Received: 10 December 2017 | Revised: 26 March 2018 | Accepted: 29 March 2018

DOI: 10.1111/1755-0998.12886

SPECIAL ISSUE: ASSOCIATION MAPPING IN NATURAL POPULATIONS

WILEY MOLECULAR ECOLOGY RESOURCES

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

C. Perrier | B. Delahaie | A. Charmantier

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

Correspondence
Charles Perrier, Centre d'Ecologie Fonctionnelle et Evolutive, CNRS-UMR5175 CEFE, Montpellier, France.
Email: charles.perrier@cefe.cnrs.fr

Funding information
ERC, Grant/Award Number: ERC-2013-StG-337365-SHE; OSU-OREME

Abstract

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

KEYWORDS
blue tit, *Cyanistes caeruleus*, genetic correlation, genomewide relatedness matrices, heritability, partitioning, pedigree, phenotype, RAD-sequencing, SNP

1 | INTRODUCTION

Estimating additive genetic variance and heritability of quantitative traits is a major aim in evolutionary biology because they are crucial components of all evolutionary models. Further, fine-scale dissection of the genomic basis of quantitative traits in the wild, using methods such as genomewide association and linkage mapping, requires validation that the traits being mapped are actually heritable. New sequencing technologies have empowered the estimation of genomic relatedness matrices (GRM) between individuals (Gienapp et al.,

2017; Huisman, 2017; Visscher, Hill, & Wray, 2008), feasibly without any social pedigree (PED-social) recorded in the field. Since in a quantitative framework, common SNPs additively explain a large part of genetic additive variance, it is possible to use such realized relatedness estimated from genomic data to measure traits' heritability in wild populations (Yang et al., 2010). Furthermore, by aligning markers against a reference genome, heritability can then be partitioned among chromosomes (or at other genomic scales) enabling testing quantitative models of increasing heritability with the number of genes per chromosomes (Yang, Manolio, et al., 2011), or alternatively, identifying chromosomes displaying higher genetic variance than expected. GRM estimates and chromosome partitioning, first applied on human data (Yang, et al., 2010; Yang, Manolio, et al., 2011), were then applied in well-known free-ranging animal populations (Great tits, *Parus major*, Robinson, Santure, DeCauwer, Sheldon, & Slate, 2013; Soay sheep, *Ovis aries*, Bérénos et al., 2015) for which thousands of individuals were genotyped at thousands of loci. There have been fewer empirical applications of these methods to smaller samples (but see Silva et al., 2017; Wenzel, James, Douglas, & Piertney, 2015), although they have great potential in the current context of genomic tools being increasingly applied to natural populations (Gienapp et al., 2017) that often rely on relatively small samples.

Three important limitations for estimating GRM and GRM-based heritability are the number of SNPs, the number of individuals and the sampling variance (Visscher & Goddard, 2015). The limitation from the number (as well as diversity and quality) of markers and their ability to resolve relatedness essentially vanished when we entered the genomic era (Csillery, 2006; Gienapp et al., 2017; Stanton-Geddes, Yoder, Briskine, Young, & Tiffin, 2013). The sample size of individuals genotyped and phenotyped, however, remains an issue. For a given data set, the number of individuals included will influence both error and median values of heritability estimates, as tested in two studies (Stanton-Geddes et al., 2013; Visscher & Goddard, 2015). Nevertheless, reasonable heritability estimates can be reached using 150–200 individuals, as showed in a study on barrelclover (*Medicago trunculata*, Stanton-Geddes et al., 2013). However, different species, demographic contexts and sampling strategies will likely perform differently for a similar number of individuals, number of markers and traits' architecture. Visscher and Goddard (2015) showed that for designs that use genetic markers to estimate relatedness among randomly sampled individuals from a population (which may be a common situation for nonmodel species in natural populations), the error in heritability estimation is inversely proportional to the squared sample size and is proportional to the effective population size. Therefore, heritability estimates based from genetic marker relatedness in extremely large populations will likely require thousands of samples (e.g., marine pelagic fish (Gagnaire & Gaggiotti, 2016)), but also thousands of loci (given the small linkage disequilibrium among loci). Nevertheless, for populations of smaller size, a sampling strategy aiming at capturing variance in realized relatedness offers possibilities to obtain robust heritability estimates bounded by reasonable standard errors.

While genomics offers promising potential to obtain quantitative genetic parameters for virtually any species or population (Gienapp et al., 2017), it has only been used in a few studies to date and more applications in smaller samples and nonmodel species are needed. The first studies to estimate genomic heritability and partition variation across the genome elsewhere than in humans were, understandably, focusing on already well-known populations and species. In particular, analysis of 2,644 individuals at 7,203 SNPs from a long-term population study of great tits was the first application of quantitative genomics to a wild population (Robinson et al., 2013). The Soay sheep (*O. aries*) also benefited from such a genomic quantitative genetic study, with 5,805 individuals genotyped at 37,037 autosomal SNPs (Bérénos, Ellis, Pilkington, & Pemberton, 2014). This study revealed that most of the additive genetic (co)variances in sheep body size traits were captured by half of the SNPs. These pioneering studies enabled the verification of the power of quantitative genomics on known populations with long-term pedigrees and characterized quantitative genetic parameters (Edwards, 2013). Other authors also estimated heritability from GRM in data sets with much less individuals, for example, 200 barrelclover (*M. trunculata*) individuals genotyped at more than 5 million SNPs. Wenzel et al. (2015) estimated genomewide heritability in 695 red grouse (*Lagopus lagopus scotica*) genotyped at 384 SNPs. Silva et al. (2017) estimated quantitative genomic parameters in 1,898 house sparrows (*Passer domesticus*) genotyped at 6,348 SNPs and in 825 collared flycatchers (*Ficedula albicolis*) genotyped at 38,689 SNPs. These examples confirmed the usefulness of the method in new cases and with fewer individuals. Although we did not perform a thorough meta-analysis, there was a trend among these studies for an increase, among and within studies, of heritability estimates' standard errors with decreasing sample size and SNP amount. Nevertheless, and surprisingly considering the number of studies performing genomic analyses since the beginning of the genomic era, there are relatively few applications of genomic data to estimate heritability, genetic correlations and chromosome partitioning, using smaller data sets.

Here, we assessed heritability, genetic correlations and chromosome partitioning for four phenotypic traits, namely tarsus length, wing length, bill length and body mass, in a collection of 494 individuals of Corsican blue tits (*Cyanistes caeruleus ogliastrae*) from three closely located sites in Corsica (Figure 1). Based on mitochondrial genetic and phenotypic data, blue tits in Corsica and Sardinia have been qualified as a subspecies, *C. caeruleus ogliastrae* (Kvist, Viiri, Dias, Rytkönen, & Orell, 2004). The focal populations are ideal for quantitative genetic study because of the availability of pedigree, phenotypic and genetic data gathered through a long-term study (Charmantier, Doutrelant, Dubuc-Messier, Fargevieille, & Szulkin, 2016). Based on microsatellite and SNP genetic data, important gene flow together with significantly low genetic differentiation has been suggested between Corsican populations (Porlier, Garant, Perret, & Charmantier, 2012; Szulkin, Gagnaire, Bierne, & Charmantier, 2016). Phenotypic variance for the four aforementioned traits has been characterized as quantitative and moderately to highly heritable (Blondel et al., 2006; Charmantier, Kruuk, Blondel, & Lambrechts,

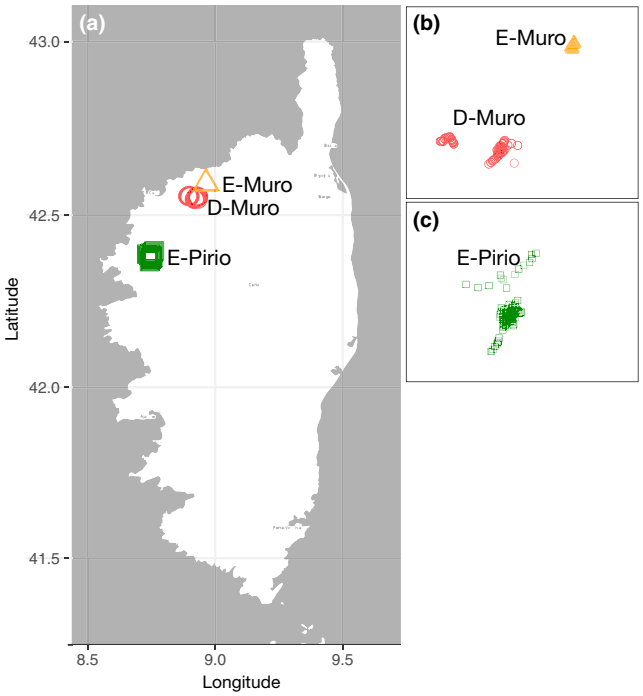

FIGURE 1 (a) Map of the sampling locations in Corsica, including a zoom-in of the area where the populations are located (b) D-Muro and E-Muro and (c) E-Pirio (see Table 1 for details) [Colour figure can be viewed at wileyonlinelibrary.com]

2004; Charmantier et al., 2016). The latest published results based on 17–27 years of monitoring provided population-specific significant heritability ranging from 0.43 to 0.57 for tarsus length, 0.20–0.33 for wing length, 0.18–0.34 for bill length and 0.22–0.32 for body mass. All trait combinations showed significant additive genetic covariance (COVA) in at least one population; in particular, COVAs between wing length and tarsus length and between wing length and body mass were significantly positive (Delahaie et al., 2017). Lastly, from 18% to 25% of young have been identified as extra-pair offspring in these Corsican blue tit populations (Charmantier, Blondel, Perret, & Lambrechts, 2004), which may lead to slight underestimation of trait heritability from a field-based social pedigree (Charmantier & Réale, 2005).

The general objective of this study was to test the effectiveness of GRM to estimate heritability and genetic correlations among traits and to perform heritability partitioning while using a relatively small sample size of Corsican blue tits. We produced a data set of 494 Corsican blue tits, from three sites (sample sizes of 110, 185 and 199), genotyped at nearly 50,000 filtered SNPs derived from RAD-seq and phenotyped for four quantitative phenotypic traits (tarsus length, wing length, bill length and body mass). We looked for erroneous links in the social pedigree (PED-social) using the genetic data, to build a corrected social pedigree (PED-corrected). We estimated heritability using both GCTA (Yang, Lee, Goddard, & Visscher, 2011) and *MCMCglmm* (Hadfield, 2010) for the 494 genotyped individuals (with the three sites pooled or separately), based on GRM, PED-social and PED-corrected. Heritability differences between PED-corrected and PED-social were computed to infer the possible effects of pedigree errors on estimates of traits heritability (notably underestimation originating from extra-pair copulations (Charmantier & Réale, 2005)). Comparisons between heritability estimated using PED-corrected and GRM aimed at defining the effect of the greater precision and information content of the GRM. We estimated the effect of the number of SNPs on tarsus length heritability estimates, in order to discuss the SNP density needed to recover much of the additive genetic variance. Finally, we estimated genetic correlations between traits and we partitioned heritability between chromosomes, although these quantitative genomic tests may challenge our sample size and probably require much larger sample size to produce accurate results. We discuss the usefulness and limitations of these heritability inferences in small samples from wild populations of non-model species.

2 | METHODS

2.1 | Study sites, bird monitoring and phenotypic data

Data were collected in three locations (E-PIRIO, D-MURO and E-MURO) in Corsica (France, Figure 1, Table 1). Approximately 6 km separate D-MURO from E-MURO and 27 km separate these sites from E-PIRIO. The landscape of these sites is dominated either by the evergreen holm oak *Quercus ilex* (E-PIRIO and E-MURO) or by the deciduous downy oak *Quercus pubescens* (D-MURO). All sites were monitored as part of a long-term research programme for 20 (E-Muro), 24 (D-Muro) and 42 years (E-Pirio) until 2017. A total of 360 nest boxes were monitored and 26,650 birds were ringed. Capture and handling of the birds was conducted under permits provided by the Centre de Recherches sur la Biologie des Populations d'Oiseaux (CRBPO) and by the Direction Départementale des Services Vétérinaires (DDSV).

Each year, nest boxes were visited at least once a week during the reproductive period (from early April to late June). Breeding blue tits were captured in nest boxes during the feeding of their young and banded (if not already banded earlier) with a unique metal ring provided by the CRBPO. Nestlings were also banded before fledging at 9–15 days old.

Four phenotypic traits were measured on male and female breeders: tarsus length (from the intertarsal joint to the most distal undivided scute on the tarsometatarsus), flattened wing length, bill length (from the anterior end of the nares to the tip of the upper mandible) and body mass. Phenotypic differences among individuals from the different sites and used in genomic heritability estimates are presented in Table 1 and Figure S1 (see Charmantier et al. (2016) for a review of phenotypic divergence in these sites). Five to 20 μl of blood was sampled from a small neck vein or from a wing vein from adult breeders for later DNA extraction. Blood was stored at 4°C in Queen's buffer (Seutin, White, & Boag, 1991).

TABLE 1 Geographic coordinates, number of genotyped individuals (*N*) and trait mean values (±*SE*) for this sample, for each of the three sites and for the entire data set. The phenotypic variance (V_p) is also given for each trait

					Tarsus length (mm)				Wing length (mm)				Bill length (mm)				Body mass (g)			
Site	Latitude	Longitude	*N*	Sex	*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	0.25	98	63.5	1.6	4.91	90	6.4	0.3	0.16	97	9.4	0.4	0.27
				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		185	62.2	2.0		172	6.5	0.4		183	9.3	0.4	
D-Muro	42.551	8.923	199	Male	95	16.6	0.4	0.24	95	63.3	1.6	4.38	92	6.4	0.3	0.16	92	9.8	0.4	0.27
				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		199	62.0	1.9		191	6.4	0.3		196	9.6	0.5	
E-Muro	42.589	8.967	110	Male	55	16.5	0.5	0.25	55	63.4	1.5	3.84	53	6.4	0.3	0.19	54	9.7	0.4	0.24
				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		109	62.1	1.8		107	6.6	0.4		109	9.6	0.4	
Total			494	Male	248	16.5	0.4	0.26	248	63.4	1.6	4.43	235	6.4	0.3	0.17	243	9.6	0.4	0.30
				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		493	62.1	1.9		470	6.5	0.3		488	9.5	0.5	

2.2 | DNA extraction, RAD-sequencing, SNP calling and data filtering

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

Library preparation using restriction site-associated DNA sequencing (RAD-seq; Baird et al., 2008) with the enzyme SbfI was done by MGX (CNRS, Montpellier). Each individual was identified using a unique 6 nucleotide tag and individuals were multiplexed in equimolar proportions by groups of 36 individuals. Each library was sequenced on one of 20 lanes of an Illumina HiSeq 2000 (libraries also included blue tit individuals from another site that was not analysed in this study).

Raw sequences were inspected with *FastQC* (Andrews, 2010) for quality controls. Potential fragments of Illumina adapters were trimmed with *Cutadapt* (Martin, 2011), allowing for a 10% mismatch in the adapter sequence. Reads were filtered for overall quality, demultiplexed and trimmed to 85 bp using *process_radtags*, *from the Stacks software pipeline Version 1.39* (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013), allowing for one mismatch in the barcode sequence. BWA-MEM 0.7.13 (Li & Durbin, 2009) was used to map individual sequences against the reference genome of the great tit (Laine et al., 2016) and to produce *sam* files using default options. Samtools 0.1.19 (Li et al., 2009) was used to build and sort *bam* files. Back in Stacks Version 1.39, we used *pstacks* to treat *bam* files, align the reads into matching stacks, infer loci and detect SNPs at each locus. We used a minimum depth of coverage (m) of 5, the SNP model and alpha = 0.05 (chi-square significance level required to call a heterozygote or homozygote). *cstacks* was used to build the catalogue of loci using $n = 3$ (number of mismatches allowed between sample loci when build the catalogue). *sstacks* was used to match loci against the catalogue. *populations*, the last Stacks program used here, genotyped individuals. Loci were retained if genotyped in at least 90% of individuals (all individuals from all sites grouped), with heterozygosity per site ≤ 0.60, and with individual minimal read depth of 10 ("na" replaced genotypes below a read depth of 10). *VCFtools* (Danecek et al., 2011) was used for further filtering of loci for a minimum average read depth of 20 across all genotypes and a maximum average read depth of 100 across all genotypes. Individuals were genotyped for at least 90% of all loci; a filter would have otherwise been implemented to remove individuals with low genotyping rate. The data set was filtered to retain loci with minimum allele frequency larger than 5% (MAF ≥ 0.05). Subsequently, the data set was pruned for linkage disequilibrium with the plink command *LD indep 50 5 2*.

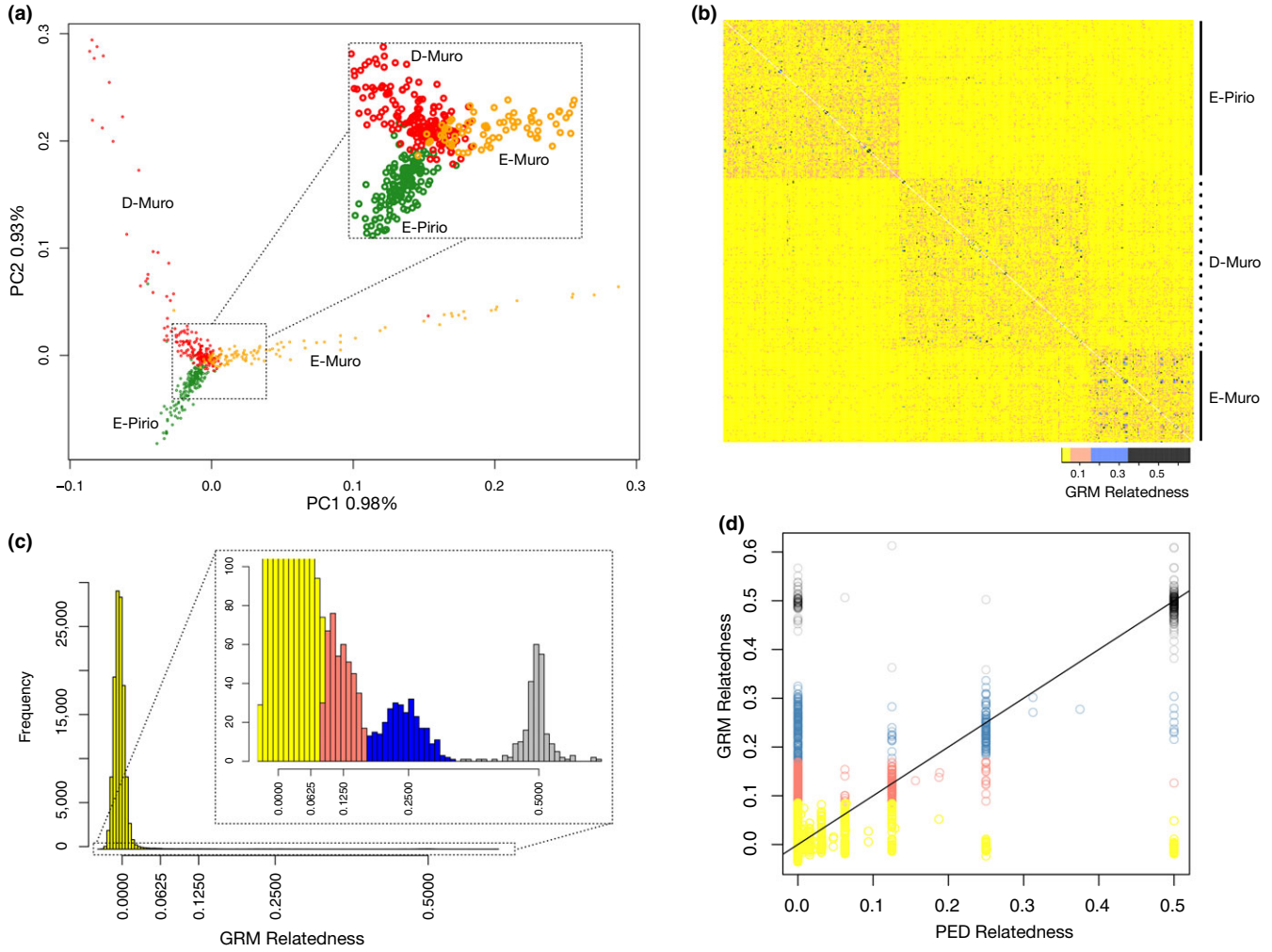

FIGURE 2 Genetic structure and relatedness among individuals. (a) Biplot of the two-first axes of a principal component analysis of genetic distances between individuals, (b) heatmap of the genomewide relatedness matrix (GRM) among the 494 genotyped individuals, (c) histogram of GRM among individuals, with a zoom-in highlighting particularly high relatedness values and (d) biplot of GRM and social pedigree relatedness from the field observations (PED-social). In (b), (c) and (d), colour-coding is based on the empirical relatedness distribution and is only indicative. PED, pedigree [Colour figure can be viewed at wileyonlinelibrary.com]

2.3 | Population genetic structure and genomewide relatedness matrix

Data were converted to several formats using the R-package *radiator* (Gosselin, 2017). Expected heterozygosity per site and F_{st} between sites were estimated using *Genodive* (Meirmans & van Tienderen, 2004). A PCA resolving genetic distance between individuals was calculated using the R-packages *gdsfmt* and *SNPRelate* (Zheng et al., 2012).

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

2.4 | Social pedigree

Social pedigrees (PED-social) were constructed using the pedantics R-package (Morrissey & Wilson, 2009), on the basis of the complete pedigree generated from the whole long-term monitoring period. We first included all ringed individuals and assigned their mother and father based on observational data at capture. Unknown parents of a given nest were coded using a dummy identity in the PED-social to preserve sibship information. PED-social was pruned for each data set studied (each population separately and the combined pedigree of the three sites pooled) to retain only genotyped individuals and their ancestors (see Table S2 for detailed characteristics of the pedigree).

2.5 | Comparison between social pedigree and GRM and creation of a corrected social pedigree

We compared the PED-social and GRM for the 494 genotyped individuals mainly to infer the potential presence of erroneous pedigree relationship (particularly parent–offspring and sibling relationship). Indeed, errors may occur in social pedigrees, essentially due to extra-pair paternities, but also observational errors such as misreading of bird ring, successive egg-laying by different females in the same nest box and brood parasitism. Such errors, notably extra-pair paternities, can downwardly bias estimates of heritability (Charmantier & Réale, 2005). We illustrated the concordances and discrepancies between PED-social and GRM using a biplot (Figure 2d). We then created corrected social pedigrees (PED-corrected) on the basis of discrepancies found between the PED-social and the GRM. For pairs of individuals having PED-social relatedness values of 0.5 (full-sibs or parent–offspring relationships) but GRM relatedness estimates below 0.1, we assumed that these relationships corresponded to erroneous parent–offspring relationship due to extra-pair paternity and assigned them a 0 in PED-corrected. For PED-social relatedness values of 0.5 but GRM estimates between 0.15 and 0.35, we assumed they corresponded to full-sibs being half-sibs and assigned them a value of 0.25 in PED-corrected. We also corrected false half-sibs in the pedigree (pairs of individuals having a 0.25 relatedness value in the PED-social but a GRM estimates smaller than 0.1) and assigned them 0 relatedness values. We acknowledge that probably some more relationships could have been corrected based on discrepancies between PED-social and GRM, but the distribution of GRM values below 0.15 was largely overlapping and it would have been difficult to assign correct PED-corrected values. Using a method such as the one implemented in the R-package *sequoia* (Huisman, 2017) would certainly have allowed the correction of these potential errors as well. Another potential limitation comes from the fact that ancestors of the genotyped individuals in the pedigree were not genotyped and therefore erroneous relationships between these individuals were impossible to correct.

2.6 | Inference of heritability

We used both a frequentist and a Bayesian method to estimate heritability of the four traits. The frequentist method was implemented in GCTA. The Bayesian framework was implemented in the *MCMCglmm* R-package (Hadfield, 2010). Bayesian inference is renowned to have a clear advantage over the other existing methods because the use of posterior distributions propagates the errors in estimates derived from animal models (Morrissey, de Villemereuil, Doligez, & Gimenez, 2014).

GCTA was used to estimate heritability for each trait, separately for each of the four GRMs. We here used best linear unbiased estimates (BLUPs) for each individual and phenotype because we could not implement more complex models in GCTA. BLUPs were estimated for the entire data set of 494 genotyped individuals using a generalized linear mixed model with the *MCMCglmm* function, integrating a random effect for the observer and permanent environment effects accounting for multiple measurements of the same individual. Differences between sites and sex were also accounted for by adding these two factors as fixed effects. We then extracted the posterior mode of the BLUPS for each individual and ran GCTA for each of the four GRMs and the four traits.

MCMCglmm was used to estimate heritability for each trait, separately for each of the four GRMs, the four PED-social and the four PED-corrected. Sex was integrated as a fixed effect in all models. Site was integrated as a fixed effect in the models that included all the individuals from the different sites. Measurer, identity and animal (corresponding either to the GRM, PED-social or PED-corrected matrix) effects were incorporated as random factors to partition the phenotypic variance into its observer, permanent environment and additive genetic components. Running models based on the GRMs, we used both BLUPs (to compare with GCTA) and alternatively all phenotypic measurements. For the models using GRMs and BLUPs, the only random factor was the animal effect (i.e., the GRM) as permanent environmental and observer effects were already taken into account in the BLUPs. For the models based on PED-social and PED-corrected, we did not use BLUPs (since it could bias the analysis and/or give anticonservative results, Hadfield, Wilson, Garant, Sheldon, & Kruuk, 2010) but instead used all phenotypic measurements. We used identical parameters, priors and iterations for each estimate. Random effects included additive genetic effects (V_A) estimated through the inclusion of pedigree data, permanent environmental effects (V_{PE}) accounting for repeated measurements of the same individual, measurer identity controlling for any potential confounding measurer effect (V_{OBS}) and residual variance (V_R). The models used can be described as follows:

$$Y = \mu + Xb + Z_A a + Z_{PE} pe + Z_O obs + e \quad (1)$$

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

2.7 | Inference of the effect of the number of SNPs on heritability estimates

We inspected whether the number of SNPs included resulted in (i) an increase in heritability estimation precision (approximated via

standard deviation and LRT (likelihood ratio test) in the case of GCTA) and (ii) an increase in heritability (median) via saturation of the GRMs by SNPs in linkage disequilibrium with loci most likely causative of the phenotypic variation (Gienapp et al., 2017). We reported the median and standard deviation values of tarsus length heritability inferred using GCTA and *MCMCglmm* for several GRMs produced using a variable number of SNPs. Using GCTA, we analysed 991 GRMs made up of a decreasing number of randomly chosen SNPs from the entire data set, by step of 50 SNPs from the total number of SNPs. To confirm GCTA results, we analysed a smaller number of 25 GRMs with *MCMCglmm* (much less than for GCTA since the Bayesian analysis was much more time-consuming than the frequentist one) made from randomly chosen SNPs, concentrated mainly between 1,000 and 10,000 SNPs (since GCTA indicated that the rate of improvement in estimates was particularly concentrated between these numbers).

2.8 | Inference of genetic correlations

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

2.9 | Chromosome partitioning

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

3 | RESULTS

3.1 | Phenotypes

Among the 494 genotyped individuals, 494, 493, 470 and 488 were measured for tarsus length, wing length, bill length and body mass, respectively (Table 1, Figure S1). The sex ratio of genotyped individuals was near 0.5, with 246 females of 494 individuals. Phenotypic traits varied slightly between sexes and between sites, justifying the inclusion of sex and site effects in the models partitioning phenotypic variance.

3.2 | SNP calling and population genetic structure

The median number of reads per individuals was 6,066,514. The median read depth per individual and per locus was 56. The Stacks program *population* outputted 52,783 loci totalling 96,009 SNPs. After MAF pruning, we retained 41,986 loci totalling 68,114 SNPs. After LD pruning, we retained 38,030 loci totalling 49,682 SNPs. 47,865 of these SNPs were on autosomes. The number of filtered SNPs per chromosome ranged from 5,523 (chromosome 2) to 56 (LGE22).

The heterozygosity was 0.205, 0.205 and 0.204 in E-PIRIO, D-MURO and E-MURO, respectively (Table S3). Genetic differentiation between sites estimated using an *Fst* index was low yet significant, ranging from 0.006 to 0.008 (p-value < .001; Table S3). The two-first axes of the PCA (Figure 2a) explaining each 0.98% and 0.93% of the genotypic variance, and the heatmap of the GRM (Figure 2b), depicted the low genetic structure between the three sites.

3.3 | Genomewide relatedness matrix and social pedigree

The GRM (Figure 2c) mostly included nonrelated (or distantly related) individuals. Zooming in the histogram (Figure 2c) shows the presence of related individuals with parent–offspring or full-sib-like links (grey), and half-sib-like links (blue). Comparing the relatedness values from the GRM to the one from the PED-social showed great consistency between the two matrices (Figure 2d). However, as expected (and shown elsewhere, Charmantier & Réale, 2005), 112 links of 858 known links presented PED-social relatedness higher than GRM relatedness by at least 0.10. Among these, 39 links exhibited PED-social relatedness of 0.5 but GRM relatedness of 0, and 36 links showed PED-social relatedness of 0.25 but GRM relatedness of 0. In addition, the GRM allowed the identification of many links that cannot be documented in the PED-social because of immigration, incomplete sampling or extra-pair paternity (Pemberton, 2008). Notably, we found 460 links with PED-social relatedness was equal to 0 but GRM relatedness larger than 0.1. In total, we corrected 83 erroneous PED-social relatedness values (corresponding to ca. 10% of the total number of links known) thanks to the GRM analysis to create the PED-corrected.

3.4 | Heritability estimates

In general, heritability measures based on PED-social, PED-corrected and GRM, using BLUPs or all the measurements, and with *MCMCglmm* or GCTA, were highly consistent for the same trait (Table 2, Figure 3; see Table S4 for variance estimates associated with the other random factors). The variations were very high among the three localities, for which sample sizes were very small. When pooling the three localities, variations between methods and credible intervals decreased.

Considering estimates from *MCMCglmm* and models fitting all phenotypic measurements, there was a consistent trend for lower

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

					Tarsus length			Wing length			Bill length			Body mass		
Relatedness	Data	Implementation	Site	*N*	*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	*MCMCglmm*	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	*MCMCglmm*	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	*MCMCglmm*	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	*MCMCglmm*	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

GRM, genomewide relatedness matrix; PED, pedigree.

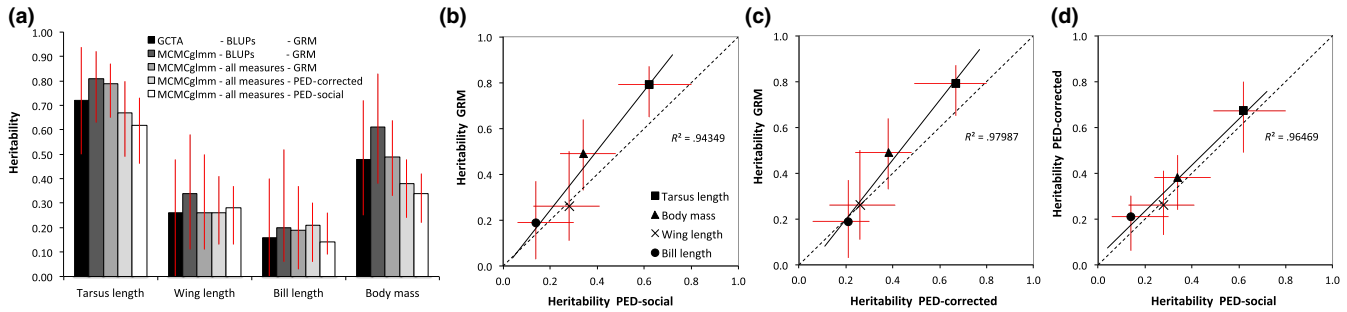

FIGURE 3 (a) Heritability estimates using the GRM, the PED-social or the PED-corrected, based on either BLUPS or the entire phenotypic measurements, and implementing models in GCTA or *MCMCglmm*, for the 494 genotyped individuals from the three populations pooled. (b, c and d) Relationship between heritability inferred for the four traits for PED-social, PED-corrected and GRM using *MCMCglmm*, using the model with all phenotypic measurements in *MCMCglmm*. Errors bars correspond to 95% credible interval for *MCMCglmm* and to 95% confidence intervals for GCTA. GRM, genomewide relatedness matrix; PED, pedigree [Colour figure can be viewed at wileyonlinelibrary.com]

heritability when using PED-social compared to GRM (Figure 3). For example, tarsus length and body mass heritability were lower by 22% and 31%, respectively, when using PED-social compared GRM. Specifically, there was a trend for lower heritability when using PED-social than PED-corrected. For example, tarsus length and body mass heritability were lower by 7% and 11%, respectively, when using PED-social compared to PED-corrected. There was also a trend for lower heritability when using PED-corrected compared to GRM. For example, tarsus length and body mass heritability were lower by 15% and 22%, respectively, when using PED-social compared to PED-corrected.

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

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

3.5 | Genetic correlations between traits

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

3.6 | Chromosome partitioning

When fitting the univariate GCTA model simultaneously on each autosome, the correlation between heritability and chromosome length was positive for tarsus length (r^2 = .26; p-value = .009), being the most heritable traits of this study. The correlation was nonsignificant for body mass (r^2 = .08; p-value = .20), wing length (r^2 = .03; p-value = .41) and bill length (r^2 = .00; p-value = 0.85). Standard error intervals for chromosome scaled heritability did not include zero only for chromosomes 1A and 3 for tarsus length (Table 4, Figure 5) and for chromosome 6 for body mass. In the case of tarsus length, no autosome had to be removed to enable the convergence of the model fitting all the autosomes simultaneously. In contrast, for wing length, bill length and body mass, the microchromosomes group had to be removed and respectively one, 12 and three additional autosomes (by order of increasing size) had to be removed to enable convergence of the models. When fitting the univariate GCTA model separately for each autosome vs. the rest of the genome, several heritabilities were likely overestimated because the total heritability was larger than with the previous method (h^2 = 1.47 vs. 0.84 for tarsus length, 0.85 vs. 0.36 for wing length, 1.11 vs. 0.31 for bill length and 0.95 vs. 0.62 for body mass).

4 | DISCUSSION

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

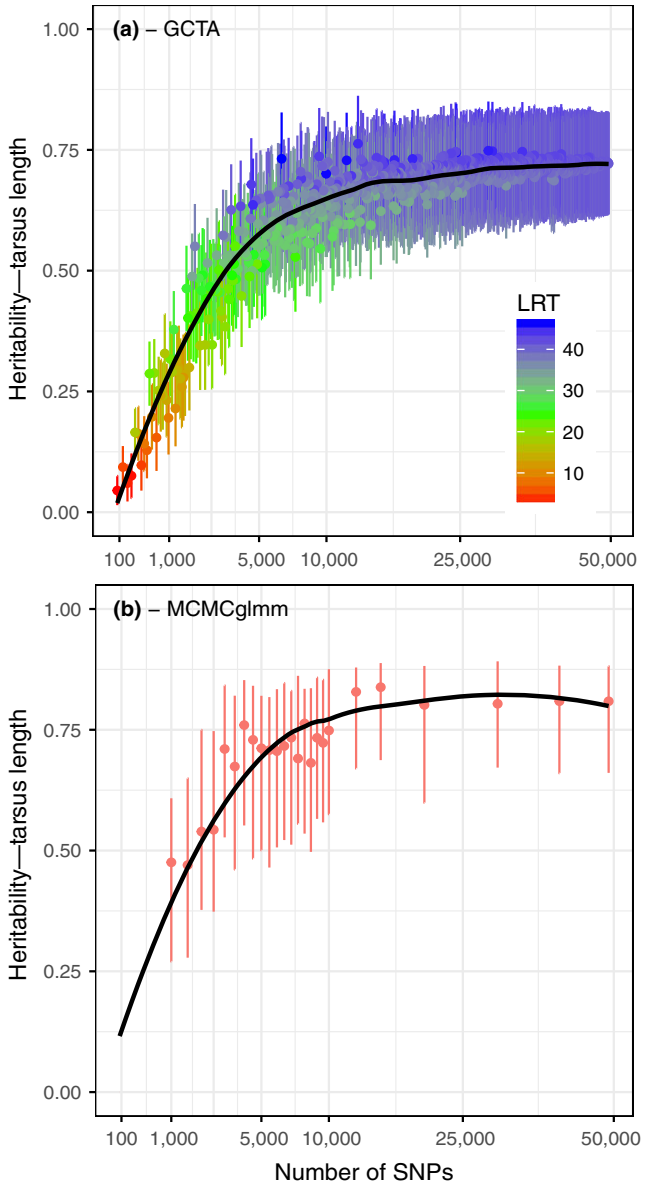

FIGURE 4 Effect of the number of SNPs used to build the genomewide relatedness matrix (GRM) on the tarsus length heritability estimated with GCTA and BLUPs (a) and *MCMCglmm* and all the phenotypic measures (b). Circles represent the median value. Bars represent standard errors for GCTA and 95% credible interval for *MCMCglmm*. Black curves correspond to LOESS fits. The log likelihood ratio of GCTA estimates is colour-coded [Colour figure can be viewed at wileyonlinelibrary.com]

sets of well-studied wild organisms (Bérénos et al., 2014; Robinson et al., 2013; Stanton-Geddes et al., 2013), we require more explorations on whether smaller data sets 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 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 data set 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 nonmodel 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 likely than heritability estimates to be very informative or robust given the large credible intervals and the number of nonsignificant estimates, most likely due to power issues (insufficient number of SNPs per chromosome in the case of the microchromosomes and insufficient number of individuals). We thereafter discuss the usefulness and limitations of inferences of quantitative genomic parameters, including heritability, in relatively small sample sizes, with a medium number of genetic markers, for wild populations of nonmodel species.

4.1 | Comparing PED-social and GRM

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

The second major difference between PED-social and GRM was that several individuals had much higher relatedness in PED-social than in the GRM, most likely indicating errors in PED-social. The most common cases consisted of wrongly assigned parentages (relatedness of 0.5 in PED-social and 0 in GRM), half-sibs most likely originating from extra-pair paternities (relatedness of 0.5 in PED-social and 0.25 in GRM) and the spread of such errors in the pedigree (other smaller relatedness in GRM than in PED-social). Using this comparison between GRM and PED-social, we created a PED-corrected in which these aforementioned erroneous links, accounting for approximately 10% of the PED-social links, were corrected. This proportion of erroneous linked corrected in our data set should be

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

Relatedness	Data	Implementation	Genetic correlation	Tarsus length Wing length	Tarsus length Bill length	Tarsus length Body mass	Wing length Bill length	Wing length Body mass	Bill length Body mass
GRM	BLUPs	GCTA	Mode	0.40	0.37	0.48	−0.06	0.22	0.13
			.05	0.05	−0.14	0.23	−0.86	−0.022	−0.49
			.95	0.75	0.88	0.73	0.86	0.67	0.76
GRM	BLUPs	*MCMCglmm*	Mode	0.41	0.32	0.52	0.10	0.21	0.03
			.05	−0.08	−0.09	0.20	−0.57	−0.28	−0.50
			.95	0.67	0.73	0.75	0.48	0.66	0.59
GRM	All measures	*MCMCglmm*	Mode	0.48	0.45	0.56	0.50	0.50	0.33
			.05	0.16	0.06	0.37	−0.08	0.05	0.05
			.95	0.81	0.80	0.76	0.87	0.71	0.88
PED-corrected	All measures	*MCMCglmm*	Mode	0.48	0.43	0.57	0.46	0.49	0.52
			.05	0.13	0.00	0.40	−0.12	0.06	0.05
			.95	0.79	0.81	0.78	0.89	0.74	0.84
PED-social	All measures	*MCMCglmm*	Mode	0.46	0.52	0.61	0.42	0.41	0.37
			.05	0.17	0.16	0.38	−0.13	0.09	0.08
			.95	0.77	0.79	0.82	0.79	0.72	0.81

GRM, genomewide relatedness matrix; PED, pedigree.

expected because 18%–25% of young have been previously identified as extra-pair offspring in these Corsican blue tit populations (Charmantier, Blondel, et al., 2004).

4.2 | Heritability estimates from PED-social, PED-corrected and GRM

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

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

Overall, when using *MCMCglmm* and models using all of the phenotypic measurements, heritability estimates were relatively similar for GRM, PED-corrected and PED-social for the same trait (Figure 3). These results are in line with previous comparisons realized with much larger sample size (Bérénos et al., 2014; Robinson et al., 2013). However, unlike these previous studies showing very small differences in heritability based on PED-social and GRM, our estimates based on GRM were larger than using PED-social and to a lesser extent, than using PED-corrected, for tarsus length and for body mass, but not for wing length and bill length. Here, wing length heritability was 0.28 using PED-social, 0.26 using PED-corrected and 0.26 using GRM (Table 2). In turn, tarsus length heritability obtained from GRM was 18% higher than from PED-corrected and the one from PED-corrected 8% higher than from PED-social. Similarly, body mass heritability obtained from GRM was 29% higher than from PED-corrected and the one from PED-corrected 12% higher than from PED-social. The difference between PED-corrected and PED-social for these two traits may be attributable to the erroneous links in PED-social. Indeed, it has been estimated that PED-social containing 5%–20% of extra-pair paternities results in underestimated heritability by up to 17%, in blue tits (Charmantier & Réale, 2005). Then, the differences observed between heritability estimated using GRM and PED-corrected are likely to originate from higher density of the GRM and the fact that GRM incorporate more variance in relatedness for a given class of kinship. Further changing the pedigree by not only correcting erroneous existing links but also informing previously unidentified links (e.g., linking previously unknown extra-pair fathers to their offspring) based on the genetic data, could also

TABLE 4 Chromosome partitioning of the heritability estimated for the four phenotypic traits using GCTA. Partitioning was achieved either (i) fitting all the chromosomes simultaneously or (ii) considering one chromosome vs. all of the other autosomes, iteratively for each chromosome. Several microchromosomes were grouped (22, 27, 28, 25LG1, 25LG2, LGE22) and referred as "micros." "na" refers to the sets of chromosomes that have been removed to enable the convergence of the model fitting simultaneously all the chromosomes (see Methods). The bold values do not include 0 in their standard error intervals

			Tarsus				Wing length				Bill length				Body mass			
			All chromosomes simultaneously		One chromosome vs. the other, iteratively		All chromosomes simultaneously		One chromosome vs. the other, iteratively		All chromosomes simultaneously		One chromosome vs. the other, iteratively		All chromosomes simultaneously		One chromosome vs. the other, iteratively	
Chromosome name	SNPs	Genes	Median	*SE*	Median	*SE*	Median	*SE*	Median	*SE*	Median	*SE*	Median	*SE*	Median	*SE*	Median	*SE*
1	4,784	1,131	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	5,523	1,287	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	4,383	1,096	**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	2,869	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	1,428	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	2,638	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	2,902	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	1,818	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	1,907	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	1,718	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	1,689	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	1,573	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	1,371	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	1,177	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	1,309	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	1,364	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	1,004	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	1,114	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	1,301	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	1,368	874	0	0.1	0.02	0.09	na	na	0.05	0.08	na	na	0	0.09	na	na	0	0.09

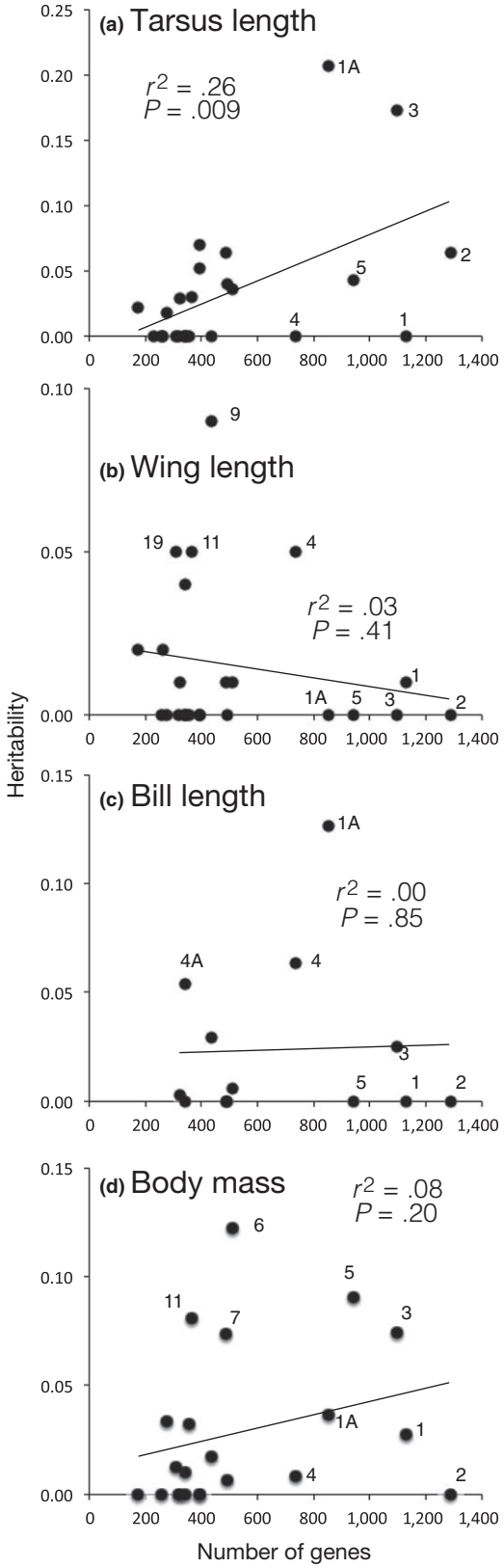

FIGURE 5 Chromosome partitioning. The univariate GCTA model was fitted simultaneously on each autosome. Linear correlation between the number of genes per chromosome and the GCTA estimates of heritability per chromosome for the four phenotypic traits. Names of the largest chromosomes, and of the chromosomes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

decrease the heritability difference between PED-corrected and GRM.

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

4.3 | Effect of the number of SNPs on the heritability estimates based on GRM

The number of SNPs had a relatively minor effect on heritability estimates, from approximately 15,000 SNPs (Figure 4). This is concordant with what has been shown in previous studies (Bérénos et al., 2014; Stanton-Geddes et al., 2013). Of course, the effect of the number of SNPs on the heritability estimation depends on several parameters (e.g., genome size, LD across the genome, LD among markers, genetic architecture of the focal trait, polymorphism of the markers, genotyping quality, missing data) and is not transposable from one study to another. These days, the number of markers may no longer be a problem, given the rise of the accessibility of many genomic tools. However, the genotyping quality and missing data occurrence may be challenging for studies using RAD-seq and GBS methods, as noted by (Gienapp et al., 2017). Here, we prioritized the quantity and the quality of the RAD markers over the number of individuals, as well as genotyping only individuals with high-quality DNA, ending up with a relatively high read depth and low missing rate. Future studies aiming to use RAD-seq-derived markers to estimate heritability should also find a trade-off between the number of genomic markers and the number of genotyped individuals. A formal simulation study determining such a trade-off between numbers of individuals and of markers and markers read depth (directly linked to quality) would be highly valuable for guiding the design of RAD-seq analyses in the context of quantitative genomics.

4.4 | Genetic correlations using PED-social, PED-corrected and GRM

Genetic correlations obtained using GRM, PED-social and PED-corrected had large credible intervals, preventing us to make robust interpretations, except that our sample size was probably too small to obtain acceptable credible intervals. This was partly expected because bivariate models are known to be data hungry. While such lack of power is a recurrent issue when estimating genetic correlations even with much larger sample sizes (Ni, Moser, Wray, & Lee, 2017; Vattikuti, Guo, &

Chow, 2012; Visscher et al., 2014), it was nevertheless a disappointment while comparing these results to the one from (Bérénos et al., 2014), in which the credible intervals were relatively small. The median values of genetic correlations for these traits were in line with previous studies in other bird species (Teplitsky et al., 2014).

4.5 | Chromosome partitioning based on GRM

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

Regarding the three other traits analysed, the chromosome partitioning results appeared much less robust because several chromosomes had to be removed (7, 18 and 9 autosomes removed, for wing length, bill length and body mass, respectively) to enable models' convergence. The need for such chromosomes removals has previously been reported (Silva et al., 2017; Wenzel et al., 2015) and has received detailed consideration by Kemppainen and Husby (2018). In addition, the relationship between the number of genes and the heritability explained by chromosomes was deviating from theoretical expectations of increasing heritability with increasing chromosome size for these polygenic traits. For comparison, Bosse et al. (2017) showed a significant increase of the proportion of additive genetic variance of bill length explained by a chromosome in relation to chromosome size in great tits. Such deviation may not have been affected by removing chromosomes (Kemppainen & Husby, 2018) but most likely originated from a limited power of the small sample size, alongside relatively small heritabilities (Kemppainen & Husby, 2018).

4.6 | Conclusion on the usefulness of GRM-based heritability measures using small samples from wild populations

Overall, our study reveals that RAD-seq data on around 50k SNPs (as advised by Bérénos et al., 2014) for around 500 phenotyped individuals can provide estimates of heritability that are close to, and probably more accurate than, estimates based on social pedigree data from 7 years monitoring for more than 1,600 individuals (as included in the PED-social used here). This opens very interesting avenues in the field of quantitative genetics because estimating heritability and genetic correlations in the wild have long been restricted to study systems where long-term monitoring is feasible. RAD-seq 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 et al., 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 appear, however, to provide insufficient power to correctly estimate either genetic covariances between traits or contribution of individual chromosomes to overall heritability.

ACKNOWLEDGEMENTS

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

DATA ARCHIVING

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

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

A.C. and C.P. conceived the study. A.C., B.D. and C.P. performed field work. C.P. performed laboratory work and bioinformatics. C.P. and B.D. performed quantitative genetic/genomics inferences. C.P. wrote the manuscript. A.C., B.D. and C.P. critically revised and improved the manuscript.

ORCID

C. Perrier http://orcid.org/0000-0001-5820-9374

REFERENCES

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

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., … Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*(10), e3376. https://doi.org/10.1371/journal.pone.0003376

Bérénos, C., Ellis, P. A., Pilkington, J. G., Lee, S. H., Gratten, J., & Pemberton, J. M. (2015). Heterogeneity of genetic architecture of body size traits in a free-living population. *Molecular Ecology*, *24*, 1810–1830. https://doi.org/10.1111/mec.13146

Bérénos, C., Ellis, P. A., Pilkington, J. G., & Pemberton, J. M. (2014). Estimating quantitative genetic parameters in wild populations: A comparison of pedigree and genomic approaches. *Molecular Ecology*, *23*, 3434–3451. https://doi.org/10.1111/mec.12827

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

Bosse, M., Spurgin, L. G., Laine, V. N., Cole, E. F., Firth, J. A., Gienapp, P., … Groenen, M. A. (2017). Recent natural selection causes adaptive evolution of an avian polygenic trait. *Science*, *358*, 365–368. https://doi.org/10.1126/science.aal3298

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

Charmantier, A., Blondel, J., Perret, P., & Lambrechts, M. (2004). Do extra-pair paternities provide genetic benefits for female blue tits *Parus caeruleus*? *Journal of Avian Biology*, *35*, 1–9.

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

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

Charmantier, A., & Réale, D. (2005). How do misassigned paternities affect the estimation of heritability in the wild? *Molecular Ecology*, *14*, 2839–2850. https://doi.org/10.1111/j.1365-294X.2005.02619.x

Csillery, K. (2006). Performance of marker-based relatedness estimators in natural populations of outbred vertebrates. *Genetics*, *173*, 2091–2101. https://doi.org/10.1534/genetics.106.057331

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

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

Edwards, S. V. (2013). Next-generation QTL mapping: Crowdsourcing SNPs, without pedigrees. *Molecular Ecology*, *22*, 3885–3887. https://doi.org/10.1111/mec.12401

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

Gienapp, P., Fior, S., Guillaume, F., Lasky, J. R., Sork, V. L., & Csilléry, K. (2017). Genomic quantitative genetics to study evolution in the wild. *Trends in Ecology & Evolution*, *32*, 897–908. https://doi.org/10.1016/j.tree.2017.09.004

Gosselin, T. (2017). Radiator: RADseq data exploration, manipulation and visualization using R. https://doi.org/10.5281/zenodo.154432

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, *33*, 1–22.

Hadfield, J. D., Wilson, A. J., Garant, D., Sheldon, B. C., & Kruuk, L. E. B. (2010). The misuse of BLUP in ecology and evolution. *The American Naturalist*, *175*, 116–125. https://doi.org/10.1086/648604

Huisman, J. (2017). Pedigree reconstruction from SNP data: Parentage assignment, sibship clustering and beyond. *Molecular Ecology Resources*, *17*, 1009–1024. https://doi.org/10.1111/1755-0998.12665

Husby, A., Hille, S. M., & Visser, M. E. (2011). Testing mechanisms of Bergmann's rule: Phenotypic decline but no genetic change in body size in three passerine bird populations. *The American Naturalist*, *178*, 202–213. https://doi.org/10.1086/660834

Jensen, H., Steinsland, I., Ringsby, T. H., & Sæther, B.-E. (2008). Evolutionary dynamics of a sexual ornament in the house sparrow (*Passer domesticus*): The role of indirect selection within and between sexes. *Evolution*, *62*, 1275–1293. https://doi.org/10.1111/j.1558-5646.2008.00395.x

Kemppainen, P., & Husby, A. (2018). Inference of genetic architecture from chromosome partitioning analyses is sensitive to genome variation, sample size, heritability and effect size distribution. *Molecular Ecology Resources*, https://doi.org/10.1111/1755-0998.12774

Kvist, L., Viiri, K., Dias, P. C., Rytkönen, S., & Orell, M. (2004). Glacial history and colonization of Europe by the blue tit *Parus caeruleus*. *Journal of Avian Biology*, *35*, 352–359. https://doi.org/10.1111/j.0908-8857.2004.03297.x

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

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*, 10–12. https://doi.org/10.14806/ej.17.1.200

Meirmans, P. G., & van Tienderen, P. H. (2004). genotype and genodive: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, *4*, 792–794. https://doi.org/10.1111/j.1471-8286.2004.00770.x

Morrissey, M. B., de Villemereuil, P., Doligez, B., & Gimenez, O. (2014). Bayesian approaches to the quantitative genetic analysis of natural populations. In A. Charmantier, D. Garant, & L. E. B. Kruuk (Eds.), *Quantitative genetics in the wild* (pp. 228–253). Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199674237.001.0001

Morrissey, M. B., & Wilson, A. J. (2009). pedantics: An r package for pedigree-based genetic simulation and pedigree manipulation, characterization and viewing. *Molecular Ecology Resources*, *10*, 711–719. https://doi.org/10.1111/j.1755-0998.2009.02817.x

Ni, G., Moser, G., Wray, N. R., & Lee, S. H. (2017). Estimation of genetic correlation using linkage disequilibrium score regression and genomic restricted maximum likelihood. *bioRxiv*, 194019.

Pemberton, J. M. (2008). Wild pedigrees: The way forward. *Proceedings of the Royal Society B-Biological Sciences*, *275*, 613–621. https://doi.org/10.1098/rspb.2007.1531

Porlier, M., Garant, D., Perret, P., & Charmantier, A. (2012). Habitat-linked population genetic differentiation in the blue tit *Cyanistes caeruleus*. *Journal of Heredity*, *103*(6), 781–791. https://doi.org/10.1093/jhered/ess064

Postma, E. (2014). Four decades of estimating heritabilities in wild vertebrate populations: Improved methods, more data, better estimates? In A. Charmantier, D. Garant, & L. E. B. Kruuk (Eds.), *Quantitative genetics in the wild* (pp. 16–33). Oxford, UK: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199674237.001.0001

Robinson, M. R., Santure, A. W., DeCauwer, I., Sheldon, B. C., & Slate, J. (2013). Partitioning of genetic variation across the genome using multimarker methods in a wild bird population. *Molecular Ecology*, *22*, 3963–3980. https://doi.org/10.1111/mec.12375

Santure, A. W., Cauwer, I., Robinson, M. R., Poissant, J., Sheldon, B. C., & Slate, J. (2013). Genomic dissection of variation in clutch size and egg mass in a wild great tit (*Parus major*) population. *Molecular Ecology*, *22*, 3949–3962. https://doi.org/10.1111/mec.12376

Santure, A. W., Poissant, J., De Cauwer, I., Oers, K., Robinson, M. R., Quinn, J. L., … Slate, J. (2015). Replicated analysis of the genetic architecture of quantitative traits in two wild great tit populations. *Molecular Ecology*, *24*, 6148–6162. https://doi.org/10.1111/mec.13452

Schielzeth, H., & Husby, A. (2014). Challenges and prospects in genome-wide quantitative trait loci mapping of standing genetic variation in natural populations. *Annals of the New York Academy of Sciences*, *1320*, 35–57. https://doi.org/10.1111/nyas.12397

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

Silva, C. N. S., McFarlane, S. E., Hagen, I. J., Rönnegård, L., Billing, A. M., Kvalnes, T., … Qvarnström, A. (2017). Insights into the genetic architecture of morphological traits in two passerine bird species. *Heredity*, *119*, 197–205. https://doi.org/10.1038/hdy.2017.29

Stanton-Geddes, J., Yoder, J. B., Briskine, R., Young, N. D., & Tiffin, P. (2013). Estimating heritability using genomic data (J Hadfield, Ed.). *Methods in Ecology and Evolution*, *4*, 1151–1158. https://doi.org/10.1111/2041-210X.12129

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

Teplitsky, C., Tarka, M., Møller, A. P., Nakagawa, S., Balbontín, J., Burke, T. A., … Gustafsson, L. (2014). Assessing multivariate constraints to evolution across ten long-term avian studies. *PLoS One*, *9*(3), e90444.

Vattikuti, S., Guo, J., & Chow, C. C. (2012). Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits (PM Visscher, Ed.). *Plos Genetics*, *8*, e1002637–e1002638. https://doi.org/10.1371/journal.pgen.1002637

Visscher, P. M., & Goddard, M. E. (2015). A general unified framework to assess the sampling variance of heritability estimates using pedigree or marker-based relationships. *Genetics*, *199*, 223–232. https://doi.org/10.1534/genetics.114.171017

Visscher, P. M., Hemani, G., Vinkhuyzen, A. A. E., Chen, G. B., Lee, S. H., Wray, N. R., … Yang, J. (2014). Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples (GS Barsh, Ed.). *Plos Genetics*, *10*, e1004269. https://doi.org/10.1371/journal.pgen.1004269

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era—Concepts and misconceptions. *Nature Reviews Genetics*, *9*, 255–266. https://doi.org/10.1038/nrg2322

Wenzel, M. A., James, M. C., Douglas, A., & Piertney, S. B. (2015). Genome-wide association and genome partitioning reveal novel genomic regions underlying variation in gastrointestinal nematode burden in a wild bird. *Molecular Ecology*, *24*, 4175–4192. https://doi.org/10.1111/mec.13313

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., … Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*, 565–569. https://doi.org/10.1038/ng.608

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*, 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Yang, J., Manolio, T. A., Pasquale, L. R., Boerwinkle, E., Caporaso, N., Cunningham, J. M., … Visscher, P. M. (2011). Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics*, *43*, 519–525. https://doi.org/10.1038/ng.823

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

SUPPORTING INFORMATION

Additional supplemental material may be found online in the Supporting Information section at the end of the article.

How to cite this article: Perrier C, Delahaie B, Charmantier A. Heritability estimates from genomewide relatedness matrices in wild populations: Application to a passerine, using a small sample size. *Mol Ecol Resour*. 2018;18:838–853. https://doi.org/10.1111/1755-0998.12886