MAF > 0.05) than in island populations (62,728–65,876 SNPs for a MAF > 0.05). The MAF spectrum showed enrichment of variants with smaller frequencies in mainland populations versus island populations (Figure S5a). LD decayed rapidly in the first 5kb and was lower in populations from the mainland (especially A1D) compared to island populations (Figure S5b). This pattern of rapid LD decay was similar between chromosomes (see, e.g., chromosomes 1, 2, and Z, Figure S6). Contemporary N_e inferred from LD varied from 142 (in A2E) to 355 (in A1D) (Ne-LD in Table 2, Figure S5c). N_e values were rather similar within pairs of E- and D- populations, although Ne was on average smaller in evergreen populations (245 on average) than in deciduous ones (261 on average). The four largest N_e (with an average of 318) were found for the four populations monitored on the long term in forest in which hundreds of artificial nest boxes have been installed for several decades (the N_e in the 6 other populations was on average of 210). In the entire dataset, 2,599 SNPs were identified as nonsynonymous while 6,751 SNPs were identified as synonymous. We observed 1.18 times lower allele frequencies for nonsynonymous variants (average MAF = 0.09) than for synonymous ones (average MAF = 0.11) (*t* test *p*-value < .000001). This overall lower frequency of nonsynonymous compared to synonymous mutations was similar between populations from Corsica (1.19) and the mainland (1.16), as well as between deciduous (1.18) and evergreen (1.19) habitats (see

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

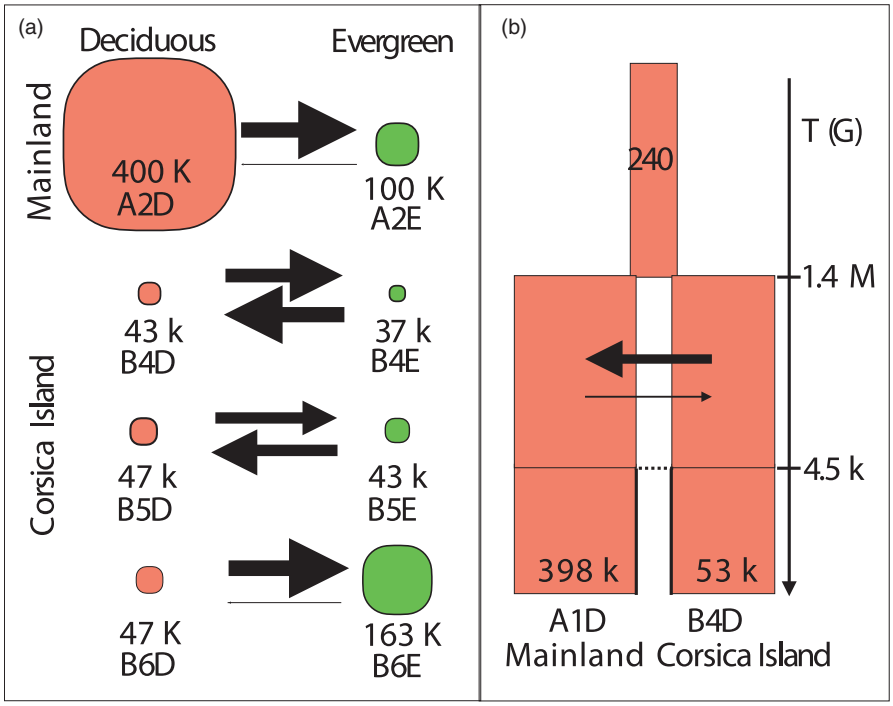


FIGURE 3 Demographic parameters (a) for each of the four pairs of blue tit populations in deciduous and evergreen habitats, estimated using EQ, and (b) for the divergence between populations on the mainland and Corsica, estimated using an ancient migration model. In panel a, circle size is proportional to effective population size (Ne) and arrow width is proportional to migration rate. In panel b, rectangle width is proportional to log10 (effective population size), arrow width is proportional to migration rate, and split time and time of ancient migration are indicated in number of generations

Table S7). The Z chromosome harbored a proportion of nonsynonymous mutations 1.83 times higher than the average for the autosomes. Finally, we observed significantly slightly longer ROH on Corsica compared to mainland ($mean_{island}$ = 44,769 kb vs. $mean_{mainland}$ = 41,586 kb, Wilcoxon rank sum test W = 32,695, p < 2e-16, Figures S7&S8). We also observed significant differences in the count of ROH among populations from the mainland versus those from the island but not among populations within the mainland or the island (ANOVA p < 2e-16; Table S8 for Tukey HSD test).

3.4 | Footprints of selection

Manhattan plots of F_{ST} per SNP revealed no clear evidence for high peaks of F_{ST} for the five comparisons considered (First panels of Figures 4 & 5). Throughout the five Bayescan tests (Corsica vs. mainland populations and the 4 D-4 population pairs), we identified 40 SNPs with a log10(BF)> 0 among which 18 SNPs had a log10(BF)> 1 (Table S9, Figure S9). None of these SNPs were detected twice among the different tests conducted with Bayescan. Among these 40 SNPs, 27 were found for the Corsica versus mainland test, 2 were found for A2D versus A2E, 9 were found for B4D versus B4E, 0 were found for B5 versus B5E, and 2 were found for B6D versus B6E. Among these 40 SNPs, 19 genes were identified (Table S9). Sliding windows of F_{ST} along the genome showed a few outlier windows of modest intensities (last panels of Figures 4 & 5), with almost no parallelism between the tests (Table S10). Expectedly, outlier windows found for the Corsica versus mainland comparison depicted larger F_{ST} than for deciduous versus evergreen comparisons. The outlier F_{ST} windows falling in the top 1% of the F_{ST} windows distribution are reported in Table S10. Using the RDA to identify outlier SNPs with extreme loading values to axes, we identified 229 outlier SNPs associated with habitat type (deciduous vs. evergreen) and 227 SNPs associated with geography (longitude and latitude, essentially representing a Corsica–mainland comparison) (Table S11, Figure S10). For the deciduous versus evergreen test, only one of the RDA outliers was also found outlier in Bayescan tests. For the Corsica–mainland test, eight of the RDA outliers were also found outliers in the Bayescan test. Genes in which the outliers were found are reported in the Table S13.

3.5 | Variation of genomic differentiation with recombination rate

F_{ST} between mainland and Corsican populations was negatively correlated to recombination rate (Figure 6a, linear model p < 2e-16; Figure S11). In contrast, F_{ST} between each pair of deciduous and evergreen populations was not correlated to recombination rate. F_{ST} outlier SNPs between Corsica and the mainland populations and identified by Bayescan or by the RDA were more often found in regions of low recombination (χ^2 test p-values <.01, Figure 6c) than observed for the entire SNPs (Figure 6b). Average recombination rate was on average twice lower for these F_{ST} outlier SNPs compared to the rest of the SNPs (t test p-values < 1e-6). In contrast, F_{ST} outlier SNPs between deciduous and evergreen populations were not preferentially found in regions of low recombination (Figure 6d).

3.6 | Genomic inversions

We detected one putative inversion on chromosome 3, spanning 2.8 Mb, from position 11,838,789 to 14,661,550, and containing

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

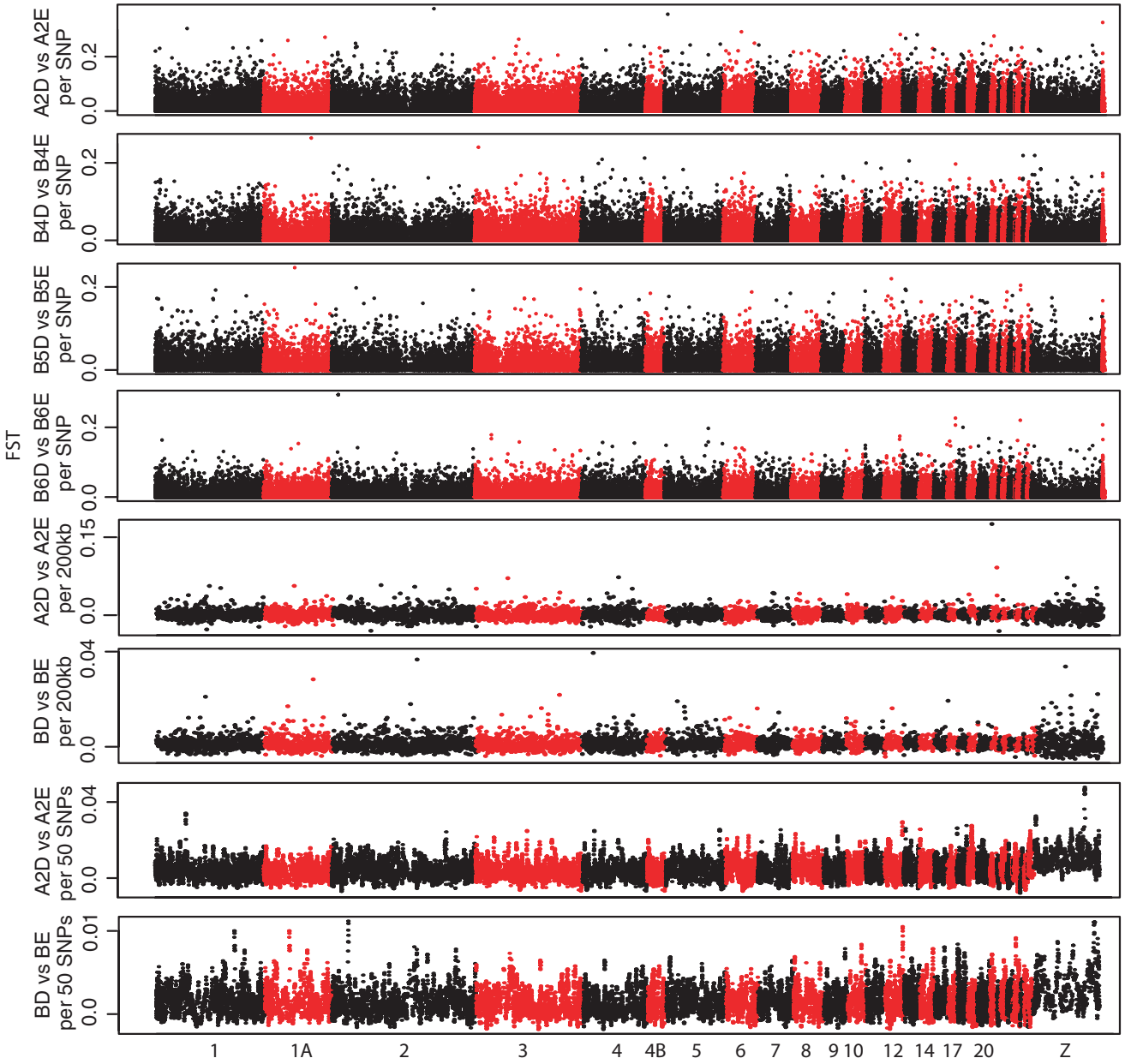


FIGURE 4 F_{ST} Manhattan plots between the four population pairs in deciduous and evergreen forests. F_{ST} are given either by SNP (four first graphs), by 200 kb windows (graphs 5 & 6), or by 50 SNPs windows (graphs 7 & 8). For graphical simplicity, the three pairs of populations in Corsica were grouped for the sliding window F_{ST} Manhattan plots (BD vs. BE). Dots alternate colors between chromosomes. F_{ST} for unplaced loci are shown at the end of each of the first four Manhattan plots

390 SNPs. The estimated recombination rate was on average nine times lower (*t* test: $p < 2.2e\text{-}16$) in this region of the chromosome compared to the rest of the chromosome (Figure 7a). However, the SNP density did not appear reduced at the location of this putative inversion (Figure 7a), suggesting a recent drop of recombination (*i.e.*, dissimilar to what is expected in a peri-centromeric location). We did not find evidence for an increase of read depth at specific loci in the putative inversion or its vicinity in inverted individuals compared to normal ones. We did however notice the absence (read depth = 0) of three loci from the inversion, in the four inverted individuals. The sliding PCA detected either two groups (individuals from Corsica and from mainland) outside of the putative inversion, or four groups at the putative inversion location, separating individuals from Corsica and three groups for mainland individuals (Figure 7b). We assumed that the inverted segment was the one with the lower frequency on the mainland and that was absent from Corsica. The Lostruct analysis confirmed for this same region the presence of four successive blocks of 100 SNPs highly discriminating individuals (Figure 7c). The average F_{ST} between the three groups of mainland individuals was much higher (*t* test: $p < 2.2e\text{-}16$, Figure 7d) for SNPs in the putative inversion ($F_{ST} = 0.26$) than for SNPs outside of this region ($F_{ST} = 5.8e\text{-}4$). π was lower (*t* test: $p = 1.9e\text{-}5$, Figure 7e) for SNPs located in the putative inversion for inverted homozygous individuals ($\pi = 2.5e\text{-}5$) compared to noninverted homozygous individuals ($\pi = 5.3e\text{-}5$). Both

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

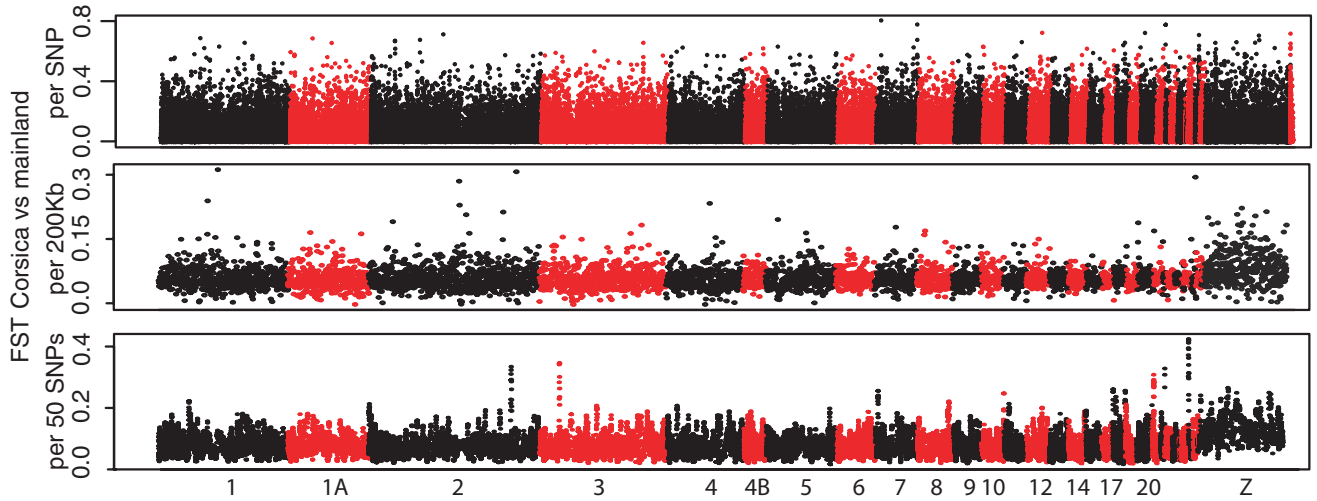


FIGURE 5 F_{ST} Manhattan plots between Corsica and the mainland. F_{ST} are given either by SNP, by 200kb windows or by 50 SNPs windows. Dots alternate colors between chromosomes. F_{ST} for unplaced loci are shown at the end of the first Manhattan plot

inverted and noninverted homozygous individuals had lower π at this inversion compared to heterozygous individuals at this region (π = 7.6e-5) and compared to the three types of individuals for the rest of the chromosome 3 (π ranging from 7.4e-5 to 7.6e-5). The PCA (Figure 7f) and the admixture analysis (Figure 7g) using the SNPs from the inversion clearly delineated inverted, noninverted, and heterozygous individuals, with heterozygous individuals falling at intermediate eigenvalue between the two categories of homozygous individuals. LD was higher (Figure 7h) for the region of the putative inversion (little LD was found when only homozygous noninverted individuals were kept in the analysis). For the region of the inversion, Da was 0.0020 between normal blue tits from the mainland and from Corsica as well as between great tits from the mainland and from Corsica. Da was 0.0044 between inverted and normal blue tits from either the mainland or from Corsica, and ranged from 0.0149 to 0.0152 for interspecific comparisons. Using the formula $T = Da/2\mu$, the inversion likely appeared approximately 220,000 generations ago (around 506,000 years ago). The tree of genetic distance illustrated the divergence of inverted blue tits from noninverted blue tits from both mainland and Corsica (Figure 7i). This region contains 21 genes listed in Table S12. Out of 113 individuals from the mainland, we found 4 inverted homozygous, 23 heterozygous, and 86 noninverted homozygous, and hence, we observed no deviation from the Hardy–Weinberg equilibrium (χ^2 test p-value > .1). The percentage of each genotype was similar across all mainland locations and both habitat types (the inverted segment was at 11% in A1D, 16% in A2D, 17% in A2E, and 14% across these three locations).

3.7 | Gene ontology

None of the gene lists gathered with the different tests (each outlier test among deciduous and evergreen environment and between Corsica and the mainland and the inversion test) yielded any significantly enriched GO term after correction multiple testing (Table S14). The most promising GO (uncorrected p-value <.0025) included the following: (a) for the deciduous–evergreen outlier tests, GO:0060385, axonogenesis involved in innervation; GO:2001013, epithelial cell proliferation; GO:0007194, negative regulation of adenylate cyclase; GO:0090647, modulation of age-related behavioral; GO:1901379, regulation of potassium ion transmembrane transport, (b) for the Corsica–mainland outlier tests, GO:0008344, adult locomotory behavior; GO:2001224, positive regulation of neuron migration, (c) for the inversion, GO:0016446, somatic hypermutation of immunoglobulin; GO:0045910, negative regulation of DNA recombination; GO:0006298, mismatch repair; GO:0008340, determination of adult lifespan.

4 | DISCUSSION

In this study, we investigated demographic history and genome-wide patterns of genetic diversity and differentiation between several populations of blue tits presenting strong phenotypic differences between deciduous versus evergreen forest habitats as well as insular versus mainland contexts. Between populations in deciduous and evergreen forest habitats, demographic modeling showed large gene flow and large effective population sizes (Ne), explaining the low genetic differentiation between these populations. Demographic modeling also revealed that each pair of deciduous and evergreen populations most likely never diverged and maintained high connectivity through migration. We identified slight and mostly unrepeated footprints of divergent selection across these evergreen and deciduous population pairs, which is consistent with their demography and the likely polygenic nature of several traits implicated in their local adaptation. In both insular and mainland populations, we found large Ne, although smaller for insular populations than mainland ones, resulting in different distributions and lengths of ROH. Gene flow between Corsica and the mainland most likely stopped at the end of the last glaciation.

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

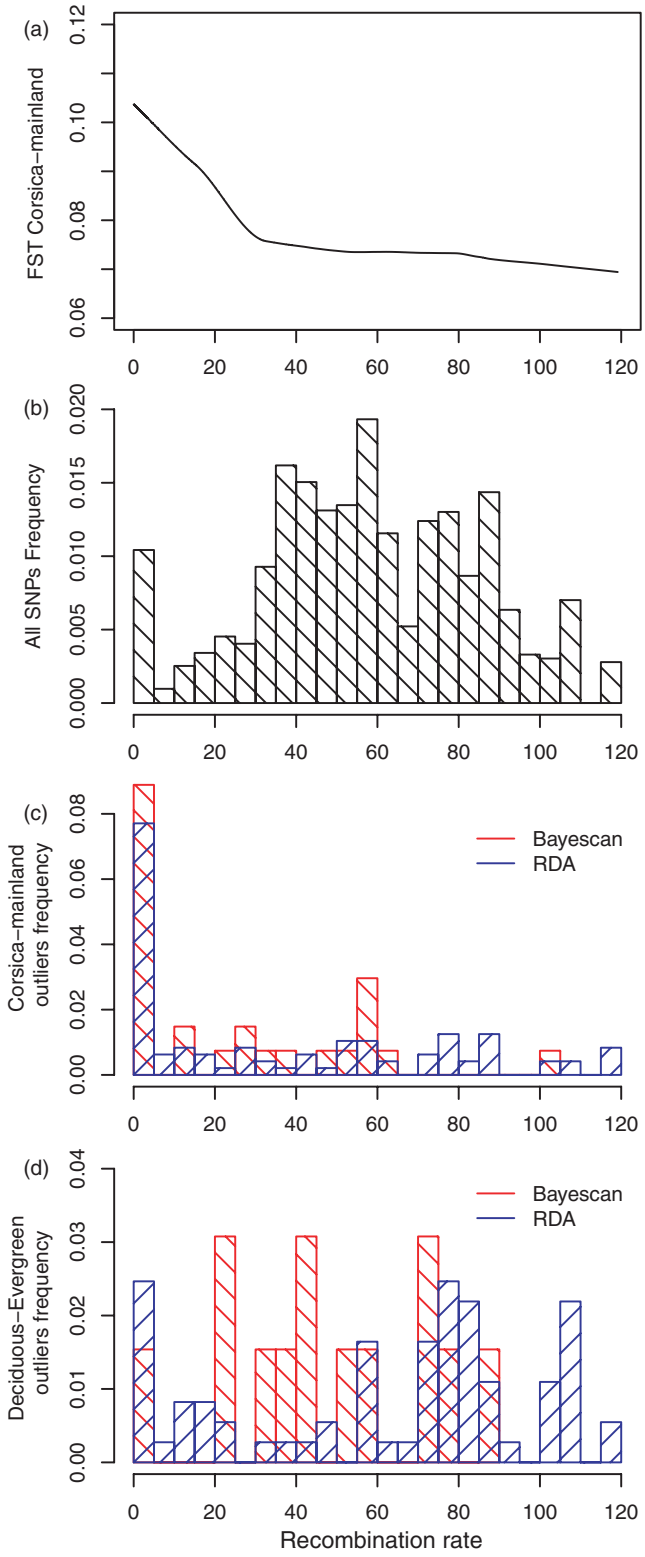


FIGURE 6 Relationship between recombination rate and divergence. (a) Correlation between local recombination rate and F_{ST} between populations on the mainland and in Corsica; (b) SNP frequency per class of recombination rate; (c) outlier SNP frequency for the mainland-Corsica comparison, per class of recombination rate; (d) outlier SNP frequency for the deciduous–evergreen comparisons, per class of recombination rate. In panels c & d, outliers are given for the Bayescan and the RDA tests

Both large Ne and gene flow until the end of the last glaciation hence explained moderate genome-wide differentiation. We identified several genomic footprints of selection, enriched in regions of low recombination only in the case of the mainland/island divergence. Finally, we identified a putative genomic inversion spanning 2.8 Mb, polymorphic in mainland populations only. We discuss these results in the context of the vast scientific knowledge acquired on these blue tit populations during the last four decades, and more generally in the context of genomics of adaptation with gene flow.

4.1 | Divergence between populations in deciduous versus evergreen habitats

Although we found a significant effect of habitat on genetic structure, the genetic differentiation between neighboring deciduous and evergreen populations was low (F_{ST} ranging from 0.0006 to 0.0079). This result is in line with the primary observations realized earlier on a smaller set of populations (Porlier, Garant, et al., 2012; Szulkin et al., 2016). Accordingly, we found high gene flow from deciduous to evergreen populations. Yet, this quantification of high gene flow and low genetic structure contrasted with the demographic knowledge collected on the Blue tit. Indeed, demographic studies suggested restricted dispersal between these populations, with four dispersal events observed between B4D and B4E (5.6 km apart) and none between the B3E and either B4D or B4E (24.1 km), among a total of 2,788 males, 2,672 females, and 25,158 nestling ringed in the three main Corsican sites between 1976 and 2018, with a mean recruitment rate of 6% (Charmantier, com pers). Our interpretation of this contrast between gene flow estimations gathered from population genomic versus recapture data is quadruple. First, dispersal estimation on the field using capture–mark–recapture is very challenging and may require more data than currently collected, to detect rare dispersal events, even though these affect population genetic parameters. Moreover, since natal dispersal in the Blue tit classically ranges between 330 m and 4 km (see Ortego, García-Navas, Ferrer, & Sanz, 2011; Tufto, Ringsby, Dhondt, Adriaensen, & Matthysen, 2005), the long-term monitoring sites in Corsica equipped with nest boxes (black circled dots in Figure 1c) are not ideally spaced to identify the origin of immigrants and the destination of emigrants, and only a small fraction of the landscape favorable to blue tit breeding is covered. Second, only a few migrants are sufficient to decrease the genetic distance between populations, measured through the F_{ST} (Cayuela et al., 2018; Marko & Hart, 2011). In that regard, our results may be compatible with the few dispersal events recorded throughout the years. Third, it is important to note that the number of migrants estimated using demographic analyses represents an average over historical time scales on the order of Ne generations and may have varied widely during contemporary times. Fourth, the large effective population size found both using coalescence and a LD method might be explained by the existence of large "meta-populations" connected by high gene flow and such large Ne

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

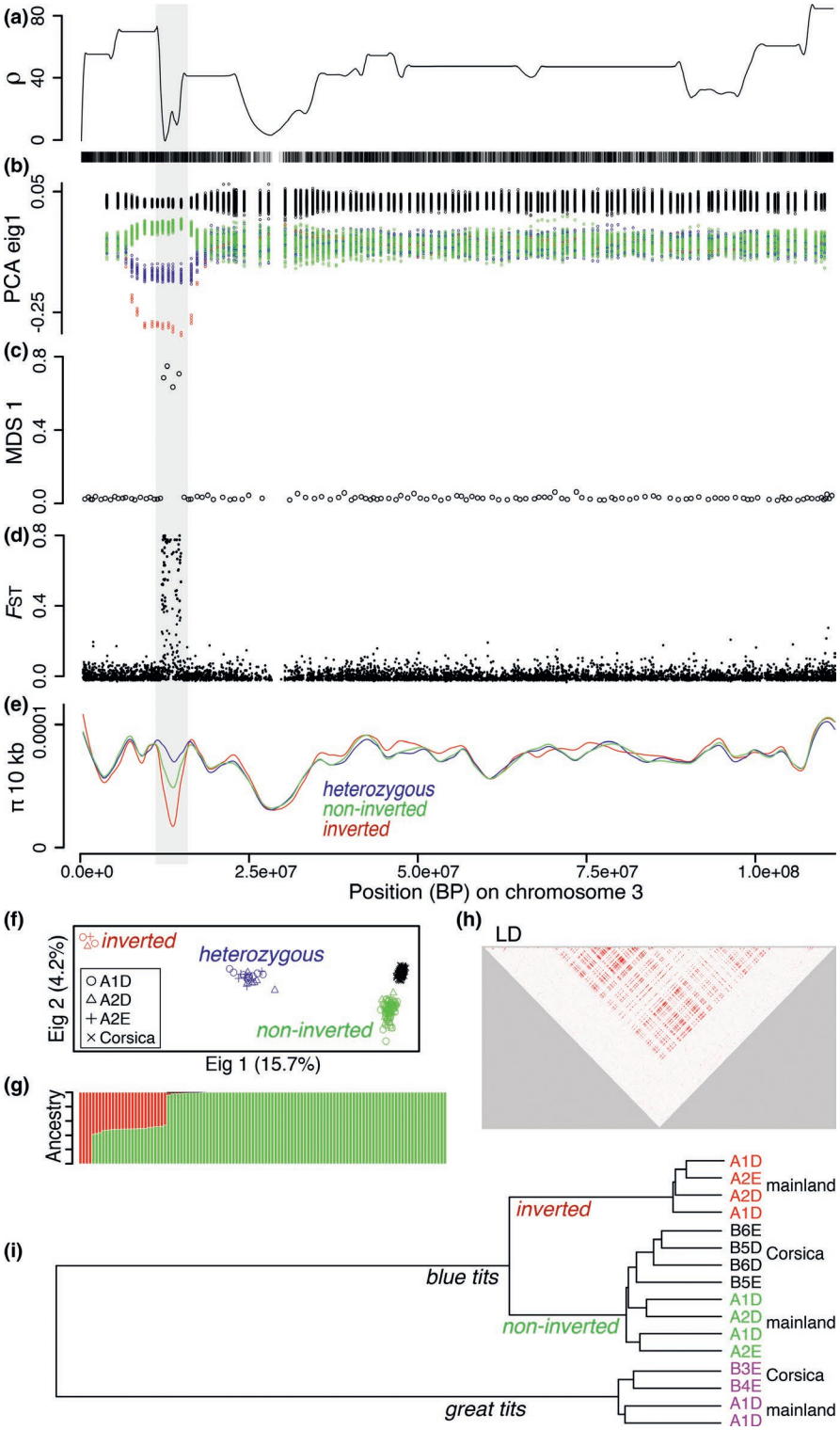


FIGURE 7 Detection and characteristics of a putative inversion on chromosome 3. (a) Recombination rate and SNP density along chromosome 3; (b) individual values for eigenvector 1 of a PCA of all the blue tit individuals from Corsica and the mainland; (c) MDS 1 values from Lostruct analysis; (d) F_{ST} between inverted, noninverted, and heterozygous individuals from the mainland; (e) nucleotide diversity for inverted (in red), noninverted (in green), and heterozygous individuals (in blue) from the mainland; (f) PCA results for the region of the putative inversion for all the blue tit individuals; (g) mainland blue tits' admixture proportions for $K = 2$ for the region of the putative inversion (individual ordered by decreasing ancestry value for the inverted cluster); (h) linkage disequilibrium for SNPs found in the putative inversion (plus 1 Mb each side of the inversion), for mainland blue tits; (i) genetic distances at the inversion, between four inverted blue tits, eight noninverted blue tits, and four great tits (sites of capture are indicated for each individual)

might largely contribute to limit the genetic divergence between populations.

Ne was on average slightly larger in deciduous than in evergreen populations. This could be explained by the higher productivity of deciduous forests compared to evergreen forests, resulting in larger clutches and more fledglings in deciduous habitats (table 1 in Charmantier et al., 2016). However, the very high gene flow between deciduous and evergreen populations and the low genetic differentiation between these populations limits further interpretations. In addition, populations monitored on the long term, for which hundreds of artificial nest boxes have been installed, tended to have larger Ne than nonmonitored populations. Breeding density in the nest box areas studied (Figure 1a) was around 1–1.3 pairs per ha (Blondel et al., 2006), which is most probably 3–5 times higher than natural densities for blue tits when these secondary hole-nesting birds rely on natural cavities only. It is hence possible that the recent

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

availability of nest boxes locally increased productivity, effective population size, and heterozygosity (Box 1).

Demographic modeling indicated that secondary gene flow between previously isolated populations, which could have eroded most genome-wide differentiation outside regions implicated in local adaptation and/or acting as reproductive barriers in particular in low recombining areas (Bierne et al., 2013), was very unlikely. Secondary contacts were strongly rejected in favor of an island model supporting ongoing gene flow without divergence (EQ model) in three out of four population pairs, the last one supporting an isolation with migration model (IM). As is often the case, our demographic estimates should be interpreted cautiously, particularly in the absence of a well-documented mutation rate in the Blue tit. The relatively narrow credible intervals around parameter estimates nevertheless provide confidence in their biological relevance. These results suggest that the focal population pairs occupying deciduous/evergreen habitats have been continuously connected by relatively high gene flow. Hence, genetic divergence between populations in each pair did not originate or build up on pre-existing genetic divergence accumulated during an allopatric phase but rather is supported as a case of genuine recent ecological divergence (Pinho & Hey, 2010; Wang & Bradburd, 2014).

We did not find strong footprints of divergent selection between any of the four pairs of deciduous and evergreen populations (Figure 4). Moreover, the observed outliers were not found repeatedly across the four deciduous–evergreen comparisons but rather each outlier was found only once, and they were not enriched in regions of low recombination. First, this result is in line with the high migration rates observed. High gene flow indeed most likely limits the potential for local adaptation since it limits the accumulation of allelic differentiation, even with high selection coefficients (Lenormand, 2002). Second, this pattern is consistent with the best demographic model being the EQ and not a SC. Indeed, the later would more often create repeated and strong outliers located in regions implicated in reproductive isolation between divergent populations (e.g., Rougemont et al., 2017). Third, this pattern is congruent with a model of local polygenic adaptation involving a transient genetic architecture with multiple alleles of small effects underlying (multiple) quantitative characters (Yeaman, 2015; Yeaman & Whitlock, 2011) and displaying low F_{ST} among loci under divergent selection. Fourth, most documented traits that are involved in the adaptation of blue tits to the deciduous versus evergreen habitat types, such as clutch size or laying date (Blondel, Maistre, Perret, Hurtrez-Boussès, & Lambrechts, 1998; Lambrechts, Blondel, Maistre, & Perret, 1997), are quantitative and most likely rely on a polygenic architecture with alleles of small effects (see, e.g., Santure et al., 2013). However, some statistical issues may limit the interpretation of this result. First, currently available genome scan methods to detect alleles of small effect are statistically limited, especially when applied to relatively small datasets (Hoban et al., 2016; Rockman, 2012). Second, a potential lack of power in detecting outliers could arise from the combination of the use of RAD sequencing and a rapid LD decay along the genome, that may result in too low resolution especially in regions with larger recombination rates (Lowry et al., 2017). Further

Box 1 Personal (Charles Perrier) reflections on my career and my collaboration with Louis Bernatchez

These are personal reflections regarding my early career at Louis Bernatchez lab. This box is hence not related to the population genomics project on blue tits.

I have many ideas to write a very scientific comment about my collaboration with Louis and how it influenced my young carrier… but in those times of confinements due to COVID-19 I am rather willing to discuss more personal and social aspects that also matter a lot in academia. During my stay in Louis laboratory, I admired especially two of Louis talents. The first one, maybe unsurprising, the outstanding leader that Louis is and the positive effect that this has on getting the job done on his laboratory. He is always enthusiastic about new ideas, permanently pushing you to pursue your scientific objectives, and constantly finding the bright side of every hard work done to get it published. I will not detail here how it helped me to grow as a scientist but it surely did in many ways! To my opinion, Louis leadership transcends his indisputable academic performance and reaches other important values, namely equity and diversity. And very importantly for me, Louis stands up and fight when necessary. I will give a small example (hoping I am authorized to): While receiving a price for his hard work from the Quebecois government, Louis had to give a speech. Instead of seeking to cast even more light on his own glory, he gave an engaged speech defending the long standing equitable and open university principles that the government of Quebec was threatening at that time by liberal reforms including the rising of tuition fees. Louis basically claimed that the previous system was the one that allowed many guys like him, coming from relatively modest families to obtain a university degree, to pursue an academic career in science, and to bring the diversity of perspectives the university and the society need… Hence arguing that such price was not only rewarding him but also a relatively open access to the university. I found that speech extremely appropriate and inspiring. In such good times as that ceremony for Louis' price, as well as in bad times as the current covid19 epidemic, I find particularly important to remember that societal and scientific "advances" or "progress" could or should also (or only?) be measured by their contribution to human equity, diversity, and fraternity. Louis, keep up doing good science, motivating young scientists, and fighting for what you believe is right for fish, science, and people! See you soon for a fishing trip in Camargue and "un café double"! Happy birthday! Charly.

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

analyses of the potential outliers found here, as well as a new analysis with higher marker density and more individuals, are needed to better document and discuss the potential genes and biological processes implicated. Particularly, quantitative genomics will be useful to establish links between the level of divergent selection on these putative targets and their effect size on phenotypic trait variation (Gagnaire & Gaggiotti, 2016; Stinchcombe & Hoekstra, 2007). This would notably reveal whether genes under divergent selection are also responsible for the observed phenotypic variation and contribute to assessing of how much of phenotypic variation is adaptive.

Last, we searched for inversions potentially associated with phenotypic variation and/or segregating in both habitats. We report multiple evidences (Figure 6) for a putative inversion on chromosome 3 in mainland populations. However, the proportion of inversions in the genomes of deciduous- versus evergreen-breeding birds did not differ, which most probably precludes a putative role of the inversion in local adaptation in these habitats. Besides, this polymorphism followed the Hardy–Weinberg equilibrium and we found 4 homozygous inverted individuals, suggesting this inversion does not involve an accumulation of lethal recessive deleterious alleles (Jay et al., 2019). The functional consequences of this inversion on variation in life-history traits in the B

lue tit call for further investigations, as achieved in other songbirds (Kim et al., 2017; Knief et al., 2017; Tuttle et al., 2016). In order to genotype this inversion in more birds and to link it to putative phenotypic variation, we could use a PCR-RFLP approach (da Silva et al., 2019).

4.2 | Divergence between blue tit populations in mainland versus insular contexts

Our demographic modeling approach of Corsican and mainland populations revealed that a model with ancient migration was the most probable, which is coherent with the history previously reconstructed for the blue tit complex. Split time of the ancestral populations in two mainland versus island populations connected by gene flow (~3 M years) was relatively coherent with the diversification time found in the literature for the entire blue tit complex (5M [Illera et al., 2011]). Gene flow between the island and the mainland likely stopped around 10,000 years, which is compatible with a gene flow break due to the rise of the flooded stretch between Corsica and the mainland during ice melting and the sea level rise after the last glaciation, during from 17,000 to 5,000 years ago (Jouet et al., 2006; Lambeck & Bard, 2000). Both the signal of population expansion during the diversification period and the gene flow five times larger from the island to the mainland than from the mainland to the island were coherent with a suspected recolonization of the mainland from the islands, as proposed by Illera et al. (2011) based on analyses of nuclear and mitochondrial DNA sequences. The larger effective population size on the mainland than on the island may be due to multiple sources of colonization on the mainland (Taberlet, Meyer, & Bouvet, 1992) and to the much larger mainland area and hence meta-population size compared to the island. Expanding this study with a sampling from the South-East of Europe, one from Sardinia and one from the Iberian Peninsula, could help determine whether the large size estimated for the mainland meta-population studied here might be explained by a mixed recolonization from these refugium (Kvist, Ruokonen, Lumme, & Orell, 1999; Kvist, Viiri, Dias, Rytkönen, & Orell, 2004) and whether insular populations consistently have reduced effective population size. It would also be interesting to explore a model allowing for multiple cycles of isolations and contacts that could have resulted from successive glacial cycles (Hewitt, 2004).

Given the demographic parameters inferred, it is not surprising that the genetic differentiation between populations from the mainland and the island was moderate (F_{ST} = 0.08). However, this differentiation value contrasted with the much larger divergence at mitochondrial DNA observed for these populations (OST = 0.67 between A1D and B4D (Kvist et al., 2004)). This discrepancy between whole nuclear genome and mitochondrial estimates of differentiation could first be explained by the typically four times smaller Ne in mitochondrial DNA compared to nuclear DNA in diploid organisms, resulting in stronger drift and divergence (Smith & Klicka, 2013). Second, accumulation of cytonuclear incompatibilities could limit its introgression (Burton & Barreto, 2012). Alternatively, gene flow may have been greater for males than females during the colonization process from Corsica to the mainland, at the end of the last glacial period. However, this hypothesis is not supported by the general observation that females disperse on average further away than males and that rare long-distance dispersals are also achieved by females (Tufto et al., 2005). Inspecting mitochondrial DNA variation for the birds studied here would be very useful to compare mitochondrial differentiation in our sample to the previous analysis which also included Corsican birds (Kvist et al., 2004). In any case, one should bear in mind that our results stem from integrating coalescent patterns observed across thousands of loci, therefore providing increased resolution to investigate the determinants of demographic divergence compared to approaches based on fewer mitochondrial and nuclear loci.

The contemporary demography inferred between the populations from Corsica and mainland France, especially the absence of gene flow for approximately 4,500 generations and the large effective population size, likely provides a suitable context for the build-up of local adaptation. Moreover, the level of divergence (F_{ST} = 0.08) was probably low enough to detect divergent outlier loci potentially implicated in local adaptation. We identified several outlier SNPs and outlier 200 kb windows showing elevated differentiation, more often in regions of lower recombination than elsewhere in the genome (contrary to the outliers found for the deciduous vs. evergreen populations). Furthermore, F_{ST} was on average higher in these regions of low recombination compared to elsewhere in the genome (which was not the case for deciduous vs. evergreen populations). These patterns of increased divergence in regions with low recombination between divergent populations are commonly observed in other species (Cruickshank & Hahn, 2014; Gagnaire et al., 2018; Nachman & Payseur, 2011), including other bird

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

species (Burri et al., 2015; Spurgin et al., 2019). It is well documented that such increased differentiation in regions with low recombination is not necessarily due to positive selection, or at least not alone, and that it is largely influenced by the effect of recombination in interaction with background selection (Burri et al., 2015; Charlesworth et al., 1993; Perrier & Charmantier, 2019). Moreover, a uniform F_{ST} sliding window, sized in Kb, is expected to dilute signatures of selection in regions of the genome where LD is fast decaying while it is expected to over-represent regions where LD decays much more slowly, hence increasing even more the potential false positives in regions with low recombination (Beissinger, Rosa, Kaeppler, Gianola, & De Leon, 2015; Perrier & Charmantier, 2019). Although perfectible, our sliding window approach using a SNP unit attempted to fix this issue and successfully captured several new outlier regions outside of deserts of recombination. This second window approach should however be improved, for example, by estimating the local neutral enveloped, which computation should integrate local variations in LD, recombination, diversity, but also the demography of populations. As mentioned for the study of adaptive divergence between deciduous and evergreen populations, complementary analyses integrating both genome scans and quantitative genomics would improve our comprehension of the genomic and phenotypic divergence observed between Corsican and mainland blue tits (Gagnaire & Gaggiotti, 2016; Stinchcombe & Hoekstra, 2007).

While the putative genomic inversion on chromosome 3 was absent from Corsica and detected in mainland individuals, its level of divergence from the noninverted sequence indicated that it was likely twice older than the beginning of divergence between blue tit populations from mainland and Corsica. This could first suggest that this polymorphism emerged in mainland blue tit populations and then did not introgress the Corsican populations, maybe due to a local disadvantage and/or genetic incompatibilities or simply due to drift coupled to little gene flow. A second hypothesis could be that this inversion was present in Corsican populations but had been purged out due to a local disadvantage. Such a disadvantage can be due to the typical accumulation of deleterious mutations in such nonrecombining inverted sequences (*e.g.*, Jay et al., 2019). Lastly, it is possible that this inversion has been acquired in mainland populations only recently, after the last period of contact between mainland and Corsican populations, *via* gene flow from a distinct refugium with which connectivity could have been enhanced after deglaciation, hence after the stop of gene flow with Corsica. The age of this inversion, its origin, its biological effects, and the potential accumulation of deleterious mutations need to be inferred more thoroughly *via* genotyping full inversion sequences from individuals in diverse locations and using advanced statistical methods (Lohse, Clarke, Ritchie, & Etges, 2015).

5 | CONCLUSION AND PERSPECTIVES

Our study demonstrated the usefulness of demographic modeling and of the analysis of the variation of genomic diversity and recombination along the genome to uncover the genetic determinants of local adaptation in a small passerine with a large distribution, and occupying different forest habitats. Especially, demographic modeling rejected the hypothesis of a secondary contact between deciduous and evergreen populations and favored a situation with continuous gene flow. These results support the idea that blue tits have adapted to their habitats despite ongoing gene flow, while contextualizing how large gene flow most probably constrained local adaptation (Lenormand, 2002) and favored its architecture based on alleles of small effects (Yeaman, 2015). The genomic modeling also refined our knowledge about the divergence between insular and mainland meta-populations, that have been likely unconnected by gene flow for the last ten thousand years. We also verified the relationship between local recombination rate and differentiation, that should probably be integrated in genome scans looking for footprints of selection (Beissinger et al., 2015; Berner & Roesti, 2017; Booker et al., 2020; Burri et al., 2015; Perrier & Charmantier, 2019). Future investigations will require increased sample sizes and marker density (Hoban et al., 2016; Lotterhos & Whitlock, 2015) in order to better detect loci with small effects that contribute to the quantitative phenotypic variation and local adaptation in blue tits. Lastly, the putative inversion found here would need further analyses since this type of structural variation is often implicated in phenotypic variation (Kirkpatrick, 2006, 2010; Stapley et al., 2017; Wellenreuther et al., 2019).

ACKNOWLEDGEMENTS

We wish to thank Louis Bernatchez for inviting our contribution in this special issue (Box 1). We wish to thank the many people who helped on the blue tit long-term project. We particularly wish to thank Philippe Perret. We also wish to thank Samuel Perret, Christophe De Franceschi, Jacques Blondel, Marta Szulkin, Monica Arias, Patricia Sourouille, Annick Lucas & Boris Delahaie. We thank the associate editor, two reviewers, and the production team for their constructive comments and edits on previous versions of this manuscript. This project was funded by the European Research Council (starting grant ERC-2013-StG-337365-SHE to AC) and the OSU-OREME.

CONFLICT OF INTEREST

None declared.

AUTHORS' CONTRIBUTIONS

CP conceived the study, conducted fieldwork, laboratory analyses, bioinformatics and statistical analyses, interpreted the results and wrote the manuscript. QR conducted bioinformatics and statistical analyses, interpreted the results and edited the manuscript. AC obtained the ERC grant, conceived the study, performed fieldwork, interpreted the results and edited the manuscript.

DATA AVAILABILITY STATEMENT

Individual sequences with metadata are available on NCBI with the SRA accession PRJNA630135. A filtered VCF file with metadata are available on Dryad with the https://doi.org/10.5061/dryad.x69p8czfg.

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ORCID

Charles Perrier https://orcid.org/0000-0001-5820-9374
Quentin Rougemont https://orcid.org/0000-0003-2987-3801
Anne Charmantier https://orcid.org/0000-0002-0691-2647

REFERENCES

Alexa, A., & Rahnenführer, J. (2009) Gene set enrichment analysis with topGO. *Bioconductor Improvement*, *27*. Retrieved from https://bioconductor.org/packages/release/bioc/vignettes/topGO/inst/doc/topGO.pdf

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

Andrews, S. (2010) FastQC: A quality control tool for high throughput sequence data. Retrieved from http://www.bioinformatics.babraham.ac.uk

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*, 1479–1495.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., … Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers (JC Fay, Ed,). *PLoS One*, *3*, e3376–e3377. https://doi.org/10.1371/journal.pone.0003376

Barbato, M., Orozco-terWengel, P., Tapio, M., & Bruford, M. W. (2015). SNeP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Frontiers in Genetics*, *6*, 1294–1297. https://doi.org/10.3389/fgene.2015.00109

Barton, N., & Bengtsson, B. O. (1986). The barrier to genetic exchange between hybridising populations. *Heredity*, *57*, 357–376. https://doi.org/10.1038/hdy.1986.135

Beissinger, T. M., Rosa, G. J. M., Kaeppler, S. M., Gianola, D., & De Leon, N. (2015). Defining window-boundaries for genomic analyses using smoothing spline techniques. *Genetics Selection Evolution*, *47*, 30–39. https://doi.org/10.1186/s12711-015-0105-9

Berner, D., & Roesti, M. (2017). Genomics of adaptive divergence with chromosome-scale heterogeneity in crossover rate. *Molecular Ecology*, *114*, 7061–7119. https://doi.org/10.1111/mec.14373

Bierne, N., Roze, D., & Welch, J. J. (2013). Pervasive selection or is it…? why are FST outliers sometimes so frequent? *Molecular Ecology*, *22*, 2061–2064.

Bierne, N., Welch, J., Loire, E., Bonhomme, F., & David, P. (2011). The coupling hypothesis: Why genome scans may fail to map local adaptation genes. *Molecular Ecology*, *20*, 2044–2072. https://doi.org/10.1111/j.1365-294X.2011.05080.x

Blanquart, F., Kaltz, O., Nuismer, S. L., & Gandon, S. (2013). A practical guide to measuring local adaptation. *Ecology Letters*, *16*, 1195–1205. https://doi.org/10.1111/ele.12150

Blondel, J., Maistre, M., Perret, P., Hurtrez-Boussès, S., & Lambrechts, M. (1998). Is the small clutch size of a Corsican blue tit population optimal? *Oecologia*, *117*, 80–89. https://doi.org/10.1007/s004420050634

Blondel, J., Thomas, D. W., Charmantier, A., Perret, P., Bourgault, P., & Lambrechts, M. M. (2006). A thirty-year study of phenotypic and genetic variation of blue tits in Mediterranean habitat mosaics. *BioScience*, *56*, 661–673. https://doi.org/10.1641/0006-3568(2006)56[661:ATSOPA]2.0.CO;2

Booker, T. R., Yeaman, S., & Whitlock, M. (2020) Variation in recombination rate affects detection of FST outliers under neutrality. *bioRxiv*. https://doi.org/10.1101/2020.02.06.937813

Bosse, M., Spurgin, L. G., Laine, V. N., Cole, E. F., Firth, J. A., Gienapp, P., … Slate, J. (2017) Recent natural selection causes adaptive evolution of an avian polygenic trait. *Science (New York, N.Y.)*, *358*, 365–368.

Broggi, J., Hohtola, E., Orell, M., & Nilsson, J.-Å. (2005). Local adaptation to winter conditions in a passerine spreading north: A common-garden approach. *Evolution*, *59*, 1600–1603. https://doi.org/10.1111/j.0014-3820.2005.tb01810.x

Burri, R., Nater, A., Kawakami, T., Mugal, C. F., Olason, P. I., Smeds, L., … 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*, 1656–1665.

Burton, R. S., & Barreto, F. S. (2012). A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? *Molecular Ecology*, *21*, 4942–4957. https://doi.org/10.1111/mec.12006

Carbonell, R., Perez-Tris, J., & Telleria, J. L. (2003). Effects of habitat heterogeneity and local adaptation on the body condition of a forest passerine at the edge of its distributional range. *Biological Journal of the Linnean Society*, *78*, 479–488. https://doi.org/10.1046/j.0024-4066.2002.00156.x

Caro, S. P., Charmantier, A., Lambrechts, M. M., Blondel, J., Balthazart, J., & Williams, T. D. (2009). Local adaptation of timing of reproduction: Females are in the driver's seat. *Functional Ecology*, *23*, 172–179. https://doi.org/10.1111/j.1365-2435.2008.01486.x

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140. https://doi.org/10.1111/mec.12354

Cayuela, H., Rougemont, Q., Prunier, J. G., Moore, J.-S., Clobert, J., Besnard, A., & Bernatchez, L. (2018). Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Molecular Ecology*, *27*, 3976–4010. https://doi.org/10.1111/mec.14848

Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, *134*, 1289–1303.

Charmantier, A., Doutrelant, C., Dubuc-Messier, G., Fargevieille, A., & Szulkin, M. (2016). Mediterranean blue tits as a case study of local adaptation. *Evolutionary Applications*, *9*, 135–152. https://doi.org/10.1111/eva.12282

Charmantier, A., Kruuk, L. E. B., Blondel, J., & Lambrechts, M. M. (2004). Testing for microevolution in body size in three blue tit populations. *Journal of Evolutionary Biology*, *17*, 732–743. https://doi.org/10.1111/j.1420-9101.2004.00734.x

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*, 3133–3157. https://doi.org/10.1111/mec.12796

Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., & François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, *25*, 410–418. https://doi.org/10.1016/j.tree.2010.04.001

Cutter, A. D., & Payseur, B. A. (2013). Genomic signatures of selection at linked sites: Unifying the disparity among species. *Nature Reviews Genetics*, *14*, 262–274. https://doi.org/10.1038/nrg3425

da Silva, V. H., Laine, V. N., Bosse, M., Spurgin, L. G., Derks, M. F. L., van Oers, K., … Groenen, M. A. M. (2019). The genomic complexity of a large inversion in great tits. *Genome Biology and Evolution*, *11*, 1870–1881. https://doi.org/10.1093/gbe/evz106

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

de Villemereuil, P., Frichot, É., Bazin, É., François, O., & Gaggiotti, O. E. (2014). Genome scan methods against more complex models: When

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

and how much should we trust them? *Molecular Ecology*, *23*, 2006–2019. https://doi.org/10.1111/mec.12705

Delahaie, B., Charmantier, A., Chantepie, S., Garant, D., Porlier, M., & Teplitsky, C. (2017). Conserved G-matrices of morphological and life-history traits among continental and island blue tit populations. *Heredity*, *119*, 76–87. https://doi.org/10.1038/hdy.2017.15

Dubuc-Messier, G., Caro, S. P., Perrier, C., van Oers, K., Réale, D., & Charmantier, A. (2018). Gene flow does not prevent personality and morphological differentiation between two blue tit populations. *Journal of Evolutionary Biology*, *23*, 960–1011. https://doi.org/10.1111/jeb.13291

Ewing, G. B., & Jensen, J. D. (2016). The consequences of not accounting for background selection in demographic inference. *Molecular Ecology*, *25*, 135–141.

Excoffier, L., Hofer, T., & Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*, 285–298. https://doi.org/10.1038/hdy.2009.74

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, *180*, 977–993. https://doi.org/10.1534/genetics.108.092221

Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., & Lasky, J. R. (2015). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, *25*, 104–120. https://doi.org/10.1111/mec.13476

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*, *27*, 2215–2233. https://doi.org/10.1111/mec.14584

Gagnaire, P.-A., & Gaggiotti, O. E. (2016). Detecting polygenic selection in marine populations by combining population genomics and quantitative genetics approaches. *Current Zoology*, *62*, 603–616. https://doi.org/10.1093/cz/zow088

Gagnaire, P.-A., Lamy, J.-B., Cornette, F., Heurtebise, S., Dégremont, L., Flahauw, E., … Lapègue, S. (2018). Analysis of genome-wide differentiation between native and introduced populations of the cupped oysters *Crassostrea gigas* and *Crassostrea angulata*. *Genome Biology and Evolution*, *10*, 2518–2534. https://doi.org/10.1093/gbe/evy194

Hansson, B., Sigeman, H., Stervander, M., Tarka, M., Ponnikas, S., Strandh, M., … Hasselquist, D. (2018). Contrasting results from GWAS and QTL mapping on wing length in great reed warblers. *Molecular Ecology Resources*, *18*, 867–876. https://doi.org/10.1111/1755-0998.12785

Hewitt, G. M. (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc London. Series B: Biol Sci*, *359*(1442), 183–195. https://doi.org/10.1098/rstb.2003.1388.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., … Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, *188*, 379–397. https://doi.org/10.1086/688018

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *Plos Genetics*, *6*, e1000862. https://doi.org/10.1371/journal.pgen.1000862

Illera, J. C., Koivula, K., Broggi, J., Päckert, M., Martens, J., & Kvist, L. (2011). A multi-gene approach reveals a complex evolutionary history in the Cyanistes species group. *Molecular Ecology*, *20*, 4123–4139. https://doi.org/10.1111/j.1365-294X.2011.05259.x

Jay, P., Chouteau, M., Whibley, A., Bastide, H., Llaurens, V., & Parrinello, H. (2019). Mutation accumulation in chromosomal inversions maintains wing pattern polymorphism in a butterfly. *bioRxiv*, 736504.

Jouet, G., Berné, S., Rabineau, M., Bassetti, M. A., Bernier, P., Dennielou, B., … Taviani, M. (2006). Shoreface migrations at the shelf edge and sea-level changes around the Last Glacial Maximum (Gulf of Lions, NW Mediterranean). *Marine Geology*, *234*, 21–42. https://doi.org/10.1016/j.margeo.2006.09.012

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*, 1225–1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

Kim, K.-W., Bennison, C., Hemmings, N., Brookes, L., Hurley, L. L., Griffith, S. C., … Slate, J. (2017). A sex-linked supergene controls sperm morphology and swimming speed in a songbird. *Nature Ecology & Evolution*, *1*, 1168–1176. https://doi.org/10.1038/s41559-017-0235-2

Kirkpatrick, M. (2010). How and why chromosome inversions evolve. *PLoS Biology*, *8*, e1000501. https://doi.org/10.1371/journal.pbio.1000501

Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, *173*, 419–434. https://doi.org/10.1534/genetics.105.047985

Knief, U., Forstmeier, W., Pei, Y., Ihle, M., Wang, D., Martin, K., … Kempenaers, B. (2017). A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nature Ecology & Evolution*, *1*, 1177–1184. https://doi.org/10.1038/s41559-017-0236-1

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*(Pt 5), 495–502. https://doi.org/10.1038/sj.hdy.6885130

Kvist, L., Viiri, K., Dias, P. C., Rytkönen, S., & Orell, M. (2004). Glacial history and colonization of Europe by the blue tit *Parus caeruleus*. *Journal of Avian Biology*, *35*, 352–359.

Laaksonen, T., Sirkiä, P. M., Calhim, S. et al (2015). Sympatric divergence and clinal variation in multiple coloration traits of Ficedulafly catchers. *Journal of Evolutionary Biology*, *28*, 779–790.

Laine, V. N., Gossmann, T. I., Schachtschneider, K. M., Garroway, C. J., Madsen, O., Verhoeven, K. J. F., … Groenen, M. A. M. (2016). Evolutionary signals of selection on cognition from the great tit genome and methylome. *Nature Communications*, *7*, 1–9. https://doi.org/10.1038/ncomms10474

Lambeck, K., & Bard, E. (2000). Sea-level change along the French Mediterranean coast for the past 30 000 years. *Earth and Planetary Science Letters*, *175*, 203–222. https://doi.org/10.1016/S0012-821X(99)00289-7

Lambrechts, M., Blondel, J., Maistre, M., & Perret, P. (1997). A single response mechanism is responsible for evolutionary adaptive variation in a bird's laying date. *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 5153–5155.

Lamichhaney, S., Han, F., Berglund, J., Wang, C., Almen, M. S., Webster, M. T., … Andersson, L. (2016). A beak size locus in Darwin's finches facilitated character displacement during a drought. *Science*, *352*, 470–474. https://doi.org/10.1126/science.aad8786

Legendre, P., & Fortin, M.-J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, *10*, 831–844. https://doi.org/10.1111/j.1755-0998.2010.02866.x

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*, 183–189. https://doi.org/10.1016/S0169-5347(02)02497-7

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Li, H., & Ralph, P. (2019). Local PCA shows how the effect of population structure differs along the genome. *Genetics*, *211*, 289–304. https://doi.org/10.1534/genetics.118.301747

Lohse, K., Clarke, M., Ritchie, M. G., & Etges, W. J. (2015). Genome-wide tests for introgression between cactophilic Drosophila implicate a role of inversions during speciation. *Evolution*, *69*, 1178–1190.

Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, *24*, 1031–1046. https://doi.org/10.1111/mec.13100

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2017). Responsible RAD: Striving for best practices in population genomic studies of adaptation. *Molecular Ecology Resources*, *17*, 366–369. https://doi.org/10.1111/1755-0998.12677

Lowry, D. B., & Willis, J. H. (2010). A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation (NH Barton, Ed,). *PLoS Biology*, *8*, e1000500-14.

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*, 3498–3514. https://doi.org/10.1111/mec.14811

Ma, J., & Amos, C. I. (2012). Investigation of inversion polymorphisms in the human genome using principal components analysis. *PLoS One*, *7*, e40224. https://doi.org/10.1371/journal.pone.0040224

Manel, S., Perrier, C., Pratlong, M., Abi-Rached, L., Paganini, J., Pontarotti, P., & Aurelle, D. (2016). Genomic resources and their influence on the detection of the signal of positive selection in genome scans. *Molecular Ecology*, *25*, 170–184. https://doi.org/10.1111/mec.13468

Marko, P. B., & Hart, M. W. (2011). The complex analytical landscape of gene flow inference. *Trends in Ecology & Evolution*, *26*, 448–456. https://doi.org/10.1016/j.tree.2011.05.007

McGinnis, S., & Madden, T. L. (2004). BLAST: At the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research*, *32*, W20–W25. https://doi.org/10.1093/nar/gkh435

McVean, G. A. T. (2004). The fine-scale structure of recombination rate variation in the human genome. *Science*, *304*, 581–584. https://doi.org/10.1126/science.1092500

Nachman, M. W., & Payseur, B. A. (2011). Recombination rate variation and speciation: Theoretical predictions and empirical results from rabbits and mice. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 409–421.

Narum, S. R., & Hess, J. E. (2011). Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, *11*, 184–194.

Oksanen, J., Kindt, R., Legendre, P., Bob O'Hara, M., Stevens, H. H., & Oksanen, M. J. (2007). The Vegan package. *Community Ecology Package*, *10*, 631–637.

Ortego, J., García-Navas, V., Ferrer, E. S., & Sanz, J. J. (2011). Genetic structure reflects natal dispersal movements at different spatial scales in the blue tit, *Cyanistes caeruleus*. *Animal Behaviour*, *82*, 131–137. https://doi.org/10.1016/j.anbehav.2011.04.007

Päckert, M., Martens, J., Tietze, D. T., Dietzen, C., Wink, M., & Kvist, L. (2007). Calibration of a molecular clock in tits (Paridae)—Do nucleotide substitution rates of mitochondrial genes deviate from the 2% rule? *Molecular Phylogenetics and Evolution*, *44*, 1–14. https://doi.org/10.1016/j.ympev.2007.03.006

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, *13*, 946–952. https://doi.org/10.1111/1755-0998.12129

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

Perrier, C., Delahaie, B., & Charmantier, A. (2018). Heritability estimates from genomewide relatedness matrices in wild populations: Application to a passerine, using a small sample size. *Molecular Ecology Resources*, *18*(4), 838–853. https://doi.org/10.1111/1755-0998.12886

Perrier, C., Ferchaud, A.-L., Sirois, P., Thibault, I., & Bernatchez, L. (2017). Do genetic drift and accumulation of deleterious mutations preclude adaptation? Empirical investigation using RADseq in a northern lacustrine fish. *Molecular Ecology*, *26*, 6317–6335. https://doi.org/10.1111/mec.14361

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Molecular Biology and Evolution*, *31*, 1929–1936. https://doi.org/10.1093/molbev/msu136

Pinho, C., & Hey, J. (2010). Divergence with gene flow: Models and data. *Annual Review of Ecology Evolution and Systematics*, *41*, 215–230. https://doi.org/10.1146/annurev-ecolsys-102209-144644

Porlier, M., Garant, D., Perret, P., & Charmantier, A. (2012). Habitat-linked population genetic differentiation in the blue tit *Cyanistes caeruleus*. *Journal of Heredity*, *103*, 781–791. https://doi.org/10.1093/jhered/ess064

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. *American Journal of Human Genetics*, *81*, 559–575. https://doi.org/10.1086/519795

Reid, N. M., Proestou, D. A., Clark, B. W., Warren, W. C., Colbourne, J. K., Shaw, J. R., … Whitehead, A. (2016). The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science*, *354*, 1305–1308. https://doi.org/10.1126/science.aah4993

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols*, *12*, 2640–2659. https://doi.org/10.1038/nprot.2017.123

Rockman, M. V. (2012). The QTN program and the alleles that matter for evolution: All that's gold does not glitter. *Evolution*, *66*, 1–17. https://doi.org/10.1111/j.1558-5646.2011.01486.x

Rougemont, Q., & Bernatchez, L. (2018). The demographic history of Atlantic Salmon (*Salmo salar*) across its distribution range reconstructed from approximate Bayesian computations. *Evolution*, 1–3.

Rougemont, Q., Gagnaire, P.-A., Perrier, C., Genthon, C., Besnard, A.-L., Launey, S., & Evanno, G. (2017). Inferring the demographic history underlying parallel genomic divergence among pairs of parasitic and nonparasitic lamprey ecotypes. *Molecular Ecology*, *26*, 142–162.

Rougemont, Q., Moore, J.-S., Leroy, T., Normandeau, E., Rondeau, E. B., Withler, R., … Bernatchez, L. (2019). Demographic history, linked selection, and recombination shape the genomic landscape of a broadly distributed Pacific salmon. *bioRxiv*, *123*, 887–937.

Roux, C., Fraïsse, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, *14*, e2000234. https://doi.org/10.1371/journal.pbio.2000234

Roux, C., Tsagkogeorga, G., Bierne, N., & Galtier, N. (2013). Crossing the species barrier: Genomic hotspots of introgression between two highly divergent Ciona intestinalis species. *Molecular Biology and Evolution*, *30*, 1574–1587. https://doi.org/10.1093/molbev/mst066

Santure, A. W., De Cauwer, I., Robinson, M. R., Poissant, J., Sheldon, B. C., & Slate, J. (2013). Genomic dissection of variation in clutch size and egg mass in a wild great tit (*Parus major*) population. *Molecular Ecology*, *22*, 3949–3962.

Santure, A. W., Gratten, J., Mossman, J. A., Sheldon, B. C., & Slate, J. (2011). Characterisation of the transcriptome of a wild great tit *Parus major* population by next generation sequencing. *BMC Genomics*, *12*, 283. https://doi.org/10.1186/1471-2164-12-283

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Publishing Group*, *14*, 807–820. https://doi.org/10.1038/nrg3522

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Schrider, D. R., Shanku, A. G., & Kern, A. D. (2016). Effects of linked selective sweeps on demographic inference and model selection. *Genetics*, *204*, 1207–1223. https://doi.org/10.1534/genetics.116.190223

Seutin, G., White, B. N., & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, *69*, 82–90. https://doi.org/10.1139/z91-013

Sinclair-Waters, M., Bradbury, I. R., Morris, C. J., Lien, S., Kent, M. P., & Bentzen, P. (2017). Ancient chromosomal rearrangement associated with local adaptation of a post-glacially colonized population of Atlantic Cod in the northwest Atlantic. *Molecular Ecology*, *27*, 339–351. https://doi.org/10.1111/mec.14442

Smith, B. T., & Klicka, J. (2013). Examining the role of effective population size on mitochondrial and multilocus divergence time discordance in a songbird. *PLoS One*, *8*(2), e55161. https://doi.org/10.1371/journal.pone.0055161

Sousa, V. C., Carneiro, M., Ferrand, N., & Hey, J. (2013). Identifying loci under selection against gene flow in isolation-with-migration models. *Genetics*, *194*, 211–233. https://doi.org/10.1534/genetics.113.149211

Spurgin, L. G., Bosse, M., Adriaensen, F., Albayrak, T., Barboutis, C., Belda, E. J., … Slate, J. (2019) The great tit HapMap project: A continental-scale analysis of genomic variation in a songbird. *bioRxiv*, 1–24.

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

Stinchcombe, J. R., & Hoekstra, H. E. (2007). Combining population genomics and quantitative genetics: Finding the genes underlying ecologically important traits. *Heredity*, *100*, 158–170.

Szulkin, M., Gagnaire, P. A., Bierne, N., & Charmantier, A. (2016). Population genomic footprints of fine-scale differentiation between habitats in Mediterranean blue tits. *Molecular Ecology*, *25*, 542–558. https://doi.org/10.1111/mec.13486

Taberlet, P., Meyer, A., & Bouvet, J. (1992). Unusual mitochondrial DNA polymorphism in two local populations of blue tit *Parus caeruleus*. *Molecular Ecology*, *1*, 27–36. https://doi.org/10.1111/j.1365-294X.1992.tb00152.x

Teplitsky, C., Tarka, M., Møller, A. P., Nakagawa, S., Balbontín, J., Burke, T. A., … Charmantier, A. (2014). Assessing multivariate constraints to evolution across ten long-term avian studies. *PLoS One*, *9*, e90444. https://doi.org/10.1371/journal.pone.0090444

Tiffin, P., & Ross-Ibarra, J. (2014). Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution*, *29*, 1–8. https://doi.org/10.1016/j.tree.2014.10.004

Tine, M., Kuhl, H., Gagnaire, P.-A., Louro, B., Desmarais, E., Martins, R. S. T., … Reinhardt, R. (2014). European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature Communications*, *5*, 1–10. https://doi.org/10.1038/ncomms6770

Todesco, M., Owens, G. L., Bercovich, N., Jean-Sebastien, L., Shaghayegh, S., Dylan, B., … Rieseberg, L. H. (2019) Massive haplotypes underlie ecotypic differentiation in sunflowers. *bioRxiv*, *19*, 337–366.

Tufto, J., Ringsby, T. H., Dhondt, A. A., Adriaensen, F., & Matthysen, E. (2005). A parametric model for estimation of dispersal patterns applied to five passerine spatially structured populations. *American Naturalist*, *165*, E13–E26. https://doi.org/10.1086/426698

Tuttle, E. M., Bergland, A. O., Korody, M. L., Brewer, M. S., Newhouse, D. J., Minx, P., … Balakrishnan, C. N. (2016). Divergence and functional degradation of a sex chromosome-like supergene. *Current Biology*, *26*(3), 344–350. https://doi.org/10.1016/j.cub.2015.11.069

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*, 5649–5662. https://doi.org/10.1111/mec.12938

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*, 1–14.

Wellenreuther, M., & Hansson, B. (2016). Detecting polygenic evolution: problems, pitfalls, and promises. *Trends in Genetics*, *32*, 155–164. https://doi.org/10.1016/j.tig.2015.12.004

Wellenreuther, M., Merot, C., Berdan, E., & Bernatchez, L. (2019). Going beyond SNPs: The role of structural genomic variants in adaptive evolution and species diversification. *Molecular Ecology*, *28*, 1203–1209. https://doi.org/10.1111/mec.15066

Yeaman, S. (2015). Local adaptation by alleles of small effect. *American Naturalist*, *186*(Suppl 1), S74–S89. https://doi.org/10.1086/682405

Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under migration-selection balance. *Evolution*, *65*, 1897–1911. https://doi.org/10.1111/j.1558-5646.2011.01269.x

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*, 3326–3328. https://doi.org/10.1093/bioinformatics/bts606

SUPPORTING INFORMATION

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

How to cite this article: Perrier C, Rougemont Q, Charmantier A. Demographic history and genomics of local adaptation in blue tit populations. *Evol Appl*. 2020;13:1145–1165. https://doi.org/10.1111/eva.13035

17524571, 2020, 6, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/eva.13035 by Cochrane France, Wiley Online Library on [01/12/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License