



Heritability estimates from genomewide relatedness matrices in wild populations: Application to a passerine, using a small sample size

C. Perrier, B. Delahaie, A. Charmantier

► To cite this version:

C. Perrier, B. Delahaie, A. Charmantier. Heritability estimates from genomewide relatedness matrices in wild populations: Application to a passerine, using a small sample size. *Molecular Ecology Resources*, Wiley/Blackwell, 2018, 18 (4), pp.838-853. 10.1111/1755-0998.12886 . hal-02326729

HAL Id: hal-02326729

<https://hal.archives-ouvertes.fr/hal-02326729>

Submitted on 22 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Heritability estimates from genome wide relatedness matrices in wild populations:**
2 **application to a passerine, using a small sample size**

3

4 **Running head: GRM in Corsican blue tits**

5

6 Perrier C*, Delahaie B, Charmantier A.

7

8 Centre d'Ecologie Fonctionnelle et Evolutive, CNRS-UMR5175 CEFE, Montpellier,
9 France

10

11 *corresponding author: charles.perrier@cefe.cnrs.fr

12

13 Molecular Ecology Resources - Special issue Association Mapping in Natural
14 Populations.

15

16 **Keywords:** heritability; genetic correlation; RAD-sequencing; GRM; pedigree;
17 partitioning; phenotype; blue tit; SNP; *Cyanistes caeruleus*

18

19 **Abstract**

20 Genomic developments have empowered the investigation of heritability in wild
21 populations directly from genome wide relatedness matrices (GRM). Such GRM based
22 approaches can in particular be used to improve or substitute approaches based on
23 social pedigree (PED-social). However, measuring heritability from GRM in the wild has
24 not been widely applied yet, especially using small samples and in non-model species.
25 Here, we estimated heritability for four quantitative traits (tarsus length, wing length, bill
26 length and body mass), using PED-social, a pedigree corrected by genetic data (PED-
27 corrected) and a GRM from a small sample ($n = 494$) of blue tits from natural
28 populations in Corsica genotyped at nearly 50,000 filtered SNPs derived from RAD-seq.
29 We also measured genetic correlations among traits and we performed chromosome
30 partitioning. Heritability estimates were slightly higher when using GRM compared to
31 PED-social, and PED-corrected yielded intermediate values, suggesting a minor
32 underestimation of heritability in PED-social due to incorrect pedigree links, including
33 extra-pair paternity, and to lower information content than the GRM. Genetic correlations
34 among traits were similar between PED-social and GRM but credible intervals were very
35 large in both cases, suggesting a lack of power for this small dataset. Although a
36 positive linear relationship was found between the number of genes per chromosome
37 and the chromosome heritability for tarsus length, chromosome partitioning similarly
38 showed a lack of power for the three other traits. We discuss the usefulness and
39 limitations of the quantitative genetic inferences based on genomic data in small
40 samples from wild populations.

41 **Introduction**

42 Estimating additive genetic variance and heritability of quantitative traits is a major aim in
43 evolutionary biology because they are crucial components of all evolutionary models.
44 Further, fine-scale dissection of the genomic basis of quantitative traits in the wild, using
45 methods such as genome wide association and linkage mapping, requires validation
46 that the traits being mapped are actually heritable. New sequencing technologies have
47 empowered the estimation of genomic relatedness matrices (GRM) between individuals
48 (Visscher *et al.* 2008; Huisman 2017; Gienapp *et al.* 2017), feasibly without any social
49 pedigree (PED-social) recorded in the field. Since in a quantitative framework, common
50 SNPs additively explain a large part of genetic additive variance, it is possible to use
51 such realized relatedness estimated from genomic data to measure traits' heritability in
52 wild populations (Yang *et al.* 2010). Furthermore, by aligning markers against a
53 reference genome, heritability can then be partitioned among chromosomes (or at other
54 genomic scales) enabling testing quantitative models of increasing heritability with the
55 number of genes per chromosomes (Yang *et al.* 2011b), or alternatively, identifying
56 chromosomes displaying higher genetic variance than expected. GRM estimates and
57 chromosome partitioning, first applied on human data (Yang *et al.* 2010; 2011b), were
58 then applied in well-known free-ranging animal populations (Great tits, *Parus major*,
59 Robinson *et al.* 2013; Soay sheep, *Ovis aries*, Bérénos *et al.* 2015) for which thousands
60 of individuals were genotyped at thousands of loci. There have been fewer empirical
61 applications of these methods to smaller samples (but see Wenzel *et al.* 2015; Silva *et*
62 *al.* 2017), although they have great potential in the current context of genomic tools

63 being increasingly applied to natural populations (Gienapp *et al.* 2017) that often rely on
64 relatively small samples.

65 Three important limitations for estimating GRM and GRM based heritability are
66 the number of SNPs, the number of individuals and the sampling variance (Visscher &
67 Goddard 2015). The limitation from the number (as well as diversity and quality) of
68 markers and their ability to resolve relatedness essentially vanished when we entered
69 the genomic era (Csillary 2006; Stanton-Geddes *et al.* 2013; Gienapp *et al.* 2017). The
70 sample size of individuals genotyped and phenotyped, however, remains an issue. For a
71 given dataset, the number of individuals included will influence both error and median
72 values of heritability estimates, as tested in two studies (Stanton-Geddes *et al.* 2013;
73 Visscher & Goddard 2015). Nevertheless, reasonable heritability estimates can be
74 reached using 150 to 200 individuals, as showed in a study on barrelclover (*Medicago*
75 *trunculata*, Stanton-Geddes *et al.* 2013). However, different species, demographic
76 contexts, and sampling strategies will likely perform differently for a similar number of
77 individuals, number of markers and traits' architecture. Visscher & Goddard (2015)
78 showed that for designs that use genetic markers to estimate relatedness among
79 randomly sampled individuals from a population (which may be a common situation for
80 non-model species in natural populations), the error in heritability estimation is inversely
81 proportional to the squared sample size and is proportional to the effective population
82 size. Therefore, heritability estimates based from genetic marker relatedness in
83 extremely large populations will likely require thousands of samples (e.g. marine pelagic
84 fish (Gagnaire & Gaggiotti 2016)), but also thousands of loci (given the small linkage
85 disequilibrium among loci). Nevertheless, for populations of smaller size, a sampling

86 strategy aiming at capturing variance in realized relatedness offers possibilities to obtain
87 robust heritability estimates bounded by reasonable standard errors.

88 While genomics offers promising potential to obtain quantitative genetic
89 parameters for virtually any species or population (Gienapp et al. 2017), it has only been
90 used in a few studies to-date and more applications in smaller samples and non-model
91 species are needed. The first studies to estimate genomic heritability and partition
92 variation across the genome elsewhere than in humans were, understandably, focusing
93 on already well-known populations and species. In particular, analysis of 2644
94 individuals at 7203 SNPs from a long term population study of great tits was the first
95 application of quantitative genomics to a wild population (Robinson et al. 2013). The
96 Soay sheep (*Ovis aries*) also benefited from such a genomic quantitative genetic study,
97 with 5805 individuals genotyped at 37,037 autosomal SNPs (Bérénos et al. 2014). This
98 study revealed that most of the additive genetic (co)variances in sheep body size traits
99 were captured by half of the SNPs. These pioneering studies enabled the verification of
100 the power of quantitative genomics on known populations with long-term pedigrees and
101 characterized quantitative genetic parameters (Edwards 2013). Other authors also
102 estimated heritability from GRM in datasets with much less individuals, e.g. 200
103 barrelclover (*Medicago trunculata*) individuals genotyped at more than 5 million SNPs.
104 Wenzel et al. (2015) estimated genome-wide heritability in 695 red grouse (*Lagopus*
105 *lagopus scotica*) genotyped at 384 SNPs. Silva et al. (2017) estimated quantitative
106 genomic parameters in 1,898 house sparrows (*Passer domesticus*) genotyped at 6,348
107 SNPs and in 825 collared flycatchers (*Ficedula albicoloris*) genotyped at 38,689 SNPs.
108 These examples confirmed the usefulness of the method in new cases and with fewer

109 individuals. Although we did not perform a thorough meta-analysis, there was a trend
110 among these studies for an increase, among and within studies, of heritability estimates'
111 standard errors with decreasing sample size and SNP amount. Nevertheless, and
112 surprisingly considering the number of studies performing genomic analyses since the
113 beginning of the genomic era, there are relatively few applications of genomic data to
114 estimate heritability, genetic correlations and chromosome partitioning, using smaller
115 datasets.

116 Here, we assessed heritability, genetic correlations, and chromosome partitioning
117 for four phenotypic traits, namely tarsus length, wing length, bill length, and body mass,
118 in a collection of 494 individuals of Corsican blue tits (*Cyanistes caeruleus ogliastrae*)
119 from three closely located sites in Corsica (Figure 1). Based on mitochondrial genetic
120 and phenotypic data, blue tits in Corsica and Sardinia have been qualified as a
121 subspecies, *Cyanistes caeruleus ogliastrae* (Kvist *et al.* 2004). The focal populations are
122 ideal for quantitative genetic study because of the availability of pedigree, phenotypic
123 and genetic data gathered through a long-term study (Charmantier *et al.* 2016). Based
124 on microsatellite and SNP genetic data, important gene flow together with significantly
125 low genetic differentiation have been suggested between Corsican populations (Porlier
126 *et al.* 2012; Szulkin *et al.* 2016). Phenotypic variance for the four aforementioned traits
127 has been characterized as quantitative and moderately to highly heritable (Charmantier
128 *et al.* 2004b; Blondel *et al.* 2006; Charmantier *et al.* 2016). The latest published results
129 based on 17 to 27 years of monitoring, provided population-specific significant
130 heritability ranging from 0.43 to 0.57 for tarsus length, 0.20-0.33 for wing length, 0.18-
131 0.34 for bill length, and 0.22-0.32 for body mass. All trait combinations showed

132 significant additive genetic covariance (COV_A) in at least one population; in particular
133 COV_As between wing length and tarsus length and between wing length and body mass
134 were significantly positive (Delahaie et al 2017). Lastly, from 18 to 25% of young have
135 been identified as extra-pair offspring in these Corsican blue tit populations (Charmantier
136 et al. 2004a), which may lead to slight underestimation of trait heritability from a field-
137 based social pedigree (Charmantier & Réale 2005).

138 The general objective of this study was to test the effectiveness of GRM to
139 estimate heritability and genetic correlations among traits and to perform heritability
140 partitioning while using a relatively small sample size of Corsican blue tits. We produced
141 a dataset of 494 Corsican blue tits, from three sites (sample sizes of 110, 185 and 199),
142 genotyped at nearly 50,000 filtered SNPs derived from RAD-seq and phenotyped for
143 four quantitative phenotypic traits (tarsus length, wing length, bill length, and body
144 mass). We looked for erroneous links in the social pedigree (PED-social) using the
145 genetic data, in order to build a corrected social pedigree (PED-corrected). We
146 estimated heritability using both GCTA (Yang et al. 2011a) and *MCMCglmm* (Hadfield
147 2010) for the 494 genotyped individuals (with the three sites pooled or separately),
148 based on GRM, PED-social and PED-corrected. Heritability differences between PED-
149 corrected and PED-social were computed in order to infer the possible effects of
150 pedigree errors on estimates of traits heritability (notably underestimation originating
151 from extra-pair copulations (Charmantier & Réale 2005)). Comparisons between
152 heritability estimated using PED-corrected and GRM aimed at defining the effect of the
153 greater precision and information content of the GRM. We estimated the effect of the
154 number of SNPs on tarsus length heritability estimates, in order to discuss the SNP

155 density needed to recover much of the additive genetic variance. Finally, we estimated
156 genetic correlations between traits and we partitioned heritability between
157 chromosomes, although these quantitative genomic tests may challenge our sample
158 size and probably require much larger sample size to produce accurate results. We
159 discuss the usefulness and limitations of these heritability inferences in small samples
160 from wild populations of non-model species.

161

162 **Methods**

163 *Study sites, bird monitoring and phenotypic data*

164 Data were collected in three locations (E-PIRIO, D-MURO and E-MURO) in Corsica
165 (France, Figure 1, Table 1). Approximately six kilometers separate D-MURO from E-
166 MURO and 27 km separate these sites from E-PIRIO. The landscape of these sites is
167 dominated either by the evergreen holm oak *Quercus ilex* (E-PIRIO and E-MURO) or by
168 the deciduous downy oak *Q. pubescens* (D-MURO). All sites were monitored as part of
169 a long-term research programme for 20 (E-Muro), 24 (D-Muro) and 42 years (E-Pirio)
170 until 2017. A total of 360 nest-boxes were monitored and 26,650 birds were ringed.
171 Capture and handling of the birds was conducted under permits provided by the Centre
172 de Recherches sur la Biologie des Populations d'Oiseaux (CRBPO) and by the Direction
173 Départementale des Services Vétérinaires (DDSV).

174 Each year, nest boxes were visited at least once a week during the reproductive
175 period (from early April to late June). Breeding blue tits were captured in nest boxes
176 during the feeding of their young, and banded (if not already banded earlier) with a

177 unique metal ring provided by the CRBPO. Nestlings were also banded before fledging
178 at 9-15 days old.

179 Four phenotypic traits were measured on male and female breeders: tarsus
180 length (from the intertarsal joint to the most distal undivided scute on the tarsometatarsus),
181 flattened wing length, bill length (from the anterior end of the nares to the tip of the upper
182 mandible), and body mass. Phenotypic differences among individuals from the different
183 sites and used in genomic heritability estimates are presented in Table 1 and
184 Supplementary material 1 (see (Charmantier *et al.* 2016) for a review of phenotypic
185 divergence in these sites). Five to 20µl of blood were sampled from a small neck vein or
186 from a wing vein from adult breeders for later DNA extraction. Blood was stored at 4°C
187 in Queen's buffer (Seutin *et al.* 1991).

188

189 *DNA extraction, RAD-sequencing, SNP calling and data filtering*

190 We selected a set of 494 individuals captured between the years 2010-2016, from which
191 we extracted DNA from collected blood. The individuals were captured from three sites
192 (E-PIRIO, n = 185; D-MURO, n = 199; E-MURO, n = 110; Table 1). Individuals were
193 chosen according to the presence of several phenotypic measurements and a blood
194 sample. DNA extraction was achieved using Qiagen DNeasy Blood & Tissue kits. DNA
195 extractions were randomized across sites. DNA was quantified using first a NanoDrop
196 ND8000 spectrophotometer and then a Qubit 2.0 fluorometer with the DNA HS assay kit
197 (Life Technologies). DNA quality was checked on agarose gels.

198 Library preparation using restriction-site-associated DNA sequencing (RAD-seq;
199 Baird *et al.* 2008) with the enzyme SbfI was done by MGX (CNRS, Montpellier). Each

200 individual was identified using a unique 6 nucleotide tag and individuals were
201 multiplexed in equimolar proportions by groups of 36 individuals. Each library was
202 sequenced on one of 20 lanes of an Illumina HiSeq 2000 (libraries also included blue tit
203 individuals from another site that was not analyzed in this study).

204 Raw sequences were inspected with *FastQC* (Andrews 2010) for quality controls.
205 Potential fragments of Illumina adapters were trimmed with *Cutadapt* (Martin 2011),
206 allowing for a 10% mismatch in the adapter sequence. Reads were filtered for overall
207 quality, demultiplexed and trimmed to 85bp using *process_radtags*, *from the Stacks*
208 *software pipeline V1.39* (Catchen et al. 2013), , allowing for one mismatch in the
209 barcode sequence. BWA-MEM 0.7.13 (Li & Durbin 2009) was used to map individual
210 sequences against the reference genome of the great tit (Laine et al. 2016) and to
211 produce *sam* files using default options. Samtools 0.1.19 (Li et al. 2009) was used to
212 build and sort *bam* files. Back in Stacks V1.39, we used *pstacks* to treat *bam* files, align
213 the reads into matching stacks, infer loci, and detect SNPs at each locus. We used a
214 minimum depth of coverage (m) of 5, the SNP model, and alpha = 0.05 (chi square
215 significance level required to call a heterozygote or homozygote). *cstacks* was used to
216 build the catalogue of loci using n = 3 (number of mismatches allowed between sample
217 loci when build the catalog). *sstacks* was used to match loci against the catalog.
218 *populations*, the last Stacks program used here, genotyped individuals. Loci were
219 retained if genotyped in at least 90% of individuals (all individuals from all sites
220 grouped), with heterozygosity per site ≤ 0.60 , and with individual minimal read depth of
221 10 (“na” replaced genotypes below a read depth of 10). *VCFtools* (Danecek et al. 2011)
222 was used for further filtering of loci for a minimum average read depth of 20 across all

223 genotypes, and a maximum average read depth of 100 across all genotypes. Individuals
224 were genotyped for at least 90% of all loci; a filter would have otherwise been
225 implemented to remove individuals with low genotyping rate. The dataset was filtered to
226 retain loci with minimum allele frequency larger than 5% ($MAF \geq 0.05$). Subsequently the
227 dataset was pruned for linkage disequilibrium with the plink command *LD indep 50 5 2*.

228

229 *Population genetic structure and genome wide relatedness matrix*

230 Data were converted to several formats using the R package *radiator* (Gosselin 2017).
231 Expected heterozygosity per site and *Fst* between sites were estimated using *Genodive*
232 (Meirmans & van Tienderen 2004). A PCA resolving genetic distance between
233 individuals was calculated using the R packages *gdsfmt* and *SNPRelate* (Zheng et al.
234 2012).

235 Genome wide relatedness matrices (GRM) were computed with GCTA (Yang et
236 al. 2010; 2011a) using the markers found on autosomes (i.e. excluding markers from
237 sex chromosomes). We estimated GRMs for each of the three sites and for the three
238 sites pooled. The GRM based on the three sites pooled was represented using a
239 heatmap and histograms in order to depict relatedness diversity in the different sites
240 (Figure 2). Finally, we computed several other GRMs for the unique purpose of
241 chromosome partitioning (detailed in the corresponding section).

242

243 *Social pedigree*

244 Social pedigrees (PED-social) were constructed using the *pedantics* R-package
245 (Morrissey & Wilson 2009), on the basis of the complete pedigree generated from the

246 whole long-term monitoring period. We first included all ringed individuals and assigned
247 their mother and father based on observational data at capture. Unknown parents of a
248 given nest were coded using a dummy identity in the PED-social to preserve sibship
249 information. PED-social were pruned for each dataset studied (each population
250 separately, and the combined pedigree of the three sites pooled) to retain only
251 genotyped individuals and their ancestors (see supplementary material 2 for detailed
252 characteristics of the pedigree).

253

254 *Comparison between social pedigree and genome wide relatedness matrix and creation*
255 *of a corrected social pedigree*

256 We compared the PED-social and GRM for the 494 genotyped individuals mainly in
257 order to infer the potential presence of erroneous pedigree relationship (particularly
258 parent-offspring and sibling relationship). Indeed, errors may occur in social pedigrees,
259 essentially due to extra-pair paternities, but also observational errors such as misreading
260 of bird ring, successive egg-laying by different females in the same nest box, and brood
261 parasitism. Such errors, notably extra-pair paternities, can downwardly bias estimates of
262 heritability (Charmantier & Réale 2005). We illustrated the concordances and
263 discrepancies between PED-social and GRM using a biplot (Figure 2D). We then
264 created corrected social pedigrees (PED-corrected) on the basis of discrepancies found
265 between the PED-social and the GRM. For pairs of individuals having PED-social
266 relatedness values of 0.5 (full-sibs or parent-offspring relationships) but GRM
267 relatedness estimates below 0.1, we assumed that these relationships corresponded to
268 erroneous parent-offspring relationship due to extra-pair paternity and assigned them a

269 0 in PED-corrected. For PED-social relatedness values of 0.5 but GRM estimates
270 between 0.15 and 0.35, we assumed they corresponded to full-sibs being half-sibs and
271 assigned them a value of 0.25 in PED-corrected. We also corrected false half-sibs in the
272 pedigree (pairs of individuals having a 0.25 relatedness value in the PED-social but a
273 GRM estimates smaller than 0.1) and assigned them a 0 relatedness values. We
274 acknowledge that probably some more relationships could have been corrected based
275 on discrepancies between PED-social and GRM, but the distribution of GRM values
276 below 0.15 were largely overlapping and it would have been difficult to assign correct
277 PED-corrected values. Using a method such as the one implemented in the R package
278 *sequoia* (Huisman 2017) would certainly have allowed the correction of these potential
279 errors as well. Another potential limitation comes from the fact that ancestors of the
280 genotyped individuals in the pedigree were not genotyped and therefore erroneous
281 relationships between these individuals were impossible to correct.

282

283 *Inference of heritability*

284 We used both a frequentist and a Bayesian method to estimate heritability of the four
285 traits. The frequentist method was implemented in GCTA. The Bayesian framework was
286 implemented in the *MCMCglmm* R-package (Hadfield 2010). Bayesian inference is
287 renowned to have a clear advantage over the other existing methods since the use of
288 posterior distributions propagates the errors in estimates derived from animal models
289 (Morrissey et al., 2014).

290 GCTA was used to estimate heritability for each trait, separately for each of the
291 four GRMs. We here used best linear unbiased estimates (BLUPs) for each individual

292 and phenotype since we could not implement more complex models in GCTA. BLUPs
293 were estimated for the entire dataset of 494 genotyped individuals using a generalized
294 linear mixed model with the *MCMCglmm* function, integrating a random effect for the
295 observer and permanent environment effects accounting for multiple measurements of
296 the same individual. Differences between sites and sex were also accounted for by
297 adding these two factors as fixed effects. We then extracted the posterior mode of the
298 BLUPs for each individual, and ran GCTA for each of the four GRMs and the four traits.

299 *MCMCglmm* was used to estimate heritability for each trait, separately for each of
300 the four GRMs, the four PED-social and the four PED-corrected. Sex was integrated as
301 a fixed effect in all models. Site was integrated as a fixed effect in the models that
302 included all the individuals from the different sites. Measurer, identity and animal
303 (corresponding either to the GRM, PED-social or PED-corrected matrix) effects were
304 incorporated as random factors in order to partition the phenotypic variance into its
305 observer, permanent environment and additive genetic components. Running models
306 based on the GRMs, we used both BLUPs (to compare with GCTA) and alternatively all
307 phenotypic measurements. For the models using GRMs and BLUPs, the only random
308 factor was the animal effect (ie. the GRM) as permanent environmental and observer
309 effects were already taken into account in the BLUPs. For the models based on PED-
310 social and PED-corrected, we did not use BLUPs (since it could bias the analysis and/or
311 give anticonservative results, Hadfield et al. 2010) but instead used all phenotypic
312 measurements. We used identical parameters, priors and iterations for each estimate.
313 Random effects included additive genetic effects (V_A) estimated through the inclusion of
314 pedigree data, permanent environmental effects (V_{PE}) accounting for repeated

315 measurements of the same individual, measurer identity controlling for any potential
316 confounding measurer effect (V_{OBS}) and residual variance (V_R). The models used can be
317 described as follow:

318
$$Y = \mu + Xb + Z_A a + Z_{PE} pe + Z_O obs + e \text{ Equation 1}$$

319 Equation 1 describes the animal models run on phenotypic traits with Y the vector of
320 phenotypic observations for all individuals and μ the vector of mean phenotypes. Xb
321 stands for the fixed effects (containing sex and site for the models on all individuals, and
322 sex only for within site models). Z_A , Z_{PE} and Z_O correspond to the random factors:
323 additive genetic (a), permanent environment (pe), and measurer (obs) random effects,
324 respectively. e is the vector of residual errors. Posterior distributions were composed of
325 1,000 values per parameter. We used 120,000 iterations per model with sampling every
326 100 steps and with 20,000 discarded burn-in iterations. We used slightly informative
327 priors to facilitate convergence, with $V = V_P / (r + 1)$, $nu = 1$, V_P being the phenotypic
328 variance, and r the number of random factors (results were quantitatively and
329 qualitatively similar using uninformative priors, $V = 1$, $nu = 0.002$, but convergence takes
330 a bit longer). We checked the models graphically. We verified that autocorrelations were
331 less than 0.05. We finally reported posterior median, posterior mode, and 95% credible
332 interval.

333

334 *Inference of the effect of the number of SNPs on heritability estimates*

335 We inspected whether the number of SNPs included resulted in i) an increase in
336 heritability estimation precision (approximated via standard deviation, and LRT

337 (Likelihood Ratio Test) in the case of GCTA) and ii) an increase in heritability (median)
338 via saturation of the GRMs by SNPs in linkage disequilibrium with loci most likely
339 causative of the phenotypic variation (Gienapp et al. 2017). We reported the median and
340 standard deviation values of tarsus length heritability inferred using GCTA and
341 *MCMCglmm* for several GRMs produced using a variable number of SNPs. Using
342 GCTA, we analyzed 991 GRMs made up of a decreasing number of randomly chosen
343 SNPs from the entire dataset, by step of 50 SNPs from the total number of SNPs. To
344 confirm GCTA results, we analyzed a smaller number of 25 GRMs with *MCMCglmm*
345 (much less than for GCTA since the Bayesian analysis was much more time consuming
346 than the frequentist one) made from randomly chosen SNPs, concentrated mainly
347 between 1000 to 10,000 SNPs (since GCTA indicated that the rate of improvement in
348 estimates was particularly concentrated between these numbers).

349

350 *Inference of genetic correlations*

351 We used bivariate models in GCTA and in *MCMCglmm* in order to estimate genetic
352 correlations between each of the four traits. These bivariate models were achieved using
353 the same data, same fixed and random effects, and same number of iterations in the
354 case of *MCMCglmm*, as for the univariate models. For *MCMCglmm*, we used slightly

355 informative priors with $V = \begin{bmatrix} \frac{V_p}{r} & 0 \\ 0 & \frac{V_p}{r} \end{bmatrix}$ and $nu = 2$ with V_p being the phenotypic variance,
356 and r the number of random factors.

357

358 *Chromosome partitioning*

359 Finally, we partitioned heritability across the chromosomes using GCTA, using two
360 methods: i) fitting the univariate GCTA model simultaneously on each GRM
361 corresponding to each autosome; ii) fitting n times, where n is the number of autosomes,
362 the univariate GCTA model simultaneously on two GRMs, one GRM computed for the
363 focal chromosome and the other GRM computed using the other autosomes pooled. In
364 both cases (i) and (ii), several microchromosomes (22, 27, 28, 25LG1, 25LG2, LGE22)
365 were pooled into one artificial autosome because they had too few SNPs. When a model
366 did not converge, we discarded the smallest autosomes and ran the model again. We
367 did so until the model would converge.

368

369 **Results**

370 *Phenotypes*

371 Among the 494 genotyped individuals, 494, 493, 470, and 488 were measured for tarsus
372 length, wing length, bill length, and body mass, respectively (Table 1, Supplementary
373 material 1). The sex ratio of genotyped individuals was near 0.5, with 246 females out of
374 494 individuals. Phenotypic traits varied slightly between sexes and between sites,
375 justifying the inclusion of sex and site effects in the models partitioning phenotypic
376 variance.

377

378 *SNP calling and population genetic structure*

379 The median number of reads per individuals was 6,066,514. The median read depth per
380 individual and per locus was 56. The Stacks program *population* outputted 52,783 loci

381 totaling 96,009 SNPs. After MAF pruning, we retained 41,986 loci totaling 68,114 SNPs.
382 After LD pruning, we retained 38,030 loci totaling 49,682 SNPs. 47,865 of these SNPs
383 were on autosomes. The number of filtered SNPs per chromosome ranged from 5523
384 (chromosome 2) to 56 (LGE22).

385 The heterozygosity was 0.205, 0.205 and 0.204 in E-PIRIO, D-MURO and E-
386 MURO, respectively (Supplementary material 3). Genetic differentiation between sites
387 estimated using an *Fst* index was low yet significant, ranging from 0.006 to 0.008 (p-
388 value < 0.001; Supplementary material 3). The two first axes of the PCA (Figure 2A)
389 explaining each 0.98% and 0.93% of the genotypic variance, and the heatmap of the
390 GRM (Figure 2B), depicted the low genetic structure between the three sites.

391

392 *Genome wide relatedness matrix and social pedigree*

393 The GRM (Figure 2C) mostly included non-related (or distantly related) individuals.
394 Zooming in the histogram (Figure 2C) shows the presence of related individuals with
395 parent-offspring or full sib-like links (grey), and half-sib-like links (blue). Comparing the
396 relatedness values from the GRM to the one from the PED-social showed great
397 consistency between the two matrices (Figure 2D). However, as expected (and shown
398 elsewhere, Charmantier and Réale 2005), 112 links out of 858 known links presented
399 PED-social relatedness higher than GRM relatedness by at least 0.10. Among these, 39
400 links exhibited PED-social relatedness of 0.5 but GRM relatedness of 0, and 36 links
401 showed PED-social relatedness of 0.25 but GRM relatedness of 0. In addition, the GRM
402 allowed the identification of many links that cannot be documented in the PED-social
403 because of immigration, incomplete sampling or extra-pair paternity (Pemberton 2008).

404 Notably, we found 460 links with PED-social relatedness was equal to 0 but GRM
405 relatedness larger than 0.1. In total, we corrected 83 erroneous PED-social relatedness
406 values (corresponding to ca. 10% of the total number of links known) thanks to the GRM
407 analysis to create the PED-corrected.

408

409 *Heritability estimates*

410 In general, heritability measures based on PED-social, PED-corrected and GRM, using
411 BLUPs or all the measurements, and with *MCMCglmm* or GCTA, were highly consistent
412 for the same trait (Table 2, Figure 3; see supplementary material 4 for variance
413 estimates associated with the other random factors). The variations were very high
414 among the three localities, for which sample sizes were very small. When pooling the
415 three localities, variations between methods and credible intervals decreased.

416 Considering estimates from *MCMCglmm* and models fitting all phenotypic
417 measurements, there was a consistent trend for lower heritability when using PED-social
418 compared to GRM (Figure 3). For example, tarsus length and body mass heritability
419 were lower by 22% and 31%, respectively, when using PED-social compared GRM.
420 Specifically, there was a trend for lower heritability when using PED-social than PED-
421 corrected. For example, tarsus length and body mass heritability were lower by 7% and
422 11%, respectively, when using PED-social compared to PED-corrected. There was also
423 a trend for lower heritability when using PED-corrected compared to GRM. For example,
424 tarsus length and body mass heritability were lower by 15% and 22%, respectively,
425 when using PED-social compared to PED-corrected.

426 Heritability estimates were slightly higher when using BLUPs instead of all the
427 phenotypic measures, together with the GRM, in *MCMCglmm* (e.g. only 3% higher for
428 tarsus length, but 24% higher for body mass). Furthermore, credible intervals were
429 slightly higher when using BLUPs. Using BLUPs and the GRM in GCTA, heritability
430 estimates were relatively similar to the one obtained using *MCMCglmm* on all the
431 phenotypic measures and either the GRM or the PED-corrected.

432 The number of SNPs included to compute GRMs influenced the tarsus length
433 heritability estimates using either GCTA or BLUPs or *MCMCglmm* on all of the
434 phenotypic measures (Figure 4). For this trait, heritability sharply increased from a
435 handful of SNPs to approximately 5,000 SNPs. Heritability increased slowly from
436 approximately 5,000 to 15,000 SNPs and then reached a plateau along which a very
437 slow increase was however apparent. It is also worth mentioning a similarly shaped
438 increase of likelihood ratio tests from GCTA.

439

440 *Genetic correlations between traits*

441 In general, genetic correlations between traits calculated with the different methods,
442 were consistent for the same pair of trait (Table 3). However, the credible intervals were
443 generally large, with 13 of the 30 intervals including 0. The largest estimates and
444 smallest credible intervals were obtained for the genetic correlation between tarsus
445 length and body mass, the most heritable traits. For each pair of traits there was a trend
446 for a higher genetic correlation when using all measures than BLUPs.

447

448 *Chromosome partitioning*

449 When fitting the univariate GCTA model simultaneously on each autosome, the
450 correlation between heritability and chromosome length was positive for tarsus length (r^2
451 = 0.26; p-value = 0.009), being the most heritable traits of this study. The correlation was
452 non-significant for body mass (r^2 = 0.08; p-value = 0.20), wing length (r^2 = 0.03; p-value
453 = 0.41), and bill length (r^2 = 0.00; p-value = 0.85). Standard error intervals for
454 chromosome scaled heritability did not include zero only for chromosomes 1A and 3 for
455 tarsus length (Table 4, Figure 5), and for chromosome 6 for body mass. In the case of
456 tarsus length, no autosome had to be removed to enable the convergence of the model
457 fitting all the autosomes simultaneously. In contrast, for wing length, bill length and body
458 mass, the microchromosomes group had to be removed and respectively one, 12 and
459 three additional autosomes (by order of increasing size) had to be removed to enable
460 convergence of the models. When fitting the univariate GCTA model separately for each
461 autosome versus the rest of the genome, several heritabilities were likely overestimated
462 since the total heritability was larger than with the previous method (h^2 = 1.47 vs 0.84
463 for tarsus length, 0.85 vs 0.36 for wing length, 1.11 vs 0.31 for bill length, and 0.95 vs
464 0.62 for body mass).

465

466 **Discussion**

467 Genomic analyses offer great opportunities for measuring relatedness precisely among
468 individuals from natural populations and hence estimating heritability and genetic
469 correlations among traits in the wild (Edwards 2013; Gienapp et al. 2017). Heritability
470 partitioning along the genome can also help in identifying regions explaining the additive
471 genetic variation for polygenic traits (Yang et al. 2011b). Determining trait heritability and

its partitioning along the genome, and understanding trait genetic covariances, are also essential first steps to detect individual loci in the genome that contribute to trait differences between individuals, using for example genome wide association or linkage mapping (Schielzeth & Husby 2014). Although such approaches have been recently tested in a few datasets of well-studied wild organisms (Robinson et al. 2013; Stanton-Geddes et al. 2013; Bérénos et al. 2014), we require more explorations on whether smaller datasets could be used for such procedures and on how they compare with the classic animal model approach based on long-term pedigrees. In this context, we used both social pedigrees (PED-social) and genome wide relatedness matrices (GRM), to compute heritability estimates, genetic correlations, and chromosome partitioning for several phenotypic quantitative traits (tarsus length, wing length, bill length, and body mass) in a rather small dataset of individuals sampled in wild populations of Corsican blue tits and genotyped at nearly 50,000 RADseq derived SNPs. We discuss the power of such relatively small GRM, which may be typically produced while studying wild populations of non-model species, to examine the aforementioned quantitative genetic parameters.

Our main findings were i) a high congruence between heritability estimated using GRM, PED-corrected and PED-social. However, the median heritability obtained using the GRM was generally slightly higher than using PED-corrected and PED-social, which may be attributed to absence of erroneous links, greater precision and higher information content of the GRM. ii) The number of markers was not an issue for computing the GRM and estimating heritability, when approximately 15,000 SNPs were reached. iii) Genetic correlations among traits and chromosome partitioning were less

495 likely than heritability estimates to be very informative or robust given the large credible
496 intervals and the number of non-significant estimates, most likely due to power issues
497 (insufficient number of SNPs per chromosome in the case of the micro-chromosomes
498 and insufficient number of individuals). We thereafter discuss the usefulness and
499 limitations of inferences of quantitative genomic parameters, including heritability, in
500 relatively small sample sizes, with a medium number of genetic markers, for wild
501 populations of non-model species.

502

503 ***Comparing PED-social and GRM***

504 The GRM obtained differed in two aspects from the PED-social (Fig 2D): higher density
505 and relatedness variance on one hand, and no pedigree errors on the other hand. The
506 first major difference was that the GRM enabled estimating relatedness among
507 individuals that were not connected in PED-social (Figure 2D; 460 GRM relatedness
508 values were higher than PED-social relatedness values by at least 0.1). This concerned
509 individuals at every degree of relatedness but obviously primarily among individuals from
510 disconnected families. Therefore, GRM increased the depth and also the size of the
511 pedigree. This illustrates the usefulness of GRM in cases for which a dense and deep
512 pedigree is difficult to obtain due to for example dispersal, large population size
513 compared to the fieldwork capacity, and family structure (Pemberton 2008).
514 Furthermore, GRM allows capturing much more variance in realized relatedness among
515 individuals, both close and more distantly related (Figure 2D), which may also increase
516 the power of such GRM-based estimates in large and open populations.

517 The second major difference between PED-social and GRM was that several
518 individuals had much higher relatedness in PED-social than in the GRM, most likely
519 indicating errors in PED-social. The most common cases consisted of wrongly assigned
520 parentages (relatedness of 0.5 in PED-social and 0 in GRM), half sibs most likely
521 originating from extra-pair paternities (relatedness of 0.5 in PED-social and 0.25 in
522 GRM), and the spread of such errors in the pedigree (other smaller relatedness in GRM
523 than in PED-social). Using this comparison between GRM and PED-social, we created a
524 PED-corrected in which these aforementioned erroneous links, accounting for
525 approximately 10% of the PED-social links, were corrected. This proportion of erroneous
526 linked corrected in our dataset should be expected since 18 to 25% of young have been
527 previously identified as extra-pair offspring in these Corsican blue tit populations
528 (Charmantier et al. 2004a).

529

530 *Heritability estimates from PED-social, PED-corrected and GRM*

531 Heritability estimated for tarsus length (0.62 to 0.81 depending on the method, for three
532 sites pooled, see Table 2) and body mass (0.34-0.61) were on the upper limit compared
533 to previous estimates in this species and for other close species (Jensen et al. 2008;
534 Husby et al. 2011; Postma 2014; Delahaie et al. 2017) while heritability for wing length
535 (0.26-0.34) and bill length (0.14-0.21) were rather on the lower ranges usually found.
536 Here again it should be noted that we did not optimize the animal models by integrating
537 factors that may have contributed to explain trait variation, for example to account for
538 daily and seasonal variation in body mass. Moreover, it has been observed that the
539 number of individuals included can affect the heritability estimates (Silva et al. 2017).

540 Credible intervals and standard errors for heritability estimates were much larger
541 than in previous studies based on larger sample size (Robinson et al. 2013; Stanton-
542 Geddes et al. 2013; Bérénos et al. 2014; Silva et al. 2017). The credible intervals per
543 site were larger than for the three sites pooled, suggesting the decrease of power with
544 decreasing sample size. Similarly, Silva et al. (2017) obtained increased standard errors
545 when decreasing sample size. This most likely suggests that the number of individuals
546 used here conferred limited power.

547 Overall, when using *MCMCglmm* and models using all of the phenotypic
548 measurements, heritability estimates were relatively similar for GRM, PED-corrected and
549 PED-social for the same trait (Figure 3). These results are in line with previous
550 comparisons realized with much larger sample size (Robinson et al. 2013; Bérénos et al.
551 2014). However, unlike these previous studies showing very small differences in
552 heritability based on PED-social and GRM, our estimates based on GRM were larger
553 than using PED-social and to a lesser extent, than using PED-corrected, for tarsus
554 length and for body mass, but not for wing length and bill length. Here, wing length
555 heritability was 0.28 using PED-social, 0.26 using PED-corrected and 0.26 using GRM
556 (Table 2). In turn, tarsus length heritability obtained from GRM was 18% higher than
557 from PED-corrected and the one from PED-corrected 8% higher than from PED-social.
558 Similarly, body mass heritability obtained from GRM was 29% higher than from PED-
559 corrected and the one from PED-corrected 12% higher than from PED-social. The
560 difference between PED-corrected and PED-social for these two traits may be
561 attributable to the erroneous links in PED-social. Indeed, it has been estimated that
562 PED-social containing 5 to 20% of extra-pair paternities result in underestimated

563 heritability by up to 17%, in blue tits (Charmantier & Réale 2005). Then, the differences
564 observed between heritability estimated using GRM and PED-corrected are likely to
565 originate from higher density of the GRM and the fact that GRM incorporate more
566 variance in relatedness for a given class of kinship. Further changing the pedigree by
567 not only correcting erroneous existing links but also informing previously unidentified
568 links (eg. linking previously unknown extra-pair fathers to their offspring) based on the
569 genetic data, could also decrease the heritability difference between PED-corrected and
570 GRM.

571 Heritability estimates, and their credible intervals, based on BLUPs were slightly
572 higher compared to estimates based on models using all phenotypic measurements
573 when using *MCMCglmm* (this was not true for GCTA). This is not surprising given that
574 BLUPs are known to artificially increase the precision of the measure and therefore
575 increase statistical significance (Hadfield et al. 2010). While the difference was low for
576 tarsus and bill length (3% and 5% higher, respectively), the difference was particularly
577 pronounced for wing length and body mass (31% and 24%, respectively). One
578 speculation is that wing length and body mass are less repeatable than tarsus length
579 and bill length and that our BLUPs did not adequately take into account individual
580 variance.

581

582 ***Effect of the number of SNPs on the heritability estimates based on GRM***

583 The number of SNPs had a relatively minor effect on heritability estimates, from
584 approximately 15,000 SNPs (Figure 4). This is concordant with what has been shown in
585 previous studies (Stanton-Geddes et al. 2013; Bérénos et al. 2014). Of course, the effect

586 of the number of SNPs on the heritability estimation depends on several parameters
587 (e.g. genome size, LD across the genome, LD among markers, genetic architecture of
588 the focal trait, polymorphism of the markers, genotyping quality, missing data) and is not
589 transposable from one study to another. These days, the number of markers may no
590 longer be a problem, given the rise of the accessibility of many genomic tools. However,
591 the genotyping quality and missing data occurrence may be challenging for studies
592 using RADseq and GBS methods, as noted by (Gienapp et al. 2017). Here, we
593 prioritized the quantity and the quality of the RAD markers over the number of
594 individuals, as well as genotyping only individuals with high quality DNA, ending up with
595 a relatively high read depth and low missing rate. Future studies aiming to use RADseq
596 derived markers to estimate heritability should also find a tradeoff between the number
597 of genomic markers and the number of genotyped individuals. A formal simulation study
598 determining such a tradeoff between numbers of individuals and of markers and markers
599 read-depth (directly linked to quality) would be highly valuable for guiding the design of
600 RADseq analyses in the context of quantitative genomics.

601

602 *Genetic correlations using PED-social, PED-corrected and GRM*

603 Genetic correlations obtained using GRM, PED-social and PED-corrected had large
604 credible intervals, preventing us to make robust interpretations, except that our sample
605 size was probably too small to obtain acceptable credible intervals. This was partly
606 expected since bivariate models are known to be data hungry. While such lack of power
607 is a recurrent issue when estimating genetic correlations even with much larger sample
608 sizes (Vattikuti et al. 2012; Visscher et al. 2014; Ni et al. 2017), it was nevertheless a

609 disappointment while comparing these results to the one from (Bérénos et al. 2014), in
610 which the credible intervals were relatively small. The median values of genetic
611 correlations for these traits were in line with previous studies in other bird species
612 (Teplitsky et al. 2014).

613

614 *Chromosome partitioning based on GRM*

615 Chromosome partitioning of heritability of tarsus length was on par with what has been
616 observed in previous studies, *i.e.* increasing heritability of quantitative traits with
617 increasing number of genes per chromosome (Yang et al. 2011b; Robinson et al. 2013;
618 Santure et al. 2013, 2015; Bérénos et al. 2015; Wenzel et al. 2015; Silva et al. 2017).
619 Similarly to several of these previous studies, there were few chromosomes for which
620 standard error did not include zero. Of particular interest, chromosomes 3 and 1A
621 showed high heritability for tarsus length. These chromosomes may be of particular
622 interest for future studies investigating the genomic bases of this quantitative trait
623 variation in blue tits. Chromosome 3 was already identified in a recent study on
624 morphology of Dutch and British great tits (Santure et al. 2015) as explaining high
625 heritability for tarsus length. In contrast, chromosome 1A explained a low proportion of
626 genetic variance in this last study. These kind of discrepancies might be congruent with
627 the fact that this same study (Santure et al. 2015) also showed very little consistency
628 between the variance explained by each chromosome for great tit populations from the
629 Netherlands and from the United-Kingdom.

630 Regarding the three other traits analyzed, the chromosome partitioning results
631 appeared much less robust since several chromosomes had to be removed (7, 18, and

632 9 autosomes removed, for wing length, bill length and body mass, respectively) to
633 enable models' convergence. The need for such chromosomes removals has previously
634 been reported (Wenzel et al. 2015; Silva et al. 2017) and has received detailed
635 consideration by Kemppainen & Husby (2018). In addition, the relationship between the
636 number of genes and the heritability explained by chromosomes were deviating from
637 theoretical expectations of increasing heritability with increasing chromosome size for
638 these polygenic traits. For comparison, Bosse et al. (2017) showed a significant increase
639 of the proportion of additive genetic variance of bill length explained by a chromosome in
640 relation to chromosome size in great tits. Such deviation may not have been affected by
641 removing chromosomes (Kemppainen & Husby 2018) but most likely originated from a
642 limited power of the small sample size, alongside relatively small heritabilities
643 (Kemppainen & Husby 2018).

644

645 *Conclusion on the usefulness of GRM based heritability measures using small samples*
646 *from wild populations*

647 Overall, our study reveals that RADseq data on around 50k SNPs (as advised by
648 Berenos et al 2014) for around 500 phenotyped individuals can provide estimates of
649 heritability that are close to, and probably more accurate than, estimates based on social
650 pedigree data from seven years monitoring for more than 1600 individuals (as included
651 in the PED-social used here). This opens very interesting avenues in the field of
652 quantitative genetics since estimating heritability and genetic correlations in the wild
653 have long been restricted to study systems where long-term monitoring is feasible.
654 RADseq data could allow the estimation of quantitative genetic parameters for a much

lower cost in terms of time (years) spent in the field, although individual capture and phenotyping will most likely remain time consuming. Using GRM on few individuals gives the power to estimate such parameters on virtually any species even though no long-term pedigree is available (Gienapp 2017). This will notably allow moving on from individual study to more comparative approaches and answer important questions such as understanding the spatial variation of evolutionary potential or the role of evolutionary constraints in phenotypic evolution. Such a genomic approach also provides the possibility to explore genetic covariance between traits as well as chromosome partitioning to confirm the polygenic architecture for phenotypic traits classically measured in birds. Less than 500 individuals appears however to provide insufficient power to correctly estimate either genetic covariances between traits or contribution of individual chromosomes to overall heritability.

667 **References**

- 668 Andrews S (2010) FastQC: A quality control tool for high throughput sequence data.
669 Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- 670 Baird NA, Etter PD, Atwood TS et al. (2008) Rapid SNP Discovery and Genetic Mapping
671 Using Sequenced RAD Markers. PLoS ONE 3(10): e3376.
672 <https://doi.org/10.1371/journal.pone.0003376>.
- 673 Bérénos C, Ellis PA, Pilkington JG, Pemberton JM (2014) Estimating quantitative
674 genetic parameters in wild populations: a comparison of pedigree and genomic
675 approaches. Molecular Ecology, 23, 3434–3451.
- 676 Bérénos C, Ellis PA, Pilkington JG et al. (2015) Heterogeneity of genetic architecture of
677 body size traits in a free-living population. Molecular Ecology, 24, 1810–1830.
- 678 Blondel J, Thomas DW, Charmantier A et al. (2006) A Thirty-Year Study of Phenotypic
679 and Genetic Variation of Blue Tits in Mediterranean Habitat Mosaics. Bioscience, 56,
680 661–673.
- 681 Bosse, M., Spurgin, L. G., Laine, V. N., Cole, E. F., Firth, J. A., Gienapp, P., ... &
682 Groenen, M. A. (2017). Recent natural selection causes adaptive evolution of an
683 avian polygenic trait. Science, 358, 365-368.
- 684 Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an
685 analysis tool set for population genomics. Molecular Ecology, 22, 3124–3140.
- 686 Charmantier A, Réale D (2005) How do misassigned paternities affect the estimation of
687 heritability in the wild? Molecular Ecology, 14, 2839–2850.
- 688 Charmantier A, Blondel J, Perret P, Lambrechts M (2004a) Do extra-pair paternities

- 689 provide genetic benefits for female blue tits *Parus caeruleus*? Journal of Avian
690 Biology, 35, 1–9.
- 691 Charmantier A, Doutrelant C, Dubuc-Messier G, Fargevieille A, Szulkin M (2016)
692 Mediterranean blue tits as a case study of local adaptation. Evolutionary
693 Applications, 9, 135–152.
- 694 Charmantier A, Kruuk LEB, Blondel J, Lambrechts MM (2004b) Testing for
695 microevolution in body size in three blue tit populations. Journal of Evolutionary
696 Biology, 17, 732–743.
- 697 Csillary K (2006) Performance of Marker-Based Relatedness Estimators in Natural
698 Populations of Outbred Vertebrates. Genetics, 173, 2091–2101.
- 699 Danecek P, Auton A, Abecasis G et al. (2011) The variant call format and VCFtools.
700 Bioinformatics, 27, 2156–2158.
- 701 Delahaie, B., Charmantier, A., Chantepie, S., Garant, D., Porlier, M., & Teplitsky, C.
702 (2017). Conserved G-matrices of morphological and life-history traits among
703 continental and island blue tit populations. Heredity, 119, 76–87.
- 704 Edwards SV (2013) Next-generation QTL mapping: crowdsourcing SNPs, without
705 pedigrees. Molecular Ecology, 22, 3885–3887.
- 706 Gagnaire, P. A., & Gaggiotti, O. E. (2016). Detecting polygenic selection in marine
707 populations by combining population genomics and quantitative genetics
708 approaches. Current Zoology, 62, 603–616.
- 709 Gienapp, P., Fior, S., Guillaume, F., Lasky, J. R., Sork, V. L., & Csilléry, K. (2017).
710 Genomic quantitative genetics to study evolution in the wild. Trends in ecology &

- 711 evolution, 32, 897–908.
- 712 Gosselin T (2017). Radiator: RADseq Data Exploration, Manipulation and Visualization
713 using R. doi: 10.5281/zenodo.154432.
- 714 Hadfield JD, Wilson AJ, Garant D, Sheldon BC, Kruuk LEB (2010) The Misuse of BLUP
715 in Ecology and Evolution. The American Naturalist, 175, 116–125.
- 716 Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed
717 models: the MCMCglmm R package. Journal of Statistical Software, 33, 1-22.
- 718 Huisman J (2017) Pedigree reconstruction from SNP data: parentage assignment,
719 sibship clustering and beyond. Molecular Ecology Resources, 17, 1009–1024.
- 720 Husby A, Hille SM, Visser ME (2011) Testing Mechanisms of Bergmann’s Rule:
721 Phenotypic Decline but No Genetic Change in Body Size in Three Passerine Bird
722 Populations. American Naturalist, 178, 202–213.
- 723 Jensen H, Steinsland I, Ringsby TH, Sæther B-E (2008) Evolutionary dynamics of a
724 sexual ornament in the house sparrow (*Passer domesticus*): the role of indirect
725 selection within and between sexes. Evolution, 62, 1275–1293.
- 726 Kemppainen, P., & Husby, A. (2018). Inference of genetic architecture from
727 chromosome partitioning analyses is sensitive to genome variation, sample size,
728 heritability and effect size distribution. Molecular Ecology Resources.
729 <https://doi.org/10.1111/1755-0998.12774>.
- 730 Kvist L, Viiri K, Dias PC, Rytkönen S, Orell M (2004) Glacial history and colonization of
731 Europe by the blue tit *Parus caeruleus*. Journal of Avian Biology, 35, 352–359.
- 732 Laine VN, Gossman TI, Schachtschneider KM et al. (2016) Evolutionary signals of
733 selection on cognition from the great tit genome and methylome. Nature

- 734 Communications, 7, 1–9.
- 735 Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler
736 transform. Bioinformatics, 25, 1754–1760.
- 737 Li H, Handsaker B, Wysoker A et al. (2009) The Sequence Alignment/Map format and
738 SAMtools. Bioinformatics, 25, 2078–2079.
- 739 Martin M (2011) Cutadapt removes adapter sequences from high-throughput
740 sequencing reads. EMBnet, 17, 10-12.
- 741 Meirmans PG, van Tienderen PH (2004) genotype and genodive: two programs for the
742 analysis of genetic diversity of asexual organisms. Molecular Ecology Notes, 4, 792–
743 794.
- 744 Morrissey, M. B., de Villemereuil, P., Doligez, B., & Gimenez, O.. (2014). Bayesian
745 approaches to the quantitative genetic analysis of natural populations. In: A.
746 Charmantier, D. Garant, and L. E. B. Kruuk, eds. *Quantitative Genetics in the Wild*,
747 pp. 228–253. Oxford University Press, Oxford, UK.
- 748 Morrissey MB, Wilson AJ (2009) pedantics: an r package for pedigree-based genetic
749 simulation and pedigree manipulation, characterization and viewing. Molecular
750 Ecology Resources, 10, 711–719.
- 751 Ni, G., Moser, G., Wray, N. R., & Lee, S. H. (2017). Estimation of genetic correlation
752 using linkage disequilibrium score regression and genomic restricted maximum
753 likelihood. bioRxiv, 194019.
- 754 Pemberton JM (2008) Wild pedigrees: the way forward. Proceedings of the Royal
755 Society B-Biological Sciences, 275, 613–621.
- 756 Porlier, M., Garant, D., Perret, P., & Charmantier, A. (2012). Habitat-linked population

757 genetic differentiation in the blue tit *Cyanistes caeruleus*. Journal of Heredity, 103(6),
758 781-791.

759 Postma E (2014) Four decades of estimating heritabilities in wild vertebrate populations:
760 improved methods, more data, better estimates? In: A. Charmantier, D. Garant,
761 and L. E. B. Kruuk, eds. *Quantitative Genetics in the Wild*, pp. 16–33. Oxford
762 University Press, Oxford, UK.

763 Robinson MR, Santure AW, DeCauwer I, Sheldon BC, Slate J (2013) Partitioning of
764 genetic variation across the genome using multimarker methods in a wild bird
765 population. Molecular Ecology, 22, 3963–3980.

766 Santure, A. W., Poissant, J., De Cauwer, I., Oers, K., Robinson, M. R., Quinn, J. L., ... &
767 Slate, J. (2015). Replicated analysis of the genetic architecture of quantitative traits
768 in two wild great tit populations. Molecular ecology, 24, 6148-6162.

769 Santure, A. W., Cauwer, I., Robinson, M. R., Poissant, J., Sheldon, B. C., & Slate, J.
770 (2013). Genomic dissection of variation in clutch size and egg mass in a wild great tit
771 (*Parus major*) population. Molecular Ecology, 22, 3949-3962.

772 Schielzeth, H., & Husby, A. (2014). Challenges and prospects in genome-wide
773 quantitative trait loci mapping of standing genetic variation in natural
774 populations. *Annals of the New York Academy of Sciences*, 1320, 35-57.

775 Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for
776 DNA analyses. Canadian Journal of Zoology-Revue Canadienne De Zoologie, 69,
777 82–90.

778 Silva, C. N. S., McFarlane, S. E., Hagen, I. J., Rønnegård, L., Billing, A. M., Kvalnes, T.,

- 779 ... & Qvarnström, A. (2017). Insights into the genetic architecture of morphological
780 traits in two passerine bird species. *Heredity*, 119, 197–205.
- 781 Stanton-Geddes J, Yoder JB, Briskine R, Young ND, Tiffin P (2013) Estimating
782 heritability using genomic data (J Hadfield, Ed.). *Methods in Ecology and Evolution*,
783 4, 1151–1158.
- 784 Szulkin M, Gagnaire PA, Bierne N, Charmantier A (2016) Population genomic footprints
785 of fine-scale differentiation between habitats in Mediterranean blue tits. *Molecular
786 Ecology*, 25, 542–558.
- 787 Vattikuti S, Guo J, Chow CC (2012) Heritability and Genetic Correlations Explained by
788 Common SNPs for Metabolic Syndrome Traits (PM Visscher, Ed.). *Plos Genetics*, 8,
789 e1002637–8.
- 790 Visscher PM, Goddard ME (2015) A General Unified Framework to Assess the Sampling
791 Variance of Heritability Estimates Using Pedigree or Marker-Based Relationships.
792 *Genetics*, 199, 223–232.
- 793 Visscher PM, Hemani G, Vinkhuyzen AAE et al. (2014) Statistical Power to Detect
794 Genetic (Co)Variance of Complex Traits Using SNP Data in Unrelated Samples (GS
795 Barsh, Ed.). *Plos Genetics*, 10, e1004269.
- 796 Visscher PM, Hill WG, Wray NR (2008) Heritability in the genomics era — concepts and
797 misconceptions. *Nature Reviews Genetics*, 9, 255–266.
- 798 Wenzel MA, James MC, Douglas A, Piertney SB (2015) Genome-wide association and
799 genome partitioning reveal novel genomic regions underlying variation in
800 gastrointestinal nematode burden in a wild bird. *Molecular Ecology*, 24, 4175–4192.

- 801 Yang J, Benyamin B, McEvoy BP et al. (2010) Common SNPs explain a large proportion
802 of the heritability for human height. *Nature Genetics*, 42, 565–569.
- 803 Yang J, Lee SH, Goddard ME, Visscher PM (2011a) GCTA: a tool for genome-wide
804 complex trait analysis. *American Journal of Human Genetics*, 88, 76–82.
- 805 Yang J, Manolio TA, Pasquale LR et al. (2011b) Genome partitioning of genetic variation
806 for complex traits using common SNPs. *Nature Genetics*, 43, 519–525.
- 807 Zheng X, Levine D, Shen J et al. (2012) A high-performance computing toolset for
808 relatedness and principal component analysis of SNP data. *Bioinformatics*, 28,
809 3326–3328.
- 810

811 **Acknowledgments**

812 We thank all the people who helped to maintain the study sites and to conduct the blue
813 tit monitoring, in particular Christophe de Franceschi, Claire Doutrelant, Arnaud
814 Grégoire, Marcel Lambrechts, Jacques Blondel and Philippe Perret. We thank Pierre
815 Alexandre Gagnaire for useful advices on genomics. We thank three anonymous
816 reviewers and Anna Santure for precious comments on earlier versions of this
817 manuscript. This project was funded by the European Research Council (Starting grant
818 ERC-2013-StG-337365-SHE to AC) and a long-term funding from OSU-OREME.

819

820 **Data archiving**

821 Data (Plink files, PED-social, and repeated and BLUPs phenotypic traits values) are
822 available from the Dryad Digital Repository: doi:10.5061/dryad.k6r1mk8

823

824 **Conflict of interest**

825 The authors declare no conflict of interest.

826

827 **Author contributions**

828 Conception of the study: AC & CP. Field work: AC, BD & CP. Lab work & bioinformatics:
829 CP. Quantitative genetic/genomics inferences: CP & BD. Writing of the manuscript: CP.
830 Critically revised and improved the manuscript: AC, BD & CP.

831

832

833 **Figure captions**

834 Figure 1. A- Map of the sampling locations in Corsica, including a zoom-in of the area
835 where the populations are located B- D-Muro and E-Muro and C- E-Pirio (see table 1 for
836 details).

837

838 Figure 2. Genetic structure and relatedness among individuals. A- Biplot of the two first
839 axes of a principal component analysis of genetic distances between individuals, B-
840 Heatmap of the genome wide relatedness matrix (GRM) among the 494 genotyped
841 individuals, C- histogram of genome wide relatedness (GRM) among individuals, with a
842 zoom-in highlighting particularly high relatedness values, and D- Biplot of genome wide
843 relatedness (GRM) and social pedigree relatedness from the field observations (PED-
844 social). In B, C and D, color-coding is based on the empirical relatedness distribution
845 and is only indicative.

846

847 Figure 3. A- Heritability estimates using the GRM, the PED-social or the PED-corrected,
848 based on either BLUPs or the entire phenotypic measurements, and implementing
849 models in GCTA or *MCMCglmm*, for the 494 genotyped individuals from the three
850 populations pooled. B, C & D- Relationship between heritability inferred for the four traits
851 for PED-social, PED-corrected and GRM using *MCMCglmm*, using the model with all
852 phenotypic measurements in *MCMCglmm*. Errors bars correspond to 95% credible
853 interval for *MCMCglmm* and to 95% confidence intervals for GCTA.

854

855 Figure 4. Effect of the number of SNPs used to build the GRM on the tarsus length
856 heritability estimated with GCTA & BLUPs (A) and *MCMCglmm* & all the phenotypic
857 measures (B). Circles represent the median value. Bars represent standard errors for
858 GCTA and 95% credible interval for *MCMCglmm*. Black curves correspond to LOESS
859 fits. The log likelihood ratio of GCTA estimates is color-coded.

860

861 Figure 5. Chromosome partitioning. The univariate GCTA model was fitted
862 simultaneously on each autosome. Linear correlation between the number of genes per
863 chromosome and the GCTA estimates of heritability per chromosome for the four
864 phenotypic traits. Names of the largest chromosomes, and of the chromosomes having
865 the highest heritability, are given.

866

867 **Tables**

868 Table 1. Geographic coordinates, number of genotyped individuals (N), and trait mean values (\pm SE) for this sample, for
 869 each of the three sites and for the entire dataset. The phenotypic variance (V_p) is also given for each trait.

Site	Latitude	Longitude	N	Sex	Tarsus length (mm)				Wing length (mm)				Bill length (mm)				Body mass (g)			
					N	mean	SE	V_p	N	mean	SE	V_p	N	mean	SE	V_p	N	mean	SE	V_p
E-Pirio	42.376	8.750	185	male	98	16.4	0.5		98	63.5	1.6		90	6.4	0.3		97	9.4	0.4	
				female	87	15.9	0.6		87	60.6	1.2		82	6.6	0.4		86	9.2	0.4	
				both	185	16.1	0.6	0.25	185	62.2	2.0	4.91	172	6.5	0.4	0.16	183	9.3	0.4	
D-Muro	42.551	8.923	199	male	95	16.6	0.4		95	63.3	1.6		92	6.4	0.3		92	9.8	0.4	
				female	104	16.1	0.6		104	60.8	1.3		99	6.5	0.3		104	9.5	0.5	
				both	199	16.3	0.5	0.24	199	62.0	1.9	4.38	191	6.4	0.3	0.16	196	9.6	0.5	
E-Muro	42.589	8.967	110	male	55	16.5	0.5		55	63.4	1.5		53	6.4	0.3		54	9.7	0.4	
				female	55	16.0	0.4		54	60.8	1.1		54	6.7	0.5		55	9.5	0.4	
				both	110	16.3	0.5	0.25	109	62.1	1.8	3.84	107	6.6	0.4	0.19	109	9.6	0.4	
Total			494	male	248	16.5	0.4		248	63.4	1.6		235	6.4	0.3		243	9.6	0.4	
				female	246	16.0	0.5		245	60.7	1.2		235	6.6	0.4		245	9.4	0.5	
				both	494	16.2	0.5	0.26	493	62.1	1.9	4.43	470	6.5	0.3	0.17	488	9.5	0.5	

870 Table 2. Heritability estimates (h^2) and confidence (GCTA) / credible (*MCMCglmm*) intervals (95%CI) for the four traits,
 871 using the GRM, the PED-social or the PED-corrected, based on either BLUPs or the entire measures, and implementing
 872 models in GCTA or *MCMCglmm*, for all the individuals or each of the three sites separately. Number of individuals used
 873 (N).

Relatedness	Data	Implementation	Site	N	Tarsus length			Wing length			Bill length			Body mass						
					N	h^2	95%CI	N	h^2	95%CI	N	h^2	95%CI	N	h^2	95%CI				
GRM	BLUPs	GCTA	3 pop pooled	494	489	0.72	0.5	0.94	493	0.26	0	0.48	469	0.16	0	0.4	488	0.48	0.25	0.72
			E-Pirio	185	182	0.64	0.28	1	185	0.25	0	0.6	171	0.11	0	0.46	183	0.37	0	0.74
			D-Muro	199	197	0.73	0.38	1	199	0.46	0.04	0.87	191	0.22	0	0.64	196	0.41	0	0.88
			E-Muro	110	110	1	0.66	1	109	0.36	0	0.84	107	0.59	0	1	109	1	0.69	1
GRM	BLUPs	<i>MCMCglmm</i>	3 pop pooled	494	489	0.81	0.63	0.92	493	0.34	0.11	0.58	469	0.2	0.06	0.52	488	0.61	0.38	0.83
			E-Pirio	185	182	0.74	0.44	0.95	185	0.31	0.08	0.66	171	0.26	0.05	0.52	183	0.58	0.19	0.78
			D-Muro	199	197	0.8	0.42	0.96	199	0.49	0.19	0.91	191	0.3	0.71	0.07	196	0.58	0.11	0.86
			E-Muro	110	110	0.93	0.65	0.98	109	0.49	0.1	0.82	107	0.35	0.11	0.95	109	0.9	0.55	0.98
GRM	All measures	<i>MCMCglmm</i>	3 pop pooled	494	489	0.79	0.65	0.87	493	0.26	0.11	0.5	469	0.19	0.03	0.37	488	0.49	0.33	0.64
			E-Pirio	185	182	0.72	0.39	0.89	185	0.2	0.06	0.42	171	0.13	0.02	0.35	183	0.39	0.09	0.57
			D-Muro	199	197	0.84	0.54	0.89	199	0.48	0.18	0.73	191	0.29	0.06	0.49	196	0.17	0.04	0.57
			E-Muro	110	110	0.86	0.61	0.91	109	0.37	0.07	0.55	107	0.36	0.05	0.6	109	0.66	0.42	0.78
PED-corrected	All measures	<i>MCMCglmm</i>	3 pop pooled	494	489	0.67	0.49	0.8	493	0.26	0.13	0.41	469	0.21	0.06	0.3	488	0.38	0.24	0.48
			E-Pirio	185	182	0.54	0.22	0.78	185	0.12	0.04	0.3	171	0.14	0.03	0.27	183	0.25	0.07	0.41
			D-Muro	199	197	0.68	0.33	0.83	199	0.5	0.19	0.61	191	0.22	0.08	0.35	196	0.32	0.09	0.44
			E-Muro	110	110	0.74	0.41	0.84	109	0.28	0.06	0.4	107	0.18	0.05	0.43	109	0.85	0.43	0.95
PED-social	All measures	<i>MCMCglmm</i>	3 pop pooled	494	489	0.62	0.46	0.73	493	0.28	0.13	0.37	469	0.14	0.06	0.23	488	0.34	0.22	0.42
			E-Pirio	185	182	0.52	0.22	0.71	185	0.12	0.03	0.26	171	0.10	0.03	0.22	183	0.22	0.06	0.4
			D-Muro	199	197	0.6	0.27	0.72	199	0.37	0.16	0.52	191	0.09	0.03	0.19	196	0.29	0.11	0.41
			E-Muro	110	110	0.68	0.38	0.78	109	0.19	0.05	0.35	107	0.11	0.04	0.37	109	0.4	0.17	0.53

874

875 Table 3. Genetic correlations (r_G), and confidence (GCTA) / credible ($MCMCglmm$) intervals (95%CI), between traits
 876 estimated for the individuals from the three site pooled, using GRM, the PED-social or the PED-corrected, based on either
 877 BLUPs or all the measures, and implementing models in GCTA or $MCMCglmm$.

878

		Tarsus length	Tarsus length	Tarsus length	Wing length	Wing length	Bill length
		Wing length	Bill length	Body mass	Bill length	Body mass	Body mass
GCTA	r_G	0.40	0.37	0.48	-0.06	0.22	0.13
BLUPs	0.05	0.05	-0.14	0.23	-0.86	-0.022	-0.49
GRM	0.95	0.75	0.88	0.73	0.86	0.67	0.76
<i>MCMCglmm</i>	r_G	0.41	0.32	0.52	0.10	0.21	0.03
BLUPs	0.05	-0.08	-0.09	0.20	-0.57	-0.28	-0.50
GRM	0.95	0.67	0.73	0.75	0.48	0.66	0.59
<i>MCMCglmm</i>	r_G	0.48	0.45	0.56	0.50	0.50	0.33
all measures	0.05	0.16	0.06	0.37	-0.08	0.05	0.05
GRM	0.95	0.81	0.80	0.76	0.87	0.71	0.88
<i>MCMCglmm</i>	r_G	0.48	0.43	0.57	0.46	0.49	0.52
all measures	0.05	0.13	0.00	0.40	-0.12	0.06	0.05
PED-corrected	0.95	0.79	0.81	0.78	0.89	0.74	0.84
<i>MCMCglmm</i>	r_G	0.46	0.52	0.61	0.42	0.41	0.37
all measures	0.05	0.17	0.16	0.38	-0.13	0.09	0.08
PED-social	0.95	0.77	0.79	0.82	0.79	0.72	0.81

879 Table 4. Chromosome partitioning of the heritability estimated for the four phenotypic traits using GCTA. Partitioning was
880 achieved either i) fitting all the chromosomes simultaneously or ii) considering one chromosome versus all of the other
881 autosomes, iteratively for each chromosome. Several micro-chromosomes were grouped (22, 27, 28, 25LG1, 25LG2,
882 LGE22) and referred as “micros”. “na” refers to the sets of chromosomes that have been removed to enable the
883 convergence of the model fitting simultaneously all the chromosomes (see methods). The underlined values do not
884 include 0 in their standard error intervals.

885

Chrom osome name	SNPs	Genes	Tarsus				Wing length				Bill length				Body mass			
			All chromosomes simultaneously		One chromosome versus the other, iteratively		All chromosomes simultaneously		One chromosome versus the other, iteratively		All chromosomes simultaneously		One chromosome versus the other, iteratively		All chromosomes simultaneously		One chromosome versus the other, iteratively	
			Media	SE	Media	SE												
1	4784	1131	0	0.16	0	0.14	0.01	0.16	0.01	0.14	0	0.15	0	0.13	0.03	0.16	0.07	0.14
2	5523	1287	0.06	0.17	0.01	0.15	0	0.17	0	0.16	0	0.16	0	0.15	0	0.19	0	0.16
3	4383	1096	0.17	0.16	0.21	0.14	0	0.15	0	0.14	0.03	0.16	0.02	0.14	0.07	0.15	0.08	0.13
4	2869	736	0	0.13	0	0.11	0.05	0.13	0.16	0.12	0.06	0.13	0.12	0.12	0.01	0.13	0.04	0.11
4A	1428	343	0	0.09	0	0.08	0	0.1	0	0.09	0.05	0.1	0.11	0.09	0.01	0.09	0.01	0.08
1A	2638	853	0.21	0.13	0.24	0.11	0	0.12	0	0.11	0.13	0.13	0.17	0.12	0.04	0.12	0.06	0.11
5	2902	940	0.04	0.12	0	0.1	0	0.12	0	0.11	0	0.13	0	0.11	0.09	0.13	0.1	0.11
6	1818	513	0.04	0.11	0.11	0.1	0.01	0.1	0	0.09	0.01	0.1	0.07	0.09	0.12	0.11	0.17	0.1
7	1907	488	0.06	0.11	0.21	0.1	0.01	0.11	0.06	0.1	0	0.12	0	0.1	0.07	0.11	0.11	0.1
8	1718	492	0.04	0.1	0.11	0.09	0	0.1	0	0.09	0	0.11	0	0.1	0.01	0.1	0.04	0.1
9	1689	434	0	0.11	0	0.09	0.09	0.1	0.16	0.1	0.03	0.1	0.05	0.09	0.02	0.1	0.01	0.09
12	1573	343	0	0.1	0	0.08	0.04	0.09	0.04	0.08	0	0.1	0.01	0.08	0	0.08	0	0.09
13	1371	322	0.03	0.08	0.08	0.08	0.01	0.08	0.02	0.08	0	0.09	0.03	0.08	0	0.09	0.08	0.08
10	1177	394	0.07	0.09	0.1	0.08	0	0.08	0	0.07	na	na	0.14	0.08	0	0.1	0	0.07
11	1309	364	0.03	0.09	0.12	0.09	0.05	0.09	0.11	0.08	na	na	0.02	0.08	0.08	0.09	0.01	0.08
14	1364	396	0.05	0.09	0.08	0.08	0	0.08	0	0.08	na	na	0.04	0.08	0	0.09	0	0.08
15	1004	358	0	0.08	0	0.07	0	0.08	0	0.07	na	na	0	0.08	0.03	0.08	0.01	0.07
17	974	277	0.02	0.08	0.16	0.08	0	0.09	0	0.08	na	na	0.07	0.07	0.03	0.08	0.05	0.07
18	1114	318	0	0.09	0	0.08	0	0.09	0	0.08	na	na	0	0.08	0	0.08	0	0.07
19	849	311	0	0.07	0	0.06	0.05	0.08	0.1	0.07	na	na	0.01	0.06	0.01	0.08	0.1	0.08
21	632	260	0	0.07	0	0.06	0	0.07	0	0.06	na	na	0	0.07	0	0.07	0	0.07
24	846	175	0.02	0.08	0.04	0.07	0.02	0.07	0.08	0.07	na	na	0	0.07	0	0.08	0	0.07
20	1301	341	0	0.09	0	0.08	0	0.09	0	0.08	na	na	0.17	0.09	na	na	0.01	0.07
26	802	262	0	0.07	0	0.06	0.02	0.07	0.04	0.07	na	na	0.01	0.06	na	na	0	0.07
23	521	232	0	0.07	0	0.06	na	na	0	0.06	na	na	0.07	0.06	na	na	0	0.06
micros	1368	874	0	0.1	0.02	0.09	na	na	0.05	0.08	na	na	0	0.09	na	na	0	0.09

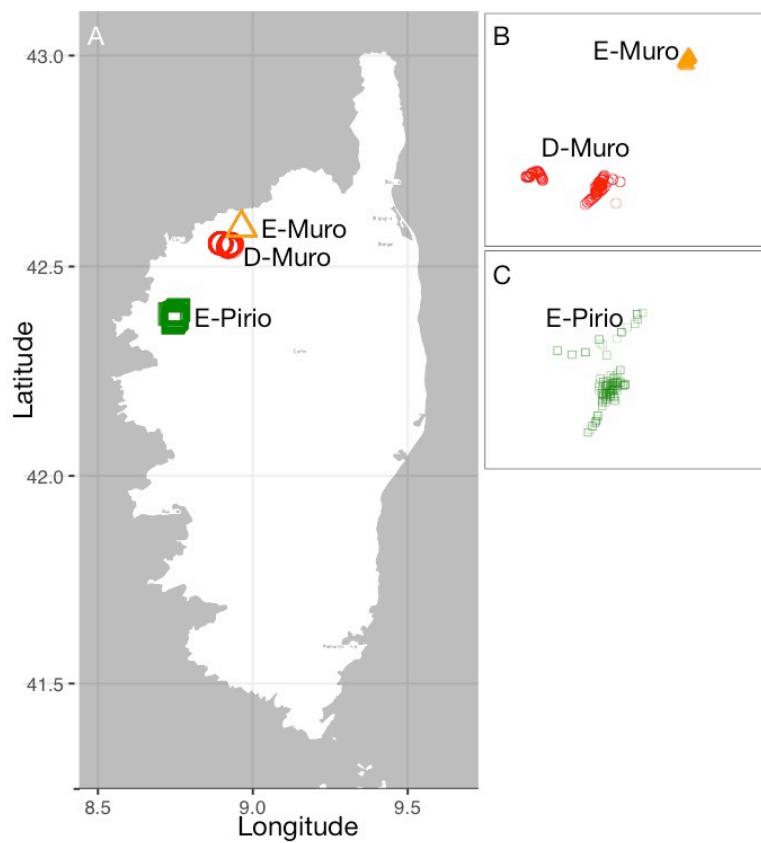


Figure 1.

893
894
895
896

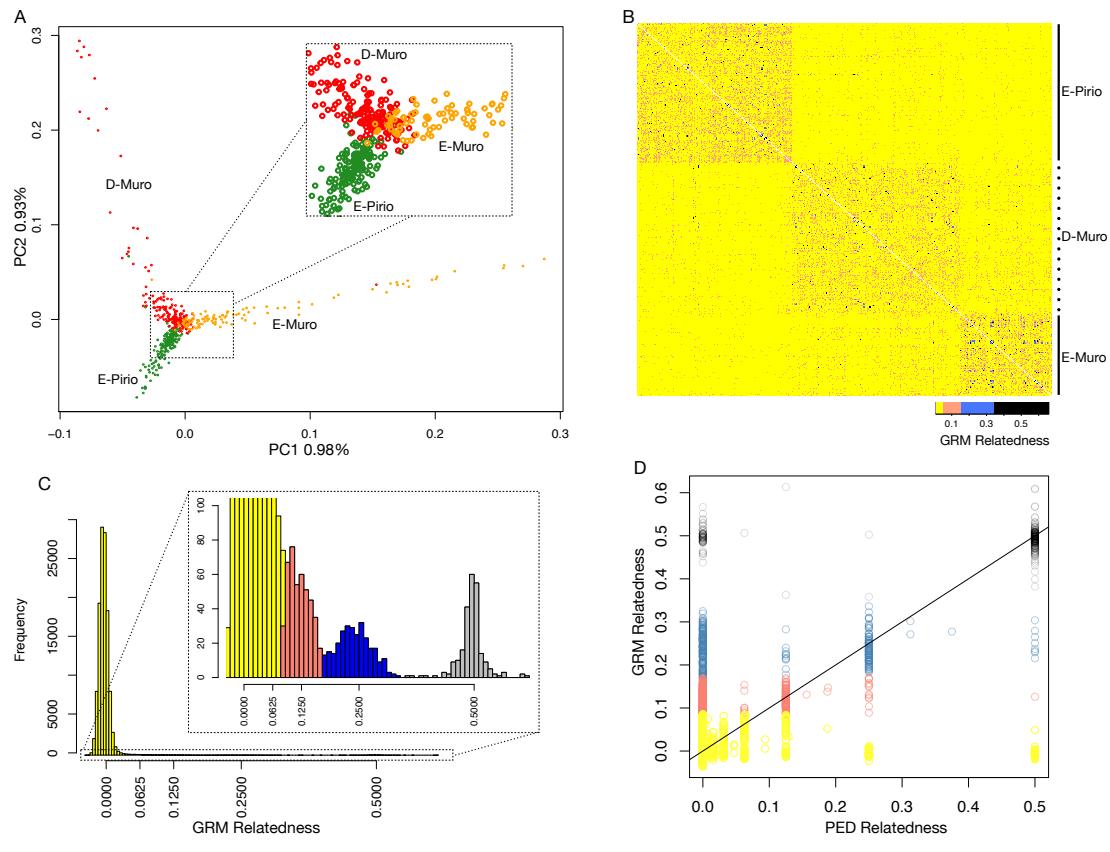
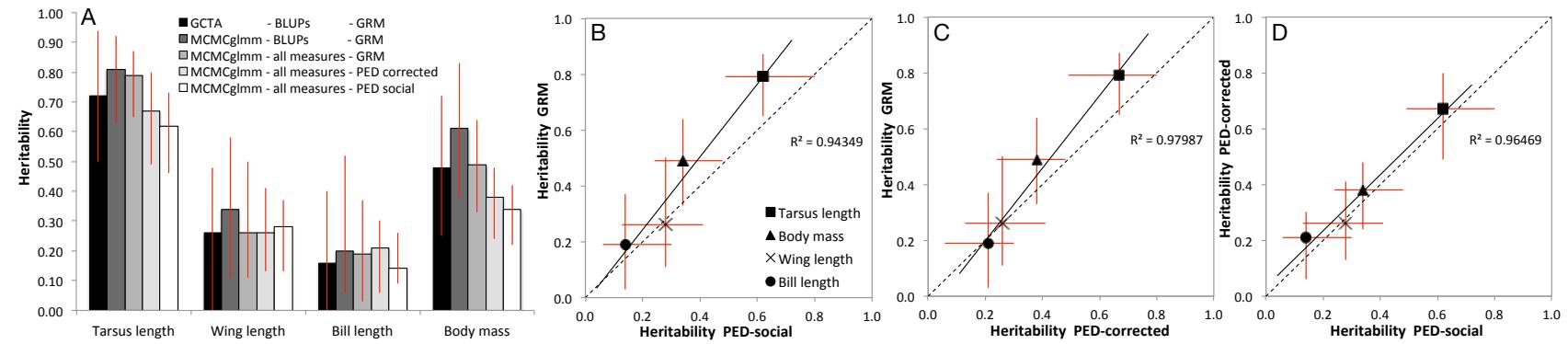


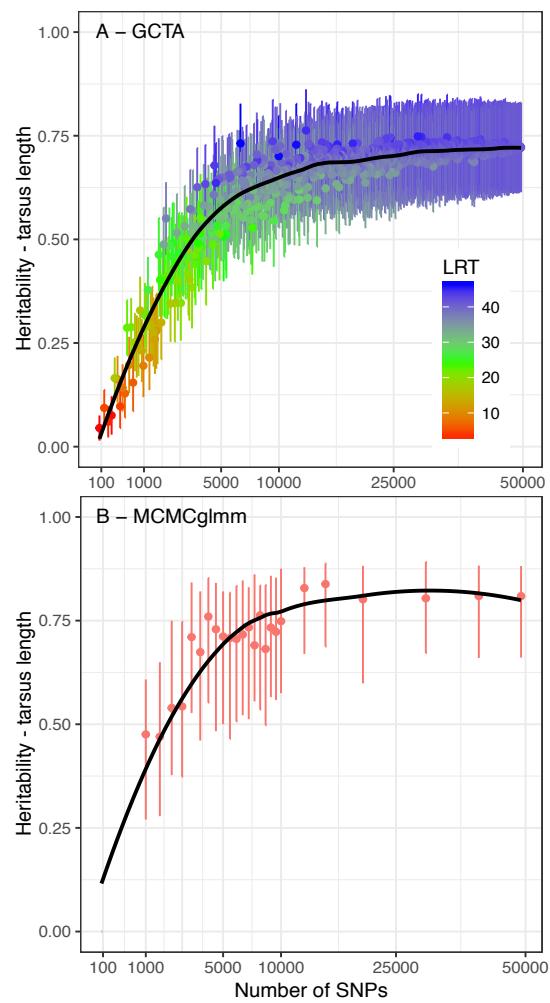
Figure 2.



897
898

Figure 3.

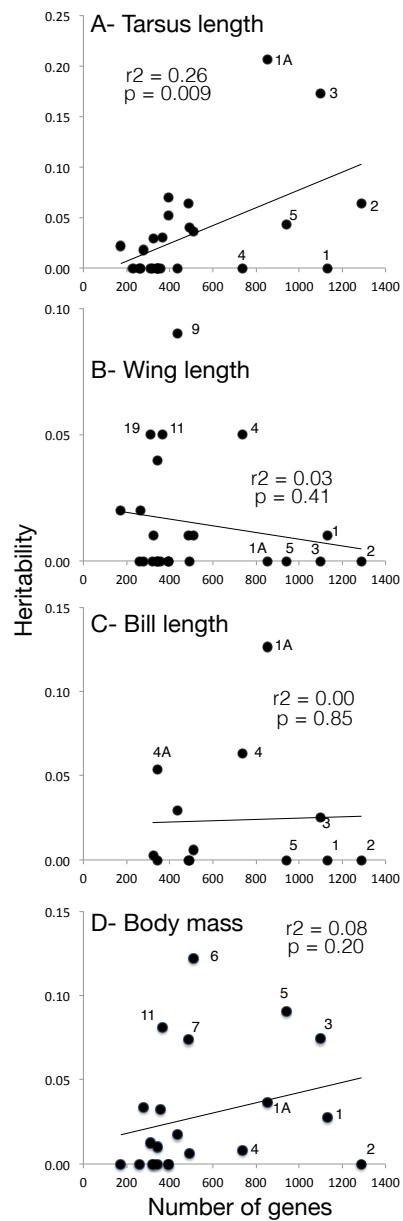
899



900

Figure 4.

901



903

904

Figure 5.

905

906 **Supplementary material**

907 Supplementary material 1. Box plots showing sex- and site-specific median and
908 variation in tarsus length, wing length, bill length and body mass for the 494
909 genotyped individuals. The first four letters code for the site and the last letter
910 codes for the sex (F: female; M: male).

911 Supplementary material 2. Social pedigree characteristics.

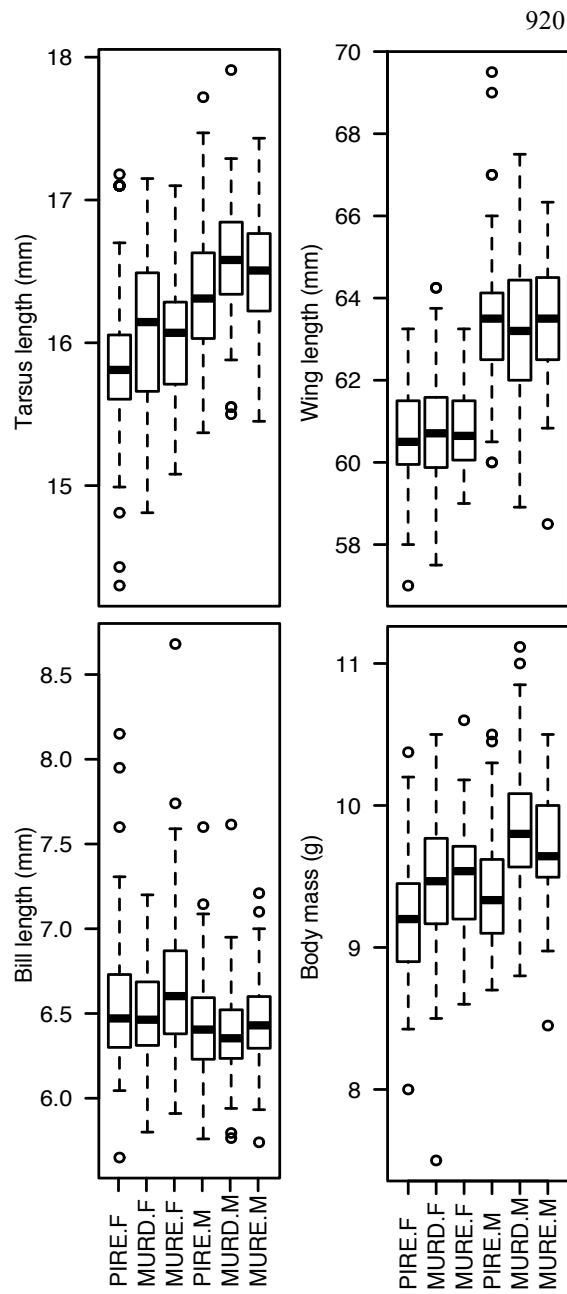
912 Supplementary material 3. Heterozygosity (H_e) in E-Piro, D-Muro and E-Muro
913 and FST among the three sites (FST above diagonal and p-value below the
914 diagonal).

915 Supplementary material 4. Variance estimates associated to the other random
916 factors.

917

918

919 Supplementary material 1.



921

922

923

924 Supplementary material 2.

	Maternities	Paternities	Full-sibs	Maternal half-sibs	Paternal half-sibs	Max Depth
Pooled sites	333	333	88	47	35	14
E-Piro	115	115	27	15	9	14
E-Muro	117	117	23	18	17	7
D-Muro	116	116	34	13	9	6

925

926

927

928

929 Supplementary material 3.

	F_{ST}			
	He	E-Piro	D-Muro	E-Muro
E-Piro	0.205	--	0.006	0.008
D-Muro	0.205	p < 0.001	--	0.007
E-Muro	0.204	p < 0.001	p < 0.001	--

930

931 **Supplementary material 4.**

932 Variance associated to the permanent environment effect.

		Tarsus				Wing				Bill				Mass			
		Vpe	95% CI	Vpe	95% CI	Vpe	95% CI	Vpe	95% CI	Vpe	95% CI	Vpe	95% CI	Vpe	95% CI		
	GRM	Pooled	0.04	0.01	0.07	0.65	0.4	1.09	0.03	0.01	0.05	0.02	0.01	0.06			
	GRM	E-Pirio	0.03	0.01	0.11	0.65	0.24	1.03	0.03	0.01	0.05	0.03	0.01	0.09			
	GRM	D-Muro	0.03	0.01	0.09	0.73	0.24	1.42	0.02	0	0.03	0.08	0.01	0.11			
	GRM	E-Muro	0.02	0.01	0.07	0.44	0.12	0.97	0.02	0.01	0.07	0.01	0	0.06			
	PED-corrected	Pooled	0.03	0.01	0.06	0.5	0.2	0.9	0.02	0.01	0.04	0.02	0.01	0.05			
	PED-corrected	E-Pirio	0.03	0.01	0.12	0.5	0.18	0.96	0.02	0	0.04	0.02	0.01	0.09			
	PED-corrected	D-Muro	0.02	0.01	0.1	0.38	0.12	1.1	0.01	0	0.03	0.02	0.01	0.08			
	PED-corrected	E-Muro	0.02	0.01	0.08	0.45	0.12	0.88	0.02	0.01	0.07	0.02	0.01	0.07			

933

934 Variance associated to the observer effect.

		Tarsus				Wing				Bill				Mass			
		Vobs	95% CI	Vobs	95% CI	Vobs	95% CI	Vobs	95% CI	Vobs	95% CI	Vobs	95% CI	Vobs	95% CI		
	GRM	Pooled	0.01	0.01	0.02	0.23	0.12	0.52	0.05	0.02	0.1	0.01	0.01	0.03			
	GRM	E-Pirio	0.01	0	0.03	0.21	0.08	0.46	0.02	0.01	0.06	0.01	0	0.03			
	GRM	D-Muro	0.01	0.01	0.04	0.36	0.15	1.05	0.07	0.03	0.16	0.02	0.01	0.05			
	GRM	E-Muro	0.01	0.01	0.04	0.26	0.11	0.72	0.04	0.01	0.09	0.02	0.01	0.08			
	PED-corrected	Pooled	0.01	0.01	0.02	0.24	0.11	0.53	0.05	0.02	0.1	0.02	0.01	0.04			
	PED-corrected	E-Pirio	0.01	0.01	0.03	0.21	0.08	0.45	0.02	0.01	0.06	0.01	0	0.03			
	PED-corrected	D-Muro	0.01	0.01	0.04	0.35	0.15	0.97	0.07	0.03	0.17	0.01	0.01	0.05			
	PED-corrected	E-Muro	0.01	0.01	0.04	0.33	0.09	0.8	0.04	0.02	0.09	0.03	0.01	0.08			

935

936 Variance associated to the residuals.

		Tarsus				Wing				Bill				Mass			
		Vresiduals	95% CI														
	GRM	Pooled	0.03	0.03	0.04	1.22	1.14	1.35	0.1	0.08	0.1	0.13	0.12	0.14			
	GRM	E-Pirio	0.03	0.03	0.04	1.68	1.46	1.95	0.11	0.09	0.13	0.14	0.13	0.17			
	GRM	D-Muro	0.03	0.03	0.04	0.9	0.77	1.04	0.06	0.05	0.07	0.14	0.12	0.16			
	GRM	E-Muro	0.02	0.02	0.03	1.12	0.97	1.39	0.09	0.07	0.11	0.09	0.07	0.1			
	PED-corrected	Pooled	0.03	0.03	0.04	1.26	1.13	1.35	0.09	0.08	0.1	0.13	0.12	0.14			
	PED-corrected	E-Pirio	0.03	0.03	0.04	1.62	1.42	1.93	0.11	0.09	0.13	0.14	0.13	0.17			
	PED-corrected	D-Muro	0.03	0.03	0.04	0.88	0.77	1.03	0.06	0.05	0.07	0.13	0.12	0.16			
	PED-corrected	E-Muro	0.02	0.02	0.03	1.18	1	1.39	0.09	0.07	0.11	0.09	0.07	0.11			

937

938

939