Title: Demographic history and genomics of local adaptation in blue tit populations

Running title: Population genomics in blue tit

Authors: Perrier Charles 1,2 , Rougemont Quentin 3 , Charmantier Anne 2

Addresses: 1. Centre d'Ecologie Fonctionnelle et Evolutive UMR 5175, Campus CNRS, 1919 route de Mende, 34293 Montpellier Cedex 5, 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 QC G1V OA6.

Correspondence: Perrier Charles charles.perier@icloud.com

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 variation in population demography and variation in recombination rate and 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} = 0.004 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 migrations, followed by isolation since the end of the last glaciation. Effective population sizes were overall large, yet smaller on the island than on the mainland. Weak and non-parallel 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 2.8Mb. These results provide insights into the demographic history and genetic architecture of local adaptation in blue tit populations at multiple geographic scales.

Introduction

Local adaptation in heterogeneous environments imposing divergent selection on counterparts of a same species has fascinated scientists for decades (Kawecki & Ebert 2004; Blanquart *et al.* 2013). However, it is only with the recent advance in sequencing technologies that numerous recent empirical studies started discovering the genomic

45 features and processes underlying local adaptation in heterogeneous environments
46 (Savolainen *et al.* 2013; Tiffin & Ross-Ibarra 2014; Manel *et al.* 2016; Hoban *et al.* 2016).
47 Especially, many studies used so-called F_{ST} genome scans between populations (Excoffier
48 *et al.* 2009; Narum & Hess 2011; de Villemereuil *et al.* 2014) to detect genes under
49 selection, and potentially implicated in local adaptations (Hohenlohe et al. 2010;
50 Lamichhaney et al. 2016; Reid et al. 2016; Perrier et al. 2017). Nonetheless, genome
51 scans need improvements for example for studying polygenic variation (Rockman 2012;
52 Wellenreuther & Hansson 2016). Genome scans also need to be interpreted in the light
53 of recombination variation along the genome (Stapley et al. 2017, Booker et al. 2020)
54 since it shapes the potential extent of diversity and of differentiation in the different
55 regions of the genome (Cutter & Payseur 2013; Tine et al. 2014; Burri et al. 2015;
56 Rougemont et al. 2019). Genome scans results need to be interpreted in the light of the
57 demographic history of populations since i) spatio-temporal variation in gene flow and
58 of effective population size affect the possibility of local adaptations (Lenormand 2002)
59 and the level of false positive due to genetic drift, and ii) regions implicated in
60 reproductive isolation accumulated during allopatry showing increased differentiation
61 upon secondary contacts can be misinterpreted as genuine footprints of recent
62 divergent selection (Bierne et al. 2011; 2013). Lastly, structural variants such as
63 inversions also require special attention, since they protect adaptive gene sets from
64 recombination (Kirkpatrick 2006; 2010; Stapley *et al.* 2017; Wellenreuther *et al.* 2019)
65 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 (Carbonell et al. 2003; Broggi et al. 2005; Laaksonen et al. 2015). In particular, several blue tit populations breeding in a heterogeneous habitat 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 et al. 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 behavioural traits (e.g. song characteristics and handling aggression) (Charmantier et al 2016). These phenotypic differences were found at two spatial scales. First, 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, neighbouring populations in deciduous and evergreen habitats are weakly genetically differentiated (Porlier et al. 2012b; Szulkin et al. 2016; Dubuc-Messier et al. 2018) 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

88 the last glacial maximum (the sea level raised of 120 m from 17,000 to 5,000 years ago
89 (Lambeck & Bard 2000; Jouet et al. 2006)), are smaller and more colourful than their
90 mainland counterparts (again, see table 1 in Charmantier et al 2016) and are listed as
91 two subspecies (*Cyanistes caeruleus caeruleus* on mainland Europe and *Cyanistes*
92 *caeruleus ogliastrae* mainly in Corsica and Sardinia). Overall, traits displaying these
93 strong phenotypic differences had heritability ranging from 0.20-0.43 (eg for lay date,
94 (Caro et al. 2009)) to 0.42-0.60 (eg for tarsus length, (Teplitsky et al. 2014; Delahaie et
95 al. 2017; Perrier et al. 2018)), are classically related with fitness, and hence could be
96 involved in a local adaptation process. In particular, the heritability of lay date and the
97 breeding time gap between populations could result in reproductive isolation by
98 breeding time, limiting gene flow and favouring local adaptation. The studied traits were
99 typically quantitative (Charmantier et al. 2016), hence probably controlled by a
100 polygenic architecture (Perrier et al. 2018) composed of many loci with small individual
101 effects, as found in similar traits for other passerine birds (Santure et al. 2013; Bosse et
102 al. 2017; Hansson et al. 2018; Lundregan et al. 2018). Overall, given their phenotypic,
103 demographic and genetic characteristics, these blue tit populations provide an ideal
104 framework to study the genomic architecture of polygenic adaptation in heterogeneous
105 environments.

106 We investigated genome wide patterns of genetic diversity and differentiation
107 and the demographic history between several populations of blue tit from Southern
108 France, in heterogeneous forest habitats (deciduous versus evergreen) and in insular
109 (Corsica island) and mainland areas (Mainland France) in order to better understand the

110 determinants of their local adaptation. The analysis was based on Resequencing of birds
111 sampled in four sites studied in the context of a long-term project (Blondel et al. 2006;
112 Charmantier et al. 2016) together with three additional pairs of deciduous and
113 evergreen forests in order to test for parallel evolution (see Figure 1A). First, we
114 investigated variation in genetic diversity and differentiation in order to verify that
115 habitat type and geographic distance explained a significant proportion of the genetic
116 structure between populations (Szulkin et al. 2016). Second, we investigated the
117 historical and contemporary demography of each deciduous and evergreen population
118 pair in order to better understand the origin of their differentiation, and the
119 demography between the populations on mainland France and on the Corsica Island in
120 order to estimate their split time and their subsequent connectivity and effective
121 population size. Third, we screened the genome for F_{ST} outlier loci and inversions
122 potentially implicated in adaptation in the two habitat types and between Corsica and
123 Mainland France. Rather than identifying the genes under selection, we wanted to test
124 whether the adaptation had a polygenic or oligogenic architecture and whether outlier
125 loci would be more frequent in regions with low recombination rate. We discuss our
126 results in the light of the recent literature in population genomics, to decipher the role
127 of genomic and demographic variations on the genetic and phenotypic divergence
128 between blue tits in deciduous *versus* evergreen habitats and in mainland *versus* insular
129 contexts.
130
131 **Methods**

132 **Sites and sampling**

133 947 blue tits (*Cyanistes caeruleus*) were captured in ten locations (Figure 1A, Table 1).
134 Three locations are located in the South of mainland France (codes staring with "A":
135 A1D, A2D, A2E) and seven in the Corsica Island (staring with "B": B3E, B4D, B4E, B5D,
136 B5E, B6D, B6E). The numbers from 1 to 6 indicated the sampling area and hence the
137 population pairs. Half of these sites were dominated by the deciduous downy oak
138 *Quercus pubescens* (indicated by "D" in the code) and the other half are dominated by
139 the evergreen holm oak *Q. ilex* (indicated by "E"). Four of these locations (A1D, B3E,
140 B4D, B4E) are monitored as part of a long-term research program started in 1976 and
141 have been described in previous studies (Blondel et al. 2006; Charmantier et al. 2016).
142 The 6 other sites (A2D, A2E, B5D, B5E, B6D & B6E) were sampled in order to obtain
143 further replicates of neighboring deciduous and evergreen populations. Capture and
144 handling of the birds was conducted under permits provided by the *Centre de*
145 *Recherches sur la Biologie des Populations d'Oiseaux* (CRBPO) and by the *Direction*
146 *Départementale des Services Vétérinaires* (DDSV). Birds were captured during the
147 reproductive period, from early April to late June, on their nesting territory. Birds were
148 captured either in nest boxes during the feeding of their young (in the four sites studied
149 on the long term), or using nets in the vicinity of their nest (in the other sites). Birds
150 were banded with a unique metal ring provided by the CRBPO. Five to 20µl of blood
151 were sampled from a neck or a wing vein from breeding blue tits. Blood was stored at
152 4°C in Queen's buffer (Seutin et al. 1991).

153 **Molecular biology and sequencing**

154 DNA extractions were achieved using Qiagen DNeasy Blood & Tissue kits and were
155 randomized across sites. DNA was quantified using first a NanoDrop ND8000
156 spectrophotometer and then a Qubit 2.0 fluorometer with the DNA HS assay kit (Life
157 Technologies). DNA quality was examined on agarose gels. Library preparation using
158 RAD-seq (restriction-site-associated DNA sequencing; (Baird et al. 2008)) with the
159 enzyme *SbfI* was done by Montpellier GenomiX (MGX) platform (CNRS, Montpellier).
160 Each individual was identified using a unique 6 nucleotides tag, individuals were
161 multiplexed in equimolar proportions by libraries of 36 individuals, and each library was
162 sequenced on a lane of an Illumina HiSeq 2000. Single-end sequencing was used to
163 produce 100bp reads. This design was used to obtain an average read depth of 50x. The
164 DNA of three individuals was extracted twice and processed in different libraries to test
165 for reliability of the genotyping process.

166 **Bioinformatics and data filtering**

167 Raw sequences were inspected with *FastQC* (Andrews 2010) for quality controls.
168 Potential fragments of Illumina adapters were trimmed with *Cutadapt* (Martin 2011),
169 allowing for a 10% mismatch in the adapter sequence. Reads were filtered for overall
170 quality, demultiplexed and trimmed to 85bp using *process_radtags,* from the *Stacks*
171 *software pipeline* 1.39 (Catchen et al. 2013), allowing for one mismatch in the barcode
172 sequence. Sequencing RAD-tags resulted in a median value of 5,449,564 reads per
173 individual. *BWA-MEM* 0.7.13 (Li & Durbin 2009) was used to map individual sequences
174 against the reference genome of the great tit *Parus major* (Laine et al. 2016) and to
175 produce *sam* files using default options. Indeed, although great tit and blue tit diverged

176 about 7 to 14My ago (Päckert et al 2007), the use of a reference genome over a de-novo
177 approach is very often highly recommended (See e.g. Rochette et Catchen 2017). It is
178 also interesting to note that synteny is typically highly conserved in birds (Backström et
179 al 2008), suggesting that aligning reads on a divergent species would still allow analyses
180 based on SNP position along the genome. On average, 93% of the raw reads were
181 mapped against the genome (Supplementary table 1). *Samtools* 0.1.19 (Li et al. 2009)
182 was used to build and sort *bam* files. We used *pstacks* to treat *bam* files, align the reads
183 as assembled loci and call SNPs in each locus. We used a minimum depth of coverage
184 (m) of 5, the SNP model, and alpha = 0.05 (chi square significance level required to call a
185 heterozygote or homozygote). *cstacks* was used to build the catalogue of loci, allowing
186 3 mismatches between sample loci when building the catalog. *sstacks* was used to
187 match loci against the catalog. Lastly, *populations* program in *Stacks* was used to
188 genotype individuals. In this program, relatively permissive filters were applied, to retain
189 SNP genotyped in at least 50% of individuals (all individuals from all sites grouped) and
190 with a minimum read depth of 4. 350,941 SNP and 947 individuals were obtained in this
191 "GeneralDataset".

192 We hence applied additional filters to this GeneralDataset using the programs
193 *VCFtools* (Danecek et al. 2011) and *Plink* (Purcell et al. 2007). We filtered for a minimum
194 average read depth of 10 and a maximum average read depth of 100 (corresponding
195 approximately to the 5%-95% distribution of read depth). SNP genotyped in less than
196 80% of the individuals were removed. We removed SNPs with observed heterozygosity
197 ≥ 0.65 among individuals on Corsica or on mainland to reduce the potential occurrence

198 of stacked paralogues. Individuals below 85% genotyping rate were removed. Identity-
199 By-State (IBS) was estimated for three replicated individuals (from the DNA extraction to
200 SNP-calling) in order to investigate reliability of the entire genotyping process. The IBS
201 measured between replicates were high, ranging from 0.9989 to 0.9999, indicating very
202 low genotyping error rate. These replicates were then removed from the dataset. Using
203 the R packages *gdsfmt* and *SNPRelate* (Zheng et al. 2012), we measured the realized
204 genomic relatedness (GRM) between individuals in order to remove highly related
205 individuals (*i.e.* full-sibs and parent-offspring). For each pair of individuals with
206 relatedness ≥ 0.35, we removed one individual. This procedure was applied in order to
207 limit biases due to highly related individuals. Indeed, since we sampled breeding birds at
208 their nests for the four sites studied on the long-term and that they tend to disperse
209 relatively close to their nests, we expected a higher percentage of related individuals
210 than there actually is in the population. We then limited the number of individuals to
211 60, chosen at random, in populations (A1D, B3E, B4D & B4E) in which a large number of
212 individuals were genotyped. This limitation was achieved in order to limit differences in
213 analyses precision between populations due to unequal sample sizes (all individuals will
214 be used in another ongoing study of the genomic architecture of quantitative adaptive
215 traits). We hence removed potential monomorphic SNPs and created the dataset
216 "FilteredDataset". 454 individuals (Table 1) and 144,773 SNPs were kept in this final
217 dataset used for all the genomic analyses excepted analyses of demographic history
218 using ABC. The median genotyping rate across all SNPs for these individuals was 0.981.
219 The median read depth across genotypes (SNPs x Individual) was 49.2x. The number of

220 SNPs per chromosome ranged from 143 (LGE22) to 15,738 (chromosome 2). Further
221 filtering (e.g. MAF) was often operated, depending on the analyses, and thereafter
222 mentioned if it was the case. To perform analyses of demographic history using ABC
223 analyses, we produced a haplotype VCF file for these 454 aforementioned individuals
224 with the populations module of Stacks and we filtered it as explained in the
225 supplementary note 1.

226 **Analysis of genetic structure and effects of environmental variables**

227 We used PCAs with the function *snpgdsPCA* from *snprelate* to depict genetic structure
228 between individuals. PCAs were run for the entire dataset, for mainland and Corsican
229 populations separately and for each pairs of deciduous and evergreen populations. We
230 inferred admixture proportions for each individual using *Admixture* 1.23 (Alexander et
231 al. 2009) with K-values ranging from 1 to 12 and 1,000 bootstraps. The different
232 clustering solutions were inferred by plotting cross-validations errors and by plotting
233 individual admixture proportion. The effect of environmental variables, forest
234 phenology (E versus D), geographic distance (latitude & longitude) and altitude, on
235 genomic differentiation was measured using a redundancy analysis (RDA) (Legendre &
236 Fortin 2010; Forester et al. 2015) as implemented in the R package *Vegan* (Oksanen et
237 al. 2007). We investigated the proportion of genetic variability explained by a
238 constraining covariance matrix consisting of phenology, latitude, longitude, and altitude
239 for each individual. We tested the global significance of the model using 1,000
240 permutations. We ran marginal effects permutations to address the significance of each
241 variable. Then we focused on the effect of phenology alone, using partial RDA to take

242 into account the effect of latitude, longitude and altitude. Significance was tested
243 running 1,000 permutations. For the PCA, the admixture analysis and the RDA, we
244 selected from the FilteredDataset the SNPs with more than 95% genotyping rate, MAF >
245 0.05, retaining one SNP per locus and we remove SNP in linkage disequilibrium using the
246 *Plink* command "indep 50 5 2". Genome wide differentiation between each sampling
247 location was measured with Weir and Cockerham's F_{ST} estimator (Weir & Cockerham
248 1984) implemented in *StAMPP* (Pembleton et al. 2013). F_{ST} was estimated for all the
249 SNPs, the ones on the autosomes and the ones on the sex chromosome Z separately.
250 Significance was assessed using 1,000 bootstraps replicates.

251 **Analysis of demographic history**

252 Alternative models of divergence history including the effects of selection at linked sites
253 affecting *Ne* and of differential introgression (*m*) were compared using an ABC
254 framework modified from Roux et al. (2016). Linked selection either under the form of
255 hitchhiking of neutral loci linked to a selective sweep (Maynard Haigh & Smith, 1974) or
256 under the form of background selection (Charlesworth et al. 1993), have strong effects
257 in regions of low recombination and have been shown to influence model choice and
258 parameter estimates (Ewing & Jensen, 2015; Schrider et al. 2016). The same is true
259 when populations accumulate reproductive incompatibilities during the divergence
260 process: the resulting barrier to gene flow reduces the effective migration rate along the
261 genome (Barton & Bengtsson, 1986) and not accounting for it can affect demographic
262 model choice and parameter estimates (Roux et al. 2014; Souza et al. 2013). Moreover,
263 including selected loci in demographic inferences can reveal the deeper origins of

264 population divergence (Bierne et al. 2013). Six scenarios were compared for the 4 pairs
265 of deciduous and evergreen populations in order to test whether the divergence
266 between different habitats was not due to a divergence in different historical refugia
267 but to a contemporary ecological divergence. We included a model of panmixia (PAN), a
268 model of equilibrium corresponding to the island model with two populations (EQ), a
269 model of isolation with migration (IM), a model of secondary contact (SC), a model of
270 divergence with migration during the first generations, *i.e.* ancestral migration (AM),
271 and a model of strict isolation (SI). The prior and details of the simulation pipeline are
272 fully described in the supplementary note 1. The PAN model assumes that the two focal
273 populations descent from a single panmictic population characterized by its effective
274 size (Ne). The EQ model (equivalent to the island model) assumes that the population is
275 subdivided into two discrete populations of sizes N_{pop1} and N_{pop2} that are connected by
276 continuous gene flow at a constant rate each generation. In this model, the divergence
277 time is not a parameter. The IM, SI, SC, and AM models all assume that an ancestral
278 population of size N_{ANC} splits at T_{split} into two daughter populations of sizes N_{pop1} and
279 N_{pop2}. Under SI no subsequent gene flow occurs. Under AM model, gene flow occurs
280 from T_{split} to T_{am} and is followed by a period without gene flow. Under IM, gene flow is
281 continuous after T_{split}. Under, the SC model, T_{split} is followed by a period of strict
282 isolation, after which a secondary contact starts T_{sc} generations. The EQ, IM, AM and SC
283 models included migration as $M = 4\ N_0.m$, with $M_{1\leftarrow 2}$ being the number of migrants from
284 population 2 to population 1 and $M_{2\leftarrow 1}$ being the reverse. The effect of linked selection
285 and barriers to gene flow were accounted for, by modeling heterogeneous population

286 size (Ne) and heterogeneous migration (m) respectively. Such heterogeneity was
287 modeled using beta distributions as hyper-prior on each of these two parameters. These
288 resulted in four alternative versions for models with gene flow EQ, AM, IM, SC
289 (NhomoMhomo, NhomoMhetero, NheteroMhomo, NheteroMhetero) and two versions
290 for PAN and SI (Nhomo and Nhetero). We used a modified ABC (Csilléry et al. 2010)
291 pipeline from (Rougemont & Bernatchez 2018) to perform model selection, compute
292 robustness and to estimate posterior probabilities of parameters.

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

304 **Genomic diversity**

305 Genome wide genetic diversity was inferred for each sampling location in the dataset by
306 measuring observed heterozygosity (Ho), proportion of polymorphic loci and MAF
307 spectrums. For each population and chromosome, and subsequently on average for the

308 entire genome, Linkage Disequilibrium (LD) decay was measured with *Plink* and
309 smoothed in R. To contrast with long term Ne estimates from coalescent simulation in
310 our ABC modeling (see "Analysis of demographic history"), we also inferred recent Ne
311 for each population using *SNeP* V1.1 (Barbato et al. 2015), which uses LD data, with a
312 MAF ≥ 0.05 filter per population. We investigated the nature of SNP variation, *i.e.*
313 synonymous or non-synonymous, blasting all rad sequences on the reference genome
314 and the transcriptome of the great tit (ftp://ftp.ncbi.nih.gov/genomes/Parus_major/;
315 Santure et al 2011) using *blastx* (McGinnis & Madden 2004). We kept hits with at least
316 90% similarity and a minimum amino acid sequence length alignment of 25. We kept
317 only SNPs for which both the alternative and reference allele yields the same score.
318 Finally, we tested for differences in the distribution of run of homozygosity (ROH)
319 between the mainland and Corsica that may have resulted from smaller Ne and larger
320 inbreeding in Corsica versus the mainland. We used plink 1.9 to estimate the length and
321 number of ROH. We required a window of 500 kb to be homozyguous in order to be
322 considered as a ROH, and with a maximum of 100 SNP using the following parameters: --
323 homozyg-density 50,--homozyg-gap 100,--homozyg-kb 500,--homozyg-snp 100,--
324 homozyg-window-het 1,--homozyg-window-snp 100,--homozyg-window-threshold 0.05,
325 and--homozyg-window-missing 20.

326 **Identification of genomic footprints of selection**

327 We used three methods to search for outlier SNPs potentially under divergent selection
328 between blue tit populations. First, we used *Bayescan* V2.1 (Foll & Gaggiotti 2008) to
329 search for SNPs potentially under divergent selection at two geographic levels: i)

330 between the mainland and Corsican populations and ii) between each local pairs of
331 evergreen (E) and deciduous (D) populations, *i.e.* A2D vs A2E, B4D vs B4E, B5D vs B5E,
332 B6D vs B6E. We filtered each of the five datasets for a minimum MAF of 0.05. We used
333 default parameters except for prior odds that were set at 10,000 in order to limit false
334 positives. We investigated the parallelism across pairs of D-E environments. Second, we
335 estimated F_{ST} along the genome using either a 200kb sliding average with *VCFtools*, or
336 the function "snpgdsSlidingWindow" from the package SNPrelate to estimate F_{ST} in
337 blocks of 50 SNPs moving by 5 SNPs. The second window strategy was used in order to
338 compensate for lower SNP density in regions of low recombination that tend to
339 exaggerate the contribution of individual SNPs in these regions and to dilute the
340 individual SNP contribution in regions of high recombination (see Perrier & Charmantier
341 2019 for a broader comment on this). Third, we used an RDA as an alternative method
342 to search for SNPs putatively implicated in multilocus adaptation i) between populations
343 in deciduous and evergreen habitats, and ii) between populations in Corsica and on the
344 mainland. As suggested by Forester et al (2018), such a multivariate method may be
345 more suitable than univariate ones to detect weaker footprints of adaptation that are
346 expected in polygenic adaptations in response to complex environmental heterogeneity.
347 Using a similar procedure as described earlier in the methods, we used two RDAs
348 constrained to investigate the effect of phenology (i) or of the geography (ii). We then
349 used a 3 standard deviation cutoff as suggested by Forester et al (2018) to list loci with
350 outlier loading scores on the first RDA axes. We compared the loci found using these

351 different methods. We reported in which genes these outliers were found (the list of
352 genes can be found together with the genome published by Laine et al 2016 on NCBI).

353 **Variation of genomic differentiation with recombination rate**

354 We investigated variation of F_{ST} with local recombination rate and whether SNP outliers
355 were more often found in regions of low recombination than elsewhere in the genome.
356 We estimated local recombination rates using a coalescent method implemented in
357 *Ldhat* (McVean 2004) using linkage disequilibrium signal. Following these authors'
358 recommendation, the dataset was split in blocks of 2,000 SNPs with 500 overlapping
359 SNPs. A MAF of 0.05 and a maximum of 5% of missing data were allowed. The local
360 recombination rate, ρ = 4Ne r, was then estimated in each block independently with the
361 Bayesian reversible jump MCMC scheme implemented in interval. We used a block
362 penalty of 5, with 30 millions MCMC iterations sampled every 5,000 iterations. The first
363 250,000 iterations were discarded. To speed up computations we used the pre-
364 computed two locus likelihood table for n = 190 and assuming theta = 0.001. We
365 estimated ρ in a composite dataset with individuals from every population, habitats and
366 sex. We tested for correlations between SNP F_{ST} and recombination rate using linear
367 models, for Corsica-mainland and for deciduous-evergreen differentiation,
368 independently, and we represented the correlation using a LOESS fit. We tested
369 whether outliers found using Bayescan and the RDA method for both Corsica-mainland
370 and deciduous-evergreen differentiation were more often found in regions of low
371 recombination than elsewhere in the genome using chi-square (χ2) tests, and we
372 represented the pattern using histograms.

373 **Detection of genomic inversions**

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

386 We detected one putative inversion. We verified, using admixture with K = 2 for
387 analyzing SNPs from the inversion, that putative heterozygous individuals had an
388 admixture ratio close to 1:1 of both putative inverted and non-inverted homozygous
389 clusters. We then inspected variations of F_{ST} (per SNP) and π (per 10kb window, using
390 *vcftools*) along the genome between individuals that were inverted homozygous, non-
391 inverted homozygous, and heterozygous for the inversion, looking for potentially
392 reduced diversity and increased differentiation at the inversion. We looked for potential
393 salient variations in read depth in and around the putative inversion. We also tested
394 whether the detected inversion was at Hardy-Weinberg equilibrium (using a common

395 χ2-test) and whether the frequency of the inversion varied geographically and between
396 evergreen and deciduous habitats.

397 To study the history of the putative inversion identified, we aimed at measuring
398 intra- and inter-specific genetic distance and absolute divergence, at the inversion and
399 for the entire genome, for several blue tits and great tits. To do that, we first generated
400 a new SNP dataset by running stacks with the same pipeline and parameters as
401 explained earlier, with 2 mainland great tits, 2 insular great tits, 4 mainland blue tits
402 homozygous for the inversion, 4 mainland blue tits not carrying the inversion, and 4
403 Corsican blue tits not carrying the inversion (the inversion was not found in Corsica).
404 Second, we selected the SNPs from the region of the inversion. Third, Dxy was measured
405 between the 5 aforementioned groups of individuals using *PopGenome* (Pfeifer et al
406 2014). Using this measure of Dxy, we estimated approximately the inversion apparition
407 time using $T = Dxy / 2\mu$, assuming a standard mutation rate of 1e-8 and with the
408 simplifying assumption of no gene flow and no introgression. Lastly, we represented the
409 divergence between these individuals using an UPGMA tree of bitwise distance using
410 the R package *poppr* 1.1.1 (Kamvar et al 2014).

411 **Gene ontology**

412 We used the R package topGO (Alexa & Rahnenführer, 2009) to investigate the potential
413 gene ontologies (GO) that were statistically enriched for the sets of genes identified
414 among outliers and the inversion, compared to the entire list of genes in which all the
415 SNPs from the entire dataset were found. For the outlier tests, we used GO analyses for
416 each gene list obtained using the different outlier identification methods but also for

417 aggregated lists of all of the gene identified for deciduous-evergreen tests or for Corsica-
418 Mainland tests. We used the GO referenced for the zebra finch, T. Guttata
419 (tguttata_gene_ensembl). We report a Fisher enrichment test p-value and a p-value
420 after applying a Benjamini and Hochberg (BH) correction for multiple testing to control
421 for false discovery rate (FDR).
422

423 **Results**

424 **Genetic structure and effects of environmental variables**

425 The admixture analysis suggested the existence of 2 main distinct genetic groups
426 corresponding to mainland and island populations (Figure 2A). Increasing K value
427 contributed to delineate the populations, showing the existence of a weak structure
428 between the 6 main areas. However, the coefficient of variation increased with K. There
429 was very little evidence for genetic structure within each of the four pairs of D- and E-
430 populations. PCAs revealed a clear structure between individuals from the mainland and
431 the island but little structure within each of these 2 groups (Supplementary Figure 1).
432 The full RDA model was significant ($p < 0.001$, Supplementary table 2, Figure 2B), as well
433 as the effects of each variable tested in the model: phenology (*i.e.* E versus D, $p = 0.006$),
434 latitude ($p < 0.001$), longitude ($p < 0.001$) and altitude ($p < 0.001$). The first axis of the
435 global RDA explained 4.54% of the variance, and was correlated mainly with latitude (-
436 0.987) and longitude (+0.999). The fourth axis of this global RDA displayed the strongest
437 correlation with phenology (-0.856), and explained 0.27% of the variance. The partial
438 RDA model conditioning for the effects of latitude, longitude and altitude, was globally

439 significant ($p < 0.001$), as well as the effect of phenology ($p < 0.001$). The axis of the
440 partial RDA explained 0.29% of the variance and was correlated with phenology (0.922).

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

450 **Demographic history**

451 When deciphering the historical demography of deciduous and evergreen population
452 pairs, the hierarchical model choice procedure strongly rejected models of strict
453 isolation (SI), ancient migration and of panmixia (PAN), which were associated with a
454 posterior probability of 0 (Supplementary table 3). Instead, models with gene flow were
455 highly supported. In three of the four pairs, the equilibrium model (EQ) received the
456 highest posterior probability with P(EQ) = 0.99 and 0.95, 0,94 in the B6D vs B6E, B5D vs
457 B5E and B4D vs B4E comparisons, respectively. In the A2D vs A2E comparisons, the best
458 supported model was the isolation with migration model, P(IM) = 0,74 with the second
459 best model being SC, with P(SC) = 0.23. Comparing the two models against each other
460 while excluding all remaining models yield unambiguous support for IM, with P(IM)

461 =0.99 (Supplementary Table S3). Across all models, comparisons with heterogeneous
462 gene flow and heterogeneous effective population size were not supported, indicating
463 that, if genetic barriers or linked selection were at play they could not be detected.
464 Demographic parameters were estimated for each pair of populations under the best
465 model (Supplementary table 4). Posterior distributions were well differentiated from
466 their prior indicating that estimated parameters were confidently estimated
467 (Supplementary Figure 2). Effective population size (N_e) where slightly higher in
468 deciduous than in evergreen habitats (*Ne*1 in Table 2, Figure 3A, Supplementary table 4)
469 and tended to be higher on the mainland than on the island. In a number of
470 comparisons, our migration estimates reached the prior upper bounds. Therefore, we
471 ran a new set of simulations with wider priors (Supplementary Note 1). Migration rates
472 were not different between deciduous and evergreen habitats: in two instances, the
473 rate of migration was higher from the deciduous to the evergreen and in two other
474 instances, the reverse was true (Supplementary Figure 3).

475 Deciphering the historical demography between Corsica and the mainland, the
476 ancient migration models (AM) strongly outperformed the strict isolation models (SI),
477 with p(AM) = 0.99 (Supplementary table 5). For the AM model, the simplest model
478 without linked selection or barriers to gene flow was the best supported, with p = 0.76.
479 Parameters estimates revealed posterior generally well differentiated from the prior
480 providing increased confidence (Supplementary Figure 4). Effective population size of
481 the mainland was 397,685 [CI = 244,375 -- 600,105], around 7.5 times greater than the
482 one of the island, 52,990 [CI = 33,065 -- 78,475] (Figure 3B, Ne1 in Table 2,

483 Supplementary table 6). Our analysis also indicated a strong population size change
484 during the process of population isolation, since the ancestral Ne was estimated at only
485 240 [CI = 65 -- 685]. Split time was estimated at 1,417,480 [CI = 740,000 -- 2,394,000]
486 generations ago, hence around 3.2M years ago (assuming a 2.3 years generation time
487 (Charmantier et al. 2004)) and gene flow subsequently stopped 4540 [CI = 2500 -- 6360]
488 generations ago, hence around 10,000 years ago, corresponding closely with the end of
489 the last glacial maximum. The ancestral migration rate between Corsica and the
490 mainland, from the split in two populations to the end of the latest gene flow between
491 these populations, was inferred asymmetric, with a 5 times larger migration rate from
492 the island to the mainland (*m ~ 1.9e-4)* than the other way around *(m ~ 3.8e-5*).
493 However, this asymmetry did not generate a strong difference in gene flow between the
494 two groups, given the lower effective population size on the island compared to the
495 mainland, (i.e. respective number of migrants of 10 and 15 from Corsica to the mainland
496 and in the reverse direction, with overlapping credible intervals, see details in
497 Supplementary table 6).

498 **Genetic diversity**

499 The mainland populations displayed significantly higher patterns of observed
500 heterozygosity (on average 0.169, table 2) than Island populations (on average 0.147),
501 all t-tests' p-values < 0.001. On the contrary there was no significant difference of
502 observed heterozygosity at smaller spatial scale among populations within the island
503 (either same or different tree phenology, E- and D-) or within the mainland (all t-tests' p-
504 value > 0.05). More SNPs were polymorphic in mainland populations (77,306 to 80,616

505 SNPs for a MAF > 0.05) than in island populations (62,728 to 65,876 SNPs for a MAF >
506 0.05). The MAF spectrum showed enrichment of variants with smaller frequencies in
507 mainland populations versus island populations (Supplementary Figure 5A). LD decayed
508 rapidly in the first 5kb and was lower in populations from the mainland (especially A1D)
509 compared to island populations (Supplementary Figure 5B). This pattern of rapid LD
510 decay was similar between chromosomes (see for example chromosomes 1, 2 and Z,
511 Supplementary Figure 6). Contemporary N_e inferred from LD varied from 142 (in A2E) to
512 355 (in A1D) (Ne2 in Table 2, Supplementary Figure 5C). N_e values were rather similar
513 within pairs of E- and D- populations, although Ne was in average smaller in evergreen
514 populations (245 in average) than in deciduous ones (261 in average). The four largest
515 N_e (with an average of 318) were found for the four populations monitored on the long
516 term in forest in which hundreds of artificial nest boxes have been installed for several
517 decades (the N_e in the 6 other populations was in average of 210). In the entire dataset,
518 2599 SNPs were identified as non-synonymous while 6751 SNPs were identified as
519 synonymous. We observed 1.18 times lower allele frequencies for non-synonymous
520 variants (average MAF = 0.09) than for synonymous ones (average MAF = 0.11) (t.test p-
521 value < 0.000001). This overall lower frequency of non-synonymous compared to
522 synonymous mutations was similar between populations from Corsica (1.19) and the
523 mainland (1.16), as well as between deciduous (1.18) and evergreen (1.19) habitats (See
524 supplementary table 7). The Z chromosome harbored a proportion of non-synonymous
525 mutations 1.83 times higher than the average for the autosomes. Finally, we observed
526 significantly slightly longer ROH on Corsica compared to mainland (mean $_{island}$ = 44,769

527 kb *vs* mean $_{mainland}$ = 41,586 kb, Wilcoxon rank sum test W = 32,695, p <2e-16,
528 Supplementary Figure 7&8). We also observed significant differences in the count of
529 ROH among populations from the mainland versus those from the island but not among
530 populations within the mainland or the island (ANOVA p <2e-16; Supplementary Table 8
531 for Tukey HSD test).

532 **Footprints of selection**

533 Manhattan plots of F_{ST} per SNP revealed no clear evidence for high peaks of F_{ST} for the
534 five comparisons considered (First panels of Figures 4 & 5). Throughout the five
535 Bayescan tests (Corsica vs mainland populations and the 4 D-4 population pairs), we
536 identified 40 SNPs with a log10(BF) > 0 among which 18 SNPs had a log10(BF) > 1
537 (Supplementary table 9, Supplementary Figure 9). None of these SNPs were detected
538 twice among the different tests conducted with Bayescan. Among these 40 SNPs, 27
539 were found for the Corsica vs mainland test, 2 were found for A2D vs A2E, 9 were found
540 for B4D vs B4E, 0 were found for B5 vs B5E and 2 were found for B6D vs B6E. Among
541 these 40 SNPs, 19 genes were identified (Supplementary table 9). Sliding windows of F_{ST}
542 along the genome showed a few outlier windows of modest intensities (last panels of
543 Figure 4 & 5), with almost no parallelism between the tests (Supplementary Table 10).
544 Expectedly, outlier windows found for the Corsica *vs* mainland comparison depicted
545 larger F_{ST} than for deciduous vs evergreen comparisons. The outlier F_{ST} windows falling
546 in the top 1% of the F_{ST} windows distribution are reported in Supplementary Table 10.
547 Using the RDA to identify outlier SNPs with extreme loading values to axes, we identified
548 229 outlier SNPs associated to habitat type (deciduous vs evergreen), and 227 SNPs

549 associated to geography (longitude & latitude, essentially representing a Corsica-
550 Mainland comparison) (Supplementary table 11, Supplementary Figure 10). For the
551 deciduous vs evergreen test, only one of the RDA outliers was also found outlier in
552 Bayescan tests. For the Corsica-mainland test, 8 of the RDA outliers were also found
553 outliers in the Bayescan test. Genes in which the outliers were found are reported in the
554 supplementary table 13.

555 **Variation of genomic differentiation with recombination rate**

556 F_{ST} between mainland and Corsican populations was negatively correlated to
557 recombination rate (Figure 6A, linear model p < 2e-16; Supplementary Figure 11). In
558 contrast, F_{ST} between each pair of deciduous and evergreen populations was not
559 correlated to recombination rate. F_{ST} outlier SNPs between Corsica and the mainland
560 populations and identified by Bayescan or by the RDA were more often found in regions
561 of low recombination ($\chi 2$ test p-values < 0.01, Figure 6C) than observed for the entire
562 SNPs (Figure 6B). Average recombination rate was on average twice lower for these F_{ST}
563 outlier SNPs compared to the rest of the SNPs (t-test p-values < 1e-6). In contrast, F_{ST}
564 outlier SNPs between deciduous and evergreen populations were not preferentially
565 found in regions of low recombination (Figure 6D).

566 **Genomic inversions**

567 We detected one putative inversion on chromosome 3, spanning 2.8Mb, from position
568 11,838,789 to 14,661,550, and containing 390 SNPs. The estimated recombination rate
569 was on average 9 times lower (t-test: p<2.2e-16) in this region of the chromosome
570 compared to the rest of the chromosome (Figure 7A). However, the SNP density did not

571 appear reduced at the location of this putative inversion (Figure 7A), suggesting a recent
572 drop of recombination (*i.e.* dissimilar to what is expected in a peri-centromeric
573 location). We did not find evidence for an increase of read depth at specific loci in the
574 putative inversion or its vicinity in inverted individuals compared to normal ones. We
575 did however notice the absence (read depth = 0) of 3 loci from the inversion, in the four
576 inverted individuals. The sliding PCA detected either 2 groups (individuals from Corsica
577 and from mainland) outside of the putative inversion, or 4 groups at the putative
578 inversion location, separating individuals from Corsica and 3 groups for mainland
579 individuals (Figure 7B). We assumed that the inverted segment was the one with the
580 lower frequency on the mainland and that was absent from Corsica. The Lostruct
581 analysis confirmed for this same region the presence of four successive blocks of 100
582 SNPs highly discriminating individuals (Figure 7C). The average F_{ST} between the three
583 groups of mainland individuals was much higher (t-test: p<2.2e-16, Figure 7D) for SNPs
584 in the putative inversion (F_{ST} = 0.26) than for SNPs outside of this region (F_{ST} = 5.8e-4). π
585 was lower (t-test: p=1.9e-5, Figure 7E) for SNPs located in the putative inversion for
586 inverted homozygous individuals (π = 2.5e-5) compared to non-inverted homozygous
587 individuals (π = 5.3e-5). Both inverted and non-inverted homozygous individuals had
588 lower π at this inversion compared to heterozygous individuals at this region (π = 7.6e-
589 5) and compared to the three types of individuals for the rest of the chromosome 3 (π
590 ranging from 7.4e-5 to 7.6e-5). The PCA (Figure 7F) and the admixture analysis (Figure
591 7G) using the SNPs from the inversion clearly delineated inverted, non-inverted and
592 heterozygous individuals, with heterozygous individuals falling at intermediate

593 eigenvalue between the two categories of homozygous individuals. LD was higher
594 (Figure 7H) for the region of the putative inversion (Little LD was found when only
595 homozygous non-inverted individuals were kept in the analysis). For the region of the
596 inversion, Dxy was 0.0020 between normal blue tits from the mainland and from Corsica
597 as well as between great tits from the mainland and from Corsica. Dxy was 0.0044
598 between inverted and normal blue tits from either the mainland or from Corsica, and
599 ranged from 0.0149 to 0.0152 for interspecific comparisons. Using the formula T = Dxy /
600 2μ, the inversion likely appeared approximately 220,000 generations ago (around
601 506,000 years ago). The tree of genetic distance illustrated the divergence of inverted
602 blue tits from non-inverted blue tits from both mainland and Corsica (Figure 7i). This
603 region contains 21 genes listed in supplementary table 12. Out of 113 individuals from
604 the mainland, we found 4 inverted homozygous, 23 heterozygous, and 86 non-inverted
605 homozygous, hence we observed no deviation from the Hardy-Weinberg equilibrium ($\chi 2$
606 test p-value > 0.1). The percentage of each genotype was similar across all mainland
607 locations and both habitat types (the inverted segment was at 11% in A1D, 16% in A2D,
608 17% in A2E, and 14% across these 3 locations).

609 **Gene ontology**

610 None of the gene lists gathered with the different tests (each outlier test among
611 deciduous and evergreen environment and between Corsica and the mainland and the
612 inversion test) yielded any significantly enriched GO term after correction multiple
613 testing (Supplementary table 14). The most promising GO (uncorrected p.value <0.0025)
614 included: i) for the deciduous-evergreen outlier tests, GO:0060385, axonogenesis

615 involved in innervation; GO:2001013, epithelial cell proliferation; GO:0007194, negative
616 regulation of adenylate cyclase; GO:0090647, modulation of age-related behavioral;
617 GO:1901379, regulation of potassium ion transmembrane transport, ii) for the Corsica-
618 Mainland outlier tests, GO:0008344, adult locomotory behavior; GO:2001224, positive
619 regulation of neuron migration, iii) for the inversion, GO:0016446, somatic
620 hypermutation of immunoglobulin; GO:0045910, negative regulation of DNA
621 recombination; GO:0006298, mismatch repair; GO:0008340, determination of adult
622 lifespan.
623

624 **Discussion**

625 In this study, we investigated demographic history and genome wide patterns of genetic
626 diversity and differentiation between several populations of blue tit presenting strong
627 phenotypic differences between deciduous *versus* evergreen forest habitats as well as
628 insular *versus* mainland contexts. Between populations in deciduous and evergreen
629 forest habitats, demographic modelling showed large gene flow and large effective
630 population sizes (Ne), explaining the low genetic differentiation between these
631 populations. Demographic modeling also revealed that each pair of deciduous and
632 evergreen populations most likely never diverged and maintained high connectivity
633 through migration. We identified slight and mostly unrepeated footprints of divergent
634 selection across these evergreen and deciduous population pairs, which is consistent
635 with their demography and the likely polygenic nature of several traits implicated in
636 their local adaptation. In both insular and mainland populations, we found large Ne,

637 although smaller for insular populations than mainland ones, resulting in different
638 distributions and lengths of ROH. Gene flow between Corsica and the mainland most
639 likely stopped at the end of the last glaciation. Both large Ne and gene flow until the end
640 of the last glaciation hence explained moderate genome-wide differentiation. We
641 identified several genomic footprints of selection, enriched in regions of low
642 recombination only in the case of the mainland/island divergence. Finally, we identified
643 a putative genomic inversion spanning 2.8Mb, polymorphic in mainland populations
644 only. We discuss these results in the context of the vast scientific knowledge acquired
645 on these blue tit populations during the last four decades and more generally in the
646 context of genomics of adaptation with gene flow.

647 **Divergence between populations in deciduous *versus* evergreen habitats**

648 Although we found a significant effect of habitat on genetic structure, the genetic
649 differentiation between neighbouring deciduous and evergreen populations was low
650 (F_{ST} ranging from 0.0006 to 0.0079). This result is in line with the primary observations
651 realized earlier on a smaller set of populations (Porlier et al. 2012b; Szulkin et al. 2016).
652 Accordingly, we found high gene flow from deciduous to evergreen populations. Yet,
653 this quantification of high gene flow and low genetic structure contrasted with the
654 demographic knowledge collected on blue tit. Indeed, demographic studies suggested
655 restricted dispersal between these populations, with 4 dispersal events observed
656 between B4D and B4E (5.6 km apart) and none between the B3E and either B4D or B4E
657 (24.1 km), among a total of 2788 males, 2672 females and 25,158 nestling ringed in the
658 three main Corsican sites between 1976 and 2018, with a mean recruitment rate of 6%

659 (Charmantier, com pers). Our interpretation of this contrast between gene flow
660 estimations gathered from population genomic *versus* recapture data is quadruple. First,
661 dispersal estimation on the field using capture-mark-recapture is very challenging and
662 may require more data than currently collected, to detect rare dispersal events, even
663 though these affect population genetic parameters. Moreover, since natal dispersal in
664 blue tit classically ranges between 330 m and 4 km (see (Tufto et al. 2005; Ortego et al.
665 2011)), the long-term monitoring sites in Corsica equipped with nest-boxes (black circled
666 dots in Figure 1C) are not ideally spaced to identify the origin of immigrants and the
667 destination of emigrants, and only a small fraction of the landscape favourable to blue
668 tit breeding is covered. Second, only a few migrants are sufficient to decrease the
669 genetic distance between populations, measured through the F_{ST} (Marko & Hart 2011;
670 Cayuela et al. 2018). In that regard, our results may be compatible with the few
671 dispersal events recorded throughout the years. Third, it is important to note that the
672 number of migrants estimated using demographic analyses represents an average over
673 historical time scales on the order of Ne generations and may have varied widely during
674 contemporary times. Fourth, the large effective population size found both using
675 coalescence and a LD method might be explained by the existence of large
676 "metapopulations" connected by high gene flow and such large Ne might largely
677 contribute to limit the genetic divergence between populations.

678 Ne was on average slightly larger in deciduous than in evergreen populations.
679 This could be explained by the higher productivity of deciduous forests compared to
680 evergreen forests, resulting in larger clutches and more fledglings in deciduous habitats

681 (Table 1 in Charmantier et al 2016). However, the very high gene flow between
682 deciduous and evergreen populations and the low genetic differentiation between these
683 populations limits further interpretations. In addition, populations monitored on the
684 long term, for which hundreds of artificial nest boxes have been installed, tended to
685 have larger Ne than non-monitored populations. Breeding density in the nest-box areas
686 studied (Figure 1A) was around 1 to 1.3 pairs per ha (Blondel et al 2006), which is most
687 probably 3 to 5 times higher than natural densities for blue tit when these secondary
688 hole-nesting birds rely on natural cavities only. It is hence possible that the recent
689 availability of nest boxes locally increased productivity, effective population size and
690 heterozygosity.

691 Demographic modeling indicated that secondary gene flow between previously
692 isolated populations, which could have eroded most genome-wide differentiation
693 outside regions implicated in local adaptation and/or acting as reproductive barriers in
694 particular in low recombining areas (Bierne et al. 2013), was very unlikely. Secondary
695 contacts were strongly rejected in favor of an island model supporting ongoing gene-
696 flow without divergence (EQ model) in 3 out of 4 population pairs, the last one
697 supporting an isolation with migration model (IM). As is often the case, our
698 demographic estimates should be interpreted cautiously, particularly in the absence of a
699 well-documented mutation rate in blue tit. The relatively narrow credible intervals
700 around parameter estimates nevertheless provide confidence in their biological
701 relevance. These results suggest that the focal population pairs occupying
702 deciduous/evergreen habitats have been continuously connected by relatively high gene

703 flow. Hence, genetic divergence between populations in each pair did not originate or
704 build up on pre-existing genetic divergence accumulated during an allopatric phase but
705 rather is supported as a case of genuine recent ecological divergence (Pinho & Hey
706 2010; Wang & Bradburd 2014).

707 We did not find strong footprints of divergent selection between any of the four
708 pairs of deciduous and evergreen populations (Figure 4). Moreover, the observed
709 outliers were not found repeatedly across the 4 deciduous-evergreen comparisons but
710 rather each outlier was found only once, and they were not enriched in regions of low
711 recombination. First, this result is in line with the high migration rates observed. High
712 gene flow indeed most likely limits the potential for local adaptation since it limits the
713 accumulation of allelic differentiation, even with high selection coefficients (Lenormand
714 2002). Second, this pattern is consistent with the best demographic model being the
715 equilibrium model and not a secondary contact. Indeed, the later would more often
716 create repeated and strong outliers located in regions implicated in reproductive
717 isolation between divergent populations (eg (Rougemont et al. 2017)). Third, this
718 pattern is congruent with a model of local polygenic adaptation involving a transient
719 genetic architecture with multiple alleles of small effects underlying (multiple)
720 quantitative characters (Yeaman & Whitlock 2011; Yeaman 2015) and displaying low F_{ST}
721 among loci under divergent selection. Fourth, most documented traits that are involved
722 in the adaptation of blue tit to the deciduous *versus* evergreen habitat types, such as
723 clutch size or laying date (Lambrechts et al. 1997; Blondel et al. 1998) are quantitative
724 and most likely rely on a polygenic architecture with alleles of small effects (see e.g.

725 (Santure et al. 2013)). However, some statistical issues may limit the interpretation of
726 this result. First, currently available genome-scan methods to detect alleles of small
727 effect are statistically limited, especially when applied to relatively small datasets
728 (Rockman 2012; Hoban *et al.* 2016). Second, a potential lack of power in detecting
729 outliers could arise from the combination of the use of RADsequencing and a rapid LD
730 decay along the genome, that may result in too low resolution especially in regions with
731 larger recombination rates (Lowry et al. 2017). Further analyses of the potential outliers
732 found here, as well as a new analysis with higher marker density and more individuals,
733 are needed to better document and discuss the potential genes and biological processes
734 implicated. Particularly, quantitative genomics will be useful to establish links between
735 the level of divergent selection on these putative targets and their effect size on
736 phenotypic trait variation (Stinchcombe & Hoekstra 2007; Gagnaire & Gaggiotti 2016).
737 This would notably reveal whether genes under divergent selection are also responsible
738 for the observed phenotypic variation, and contribute to assessing of how much of
739 phenotypic variation is adaptive.

740 Last, we searched for inversions potentially associated with phenotypic variation
741 and/or segregating in both habitats. We report multiple evidences (Figure 6) for a
742 putative inversion on chromosome 3 in mainland populations. However, the proportion
743 of inversions in the genomes of deciduous- *versus* evergreen-breeding birds did not
744 differ, which most probably precludes a putative role of the inversion in local adaptation
745 in these habitats. Besides, this polymorphism followed the Hardy-Weinberg equilibrium
746 and we found 4 homozygous inverted individuals, suggesting this inversion does not

747 involve an accumulation of lethal recessive deleterious alleles (Jay et al. 2019). The
748 functional consequences of this inversion on variation in life history traits in blue tit call
749 for further investigations, as achieved in other songbirds (Tuttle et al. 2016; Knief et al.
750 2017; Kim et al. 2017). In order to genotype this inversion in more birds and to link it to
751 putative phenotypic variation, we could use a PCR-RFLP approach (da Silva et al. 2019).

752 **Divergence between blue tit populations in mainland *versus* insular contexts**

753 Our demographic modelling approach of Corsican and mainland populations revealed
754 that a model with ancient migration was the most probable, which is coherent with the
755 history previously reconstructed for the blue tit complex. Split time of the ancestral
756 populations in two mainland *versus* island populations connected by gene flow (~3M
757 years) was relatively coherent with the diversification time found in the literature for
758 the entire blue tit complex (5M (Illera et al. 2011)). Gene flow between the island and
759 the mainland likely stopped around 10,000 years, which is compatible with a gene flow
760 break due to the rise of the flooded stretch between Corsica and the mainland during
761 ice melting and the sea level rise after the last glaciation, during from 17,000 to 5,000
762 years ago (Lambeck & Bard 2000; Jouet et al. 2006). Both the signal of population
763 expansion during the diversification period, and the gene flow five times larger from the
764 island to the mainland than from the mainland to the island, were coherent with a
765 suspected recolonisation of the mainland from the islands, as proposed by Illera et al.
766 (2011) based on analyses of nuclear and mitochondrial DNA sequences. The larger
767 effective population size on the mainland than on the island may be due to multiple
768 sources of colonisation on the mainland (Taberlet et al. 1992) and to the much larger

769 mainland area and hence metapopulation size compared to the island. Expanding this
770 study with a sampling from the South-East of Europe, one from Sardinia and one from
771 the Iberian Peninsula, could help determine whether the large size estimated for the
772 mainland meta-population studied here might be explained by a mixed recolonisation
773 from these refugium (Kvist et al. 1999; 2004) and whether insular populations
774 consistently have reduced effective population size. It would also be interesting to
775 explore a model allowing for multiple cycles of isolations and contacts that could have
776 resulted from successive glacial cycles (Hewitt 2004).

777 Given the demographic parameters inferred, it is not surprising that the genetic
778 differentiation between populations from the mainland and the island was moderate
779 (F_{ST} = 0.08). However, this differentiation value contrasted with the much larger
780 divergence at mitochondrial DNA observed for these populations (OST = 0.67 between
781 A1D and B4D (Kvist et al. 2004)). This discrepancy between whole nuclear genome and
782 mitochondrial estimates of differentiation could first be explained by the typically four
783 times smaller Ne in mitochondrial DNA compared to nuclear DNA in diploid organisms,
784 resulting in larger drift and divergence (Smith & Klicka 2013). Second, accumulation of
785 cytonuclear incompatibilities could limit its introgression (Burton & Barreto 2012).
786 Alternatively, gene flow may have been greater for males than females during the
787 colonisation process from Corsica to the mainland, at the end of the last glacial period.
788 However, this hypothesis is not supported by the general observation that females
789 disperse on average further away than males and that rare long distance dispersals are
790 also achieved by females (Tufto et al. 2005). Inspecting mitochondrial DNA variation for

791 the birds studied here would be very useful to compare mitochondrial differentiation in
792 our sample to the previous analysis which also included Corsican birds (Kvist et al. 2004).
793 In any case, one should bear in mind that our results stem from integrating coalescent
794 patterns observed across thousands of loci, therefore providing increased resolution to
795 investigate the determinants of demographic divergence compared to approaches
796 based on fewer mitochondrial and nuclear loci.

797 The contemporary demography inferred between the populations from Corsica
798 and mainland France, especially the absence of gene flow for approximately 4500
799 generations and the large effective population size, likely provides a suitable context for
800 the build-up of local adaptation. Moreover, the level of divergence (F_{ST} = 0.08) was
801 probably low enough to detect divergent outlier loci potentially implicated in local
802 adaptation. We identified several outlier SNPs and outlier 200 kb windows showing
803 elevated differentiation, more often in regions of lower recombination than elsewhere
804 in the genome (contrary to the outliers found for the deciduous vs evergreen
805 populations). Furthermore, F_{ST} was on average higher in these regions of low
806 recombination compared to elsewhere in the genome (which was not the case for
807 deciduous vs evergreen populations). These patterns of increased divergence in regions
808 with low recombination between divergent populations are commonly observed in
809 other species (Nachman & Payseur 2011; Cruickshank & Hahn 2014; Gagnaire et al.
810 2018), including other bird species (Burri et al. 2015; Spurgin et al. 2019). It is well
811 documented that such increased differentiation in regions with low recombination is
812 not necessarily due to positive selection, or at least not alone, and that it is largely

813 influenced by the effect of recombination in interaction with background selection
814 (Charlesworth et al. 1993; Burri et al. 2015; Perrier & Charmantier 2019). Moreover, a
815 uniform F_{ST} sliding window, sized in Kb, is expected to dilute signatures of selection in
816 regions of the genome where LD is fast decaying while it is expected to over-represent
817 regions where LD decays much more slowly, hence increasing even more the potential
818 false positives in regions with low recombination (Beissinger *et al.* 2015; Perrier &
819 Charmantier 2019). Although perfectible, our sliding window approach using a SNP unit
820 attempted to fix this issue and successfully captured several new outlier regions outside
821 of deserts of recombination. This second window approach should however be
822 improved, for example by estimating the local neutral enveloped, which computation
823 should integrate local variations in LD, recombination, diversity, but also the
824 demography of populations. As mentioned for the study of adaptive divergence
825 between deciduous and evergreen populations, complementary analyses integrating
826 both genome scans and quantitative genomics would improve our comprehension of
827 the genomic and phenotypic divergence observed between Corsican and mainland blue
828 tits (Stinchcombe & Hoekstra 2007; Gagnaire & Gaggiotti 2016).

829 While the putative genomic inversion on chromosome 3 was absent from Corsica
830 and detected in mainland individuals, its level of divergence from the non-inverted
831 sequence indicated that it was likely twice older than the beginning of divergence
832 between blue tit populations from mainland and Corsica. This could first suggest that
833 this polymorphism emerged in mainland blue tit populations and then did not introgress
834 the Corsican populations, maybe due to a local disadvantage and/or genetic

835 incompatibilities or simply due to drift coupled to little gene flow. A second hypothesis
836 could be that this inversion was present in Corsican populations but had been purged
837 out due to a local disadvantage. Such a disadvantage can be due to the typical
838 accumulation of deleterious mutations in such non-recombining inverted sequences
839 (*e.g.* Jay et al 2019). Lastly, it is possible that this inversion has been acquired in
840 mainland populations only recently, after the last period of contact between mainland
841 and Corsican populations, *via* gene flow from a distinct refugium with which
842 connectivity could have been enhanced after deglaciation, hence after the stop of gene
843 flow with Corsica. The age of this inversion, its origin, its biological effects and the
844 potential accumulation of deleterious mutations need to be inferred more thoroughly
845 *via* genotyping full inversion sequences from individuals in diverse locations and using
846 advanced statistical methods (Lohse *et al.* 2015).

847 **Conclusion and perspectives**

848 Our study demonstrated the usefulness of demographic modeling and of the analysis of
849 the variation of genomic diversity and recombination along the genome to uncover the
850 genetic determinants of local adaptation in a small passerine with a large distribution,
851 and occupying different forest habitats. Especially, demographic modeling rejected the
852 hypothesis of a secondary contact between deciduous and evergreen populations and
853 favoured a situation with continuous gene flow. These results support the idea that blue
854 tits have adapted to their habitats despite ongoing gene flow, while contextualising how
855 large gene flow most probably constrained local adaptation (Lenormand 2002) and
856 favoured its architecture based on alleles of small effects (Yeaman 2015). The genomic

857 modelling also refined our knowledge about the divergence between insular and
858 mainland meta-populations, that have been likely unconnected by gene flow for the last
859 ten thousand years. We also verified the relationship between local recombination rate
860 and differentiation, that should probably be integrated in genome scans looking for
861 footprints of selection (Beissinger *et al.* 2015; Burri *et al.* 2015; Berner & Roesti 2017;
862 Perrier & Charmantier 2019, Booker et al 2020). Future investigations will require
863 increased sample sizes and marker density (Lotterhos & Whitlock 2015; Hoban *et al.*
864 2016) in order to better detect loci with small effects that contribute to the quantitative
865 phenotypic variation and local adaptation in blue tit. Lastly, the putative inversion found
866 here would need further analyses since this type of structural variation is often
867 implicated in phenotypic variation (Kirkpatrick 2006; 2010; Stapley *et al.* 2017;
868 Wellenreuther *et al.* 2019).

869 **Tables**

870 **Table 1. Study populations' names, abbreviations, habitat types, locations, approximate forest area (ha), altitude (m), and sample**

871 **sizes.**

Name	Abbreviation	Region	Site	Type of oaks	Latitude	Longitude	Forest area	Altitude	N-GeneralDataset	N-FilteredDataset
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	1027	44	44
E-OlmiCappella	B6E	Corsica	OlmiCappella	Evergreen	42.529414	9.021652	25	881	43	43
Total/average									947	454

872

873 **Table 2. Diversity and differentiation across populations.** Fixation index between populations (F_{ST}), effective population size (Ne)

874 estimated with ABC (Ne1) using models for Deciduous-Evergreen population pairs or for Corsica-Mainland (A1D-B4D), Ne estimated

875 with a LD method (Ne2), and observed heterozygosity (Ho). Color gradients indicate differentiation intensity (Darker red for higher

876 differentiation and darker blue for lowest differentiation). Underlined values indicate deciduous and evergreen pairs.

877

	F_{ST}										Ne1		Ne2	Ho
	A1D	A2D	A2E	B3E	B4D	B4E	B5D	B5E	B6D	B6E	Deciduous-Evergreen	Corsica-Mainland		
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

878 **Figures captions**

879 **Figure 1. Map of the sampling locations of blue tit populations on the mainland and in**
880 **Corsica (A) and phenotypic differences (B).** In red, habitats dominated by deciduous
881 oaks and in green, habitats dominated by evergreen oaks. Sites with long-term
882 monitoring are circled in black on the map (A1D is D-Rouvière, B3E is D-Pirio, D4D is D-
883 Muro and B4E is E-Muro). On figures B & C are presented illustrative phenotypic
884 differences (with standard deviations) between the four main populations with long
885 term monitoring. Are shown laying date (days), clutch size and male tarsus length (mm)
886 (data from Table 1 in Charmantier et al 2016 Evol Appl, see this reference for detailed
887 information for these measures and for differences in many other traits).

888

889 **Figure 2. Population structure between populations. A)** Admixture for K = 2, 6 and 10
890 (Vertical lines correspond to individual admixture; colors correspond to distinct genetic
891 groups). **B)** RDA showing the influence of latitude, longitude, altitude and phenology on
892 genetic structure.

893

894 **Figure 3. Demographic parameters A)** for each of the four pairs of blue tit populations
895 in deciduous and evergreen habitats, estimated using equilibrium models, and **B)** for the
896 divergence between populations on the mainland and Corsica, estimated using an
897 ancient migration model. In panel A, circle size is proportional to effective population
898 size (Ne) and arrow width is proportional to migration rate. In panel B, rectangle width is
899 proportional to log10 (effective population size), arrow width is proportional to

900 migration rate, and split time and time of ancient migration are indicated in number of
901 generations.
902
903 **Figure 4. F_{ST} Manhattan plots between the 4 population pairs in deciduous and**
904 **evergreen forests.** F_{ST} are given either by SNP (four first graphs), by 200kb windows
905 (graphs 5 & 6) or by 50 SNPs windows (graphs 7 & 8). For graphical simplicity, the 3 pairs
906 of populations in Corsica were grouped for the sliding window F_{ST} Manhattan plots (BD
907 *vs* BE). Dots alternate colors between chromosomes. F_{ST} for unplaced loci are shown at
908 the end of each of the first four Manhattan plots.
909
910 **Figure 5. F_{ST} Manhattan plots between Corsica and the mainland.** F_{ST} are given either
911 by SNP, by 200kb windows or by 50 SNPs windows. Dots alternate colors between
912 chromosomes. F_{ST} for unplaced loci are shown at the end of the first Manhattan plot.
913
914 **Figure 6. Relationship between recombination rate and divergence. A)** Correlation
915 between local recombination rate and F_{ST} between populations on the mainland and in
916 Corsica; **B)** SNP frequency per class of recombination rate; **C)** Outlier SNP frequency for
917 the mainland-Corsica comparison, per class of recombination rate; **D)** Outlier SNP
918 frequency for the deciduous-evergreen comparisons, per class of recombination rate. In
919 panels C & D, outliers are given for the Bayescan and the RDA tests.
920

921 **Figure 7. Detection and characteristics of a putative inversion on chromosome 3.** A)
922 Recombination rate and SNP density along chromosome 3; B) individual values for
923 eigenvector 1 of a PCA of all the blue tit individuals from Corsica and the mainland; C)
924 MDS 1 values from Lostruct analysis; D) F_{ST} between inverted, non-inverted and
925 heterozygous individuals from the mainland; E) Nucleotide diversity for inverted (in red),
926 non-inverted (in green) and heterozygous individuals (in blue) from the mainland; F)
927 PCA results for the region of the putative inversion for all the blue tit individuals; G)
928 Mainland blue tits' admixture proportions for K = 2 for the region of the putative
929 inversion (individual ordered by decreasing ancestry value for the inverted cluster); H)
930 Linkage disequilibrium for SNPs found in the putative inversion (plus 1 Mb each side of
931 the inversion), for mainland blue tit; I) genetic distances at the inversion, between 4
932 inverted blue tits, 8 non-inverted blue tits, and 4 great tits (sites of capture are indicated
933 for each individual).
934

Acknowledgments

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. This project was funded by the European Research Council (Starting grant ERC-2013-StG-337365-SHE to AC) and the OSU-OREME.

Authors' contributions:

CP conceived the study, conducted field-work, lab analyses, bioinformatics and statistical analyses, interpreted the results and wrote the manuscript. QR conducted bioinformatics and statistical analyses (especially demographic analyses using ABC), interpreted the results and edited the manuscript. AC obtained the ERC grant, conceived the study, performed field-work, interpreted the results and edited the manuscript.

Data accessibility:

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 doi:10.5061/dryad.x69p8czfg.

954 **References**

955 Alexa A., & Rahnenführer J. (2009) Gene set enrichment analysis with topGO.
956 Bioconductor Improv, 27.

957 Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in
958 unrelated individuals. Genome Research, 19, 1655–1664.

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

961 Baird NA, Etter PD, Atwood TS et al. (2008) Rapid SNP Discovery and Genetic Mapping
962 Using Sequenced RAD Markers (JC Fay, Ed,). PlosOne 3, e3376–7.

963 Barbato M, Orozco-terWengel P, Tapio M, Bruford MW (2015) SNeP: a tool to estimate
964 trends in recent effective population size trajectories using genome-wide SNP data.
965 Frontiers in Genetics, 6, 1294–7.

966 Beissinger TM, Rosa GJM, Kaeppler SM, Gianola D, De Leon N (2015) Defining window-
967 boundaries for genomic analyses using smoothing spline techniques. Genetics
968 Selection Evolution, 47, 30–9.

969 Berner D, Roesti M (2017) Genomics of adaptive divergence with chromosome-scale
970 heterogeneity in crossover rate. Molecular Ecology, 114, 7061–19.

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

973 Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why
974 genome scans may fail to map local adaptation genes. Molecular Ecology, 20, 2044–
975 2072.

976 Backström N., Karaiskou N., Leder E. H., Gustafsson L., Primmer C. R., Qvarnström A., &
977 Ellegren H. (2008). A gene-based genetic linkage map of the collared flycatcher
978 (*Ficedula albicollis*) reveals extensive synteny and gene-order conservation during
979 100 million years of avian evolution. Genetics, 179, 1479-1495.

980 Barton N., and Bengtsson B.O. (1986) The barrier to genetic exchange between
981 hybridising populations. Heredity 57, 357–376.

982 Blanquart F, Kaltz O, Nuismer SL, Gandon S (2013) A practical guide to measuring local

983 adaptation (D Ebert, Ed,). Ecology Letters, 16, 1195–1205.
984 Blondel J, Maistre M, Perret P, Hurtrez-Boussès S, Lambrechts M (1998) Is the small
985 clutch size of a Corsican blue tit population optimal? Oecologia, 117, 80–89.
986 Blondel J, Thomas DW, Charmantier A et al. (2006) A Thirty-Year Study of Phenotypic
987 and Genetic Variation of Blue Tits in Mediterranean Habitat Mosaics. Bioscience, 56,
988 661–673.
989 Booker TR, Yeaman S, & Whitlock M (2020) Variation in recombination rate affects
990 detection of FST outliers under neutrality. *bioRxiv*.
991 https://doi.org/10.1101/2020.02.06.937813.
992 Bosse M, Spurgin LG, Laine VN et al. (2017) Recent natural selection causes adaptive
993 evolution of an avian polygenic trait. Science (New York, N.Y.), 358, 365–368.
994 Broggi J, Hohtola E, Orell M, Nilsson J-Å (2005) Local adaptation to winter conditions in a
995 passerine spreading north: a common‐garden approach. Evolution, 59, 1600–
996 1603.
997 Burri R, Nater A, Kawakami T et al. (2015) Linked selection and recombination rate
998 variation drive the evolution of the genomic landscape of differentiation across the
999 speciation continuum of Ficedula flycatchers. Genome Research, 25, 1656–1665.
1000 Burton RS, Barreto FS (2012) A disproportionate role for mtDNA in Dobzhansky-Muller
1001 incompatibilities? Molecular Ecology, 21, 4942–4957.
1002 Carbonell R, Perez-Tris J, Telleria JL (2003) Effects of habitat heterogeneity and local
1003 adaptation on the body condition of a forest passerine at the edge of its
1004 distributional range. Biological journal of the linnean society, 479–488.
1005 Caro SP, Charmantier A, Lambrechts M et al. (2009) Local adaptation of timing of
1006 reproduction: females are in the driver's seat. Functional Ecology, 23, 172–179.
1007 Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis
1008 tool set for population genomics. Molecular Ecology, 22, 3124–3140.
1009 Cayuela H, Rougemont Q, Prunier JG et al. (2018) Demographic and genetic approaches
1010 to study dispersal in wild animal populations: A methodological review. Molecular
1011 Ecology, 27, 3976–4010.

1012 Charlesworth B, Morgan MT, Charlesworth D (1993) The Effect of Deleterious Mutations
1013 on Neutral Molecular Variation. Genetics, 134, 1289–1303.
1014 Charmantier A, Doutrelant C, Dubuc-Messier G, Fargevieille A, SZULKIN M (2016)
1015 Mediterranean blue tits as a case study of local adaptation. Evolutionary
1016 Applications 9, 135–152.
1017 Charmantier A, Kruuk LEB, Blondel J, Lambrechts MM (2004) Testing for microevolution
1018 in body size in three blue tit populations. Journal of Evolutionary Biology, 17, 732–
1019 743.
1020 Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of speciation
1021 are due to reduced diversity, not reduced gene flow. Molecular Ecology, 23, 3133–
1022 3157.
1023 Csilléry K, Blum MGB, Gaggiotti OE, François O (2010) Approximate Bayesian
1024 Computation (ABC) in practice. Trends in Ecology & Evolution, 25, 410–418.
1025 Cutter AD, Payseur BA (2013) Genomic signatures of selection at linked sites: unifying
1026 the disparity among species. Nature Reviews Genetics, 14, 262–274.
1027 da Silva VH, Laine VN, Bosse M et al. (2019) The Genomic Complexity of a Large
1028 Inversion in Great Tits. (J Mank, Ed,). Genome Biology and Evolution, 11, 1870–
1029 1881.
1030 Danecek P, Auton A, Abecasis G et al. (2011) The variant call format and VCFtools.
1031 Bioinformatics, 27, 2156–2158.
1032 de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan
1033 methods against more complex models: when and how much should we trust
1034 them? Molecular Ecology, 23, 2006–2019.
1035 Delahaie B, Charmantier A, Chantepie S et al. (2017) Conserved G-matrices of
1036 morphological and life-history traits among continental and island blue tit
1037 populations. Heredity ,119, 76-87.
1038 Dubuc-Messier G, Caro SP, Perrier C et al. (2018) Gene flow does not prevent
1039 personality and morphological differentiation between two blue tit populations.
1040 Journal of Evolutionary Biology, 23, 960–11.

1041 Ewing, G. B., and J. D. Jensen. (2016) The consequences of not accounting for
1042 background selection in demographic inference. Mol. Ecol. 25, 135–141
1043 Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically
1044 structured population. Heredity, 103, 285–298.
1045 Foll M, Gaggiotti O (2008) A Genome-Scan Method to Identify Selected Loci Appropriate
1046 for Both Dominant and Codominant Markers: A Bayesian Perspective. Genetics, 180,
1047 977–993.
1048 Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR (2015) Detecting spatial genetic
1049 signatures of local adaptation in heterogeneous landscapes. Molecular Ecology, 25,
1050 104–120.
1051 Forester B.R., Lasky J.R., Wagner H.H., & Urban D.L. (2018) Comparing methods for
1052 detecting multilocus adaptation with multivariate genotype–environment
1053 associations. Molecular Ecology, 27, 2215-2233.
1054 Gagnaire P-A, Gaggiotti OE (2016) Detecting polygenic selection in marine populations
1055 by combining population genomics and quantitative genetics approaches. Current
1056 Zoology, 62, 603–616.
1057 Gagnaire P-A, Lamy J-B, Cornette F et al. (2018) Analysis of Genome-Wide
1058 Differentiation between Native and Introduced Populations of the Cupped Oysters
1059 Crassostrea gigas and Crassostrea angulata. Genome Biology and Evolution, 10,
1060 2518–2534.
1061 Hansson B, Sigeman H, Stervander M et al. (2018) Contrasting results from GWAS and
1062 QTL mapping on wing length in great reed warblers. Molecular Ecology Resources,
1063 18, 867–876.
1064 Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary (KJ
1065 Willis, KD Bennett, D Walker, Eds,). Philosophical Transactions of the Royal Society
1066 B: Biological Sciences, 359, 183–195.
1067 Hoban S, Kelley JL, Lotterhos KE et al. (2016) Finding the Genomic Basis of Local
1068 Adaptation: Pitfalls, Practical Solutions, and Future Directions. The American
1069 Naturalist, 188, 379–397.

1070 Hohenlohe PA, Bassham S, Etter PD et al. (2010) Population genomics of parallel
1071 adaptation in threespine stickleback using sequenced RAD tags. Plos Genetics, 6,
1072 e1000862–e1000862.
1073 Illera J-C, KOIVULA K, Broggi J et al. (2011) A multi-gene approach reveals a complex
1074 evolutionary history in the Cyanistes species group. Molecular Ecology, 20, 4123–
1075 4139.
1076 Jay P, Chouteau M, Whibley A, Bastide H, Llaurens V, Parrinello H, & Joron M (2019)
1077 Mutation accumulation in chromosomal inversions maintains wing pattern
1078 polymorphism in a butterfly. *bioRxiv*, 736504.
1079 Jouet G, Berné S, Rabineau M et al. (2006) Shoreface migrations at the shelf edge and
1080 sea-level changes around the Last Glacial Maximum (Gulf of Lions, NW
1081 Mediterranean). Marine Geology, 234, 21–42.
1082 Kamvar Z.N., Tabima J.F., & Grünwald N.J. (2014) Poppr: an R package for genetic
1083 analysis of populations with clonal, partially clonal, and/or sexual reproduction.
1084 PeerJ, 2, e281.
1085 Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. Ecology Letters, 7,
1086 1225–1241.
1087 Kim K-W, Bennison C, Hemmings N et al. (2017) A sex-linked supergene controls sperm
1088 morphology and swimming speed in a songbird. Nature Ecology & Evolution, 1,
1089 1168–1176.
1090 Kingman JFC (1982) The coalescent. Stochastic Processes and their Applications, 13,
1091 235–248.
1092 Kirkpatrick M (2006) Chromosome Inversions, Local Adaptation and Speciation.
1093 Genetics, 173, 419–434.
1094 Kirkpatrick M (2010) How and Why Chromosome Inversions Evolve. PLoS Biology, 8,
1095 e1000501–5.
1096 Knief U, Forstmeier W, Pei Y et al. (2017) A sex-chromosome inversion causes strong
1097 overdominance for sperm traits that affect siring success. Nature Ecology &
1098 Evolution, 1–8.

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.

Kvist L, Viiri K, Dias PC, 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 PM, Calhim S et al. (2015) Sympatric divergence and clinal variation in multiple coloration traits of Ficedulaflycatchers. Journal of Evolutionary Biology, 28, 779–790.

Laine VN, Gossmann TI, Schachtschneider KM et al. (2016) Evolutionary signals of selection on cognition from the great tit genome and methylome. Nature Communications, 7, 1–9.

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.

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 (2016) A beak size locus in Darwin's finches facilitated character displacement during a drought. Science, 352, 470-474.

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.

Lenormand T (2002) Gene flow and the limits to natural selection. Trends in Ecology & Evolution, 17, 183–189.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics, 25, 1754–1760.

Li H, Ralph P (2019) Local PCA Shows How the Effect of Population Structure Differs Along the Genome. Genetics, 211, 289–304.

Li H, Handsaker B, Wysoker A et al. (2009) The Sequence Alignment/Map format and

1128 SAMtools. Bioinformatics, 25, 2078–2079.
1129 Lohse K, Clarke M, Ritchie MG, Etges WJ (2015) Genome-wide tests for introgression
1130 between cactophilic Drosophila implicate a role of inversions during speciation.
1131 Evolution, 69, 1178–1190.
1132 Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local
1133 adaptation depends on sampling design and statistical method. Molecular Ecology,
1134 24, 1031–1046.
1135 Lowry DB, Willis JH (2010) A Widespread Chromosomal Inversion Polymorphism
1136 Contributes to a Major Life-History Transition, Local Adaptation, and Reproductive
1137 Isolation (NH Barton, Ed,). PLoS Biology, 8, e1000500–14.
1138 Lowry DB, Hoban S, Kelley JL et al. (2017) Responsible RAD: Striving for best practices in
1139 population genomic studies of adaptation. Molecular Ecology Resources, 17, 366–
1140 369.
1141 Lundregan SL, Hagen IJ, Gohli J et al. (2018) Inferences of genetic architecture of bill
1142 morphology in house sparrow using a high-density SNP array point to a polygenic
1143 basis. Molecular Ecology, 27, 3498–3514.
1144 Ma J, Amos CI (2012) Investigation of Inversion Polymorphisms in the Human Genome
1145 Using Principal Components Analysis (F Emmert-Streib, Ed,). PLOS ONE, 7, e40224–
1146 12.
1147 Manel S, Perrier C, Pratlong M et al. (2016) Genomic resources and their influence on
1148 the detection of the signal of positive selection in genome scans. Molecular Ecology,
1149 25, 170–184.
1150 Marko PB, Hart MW (2011) The complex analytical landscape of gene flow inference.
1151 Trends in Ecology & Evolution, 26, 448–456.
1152 McGinnis S, Madden TL (2004) BLAST: at the core of a powerful and diverse set of
1153 sequence analysis tools. Nucleic Acids Research, 32, W20–W25.
1154 McVean GAT (2004) The Fine-Scale Structure of Recombination Rate Variation in the
1155 Human Genome. Science, 304, 581–584.
1156 Nachman MW, Payseur BA (2011) Recombination rate variation and speciation:

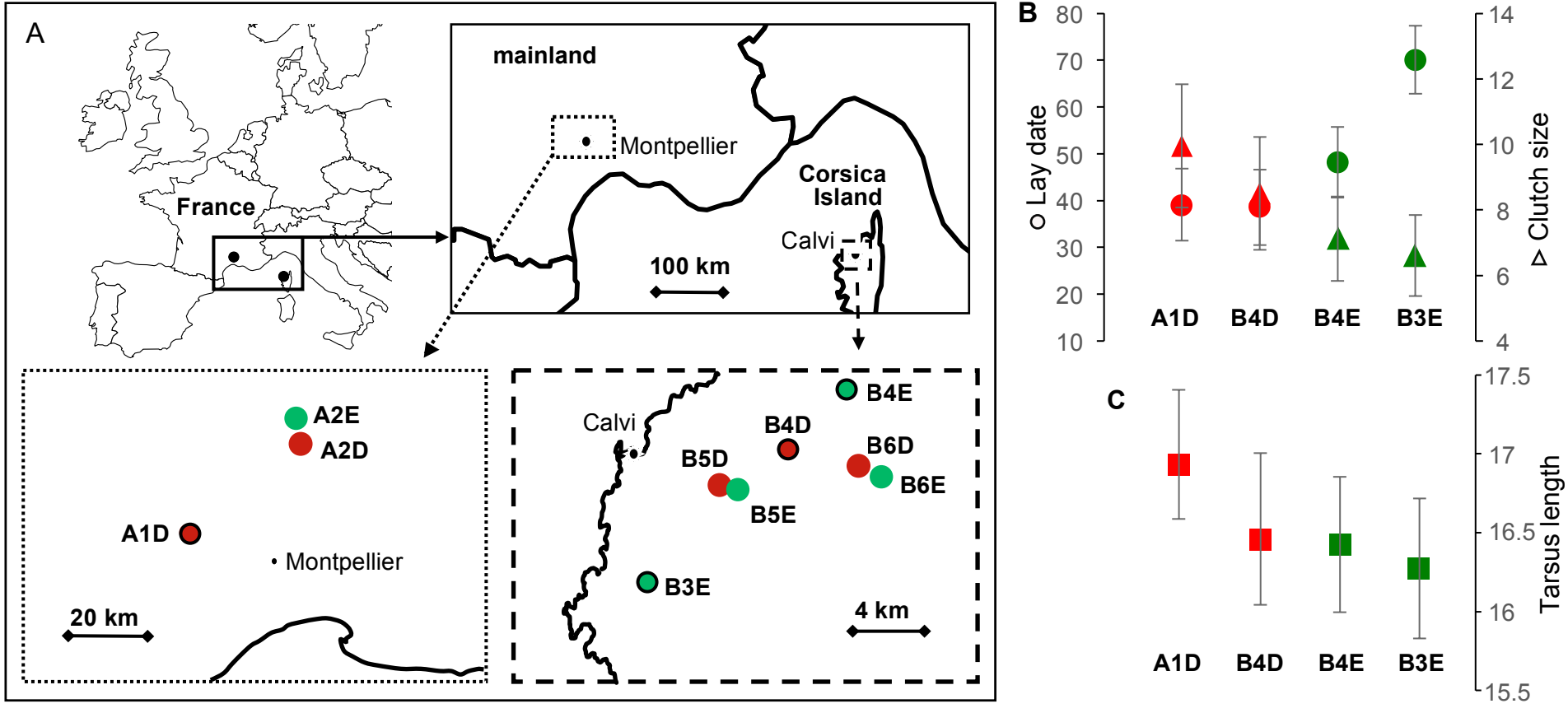
1157 theoretical predictions and empirical results from rabbits and mice. Philosophical
1158 Transactions of the Royal Society B: Biological Sciences, 367, 409–421.
1159 Narum SR, Hess JE (2011) Comparison of FST outlier tests for SNP loci under selection.
1160 Molecular Ecology Resources, 11, 184–194.
1161 Oksanen J, Kindt R, Legendre P et al. (2007) The vegan package. Community ecology
1162 package, 10, 631–637.
1163 Ortego J, García-Navas V, Ferrer ES, Sanz JJ (2011) Genetic structure reflects natal
1164 dispersal movements at different spatial scales in the blue tit, Cyanistes caeruleus.
1165 Animal Behaviour, 82, 131–137.
1166 Päckert M., Martens J., Tietze D.T., Dietzen C., Wink M., & Kvist L. (2007). Calibration of
1167 a molecular clock in tits (Paridae)—Do nucleotide substitution rates of
1168 mitochondrial genes deviate from the 2% rule? Molecular phylogenetics and
1169 evolution, 44, 1-14.
1170 Pembleton LW, Cogan NOI, Forster JW (2013) StAMPP: an R package for calculation of
1171 genetic differentiation and structure of mixed-ploidy level populations. Molecular
1172 Ecology Resources, 13, 946–952.
1173 Perrier C, Charmantier A (2019) On the importance of time scales when studying
1174 adaptive evolution. Evolution Letters 3, 240-247.
1175 Perrier C, Delahaie B, Charmantier A (2018) Heritability estimates from genomewide
1176 relatedness matrices in wild populations: Application to a passerine, using a small
1177 sample size. Molecular Ecology Resources, 3, e3376–16.
1178 Perrier C, Ferchaud A-L, Sirois P, Thibault I, Bernatchez L (2017) Do genetic drift and
1179 accumulation of deleterious mutations preclude adaptation? Empirical investigation
1180 using RADseq in a northern lacustrine fish. Molecular Ecology, 26, 6317–6335.
1181 Pfeifer B., Wittelsbürger U., Ramos-Onsins S.E., & Lercher M.J. (2014) PopGenome: an
1182 efficient Swiss army knife for population genomic analyses in R. *Molecular biology*
1183 *and evolution*, *31*, 1929-1936.
1184 Pinho C, Hey J (2010) Divergence with Gene Flow: Models and Data. Annual Review of
1185 Ecology Evolution and Systematics, 41, 215–230.

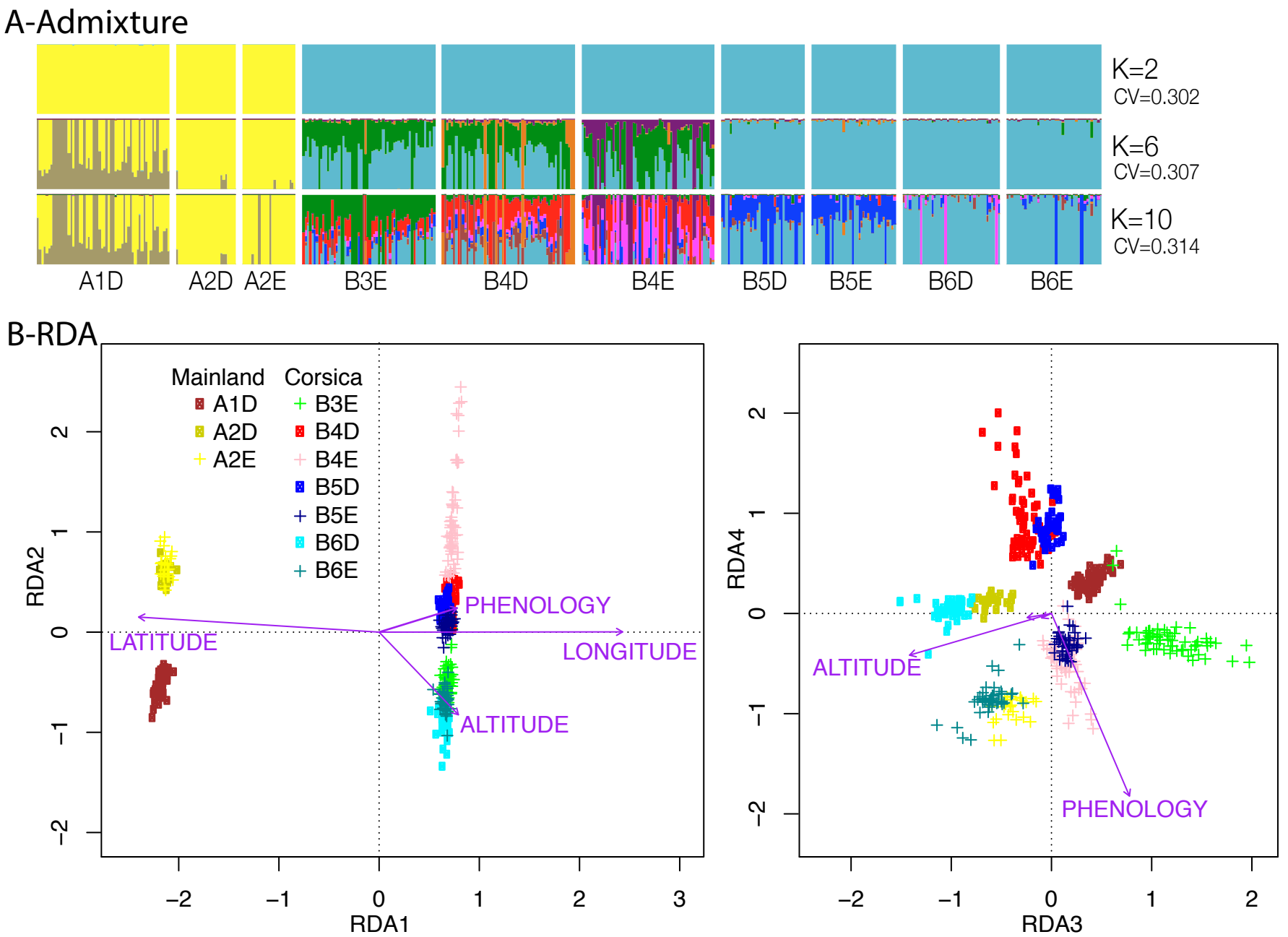
1186 Porlier M, Charmantier A, Bourgault P et al. (2012a) Variation in phenotypic plasticity
1187 and selection patterns in blue tit breeding time: between- and within-population
1188 comparisons. Journal of Animal Ecology, 81, 1041–1051.
1189 Porlier M, Garant D, Perret P, Charmantier A (2012b) Habitat-Linked Population Genetic
1190 Differentiation in the Blue Tit Cyanistes caeruleus. Journal of Heredity, 103, 781–
1191 791.
1192 Purcell S, Neale B, Todd-Brown K et al. (2007) PLINK: a tool set for whole-genome
1193 association and population-based linkage analyses. American Journal of Human
1194 Genetics, 81, 559–575.
1195 Reid NM, Proestou DA, Clark BW et al. (2016) The genomic landscape of rapid repeated
1196 evolutionary adaptation to toxic pollution in wild fish. Science, 354, 1305–1308.
1197 Rochette N.C., & Catchen J.M. (2017). Deriving genotypes from RAD-seq short-read data
1198 using Stacks. Nature Protocols, 12, 2640-2659
1199 Rockman MV (2012) The QTN program and the alleles that matter for evolution: all
1200 that's gold does not glitter. Evolution, 66, 1–17.
1201 Rougemont Q, Bernatchez L (2018) The demographic history of Atlantic Salmon (Salmo
1202 salar) across its distribution range reconstructed from Approximate Bayesian
1203 Computations. Evolution, 1–3.
1204 Rougemont Q, Gagnaire P-A, Perrier C et al. (2016) Inferring the demographic history
1205 underlying parallel genomic divergence among pairs of parasitic and nonparasitic
1206 lamprey ecotypes. Molecular Ecology, 26, 142–162.
1207 Rougemont Q, Moore J-S, Leroy T et al. (2019) Demographic history, linked selection,
1208 and recombination shape the genomic landscape of a broadly distributed Pacific
1209 salmon. 123, 887–37.
1210 Roux C, Fraïsse C, Romiguier J et al. (2016) Shedding Light on the Grey Zone of
1211 Speciation along a Continuum of Genomic Divergence (C Moritz, Ed,). PLoS Biology,
1212 14, e2000234–23.
1213 Roux C., Tsagkogeorga G., Bierne N., and Galtier N. (2013) Crossing the species barrier:
1214 genomic hotspots of introgression between two highly divergent Ciona intestinalis

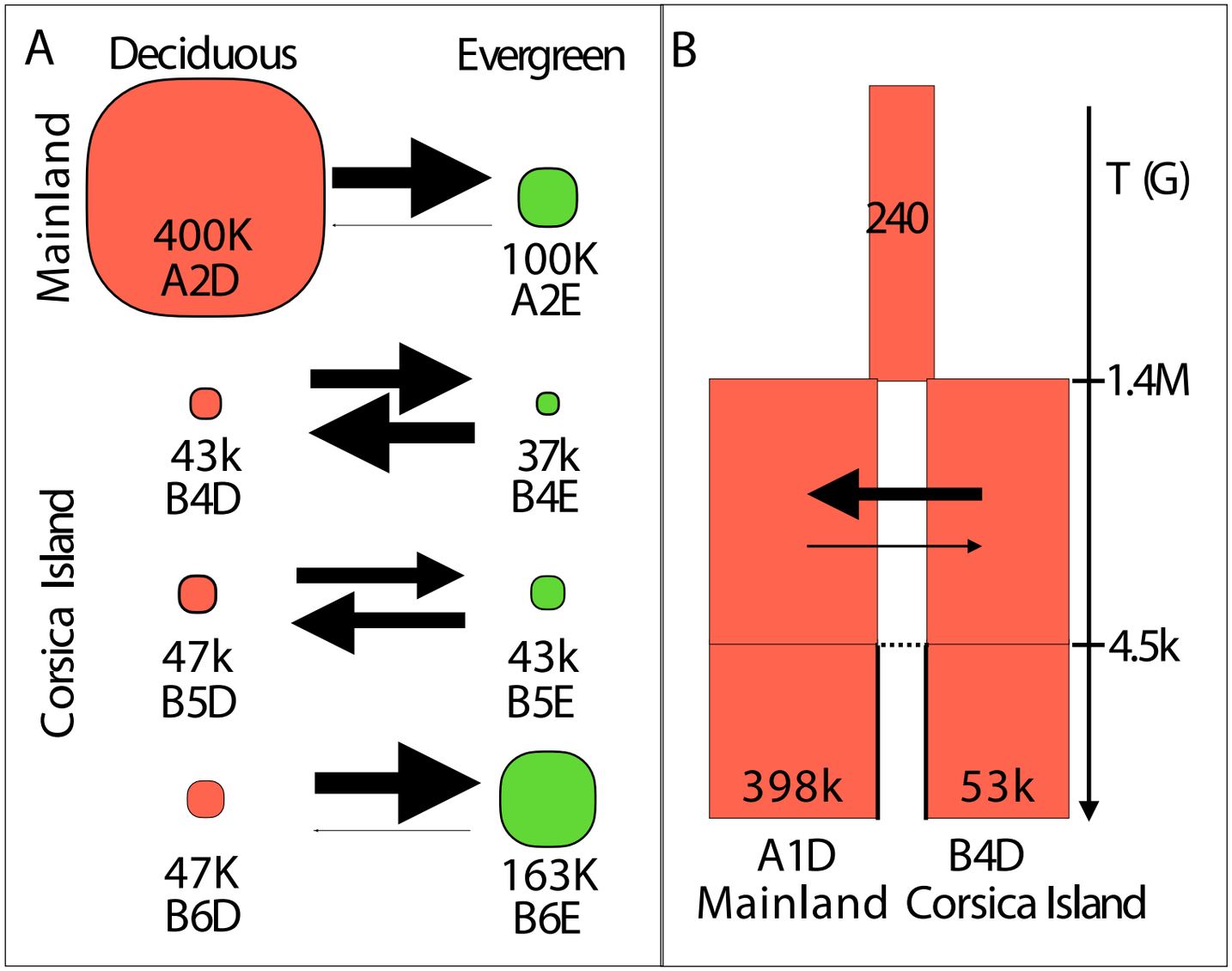
1215 species. Mol. Biol. Evol. 30, 1574–1587.
1216 Santure AW, Gratten J, Mossman JA, Sheldon BC, Slate J. (2011) Characterisation of the
1217 transcriptome of a wild great tit Parus major population by next generation
1218 sequencing. BMC Genomics 12:283. doi:10.1186/1471-2164-12-283.
1219 Santure AW, Cauwer I, Robinson MR et al. (2013) Genomic dissection of variation in
1220 clutch size and egg mass in a wild great tit (Parus major) population. Molecular
1221 Ecology, 22, 3949–3962.
1222 Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation.
1223 Nature Publishing Group, 14, 807–820.
1224 Schrider D.R., Shanku A.G., and Kern A.D. (2016) Effects of linked selective sweeps on
1225 demographic inference and model selection. Genetics 204, 1207–1223.
1226 Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for
1227 DNA analyses. Canadian Journal of Zoology-Revue Canadienne De Zoologie, 69, 82–
1228 90.
1229 Sinclair-Waters M, Bradbury IR, Morris CJ et al. (2017) Ancient chromosomal
1230 rearrangement associated with local adaptation of a post-glacially colonized
1231 population of Atlantic Cod in the northwest Atlantic. Molecular Ecology, 27, 339-
1232 351.
1233 Smith B.T., & Klicka J. (2013) Examining the role of effective population size on
1234 mitochondrial and multilocus divergence time discordance in a songbird. PloS one,
1235 8(2), e55161.
1236 Sousa V.C., Carneiro M., Ferrand N., and Hey J. (2013) Identifying loci under selection
1237 against gene flow in isolation-with-migration models. Genetics 194, 211–233.
1238 Spurgin LG, Bosse M, Adriaensen F et al. (2019) The great tit HapMap project: a
1239 continental-scale analysis of genomic variation in a songbird. bioRxiv, 1–24.
1240 Stapley J, Feulner PGD, Johnston SE, Santure AW, Smadja CM (2017) Variation in
1241 recombination frequency and distribution across eukaryotes: patterns and
1242 processes. Philosophical Transactions of the Royal Society B: Biological Sciences,
1243 372, 20160455–18.

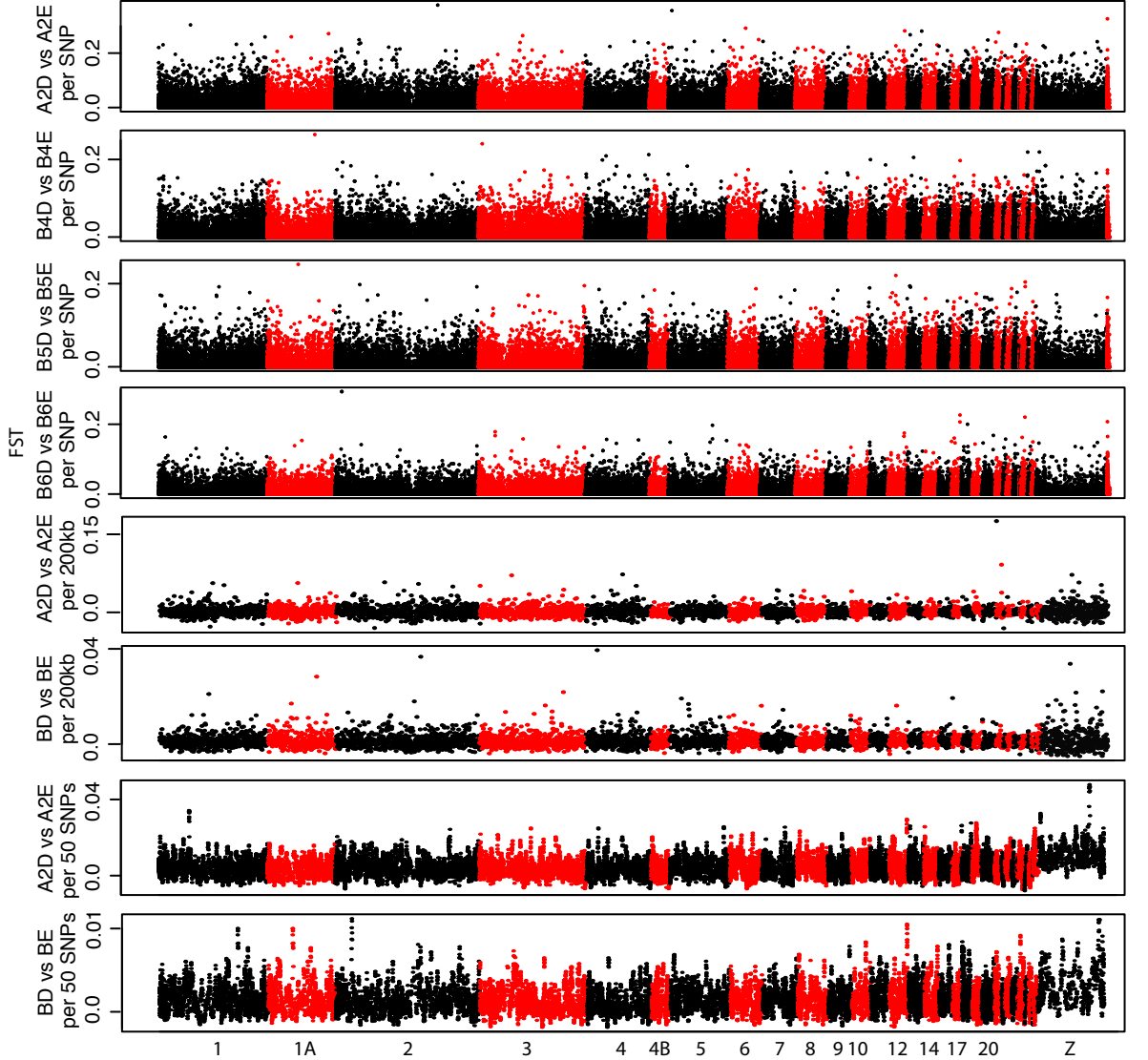
1244 Stinchcombe JR, Hoekstra HE (2007) Combining population genomics and quantitative
1245 genetics: finding the genes underlying ecologically important traits. Heredity, 100,
1246 158–170.
1247 Szulkin M, Gagnaire PA, Bierne N, Charmantier A (2016) Population genomic footprints
1248 of fine-scale differentiation between habitats in Mediterranean blue tits. Molecular
1249 Ecology, 25, 542–558.
1250 Taberlet P, Meyer A, Bouvet J (1992) Unusual mitochondrial DNA polymorphism in two
1251 local populations of blue tit Parus caeruleus. Molecular Ecology, 1, 27–36.
1252 Teplitsky C, Tarka M, Møller AP et al. (2014) Assessing Multivariate Constraints to
1253 Evolution across Ten Long-Term Avian Studies. 9, e90444–15.
1254 Tiffin P, Ross-Ibarra J (2014) Advances and limits of using population genetics to
1255 understand local adaptation. Trends in Ecology & Evolution, 29, 1–8.
1256 Tine M, Kuhl H, Gagnaire P-A et al. (2014) European sea bass genome and its variation
1257 provide insights into adaptation to euryhalinity and speciation. Nature
1258 Communications, 5, 1–10.
1259 Todesco M, Owens GL, Bercovich N et al. (2019) Massive haplotypes underlie ecotypic
1260 differentiation in sunflowers. bioRxiv, 19, 337–66.
1261 Tufto J, Ringsby TH, Dhondt AA, Adriaensen F, Matthysen E (2005) A parametric model
1262 for estimation of dispersal patterns applied to five passerine spatially structured
1263 populations. American Naturalist, 165, E13–26.
1264 Tuttle EM, Bergland AO, Korody ML et al. (2016) Divergence and Functional Degradation
1265 of a Sex Chromosome-like Supergene. Current Biology, 1–8.
1266 Wang IJ, Bradburd GS (2014) Isolation by environment. Molecular Ecology, 23, 5649–
1267 5662.
1268 Waples RS, Do C (2008) ldne: a program for estimating effective population size from
1269 data on linkage disequilibrium. Molecular Ecology Resources, 8, 753–756.
1270 Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population
1271 Structure. Evolution, 38, 1–14.
1272 Wellenreuther M, Hansson B (2016) Detecting Polygenic Evolution: Problems, Pitfalls,

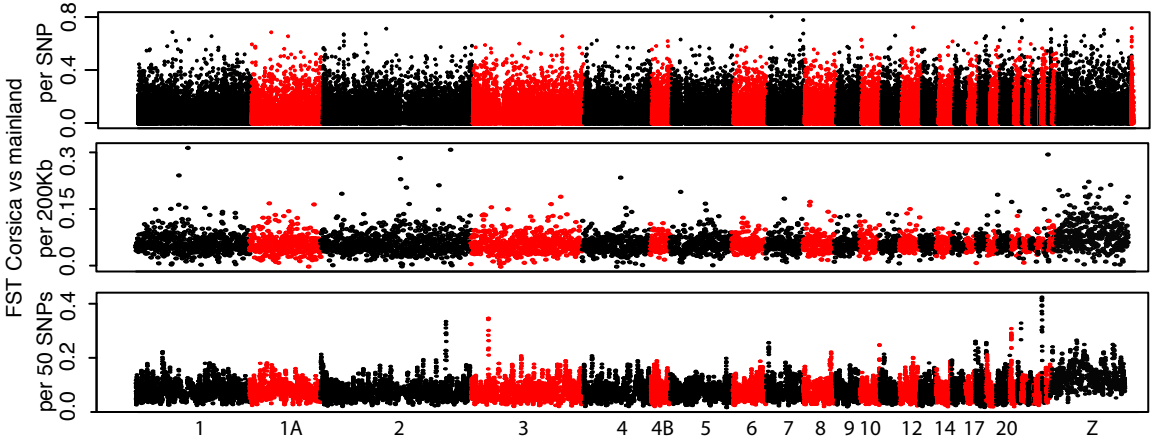
1273 and Promises. Trends in Genetics, 32, 155–164.
1274 Wellenreuther M, Merot C, Berdan E, Bernatchez L (2019) Going beyond SNPs: The role
1275 of structural genomic variants in adaptive evolution and species diversification.
1276 Molecular Ecology, 28, 1203–1209.
1277 Yeaman S (2015) Local Adaptation by Alleles of Small Effect. American Naturalist, 186
1278 Suppl 1, S74–89.
1279 Yeaman S, Whitlock MC (2011) The genetic architecture of adaptation under migration-
1280 selection balance. Evolution, 65, 1897–1911.
1281 Zheng X, Levine D, Shen J et al. (2012) A high-performance computing toolset for
1282 relatedness and principal component analysis of SNP data. Bioinformatics, 28,
1283 3326–3328.
1284

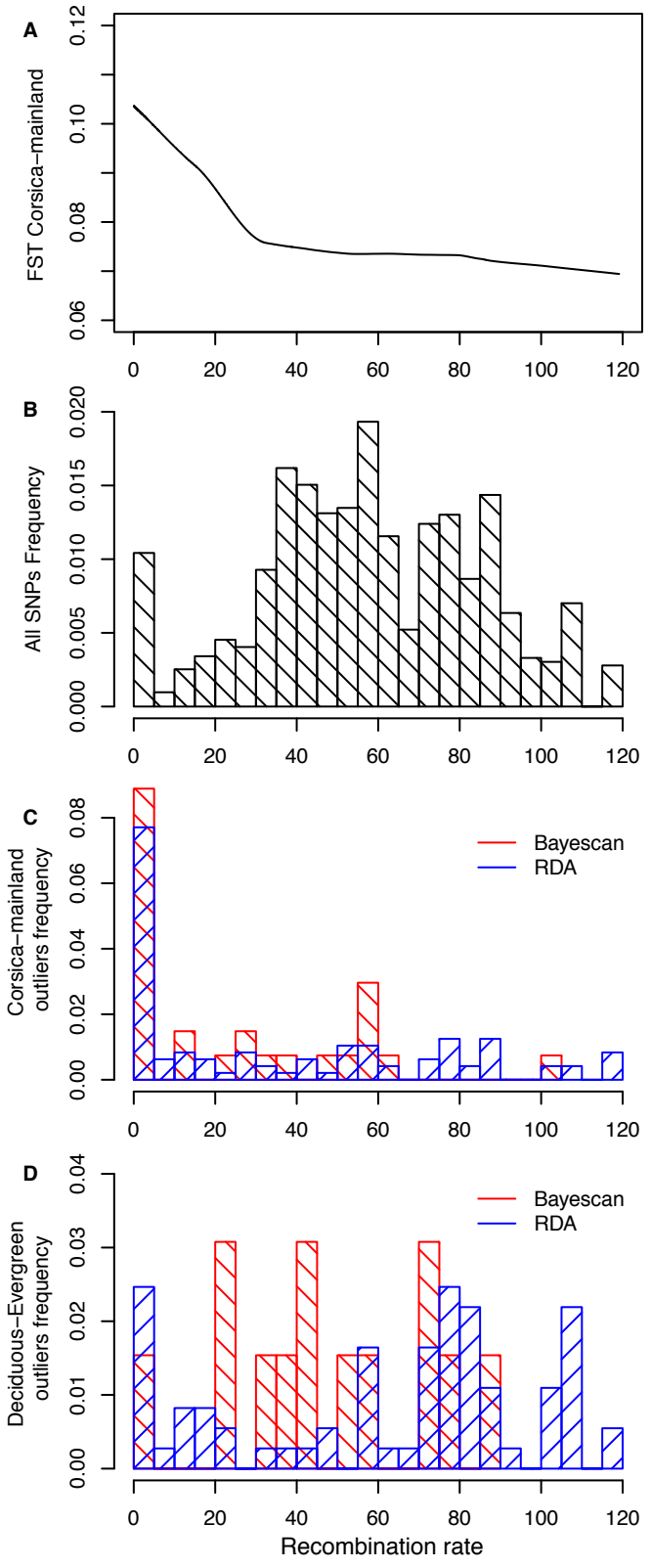

A
France
mainland
Montpellier
Corsica
Island
Calvi
100 km
A2E
A2D
A1D
Montpellier
20 km
Calvi
B4E
B4D
B6D
B5D
B6E
B5E
B3E
4 km
B
Lay date
Clutch size
80
70
60
50
40
30
20
10
14
12
10
8
6
4
A1D
B4D
B4E
B3E
C
Tarsus length
17.5
17
16.5
16
15.5
A1D
B4D
B4E
B3E

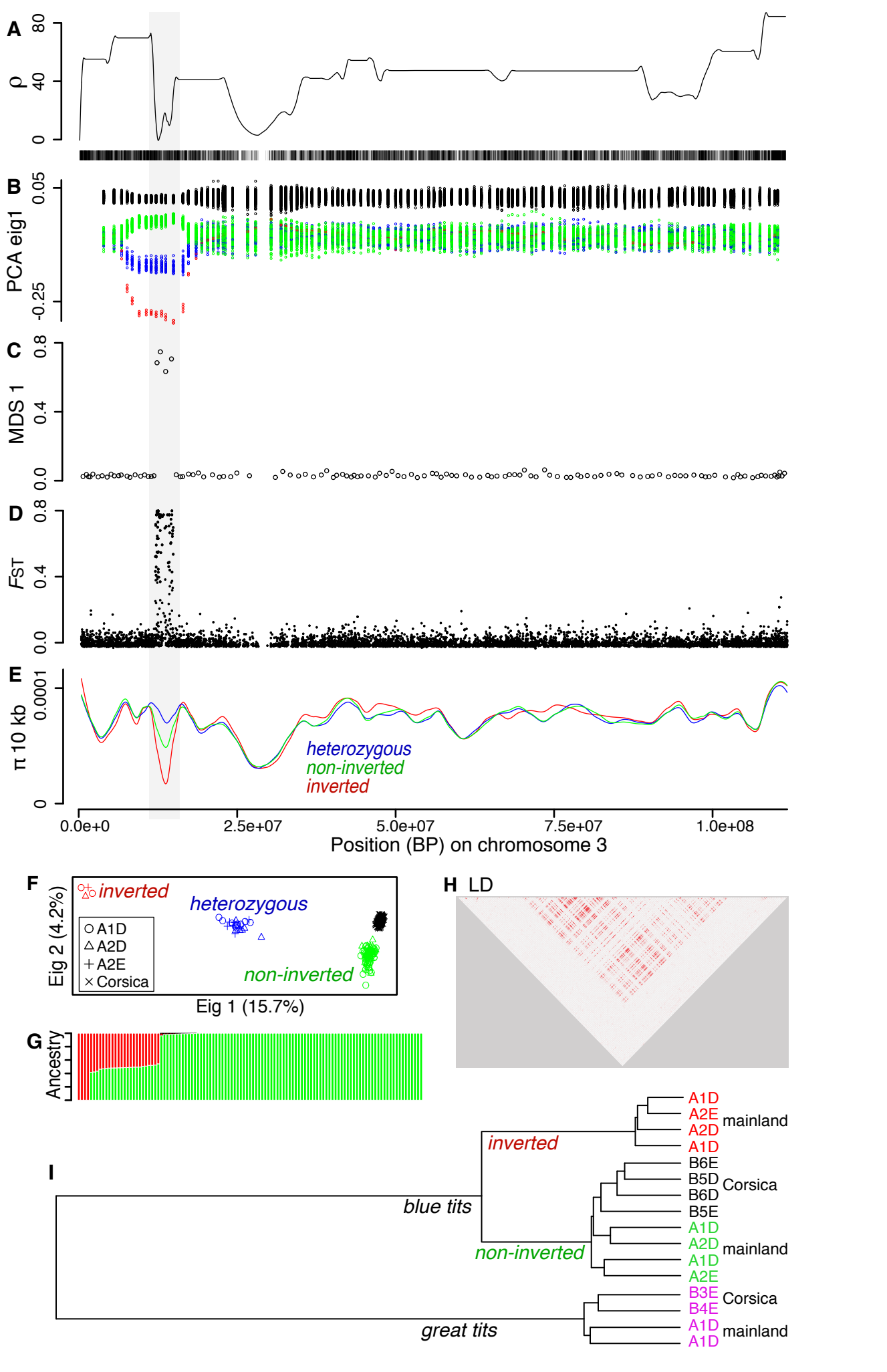

A-Admixture
K=2
CV=0.302
K=6
CV=0.307
K=10
CV=0.314
A1D
A2D
A2E
B3E
B4D
B4E
B5D
B5E
B6D
B6E
B-RDA
Mainland
A1D
A2D
A2E
Corsica
B3E
B4D
B4E
B5D
B5E
B6D
B6E
PHENOLOGY
LATITUDE
LONGITUDE
ALTITUDE
RDA1
RDA2
RDA3
RDA4
ALTITUDE
PHENOLOGY


A
Deciduous
Evergreen
Mainland
400K
A2D
100K
A2E
43k
B4D
37k
B4E
Corsica Island
47k
B5D
43k
B5E
47K
B6D
163K
B6E
B
240
T (G)
1.4M
4.5k
398k
53k
A1D
Mainland
B4D
Corsica Island


FST
A2D vs A2E per SNP
B4D vs B4E per SNP
B5D vs B5E per SNP
B6D vs B6E per SNP
A2D vs A2E per 200kb
BD vs BE per 200kb
A2D vs A2E per 50 SNPs
BD vs BE per 50 SNPs
0.0
0.2
0.15
0.04
0.01
1
1A
2
3
4
4B
5
6
7
8
9
10
12
14
17
20
Z


FST Corsica vs mainland
per SNP
0.0
0.4
0.8
per 200Kb
0.0
0.15
0.3
per 50 SNPs
0.0
0.2
0.4
1
1A
2
3
4
4B
5
6
7
8
9
10
12
14
17
20
Z


A
FST Corsica−mainland
0.06
0.08
0.10
0.12
0
20
40
60
80
100
120
B
All SNPs Frequency
0.000
0.005
0.010
0.015
0.020
C
Corsica−mainland outliers frequency
0.00
0.02
0.04
0.06
0.08
Bayescan
RDA
D
Deciduous−Evergreen outliers frequency
0.00
0.01
0.02
0.03
0.04
Recombination rate


A
ρ
80
40
0
B
PCA eig1
0.05
-0.25
C
MDS 1
0.8
0.4
0.0
D
FST
0.8
0.4
0.0
E
π 10 kb
0.0001
0
heterozygous
non-inverted
inverted
0.0e+0
2.5e+07
5.0e+07
7.5e+07
1.0e+08
Position (BP) on chromosome 3
F
inverted
heterozygous
non-inverted
A1D
A2D
A2E
Corsica
Eig 2 (4.2%)
Eig 1 (15.7%)
G
Ancestry
H LD
I
inverted
blue tits
non-inverted
great tits
A1D
A2E
A2D
A1D
mainland
B6E
B5D
B6D
B5E
Corsica
A1D
A2D
A1D
A2E
mainland
B3E
B4E
Corsica
A1D
A1D
mainland