

## ORIGINAL ARTICLE

# Conserved G-matrices of morphological and life-history traits among continental and island blue tit populations

B Delahaie<sup>1</sup>, A Charmantier<sup>1</sup>, S Chantepie<sup>2</sup>, D Garant<sup>3</sup>, M Porlier<sup>3</sup> and C Teplitsky<sup>1</sup>

The genetic variance–covariance matrix (**G**-matrix) summarizes the genetic architecture of multiple traits. It has a central role in the understanding of phenotypic divergence and the quantification of the evolutionary potential of populations. Laboratory experiments have shown that **G**-matrices can vary rapidly under divergent selective pressures. However, because of the demanding nature of **G**-matrix estimation and comparison in wild populations, the extent of its spatial variability remains largely unknown. In this study, we investigate spatial variation in **G**-matrices for morphological and life-history traits using long-term data sets from one continental and three island populations of blue tit (*Cyanistes caeruleus*) that have experienced contrasting population history and selective environment. We found no evidence for differences in **G**-matrices among populations.

Interestingly, the phenotypic variance–covariance matrices (**P**) were divergent across populations, suggesting that using **P** as a substitute for **G** may be inadequate. These analyses also provide the first evidence in wild populations for additive genetic variation in the incubation period (that is, the period between last egg laid and hatching) in all four populations. Altogether, our results suggest that **G**-matrices may be stable across populations inhabiting contrasted environments, therefore challenging the results of previous simulation studies and laboratory experiments.

*Heredity* (2017) **119**, 76–87; doi:10.1038/hdy.2017.15; published online 12 April 2017

## INTRODUCTION

A central goal in evolutionary biology is to predict if and how natural populations will adapt to changing environmental conditions. This has become increasingly important in the context of global changes imposing strong selection on wild populations (Palumbi, 2001). The ability of a population to respond to selection will not only depend on the quantity of genetic variation for each adaptive phenotypic trait but also on the strength and direction of genetic correlations among traits (Lande, 1979; Arnold *et al.*, 2001). The genetic architecture of traits linked by various genetic processes is summarized by the genetic variance–covariance matrix, also known as the **G**-matrix (or **G**), that corresponds to a symmetric matrix with additive genetic variance for the different traits contained in the diagonal of the matrix and the genetic covariances between traits in its off-diagonal elements (Lande, 1979; Lynch and Walsh, 1998). Depending on the direction of selection, genetic covariances may facilitate, constrain or even prevent evolution (Arnold, 1992; Agrawal and Stinchcombe, 2009; Morrissey *et al.*, 2012; Teplitsky *et al.*, 2014b). As such, **G** plays a central role in modern quantitative genetic theory dedicated to the evolution of phenotypic traits, as it is one of the main elements of the multivariate breeder's equation widely used for predicting phenotypic evolution (Lande, 1979). However, so far, attempts to understand adaptive responses in wild populations based on such quantitative genetics estimates have yielded equivocal results (Merilä *et al.*, 2001; Kruuk *et al.*, 2008; Walsh and Blows, 2009). Among the potential limitations to this approach is the general assumption that **G** remains constant over time and/or space (Turelli, 1988), yet we still have too few

temporal and spatial replicates to properly understand the extent of variation of **G** in nature (Arnold *et al.*, 2008; Aguirre *et al.*, 2014).

Using simulation-based approaches, previous studies have shown that several processes influence the constancy of **G**. First, genetic drift is expected to affect all aspects of the structure of **G**, such as the quantity of genetic variance for each trait and the genetic correlations between traits defining the geometry of **G** (that is, its size, shape and orientation), especially in small and/or isolated populations (Turelli, 1988; Jones *et al.*, 2003; Griswold *et al.*, 2007). Second, directional and/or stabilizing selection is expected to trigger nonproportional changes in the structure of **G** (Roff, 2000), whereas correlational selection should progressively modify the orientation of **G**-matrices until they are aligned with the adaptive landscape (Jones *et al.*, 2007; Revell, 2007). Third, high migration rates are expected to homogenize all aspects of **G** (Guillaume and Whitlock, 2007).

Using controlled laboratory experiments, the effects of these processes have been confirmed to modify the structure of **G** (see, for example, Phillips *et al.*, 2001 for genetic drift and Careau *et al.*, 2015 for selection). Therefore, wild populations experiencing different environments, and thus different selective pressures, or having undergone different demographic history, can be expected to differ in terms of **G**-matrices. At the univariate level, several studies have indeed shown that additive genetic variance is sometimes variable across populations (see, for example, Husby *et al.*, 2010; reviewed in Charmantier and Garant, 2005). However, the rare empirical studies investigating the spatial stability of the **G**-matrix in natural populations have yielded equivocal results: whereas some studies have found

<sup>1</sup>Centre d'Ecologie Fonctionnelle et Evolutive, CNRS-UMR5175 CEFE, Montpellier, France; <sup>2</sup>Laboratoire d'Écologie Alpine, Université Grenoble Alpes, Unité Mixte de Recherche 5533 CNRS, Grenoble, France and <sup>3</sup>Département de biologie, Université de Sherbrooke, Sherbrooke, Quebec, Canada

Correspondence: Dr B Delahaie, Centre d'Ecologie Fonctionnelle et Evolutive, CNRS-UMR5175 CEFE, 1919, Route de Mende, 34293 Montpellier Cedex 5, France.

E-mail: borisdelaiaie@gmail.com

Received 2 September 2016; revised 25 January 2017; accepted 26 January 2017; published online 12 April 2017

relatively stable G-matrices among populations (see, for example, Arnold and Phillips, 1999; Ashman, 2003; Puentes *et al.*, 2016; reviewed in Arnold *et al.*, 2008; see Garant *et al.*, 2008 for a temporal stability example in a wild population), others have found that G-matrices could vary among populations (see, for example, Roff *et al.*, 2004; Johansson *et al.*, 2011; reviewed in Wood and Brodie, 2015; see Björklund *et al.*, 2013 for a temporal variation example in a wild population).

Most of these studies used individuals from different populations subsequently bred under artificial laboratory conditions to estimate quantitative genetic parameters and thus very few have been conducted directly in wild populations. To our knowledge, Roff *et al.* (2004) is the only study that has investigated the variability of G among populations and where the phenotypes have been measured on individuals in their natural environments. This scarcity may be problematic given that there is growing evidence that laboratory experiments may not be generalizable to wild populations that are experiencing true natural conditions (Kruuk *et al.*, 2008; Charmantier *et al.*, 2014; Sniegula *et al.*, 2016). Compared with laboratory experiments, the drawback of working with wild populations is that genotype by environment (G × E) interactions can confound tests of differences in G-matrices when estimated in the wild (Wood and Brodie, 2015). However, in the wild, the estimated evolutionary potential is expressed in the environment where the individuals are selected so that comparing G-matrices in populations differing in their adaptive optimum offers the opportunity to test whether different selection pressures trigger differentiation in G-matrices (McGuigan, 2006; Johansson *et al.*, 2011).

Overall, information about the spatial variability of evolutionary potential in wild populations is still very scarce. Therefore, there is now a crucial need for empirical studies aiming at understanding whether the genetic architecture of traits is actually variable in the wild and at which spatial scale (Steppan *et al.*, 2002; Arnold *et al.*, 2008; Aguirre *et al.*, 2014). Because of the very demanding nature of estimating G-matrices in the wild, substituting G by phenotypic variance–covariance matrices (P), which are much easier to estimate, is a common practice (see, for example, Berner *et al.*, 2008). However, this substitution is useful only if P correctly reflects the underlying genetic architecture, and this may not always be the case (Willis *et al.*, 1991; Hadfield *et al.*, 2007). Furthermore, although comparison tools exist, until recently there was no statistical framework available to compare more than two matrices at the same time and to integrate the uncertainties around G-matrix estimates that are often large (Aguirre *et al.*, 2014). This stresses the need for more empirical studies in the wild aiming at reporting G variability for different kinds of traits and its consequences on the evolutionary potential of populations.

Here, we assess the extent and the nature of spatial variation in the G-matrix for a set of morphological and reproductive traits in four wild populations displaying strong phenotypic differentiation. We use long-term data collected from four Mediterranean blue tit (*Cyanistes caeruleus*) populations ideally suited for quantitative genetic studies because of the availability of extensive pedigree information and phenotypic data. More specifically, we estimate G-matrices for four morphological traits representing body size and shape (tarsus, wing and beak length as well as body mass) and for three reproductive traits (laying date, clutch size and incubation period). We focus particularly on incubation period, as, to our knowledge, the present study is only the second one to test for genetic variance associated with this trait in natural populations (Husby *et al.*, 2012).

Our study populations are located in a typical Mediterranean landscape, in southern France mainland and on the island of Corsica,

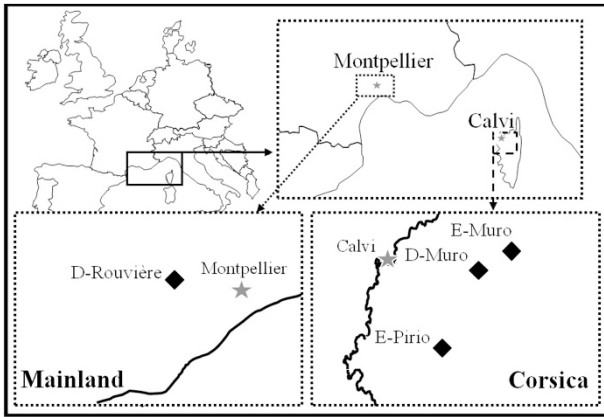
composed of patchy mosaics of heterogeneous habitats, with forest covers either dominated by deciduous oaks (*Quercus pubescens*) or evergreen oaks (*Quercus ilex*). This habitat variation translates into marked phenotypic differences in morphological and life-history traits (Blondel *et al.*, 2006; Charmantier *et al.*, 2016). For instance, in the evergreen habitat, birds lay fewer eggs and up to 1 month later in spring compared with birds in deciduous forests. Moreover, birds from the mainland are smaller (ca. 15%) than birds from Corsica (Charmantier *et al.*, 2016). These differences are likely because of differences in selective pressures between the different habitats. In this respect, Porlier *et al.* (2012a) showed that selection is significantly different among populations for clutch size and laying date (mainly in the strength of the linear selection gradients). Charmantier *et al.* (2004b) also showed that selection on chick morphology differs between sites (also see Charmantier *et al.*, 2016 for a thorough review of evidence for local adaptation). Based on mitochondrial genetic data and morphological data, Corsican populations have been attributed to a different subspecies *Cyanistes caeruleus ogliastreae*, occurring as well in the Iberian Peninsula, whereas the mainland population belongs to the subspecies *C. caeruleus caeruleus* (Kvist *et al.*, 2004). Population genetic analyses have recently confirmed the limited gene flow between mainland and Corsica (Porlier *et al.*, 2012b; Szulkin *et al.*, 2016).

Taking advantage of this well-defined ecospatial context, this system provides an excellent setting to evaluate the effects of contrasted habitats and population history on the G-matrix and their consequences on population evolutionary potential. First, we seek to estimate G in each population for the two sets of traits (morphology and reproduction) using Bayesian multivariate animal models and to describe precisely the differences observed between populations. To do so, we use a fourth-order genetic covariance tensor analysis (Hine *et al.*, 2009), a method that can detect subtle differences in variance–covariance structure among a set of matrices. Second, we aim to explore whether the potential differences are explained by the contrast in habitat and/or by geographic distance (mainland vs island). We predict that G-matrices may have diverged between mainland and island because of limited gene flow and different population histories (that is, neutral processes). We also predict that G-matrices may differ between evergreen and deciduous habitats because of divergent selection pressures in these two environments. However, the existing theoretical framework does not provide any quantitative predictions regarding the direction and magnitude of these differences. We also aim to assess whether these differences are dependent on the suite of traits studied (life history vs morphology). Following Arnold *et al.* (2001), we predict stronger divergences between life-history G-matrices than morphological G-matrices as the former traits may experience stronger and more variable selective pressures. Finally, we compare P-matrices among populations for the two sets of characters to assess what would have been inferred if the data prevented us from estimating G-matrices.

## MATERIALS AND METHODS

### Study system

Data were collected in four blue tit populations, with landscapes dominated either by the deciduous downy oak *Q. pubescens* (site names starting with D-) or by the evergreen holm oak *Q. ilex* (site names starting with E-). One study site was located in southern France (D-Rouvière), and the other three sites were located in north-western Corsica (E-Pirio, D-Muro and E-Muro) with 6 to 30 km separating them and distant to the mainland site by 440 km (Figure 1). All populations were monitored as part of a long-term research programme for 17 to 27 years (see Table 1 for sampling years). Capture and handling of the birds was conducted under permits provided by the Centre de Recherches sur



**Figure 1** Blue tit study sites in the Mediterranean region. D-Rouvière population is located in southern France, and E-Pirio, E-Muro and D-Muro in north-western Corsica, ~440 km from D-Rouvière. In Corsica, E-Muro is separated from D-Muro by ~6 km, and from E-Pirio by ~30 km.

la Biologie des Populations d'Oiseaux (CRBPO) and by the Direction Départementale des Services Vétérinaires (DDSV).

### Phenotypic data

Each year, nest boxes were visited at least once a week during the reproductive period, where data on laying date (date of first egg laid with 1 = 1st March) and clutch size (number of eggs) were collected. Incubation period was calculated as the number of days between the laying of the last egg and the hatching date of the first chick. Thus, what we call incubation period does not correspond to the physiological incubation time *sensu stricto*. This period can be longer than the physiological incubation time (14 days, Vedder, 2012) if the female adds a waiting time after her last egg is laid (typically in a relatively cold spring); it can also be shorter if the female begins to incubate before the last egg is laid (typically in warm springs). In these populations, most females lay only one clutch per year (0 to 1.5% of females lay second clutches in our populations), and upon breeding failure replacement clutches are sometimes laid. Second and replacement clutches were excluded from the data set. Breeding blue tits were captured in nest boxes during the feeding of their young, banded with a unique metal ring provided by the CRBPO. Nestlings were also uniquely banded before fledging. Four different morphological traits were measured on male and female breeders: body mass, tarsus length (from the intertarsal joint to the most distal undivided scute on the tarsometatarsus), flattened wing length and bill length (from the anterior end of the nares to the tip of the upper mandible). Phenotypic differences among the different populations as well as their locations are presented in Figure 1 and Table 1 (see Charmanter *et al.*, 2016 for a thorough review of phenotypic divergence in this system).

### Pedigree construction

Pedigrees were constructed in a similar way for each population. We first included all ringed individuals and assigned their mother and father based on observational data (caught in the nest box while feeding the brood). For a subset of individuals ( $N=161$ ), molecular assignation of the biological father using microsatellite data was performed (Charmanter *et al.*, 2004a), and observational pedigree data were corrected accordingly. In populations of this study system, 14 to 25% of nestlings are sired by an extra-pair mate (Charmanter and Blondel, 2003; Charmanter *et al.*, 2004a), a proportion that could slightly underestimate but is unlikely to greatly bias our genetic variance estimates considering the sample sizes involved (Charmanter and Réale, 2005; Firth *et al.*, 2015). Unknown parents of a given nest were coded using a dummy identity in the pedigree to preserve sibship information. Pedigrees were pruned for each data set studied to retain only phenotyped individuals and their ancestors using the *pedantics* R-package (Morrissey and Wilson, 2010). Pruned pedigree information for each population and trait type is presented in Table 1.

### Quantitative genetic analyses

**G- and P-matrix estimation.** G-matrices were estimated in each population by using multivariate animal models (Lynch and Walsh, 1998; Kruuk, 2004). Two different G-matrices were constructed for each population: one for each type of traits (morphological and life-history traits). Random effects included additive genetic effects ( $V_A$ ) estimated through the inclusion of multi-generational pedigree data, permanent environmental effects ( $V_{PE}$ ) accounting for repeated measurements of the same individual (see Table 1 for details), year effect ( $V_Y$ ) and residual variance ( $V_R$ ). In addition, for morphological traits, we added a random effect for measurer identity in order to control for any potential confounding measurer effect ( $V_{OBS}$ , the number of measurer varying between 15 and 18 depending on the locality). For morphological models, all parents caught during the breeding period were included in the analyses, whereas models for life-history traits were female specific. Several fixed effects (age, sex and ordinal date) were added to account for potentially confounding effects (see Supplementary Appendix 1 and Supplementary Table S1 for more details). Ordinal date was included in order to remove the effect of potential trait changes within year (for example, seasonal variation of body mass or variation of wing length due to feather abrasion; Supplementary Appendix 1).

As the main purpose of our study was matrix comparison for several traits, we attached particular attention to trait standardization as scale differences between variables may potentially affect our conclusions (Hansen *et al.*, 2011). Trait standardization is essentially done to remove any effects of differential measurement scale among traits and to compare traits with different orders of magnitude. All but one trait analysed were on a ratio scale (such as measurements and counts, that is, positive real numbers that are multipliable by real numbers without affecting the relationships among the measurements and possess a unique and nonarbitrary 0); laying date was on an interval scale (that is, a variable for which there is no natural 0). Transformations that preserve relationships among the measurements for variables on a ratio scale are multiplications by a constant. For variables on an interval scale, addition of a constant and multiplication by a constant are both permissible transformations (Houle *et al.*, 2011). We chose to apply mean standardization within each population for all the phenotypic variables we used (that is, dividing raw trait values by the population mean) as variance scaling is not suitable when one is interested in evolutionary potential (Houle, 1992; Hansen and Houle, 2008; Houle *et al.*, 2011; Hansen *et al.*, 2011). Thus, mean scaling removes the disproportionate effect of traits with larger means, thereby avoiding multiple problems of interpretation (Houle *et al.*, 2011; Hansen *et al.*, 2011). Laying date was the only variable for which mean standardization was not relevant, and hence we accounted for differences in mean laying date between populations by subtracting the median laying date of each population from the raw values and then added 30 days in order to avoid negative values. This removed differences in variance only due to mean differences between populations while still allowing comparison with other phenotypic traits having a true natural zero.

The models we used can be described as follow:

$$Y_{\text{morpho}} = \mu + Xb + Z_a a + Z_{\text{pe}} pe + Z_{\text{yr}} yr + Z_{\text{obs}} obs + e \quad (1a)$$

$$Y_{\text{LHT}} = \mu + Xb + Z_a a + Z_{\text{pe}} pe + Z_{\text{yr}} yr + e \quad (1b)$$

Equation (1a) describes the animal models run on morphological traits with  $Y_{\text{morpho}}$  the vector of mean standardized morphological observations for all individuals and  $\mu$  the vector of mean phenotypes.  $X$  is the design matrix relating to fixed effects, and  $b$  is the vector of fixed effects to be fitted.  $Z_a$ ,  $Z_{\text{pe}}$ ,  $Z_{\text{yr}}$  and  $Z_{\text{obs}}$  are the incidence matrices relating to additive genetic (a), permanent environment (pe), year (yr) and measurer (obs) random effects, respectively, and  $e$  is the vector of residual error. Equation (1b) describes the life-history animal models with  $Y_{\text{LHT}}$  the vector of mean standardized life-history traits observations for all individuals and  $\mu$  the vector of mean phenotypes.  $X$  is the design matrix relating to fixed effects, and  $b$  is the vector of fixed effects to be fitted.  $Z_a$ ,  $Z_{\text{pe}}$  and  $Z_{\text{yr}}$  are the incidence matrices relating to additive genetic (a), permanent environment (pe) and year (y) random effects, respectively, and  $e$  is the vector of residual error.

All these analyses were performed in a Bayesian framework using the *MCMCgmm* R-package (Hadfield, 2010). As the use of posterior distributions propagates the errors in all estimates derived from animal models, Bayesian

**Table 1** Summary of the characteristics of the four blue tit populations: sample years, means, s.d. values and number of records for each trait, pruned pedigree depth (number of generations) and sample size

Population	Sample years	Average trait value (s.d.), number of records										Pruned pedigree depth		Pruned pedigree sample size	
		Wing length (mm)	Body mass (g)	Tarsus length (mm)	Bill length (mm)	Incubation period (days)	Laying date (1 = 1st March)	Clutch size	Morphology	Life history	Morphology	Life history			
D-Rouvière	1991–2014	65.80 (2.18) 2841	11.01 (0.64) 2869	16.70 (0.50) 2896	6.62 (0.37) 2655	14.53 (2.50) 1159	38.89 (7.64) 1168	10.09 (1.65) 1166	14	13	14	13	1683	1047	
E-Pirio	1988–2014	62.18 (2.03) 2160	9.27 (0.48) 2201	16.05 (0.49) 2209	6.59 (0.35) 1856	13.82 (1.82) 1048	70.69 (6.76) 1049	6.59 (1.20) 1049	13	14	13	14	1306	795	
E-Muro	1998–2014	62.13 (1.92) 634	9.54 (0.45) 635	16.19 (0.52) 607	6.74 (0.34) 638	13.13 (1.58) 406	47.60 (7.38) 407	7.08 (1.27) 407	7	7	7	7	471	312	
D-Muro	1993–2014	62.03 (2.05) 1637	9.71 (0.48) 1625	16.28 (0.51) 1588	6.66 (0.40) 1632	13.62 (1.93) 992	37.92 (8.12) 995	8.59 (1.58) 995	7	7	7	7	1240	820	

inference has a clear advantage over the other existing methods (Morrissey *et al.*, 2014). Posterior distributions were composed of 1000 values for each parameter. We used 1 200 000 iterations per animal model with sampling every 1000 steps. The 200 000 first iterations were discarded as burn-in. To facilitate

convergence, we used slightly informative priors with  $V = \begin{bmatrix} \frac{V_p}{r} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{V_p}{r} \end{bmatrix}$

and  $\text{nu} = n$  with  $V_p$  being the phenotypic variance,  $r$  the number of random factors and  $n$  the number of traits. We assessed convergence of the models by graphically checking the posterior estimates and ensuring that autocorrelation of all the parameter estimates was lower than 0.05.

**G**-matrices thus correspond to the variance–covariance matrix associated with the additive genetic random effect. **P**-matrices were calculated as the sum of all the variance–covariance matrices associated with the different random effects but the observer effect, that is, summing  $a$ ,  $p_e$ ,  $y_r$  and  $e$  in equations (1a and 1b). Heritabilities and the proportion of phenotypic covariances explained by genetic covariances were estimated by standardizing the **G**-matrix by the **P**-matrix, that is, result of the product  $\mathbf{P}^{-1/2} \mathbf{G} \mathbf{P}^{-1/2}$ .

**Matrix description: size and shape.** The additive genetic variance associated with each trait was characterized with  $I_A$  evolvabilities.  $I_A$  estimates can be interpreted as the expected percentage of trait change per generation if it were submitted to selection as strong as on fitness itself (Hansen and Houle, 2008). In order to describe accurately every estimated **G**- and **P**-matrices, we also computed a series of descriptors of their size and shape. Total volume ( $V_{\text{tot}}$ ,  $V_{\text{tot,G}}$  and  $V_{\text{tot,P}}$  for **G** and **P**, respectively), a measure of the variance contained in each matrix, was calculated as the sum of the eigenvalues (Equation (2)). Eccentricity ( $\Omega$ ,  $\Omega_G$  and  $\Omega_P$  for **G** and **P**, respectively), a measure of the shape of the matrix, was calculated as the ratio of the first eigenvalue over the sum of all eigenvalues (Equation (3), Jones *et al.*, 2003).

$$V_{\text{tot}} = \sum_{i=1}^n \lambda_i \quad (2)$$

$$\Omega = \frac{\lambda_1}{V_{\text{tot}}} \quad (3)$$

with  $\lambda_i$  the  $i$ th eigenvalues of a given **G**- or **P**-matrix, and  $n$  the number of traits. To assess the credible interval of the shape and size metrics, we estimated these descriptors for each of the 1000 estimates that compose the distribution of **G** and **P**. We then computed the mode of their distribution and the 95% credible interval calculated as the 95% highest posterior density interval with the R-package (MCMCglmm; Hadfield, 2010).

**Comparing G-matrices among populations.** Comparison of **G**-matrices has long been a statistical challenge and numerous methods have been applied (Steppan *et al.*, 2002; Roff *et al.*, 2012; Aguirre *et al.*, 2014). However, most of these methods were not directly related to evolutionary theory (Hansen and Houle, 2008; Aguirre *et al.*, 2014). Methods were usually designed to compare matrices element by element or based on criteria that were hardly interpretable in terms of change of genetic variance or response to selection (Aguirre *et al.*, 2014). Here we used the fourth-order genetic covariance tensor method (Hine *et al.*, 2009), recently highlighted by Aguirre *et al.* (2014) as it is directly related to the evolutionary consequences of divergence in **G** for populations, can detect subtle differences between variance–covariance matrices and allows to compare more than two matrices at the same time and to integrate uncertainty in estimates.

Tensors are geometric objects that describe linear relationships between scalars, vectors, matrices and other tensors. The order of a tensor reflects the number of indices required to define its elements (De Lathauwer *et al.*, 2004). For example, a linear map is represented by a matrix (a two-dimensional array), and therefore is a second-order tensor. Here, in order to describe the variation among several additive genetic variances, two indices are required (for example, the genetic covariance between trait  $x$  and  $y$  is denoted  $\mathbf{G}_{xy}$ ). The **G**-matrix is thus a second-order tensor. Multilinear algebra extends tensors to higher order structures that can be used to describe variation in lower order structures such as matrices. The complete description of the variation among **G**-matrices requires to estimate the relationships between the (co)variances among



populations, that is, the (co)variances between genetic (co)variances, for example, the variances among populations of the covariances between traits  $x$  and  $y$  ( $\mathbf{G}_{xy}$ ) or the covariance of  $\mathbf{G}_{xy}$  and  $\mathbf{G}_{zw}$  (the covariance between traits  $z$  and  $w$ ). Such elements are denoted  $\Sigma_{\mathbf{G}_{xy}, \mathbf{G}_{zw}}$  and require the use of four indices. Therefore, the variation among G-matrices (that is, second-order tensors) can be described using a fourth-order genetic covariance tensor ( $\Sigma$ ) (Hine *et al.*, 2009).

To simplify  $\Sigma$ , it can be represented as a second-order tensor that corresponds to a covariance matrix ( $\mathbf{S}$ ) of dimension  $n(n+1)/2$  with  $n$  the number of traits (Basser and Pajevic, 2007). The eigen-analysis of the  $\mathbf{S}$ -matrix mirrors the spectral decomposition of  $\Sigma$  and describes accurately variation among the different G-matrices using eigenvalues and second-order eigentensors ( $\mathbf{E}$ ). These eigentensors are higher level equivalent of eigenvectors for matrices and describe independent aspects of variation between the different G-matrices and their eigenvalues reflect the amount by which an eigentensor contributes to variation among G-matrices. These independent eigentensors may be interpreted as biologically relevant axes of differentiation representing different sources of selective pressures for example (Hine *et al.*, 2009). Large values in the diagonal elements but small values in the off-diagonal elements of  $\mathbf{E}$  would indicate differences among G mainly in the quantity of variance for individual traits, the relative magnitude of these values indicating the relative contribution of each trait to these differences. On the other hand, the magnitude of the off-diagonal elements of  $\mathbf{E}$  indicates differences in trait covariances between G-matrices. Practically, each second-order eigentensor can be decomposed into its eigenvalues and eigenvectors ( $e$ ) that describe genetically independent linear combinations of traits that show differences in genetic variance and covariance. The distribution of the eigenvalues of an eigentensor indicates the number of independent genetic directions in which the G-matrices differ (within the space of independent change represented by the eigentensor). For example, if the first eigenvector of an eigentensor accounts for most of the variance (its eigenvalue is large compared with the others), the G-matrices differ likely only in a single combination of traits that corresponds to the trait loadings on this eigenvector. We strongly encourage the reader to refer to Hine *et al.* (2009) for more details on theoretical aspects of the genetic covariance tensor method.

We applied the fourth-order genetic covariance tensor in a Bayesian framework following Aguirre *et al.* (2014). To do so, we simply estimated the  $\mathbf{S}_i$  matrix for every  $i$ th sample ( $i$  varied from 1 to 1000) of the set of posterior distributions for the different G-matrices. We then calculated the posterior modes ( $\bar{\mathbf{S}}$ ) of the elements of  $\mathbf{S}_i$  and their associated eigentensors. We then projected every  $i$ th sample of  $\mathbf{G}$  on the different eigentensors to obtain the distribution of the position of the different G-matrices on each eigentensor. We encourage the reader to refer to Aguirre *et al.* (2014) and Careau *et al.* (2015) for concrete examples and R scripts to use these methods. In order to test for differences between G-matrices, we tested whether the credible interval of their positions along the eigentensors accounting for most of the total variance overlapped. In order to explore the robustness of these results we used several confidence intervals (0.95, 0.925 and 0.9). As complementary approaches, to assess the robustness of our results to different statistical methods, we also tested these differences using a method based on the generation of null matrices developed by Aguirre *et al.* (2014) (see Supplementary Appendix 2) and a method developed by Ovaskainen *et al.* (2008) (see Supplementary Appendix 3).

In order to compare the evolutionary potential of each population, we also calculated the multivariate average unconditional and conditional evolvabilities for the different estimated G-matrices following the approach of Hansen and Houle (2008). Univariate unconditional evolvability is the ability of a trait to respond to selection. It thus mainly depends on the standing variation available in the population and corresponds to the expected evolutionary response per generation in the direction of a linear selection gradient of unit strength (Hansen and Houle, 2008). The multivariate version of unconditional evolvability corresponds to the amount of predicted evolutionary response occurring in the direction of selection. To obtain an overall estimation of the evolutionary potential associated with the G-matrix, average unconditional evolvability over random selection gradients can be calculated following Hansen

and Houle (2008):

$$\bar{e} = E[\lambda] \quad (4)$$

where  $E[\lambda]$  denotes the expectation of  $\lambda$  (corresponding to the eigenvalues of  $\mathbf{G}$ ).

As the average unconditional evolvability  $\bar{e}$  is unaffected by covariances between traits, we also calculated the average conditional evolvability  $\bar{c}$ . The conditional evolvability of a trait corresponds to the response of this trait to a unit directional selection when a set of constraining (that is, covarying) traits is not allowed to change. When averaged over all directions in the phenotypic space, it provides a global measure of genetic constraints (low constraint for high value of  $\bar{c}$ ). It can be calculated following Hansen and Houle (2008):

$$\bar{c} \approx H[\lambda] + (1 + \frac{I[1/\lambda]}{k+2}) \quad (5)$$

where  $H[\lambda]$  is the harmonic mean of  $\lambda$  and  $I[x] \equiv \text{var}[x]/E[x]^2$  denotes the mean-standardized variance.

These different quantities and their associated credible intervals were compared among populations for each type of traits.

**Are P-matrices different among populations?** In order to know what would have been inferred if the pedigree data to estimate G-matrices was not available, we used the covariance tensor analysis on P-matrices (for the two set of characters) to test whether they differed between populations. The rationale of this analysis is that finding homogeneous P-matrices across populations is often interpreted as an indication that G-matrices are probably homogeneous as well (Arnold *et al.*, 2001). Besides, we also estimated the eigenvalues of  $\mathbf{P}^{-1/2}\mathbf{G}\mathbf{P}^{-1/2}$  for life-history and morphological traits for each population. A flat distribution of the eigenvalues of  $\mathbf{P}^{-1/2}\mathbf{G}\mathbf{P}^{-1/2}$  indicates that all regions of the phenotypic space are equally heritable and the G- and P- matrices are proportional, whereas a skew in the distribution of the eigenvalues indicates that there are regions of phenotypic space that are more heritable than others, so that P and G are not proportional.

## RESULTS

### G-matrices among trait types and populations

We found substantial levels of additive genetic variation (posterior distributions disjunct from 0, see Supplementary Figure S1 for details on distributions) in all of the 28 trait/population combinations (Tables 2a and 3). All morphological traits displayed lower levels of  $I_A$  evolvabilities (ranging from 0.012 to 0.095; Table 2a) than life-history traits (ranging from 0.109 to 1.533; Table 3). As a result, G-matrices for life-history traits displayed substantially higher  $V_{\text{tot}}$  (ranging from 1.457 to 3.242; details in Supplementary Table S2) than G-matrices for morphological traits (ranging from 0.177 to 0.232; details in Supplementary Table S2). However, when converted into heritabilities (that is, divided by phenotypic variance estimates), values for morphological traits were on average higher than for life-history ones (Tables 2a and 3). In terms of trait associations, the posterior modes of genetic covariances between morphological traits were all positive and presented similar values in each population (Table 2b, see Supplementary Table S3a for genetic correlation estimates). All genetic covariances were significant for D-Rouvière and E-Muro, except for genetic covariances between bill length and the other morphological traits in E-Muro. For E-Pirio and D-Muro, all genetic covariances between bill length and the other morphological traits were positive but only significant with body mass. For life-history traits, the genetic covariance between laying date and clutch size was negative in all populations but significant only for D-Muro (Table 3; see Supplementary Table S3b for genetic correlation estimates). All other genetic covariances for life-history traits were nonsignificant. In addition, estimates of the other variance components are provided in Supplementary Table S4.

**Table 2a** Posterior modes of mean standardized traits genetic and phenotypic variances (estimated variance  $\times 100$ , corresponding to  $I_A$  evolvabilities for genetic ones) for morphological G- and P-matrices with their 95% credible interval for the four blue tit populations

	Wing length	Body mass	Tarsus length	Bill length
<i>D-Rouvière</i>				
<b>G</b>	0.018 (0.015; 0.022)	0.095 (0.075; 0.112)	0.040 (0.031; 0.048)	0.075 (0.062; 0.097)
<b>P</b>	0.048 (0.045; 0.054)	0.292 (0.264; 0.317)	0.066 (0.062; 0.072)	0.216 (0.202; 0.235)
<b>P<sup>-1/2</sup>GP<sup>-1/2</sup></b>	0.331 (0.284; 0.399)	0.293 (0.242; 0.359)	0.571 (0.475; 0.648)	0.341 (0.278; 0.414)
<i>E-Pirio</i>				
<b>G</b>	0.012 (0.009; 0.018)	0.092 (0.058; 0.119)	0.037 (0.025; 0.052)	0.041 (0.027; 0.06)
<b>P</b>	0.050 (0.047; 0.056)	0.267 (0.25; 0.294)	0.084 (0.076; 0.091)	0.229 (0.207; 0.253)
<b>P<sup>-1/2</sup>GP<sup>-1/2</sup></b>	0.201 (0.142; 0.301)	0.320 (0.201; 0.392)	0.427 (0.294; 0.565)	0.181 (0.111; 0.247)
<i>E-Muro</i>				
<b>G</b>	0.019 (0.011; 0.026)	0.079 (0.042; 0.113)	0.05 (0.021; 0.067)	0.047 (0.022; 0.08)
<b>P</b>	0.053 (0.047; 0.066)	0.229 (0.192; 0.265)	0.092 (0.078; 0.108)	0.221 (0.198; 0.279)
<b>P<sup>-1/2</sup>GP<sup>-1/2</sup></b>	0.319 (0.203; 0.431)	0.309 (0.165; 0.432)	0.451 (0.246; 0.644)	0.195 (0.09; 0.319)
<i>D-Muro</i>				
<b>G</b>	0.018 (0.012; 0.024)	0.057 (0.027; 0.083)	0.04 (0.03; 0.054)	0.065 (0.035; 0.085)
<b>P</b>	0.058 (0.051; 0.067)	0.240 (0.218; 0.261)	0.082 (0.074; 0.091)	0.242 (0.223; 0.279)
<b>P<sup>-1/2</sup>GP<sup>-1/2</sup></b>	0.274 (0.173; 0.37)	0.221 (0.102; 0.308)	0.509 (0.39; 0.633)	0.211 (0.133; 0.324)

Posterior modes of heritabilities calculated as the diagonal elements of  $\mathbf{P}^{-1/2}\mathbf{G}\mathbf{P}^{-1/2}$  are reported with their 95% credible interval for each trait and population.

All morphological traits loaded in the same direction on the first eigenvector of  $\mathbf{G}$ , with body mass having the strongest statistical weight (Supplementary Table S5). For morphological G-matrices, eccentricity (that is, the proportion of variance accounted for by the first eigenvector of a matrix) ranged from 55.6 to 66.4% (Supplementary Table S2). For life-history traits, laying date and clutch size had opposed loadings on the first eigenvector (Supplementary Table S5). Incubation period had little influence on the first eigenvector. Life-history G-matrices showed similar eccentricity values to those obtained for morphological traits (ranging from 55 to 81.1% for life-history G; Supplementary Table S2). As eccentricity corresponds to the quantity of variance contained in the main axis of variation over the total variance contained in the matrix, it provides an indication on the shape of the G-matrix and is bounded between  $V_{\text{tot}}/n$  (sphere shape) and 1 (line). Our values of eccentricity thus indicated ellipsoid G-matrices.

### Comparisons of G-matrices between populations

In the covariance tensor analyses, the first and second eigentensors explained 90% of the variance among matrices for life-history traits and 70% for morphological traits (Supplementary Table S6). Along the first eigentensors ( $\mathbf{E}_1$  and  $\mathbf{E}_2$ , respectively, the first and second eigentensors), the 95% confidence intervals of posterior distributions of the position of the G-matrices overlapped (Figures 2a and b). As confidence intervals are usually large with G-matrix estimates, we computed less stringent confidence intervals (0.925, 0.9) to test whether our results were robust to less stringent criteria. For morphological traits, we obtained the same results. For life-history traits, 90% confidence intervals did not overlap on  $\mathbf{E}_1$  between E-Pirio and D-Muro. However, note that their relative positions on the tensors were very close.

In line with this, evolutionary potential estimates indicated weak differentiation. The different measures of evolutionary potential from each G-matrix also largely overlapped between the different populations apart from conditional evolvability of the morphological G-matrix between E-Pirio and D-Rouvière (Figure 3a). This

may indicate that the G-matrices from D-Rouvière and E-Pirio were very weakly differentiated, a trend perceptible in Figure 2a (along  $\mathbf{E}_2$ , see also Supplementary Appendix 3 and Supplementary Figure S2). Moreover, using the method developed by Aguirre *et al.* (2014) to test differences with the tensor analysis, we did not detect any difference between G-matrices for both types of traits (Supplementary Appendix 2, Supplementary Table S7 and Supplementary Figure S3a). Therefore, the only slight differences we found were inconsistent across methods suggesting weak global pattern. Thus, we found little evidence for differences among G-matrices whatever the type of trait, the type of habitat or the geographic position of the population considered (mainland vs island). Note however that a valid test of an insular effect on the G-matrix would require at least replicate mainland populations. Apart from being nonsignificant with 95% confidence intervals, the differences in the posterior modes of the different metrics computed represent negligible differences in evolutionary potential. Indeed, although evolvability values showed that life-history traits harbour slightly more genetic variance in D-Muro, conditional evolvability values, a measure that takes the genetic architecture into account, were similar for the four blue tit populations (Figure 3).

### P-matrix comparison

At the phenotypic level, covariances between morphological traits were similar in terms of magnitude to the genetic covariances. However, in contrast with the nonsignificant genetic correlations, tarsus length and bill length phenotypes covaried significantly at the phenotypic level in all populations (Table 2b; see Supplementary Table S3a for correlation estimates). In addition, phenotypic covariances for life-history traits tended to be significant, especially for covariance involving clutch size (Table 3 and Supplementary Figure S4; see Supplementary Table S3b for genetic correlation estimates). For both types of traits, the distributions of the eigenvalues of  $\mathbf{P}^{-1/2}\mathbf{G}\mathbf{P}^{-1/2}$  were skewed indicating nonproportionality between  $\mathbf{G}$  and  $\mathbf{P}$  (Figure 4 and Supplementary Figure S5).

Table 2b Posterior modes of covariances ( $\times 100$ ) for morphological G, P and  $P^{-1/2}GP^{-1/2}$  matrices with their 95% credible interval for the four blue tit populations

	Wing length–body mass	Wing length–tarsus length	Wing length–bill length	Body mass–tarsus length	Body mass–bill length	Tarsus length–bill length
<i>D-Rouvière</i>						
<b>G</b>	<b>0.018 (0.011; 0.023)</b>	<b>0.006 (0.002; 0.01)</b>	<b>0.010 (0.004; 0.015)</b>	<b>0.021 (0.014; 0.033)</b>	<b>0.026 (0.019; 0.045)</b>	<b>0.010 (0.003; 0.02)</b>
<b>P</b>	<b>0.027 (0.018; 0.034)</b>	<b>0.009 (0.007; 0.014)</b>	<b>0.010 (0.004; 0.016)</b>	<b>0.037 (0.028; 0.045)</b>	<b>0.013 (0; 0.031)</b>	<b>0.015 (0.008; 0.021)</b>
<b><math>P^{-1/2}GP^{-1/2}</math></b>	0.055 (0.023; 0.111)	–0.005 (–0.049; 0.049)	0.026 (–0.007; 0.083)	0.065 (0; 0.106)	0.090 (0.053; 0.14)	0.013 (–0.044; 0.07)
<i>E-Pirio</i>						
<b>G</b>	<b>0.012 (0.005; 0.024)</b>	<b>0.010 (0.003; 0.015)</b>	0.005 (–0.002; 0.011)	<b>0.030 (0.015; 0.048)</b>	<b>0.027 (0.008; 0.043)</b>	0.012 (–0.001; 0.023)
<b>P</b>	<b>0.022 (0.017; 0.032)</b>	<b>0.016 (0.01; 0.018)</b>	0.006 (–0.001; 0.014)	<b>0.044 (0.034; 0.053)</b>	<b>0.031 (0.018; 0.05)</b>	<b>0.019 (0.013; 0.031)</b>
<b><math>P^{-1/2}GP^{-1/2}</math></b>	0.040 (–0.022; 0.103)	0.053 (–0.026; 0.125)	0.015 (–0.038; 0.066)	0.081 (0.008; 0.188)	0.084 (0.005; 0.124)	0.019 (–0.058; 0.09)
<i>E-Muro</i>						
<b>G</b>	<b>0.014 (0.001; 0.024)</b>	<b>0.010 (0.002; 0.02)</b>	0.012 (–0.003; 0.018)	<b>0.024 (0.005; 0.049)</b>	0.015 (–0.006; 0.042)	0.008 (–0.01; 0.03)
<b>P</b>	<b>0.018 (0.004; 0.033)</b>	<b>0.012 (0.001; 0.019)</b>	0.003 (–0.011; 0.018)	<b>0.036 (0.021; 0.054)</b>	0.023 (–0.003; 0.055)	<b>0.026 (0.006; 0.042)</b>
<b><math>P^{-1/2}GP^{-1/2}</math></b>	0.055 (–0.038; 0.133)	0.101 (–0.024; 0.178)	0.052 (–0.026; 0.145)	0.069 (–0.037; 0.185)	0.035 (–0.046; 0.131)	–0.007 (–0.104; 0.13)
<i>D-Muro</i>						
<b>G</b>	<b>0.012 (0.002; 0.022)</b>	<b>0.007 (0.002; 0.016)</b>	0.007 (–0.003; 0.015)	<b>0.020 (0.001; 0.032)</b>	<b>0.029 (0.007; 0.048)</b>	0.012 (–0.005; 0.022)
<b>P</b>	<b>0.022 (0.012; 0.031)</b>	<b>0.012 (0.007; 0.018)</b>	0.006 (–0.004; 0.017)	<b>0.037 (0.028; 0.047)</b>	<b>0.042 (0.024; 0.06)</b>	<b>0.019 (0.009; 0.031)</b>
<b><math>P^{-1/2}GP^{-1/2}</math></b>	0.033 (–0.036; 0.111)	0.058 (–0.04; 0.128)	0.028 (–0.037; 0.093)	0.017 (–0.059; 0.124)	0.045 (–0.013; 0.131)	0.001 (–0.075; 0.086)

Bold values indicate that the 95% credible interval do not overlap with 0.

Using the covariance tensor analysis to compare the different **P**-matrices among populations, we found that **P**-matrices differed significantly among populations for morphological and life-history traits. For morphological traits, the first covariance tensor ( $E_1$ ), which accounted for 55.4% of the variance among matrices, described significant variation between the **P**-matrices for D-Muro, E-Muro and D-Rouvière (posterior distributions of their position on  $E_2$  did not overlap, Figure 2c; although confidence intervals slightly overlapped in additional analyses, see Supplementary Appendix 2, Supplementary Figure S3 and Supplementary Table S7). For life-history trait **P**-matrices, we found that the second covariance tensor ( $E_2$ ), which accounted for 29.4% of the variance among matrices, described significant variation between the **P**-matrices for D-Rouvière and the three Corsican populations (that is, posterior distributions of their position on  $E_2$  did not overlap, Figure 2d, Supplementary Figure S3 and Supplementary Table S7). Laying date and clutch size had opposing loadings (respective loadings: 0.750, –0.604) on the first eigenvector of the second eigentensor ( $e_{21}$ ; which explained 55.9 % of the variation captured by  $E_2$ , Supplementary Table S6) meaning that the differences between these matrices lie mainly in the quantity of variance contained in these two traits.

## DISCUSSION

Using detailed pedigrees and multivariate animal models on long-term phenotypic datasets of blue tits, we estimated **G**-matrices for morphological and life-history traits in four wild populations. Although these populations belonged to two different subspecies, experienced contrasted environments imposing different selection pressures and exhibited strong phenotypic divergence on many traits (Porlier *et al.*, 2012b; Charmantier *et al.*, 2016), we found no evidence for differences among populations in **G**-matrices for both set of traits. In contrast, although **P**-matrices for morphological traits were quite conserved among populations, **P**-matrices for life-history traits were different, thus highlighting that **P**-matrices can be poor surrogates of **G**-matrices.

### Genetic architecture of life-history and morphological traits

All studied traits displayed substantial levels of additive genetic variance in all populations (Tables 2a and 3 and Supplementary Figure S1). In birds, numerous studies have previously found significant  $V_A$  for morphological traits (reviewed in Postma, 2014). For life-history traits, results from previous studies are more equivocal, yet clutch size and laying date usually display low but significant heritable variation (Postma, 2014). In terms of magnitude, morphological traits show substantially (10 times) lower  $I_A$  than life-history traits. Traits closely associated with fitness are predicted to harbour little genetic variance (Fisher, 1930; Falconer and Mackay, 1996), and in line with this, life-history traits have been shown to have low heritability (Kruuk *et al.*, 2000; McCleery *et al.*, 2004; Teplitsky *et al.*, 2009). However, these results are likely explained by the general use of variance scaling that leads traits with large phenotypic variance, such as life-history traits, to present relatively small estimated heritabilities (Houle, 1992). For this reason, several authors have advocated the use of mean scaling that better reflects the potential of life-history traits for rapid evolution (Hansen and Houle, 2008; Hansen *et al.*, 2011).

In contrast to the large literature on clutch size and laying date, our study is, to our knowledge, only the second one to test for genetic variance in incubation period in natural populations (Husby *et al.*, 2012) and the first one to show evidence for additive genetic variation in incubation period. Note that incubation period measured here is not the strict physiological duration of egg incubation by the female,

Table 3 Posterior modes of mean standardized traits genetic and phenotypic variances (estimated variance  $\times 100$ , corresponding to  $I_A$  evolvabilities for genetic ones) and covariances ( $\times 100$ ) for life-history traits G and P matrices with their 95% credible interval for the four blue tit populations

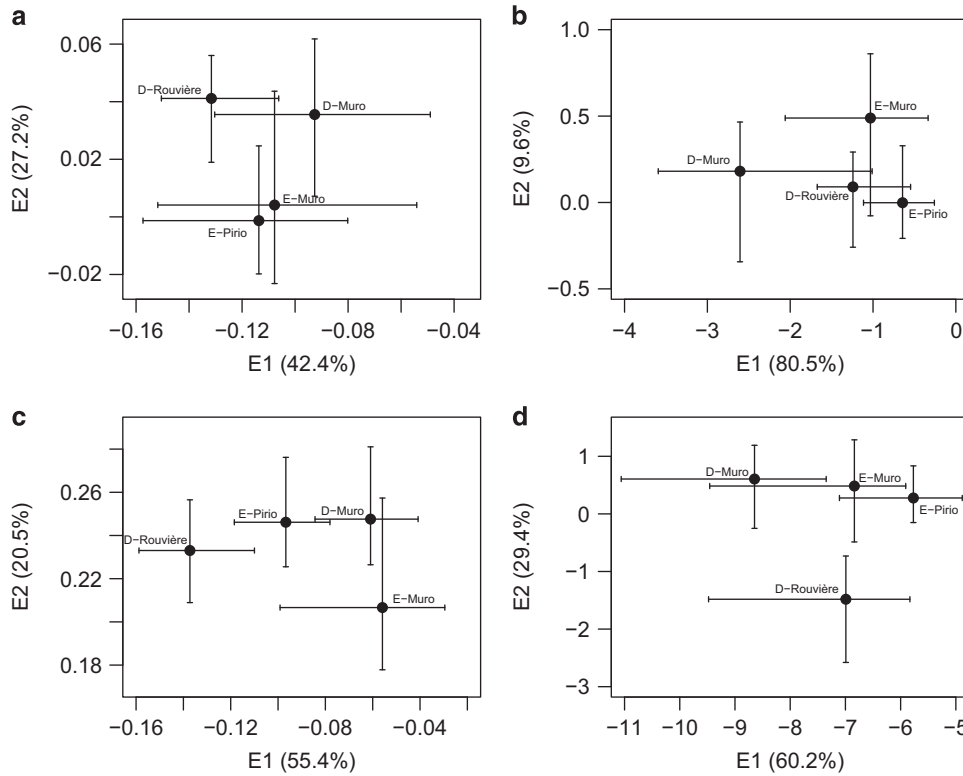
	Variance			Covariance		
	Incubation time	Laying date	Clutch size	Incubation time-laying date	Incubation time-clutch size	Laying date-clutch size
<i>D-Rouvière</i>						
G	0.428 (0.191; 0.63)	0.724 (0.396; 1.252)	0.855 (0.363; 1.227)	-0.094 (-0.348; 0.084)	0.032 (-0.235; 0.2)	-0.302 (-0.65; 0.042)
P	3.072 (2.718; 3.654)	6.766 (5.227; 9.227)	2.846 (2.504; 3.188)	<b>-0.859 (-1.522; -0.046)</b>	<b>-0.376 (-0.678; -0.08)</b>	<b>-1.078 (-1.579; -0.457)</b>
P-1/2GP-1/2	0.133 (0.061; 0.195)	0.119 (0.054; 0.18)	0.270 (0.129; 0.409)	0.009 (-0.049; 0.037)	0.011 (-0.063; 0.084)	-0.005 (-0.094; 0.044)
<i>E-Pirio</i>						
G	0.109 (0.069; 0.223)	0.537 (0.279; 0.987)	0.383 (0.224; 0.999)	-0.018 (-0.198; 0.064)	0.010 (-0.115; 0.131)	-0.071 (-0.435; 0.182)
P	1.845 (1.682; 2.06)	4.905 (4.271; 6.409)	3.452 (3.114; 3.847)	-0.274 (-0.608; 0.031)	-0.156 (-0.375; 0.005)	<b>-1.449 (-1.903; -1.025)</b>
P-1/2GP-1/2	0.074 (0.036; 0.115)	0.125 (0.059; 0.199)	0.141 (0.067; 0.297)	0.003 (-0.048; 0.031)	0.009 (-0.034; 0.063)	0.019 (-0.04; 0.102)
<i>E-Muro</i>						
G	0.153 (0.076; 0.295)	1.082 (0.408; 2.082)	0.594 (0.254; 1.131)	0.019 (-0.221; 0.228)	-0.049 (-0.246; 0.096)	-0.178 (-0.754; 0.265)
P	1.698 (1.409; 2.053)	6.315 (5.106; 8.526)	3.499 (3.077; 4.251)	-0.298 (-0.783; 0.387)	<b>-0.387 (-0.79; -0.131)</b>	<b>-1.676 (-2.426; -0.853)</b>
P-1/2GP-1/2	0.078 (0.046; 0.163)	0.146 (0.074; 0.31)	0.169 (0.085; 0.313)	0.004 (-0.051; 0.077)	0.012 (-0.057; 0.066)	0.027 (-0.075; 0.108)
<i>D-Muro</i>						
G	0.230 (0.096; 0.374)	1.533 (0.603; 2.412)	1.340 (0.564; 1.981)	-0.048 (-0.367; 0.194)	-0.043 (-0.294; 0.201)	<b>-1.071 (-1.596; -0.182)</b>
P	2.062 (1.925; 2.41)	7.440 (6.222; 9.986)	3.763 (3.326; 4.104)	-0.447 (-0.978; 0.032)	<b>-0.247 (-0.48; -0.035)</b>	<b>-2.457 (-3.08; -1.899)</b>
P-1/2GP-1/2	0.107 (0.043; 0.165)	0.171 (0.069; 0.267)	0.347 (0.135; 0.456)	-0.020 (-0.057; 0.055)	0.020 (-0.085; 0.069)	-0.105 (-0.157; 0.029)

For covariances, bold values indicate that the 95% credible interval do not overlap with 0. Posterior modes of heritabilities calculated as the diagonal elements of  $P^{-1/2}GP^{-1/2}$  are reported with their 95% credible interval for each trait and population.

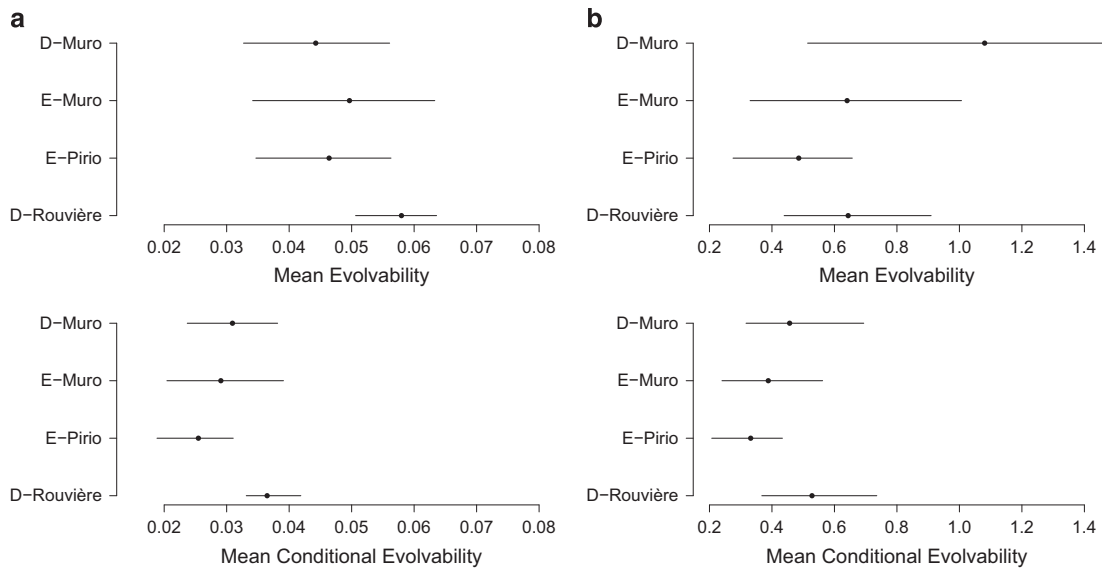
but rather the time interval between the last egg laid and hatching. This interval is a crucial period for insectivorous birds in temperate forests that need to adjust their breeding phenology in order to synchronize the late nestling stage with the food peak period (Dunn, 2004; Visser *et al.*, 2006). Although it is well known that females can adjust their laying date via individual phenotypic plasticity (see, for example, Both *et al.*, 2004; Charmantier and Gienapp, 2014), the fine tuning of breeding phenology can also be modulated through variation in this incubation period, either by: (1) a delay of the start of incubation after clutch completion (Cresswell and McCleery, 2003; Klueen *et al.*, 2011), (2) a start of incubation before the last egg is laid leading to asynchronous hatching (Kontinen *et al.*, 2010; Vedder, 2012) or (3) an adjustment of the efficiency of incubation following the completion of the clutch (Haftorn, 1988; Hadfield *et al.*, 2013). Both hatching delay and asynchrony have been shown to have substantial impact on fitness (Kontinen *et al.*, 2010; Klueen *et al.*, 2011; Hadfield *et al.*, 2013). Asynchronous hatching produces generally a size hierarchy in the brood staggering the food demand over time. In turn, this hierarchy may help sibling negotiation, reduce rivalry and increase offspring survival (see, for example, Kontinen *et al.*, 2010). Thereby, evidence for additive genetic variation in incubation period implies that this trait could not only be adjusted through phenotypic plasticity, but also that it could evolve in the long term if targeted by natural selection. This is particularly important in the actual context of global environmental changes, as it could further help avian populations to track the advancing phenology of their food resources because of warming spring temperatures (Both and Visser, 2005; Visser *et al.*, 2006; Visser, 2008).

In our blue tit study, genetic covariances between morphological traits were all positive and significant as found in numerous other studies (see, for example, Roff, 1996; Kruuk *et al.*, 2008) denoting classic positive allometric relationships between morphological traits linked to body size. For life-history traits, significant negative phenotypic correlations between laying date and clutch size were found in all populations, but the genetic covariance between these two traits was significant only in D-Muro. Larger clutches for earlier laying dates are very common in birds (see, for example, Winkler and Allen, 1996; Husby *et al.*, 2010), and this correlation corresponds to a classic tradeoff in life-history theory (Klomp, 1970). However, whether this correlation has a strong genetic basis remains unclear as significant genetic correlations were found in several previous studies on insectivorous passerines (Sheldon *et al.*, 2003; Garant *et al.*, 2008 and in this study for D-Muro), but not in others (Gienapp *et al.*, 2008; Husby *et al.*, 2010, and here D-Rouvière, E-Pirio and E-Muro). As genetic correlations often suffer from large sampling variances (Falconer and Mackay, 1996; Roff, 1996; Kruuk *et al.*, 2008), it is difficult to understand the causes of such variability in estimates. In our case, we found contrasted results across the three populations with well-detailed pedigree (D-Muro, E-Pirio and D-Rouvière). Hence, it is unlikely that the nonsignificance of genetic correlations is due to power limitations. Thus, the reason why early breeding birds are laying larger clutches could be because of a plastic response to environmental variation, whereby females may lay fewer eggs in late spring in order to adjust to the decreasing food availability (Perrins, 1970; Goodenough *et al.*, 2009).





**Figure 2** Posterior modes and 95% credible intervals of the position of each matrix along the two first eigentensors of each fourth-order covariance tensor analysis. (a) Comparison for morphological **G**-matrices; (b) comparison for life-history **G**-matrices; (c) comparison for morphological **P**-matrices and (d) comparison for life-history **P**-matrices.

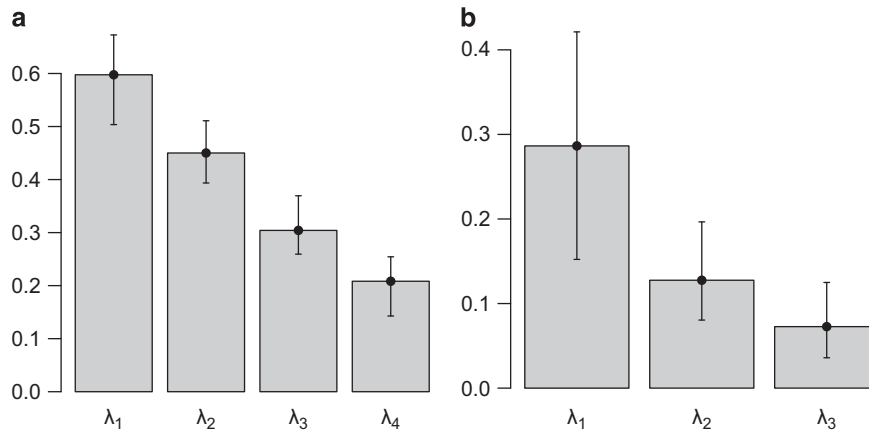


**Figure 3** Measures of average evolutionary potential ( $\bar{\epsilon}$  and  $\bar{\epsilon}_c$ ) for each **G**-matrix: posterior modes and their associated 95% credible intervals. Left (a) represents morphological **G**-matrices. Right (b) represents life-history **G**-matrices.

### No difference in G-matrix between populations

Simulation studies have shown that genetic drift or divergent selection can change **G** characteristics from one generation to another (see, for example, Roff, 2000; Jones *et al.*, 2003). It is thus interesting that we found no significant **G**-matrix differences between the four blue tits populations, as they differ in many respects (habitats, phenotypic distributions and selective pressures; Charmantier *et al.*, 2004b, 2016;

Porlier *et al.*, 2012a). A lack of difference in **G** between populations also indicates a lack of evidence for any  $\mathbf{G} \times \mathbf{E}$  interactions. Nevertheless, these results are in line with other studies investigating **G**-matrix variation among populations highlighting strong similarity of **G**-matrices (see, for example, Roff *et al.*, 2004; review in Arnold *et al.*, 2008). Compelling examples of divergent **G**-matrices across populations exist but they generally used particular experimental



**Figure 4** Distribution of the eigenvalues of  $P^{-1/2}GP^{-1/2}$ . Left (a) represents morphological variance-covariance matrices in the D-Rouvière population. Right (b) represents life-history matrices in the D-Rouvière population. The distributions were similar for the other populations and are displayed in Supplementary Figure S5.

designs with individuals collected in the field and subsequently bred in the laboratory in contrasting environmental conditions, thus complicating the interpretation for **G**-matrix evolution (see, for example, Cano *et al.*, 2004; Doroszuk *et al.*, 2008). Moreover, for many of these studies, it is difficult to evaluate the robustness of the different tests because of (1) lack of inclusion or biased estimation of error estimates (Ovaskainen *et al.*, 2008) and (2) inadequate pedigree structure (for example, few related individuals; problems already highlighted in the field; Charmantier and Garant, 2005). Therefore, interestingly, although it seems relatively easy to modify **G** by applying artificial selection (see, for example, Careau *et al.*, 2015), in nature, **G**-matrices seem stable, even between populations exposed to strong habitat differences (Arnold *et al.*, 2008 but see Wood and Brodie, 2015). Several nonmutually exclusive hypotheses may explain this observation.

First, in nature, conspecific populations are generally not genetically isolated. Migration between populations can potentially counteract the diversifying effect of divergent selective pressures and thus tend to homogenize **G**-matrices (Guillaume and Whitlock, 2007). In our system, although significant genetic structure has been revealed within Corsica (Porlier, *et al.*, 2012b; Szulkin *et al.*, 2016), dispersal between the three Corsican blue tit populations occurs and might be a powerful homogenizing force. However, dispersal has a limited role regarding island/mainland differentiation as these populations have probably evolved in allopatry and belong to two different mitochondrial lineages (Kvist *et al.*, 2004).

Second, artificial selection pressures applied in experimental studies among populations are often strong and designed to be contrasted (see, for example, Doroszuk *et al.*, 2008). In nature, although spatial variation of selection seems common, strength variability appears as the main component of variation in selection across populations, whereas direction is relatively conserved (Siepielski *et al.*, 2013). This pattern seems to hold in our study system for laying date and clutch size (Porlier *et al.*, 2012a). In laboratory experiments, the selection pressures applied to different treatments might also differ in direction rather than in strength alone. In this respect, an interesting study by Bolstad *et al.* (2015) showed that an extremely conserved allometric relationship in *Drosophila* can be modified by applying specific selective pressures. However, once the selective pressures were released, the population rapidly recovered the original allometric slope. They concluded that the allometric slope change generated

deleterious pleiotropic responses because of links with unknown traits. Thus, it is possible that the reason why we rarely observe changes is because of strong pleiotropic links with unmeasured traits that prevent structural changes of **G**-matrices unless particularly new selective gradients are applied such as in laboratory experiments (Bolstad *et al.*, 2015). In nature, if the components of selection influencing the most **G**-matrices are not variable across populations, spatial stability may actually not be surprising. In this respect, Roff and Fairbairn (2012) showed that correlational selection seems to be responsible for the eigenstructure of **G**-matrices. It is possible that this type of selection is remarkably stable across populations and environments. Unfortunately, very little is known about spatial variation of some components of selection such as nonlinear (quadratic and correlational) selection (Siepielski *et al.*, 2013, but see Garant *et al.*, 2007), and hence further investigations in this sense will be of interest.

#### Using **P**-matrices as substitutes for **G**-matrices

**P** may differ substantially from **G** for various reasons (Willis *et al.*, 1991). However, data sets allowing the estimation of **G**-matrices in the wild are still scarce. Therefore, in numerous studies, **P** is used as a surrogate of **G** either for comparative or for predictive purposes (see, for example, Berner *et al.*, 2008). In this study, we showed that morphological **P**-matrices displayed similar patterns compared with **G**-matrices. In this respect, they may appear as an adequate measure of the underlying **G**-matrices, but this conclusion is antagonized by the fact that **P** and **G** were not proportional for morphological traits (Figure 4). For life-history traits, the patterns observed for **P**-matrices were different from those observed for **G**-matrices, indicating that basing our population comparisons on **P**-matrices could have led to erroneous conclusions in this case. The more heritable a trait is, the more similar phenotypic correlations are to genetic ones (Cheverud, 1988; Hadfield *et al.*, 2007). Therefore, using **P**- instead of **G**-matrices is likely to give more reliable results for traits having substantial heritabilities, such as morphological traits, but not for traits showing low heritabilities, such as life-history traits (Roff, 1996). However, based on the general nonproportionality of **P**- and **G**-matrices, great caution should be taken.

An issue that may affect our conclusions is the difference in statistical power generally observed between **G** and **P** estimates. Phenotypic (co)variances are easier to estimate with small sample sizes than genetic (co)variances (Cheverud, 1988). This leads genetic

estimates to be often more imprecise and have wider confidence intervals. Therefore, one reason why we observe differences between **P**- but not **G**-matrices could be because of lower statistical power to estimate **G**. We are nevertheless confident that our results are robust as the **G**-matrices were estimated using a large number of individuals and good pedigree characteristics that seem sufficient to detect significant differences (see simulations by Teplitsky *et al.*, 2014a). Moreover, even if the **G**-matrices were weakly differentiated in our study system, these differences are unlikely to result in strong differences in evolutionary potential as evolvability values were similar among the four populations (Figure 3). We also used three different statistical methods commonly used in the literature to compare our results with other studies. Combining the analysis of **P**-, **G**- and **E**- (that is, the environmental variance-covariance matrix) matrices with selection and divergence estimates will definitely help us to better depict the origin of phenotypic divergence in this system.

## CONCLUSIONS

Although the populations studied here differ in many respects, they have similar evolutionary potential based on estimated **G**-matrices for morphological and life-history traits. Future comparative genomic data analyses performed in these populations will help us understand the genetic basis of these different characters and whether the same genes underlie traits in different populations (Santure *et al.*, 2015). As most estimations of **G** obtained from wild populations are limited to a few model species and populations, further studies exploring spatial variation of **G** across the whole range of the species will be necessary to test whether our results can be extrapolated to other populations. If **G**-matrices are conserved across the entire species range, it will open interesting perspectives regarding the predictive power of such estimations. Our findings finally highlight the value of long-term data sets in several wild populations of the same species in heterogeneous environments to better understand the complexity of evolutionary dynamics across space and time.

## DATA ARCHIVING

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.03mn0>.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We thank all the people who helped to maintain the study sites and conduct the blue tit monitoring since 1988, in particular Jacques Blondel, Philippe Perret, Marcel Lambrechts, Arnaud Grégoire, Claire Doutrelant and Christophe de Franceschi. We thank Loeske Kruuk, Luis-Miguel Chevin, Charles Perrier and four anonymous reviewers for useful advice and comments on earlier drafts. This project was funded by the Agence Nationale de la Recherche (BioAdapt Grant ANR-12-ADAP-0006-02-PEPS to CT and AC), the European Research Council (Starting grant ERC-2013-StG-337365-SHE to AC), the Natural Sciences and Engineering Research Council of Canada (NSERC discovery grant to DG) and the OSU-OREME.

Agrawal AF, Stinchcombe JR (2009). How much do genetic covariances alter the rate of adaptation? *Proc R Soc Lond B Biol Sci* **276**: 1183–1191.  
Aguirre JD, Hine E, McGuigan K, Blows MW (2014). Comparing G: multivariate analysis of genetic variation in multiple populations. *Heredity* **112**: 21–29.  
Arnold SJ (1992). Constraints on phenotypic evolution. *Am Nat* **140**(Suppl 1): S85–107.  
Arnold SJ, Bürger R, Hohenlohe PA, Ajie BC, Jones AG (2008). Understanding the evolution and stability of the G-matrix. *Evolution* **62**: 2451–2461.

Arnold SJ, Pfrender ME, Jones AG (2001). The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* **112–113**: 9–32.  
Arnold SJ, Phillips PC (1999). Hierarchical comparison of genetic variance-covariance matrices. II. Coastal-inland divergence in the garter snake, *Thamnophis elegans*. *Evolution* **53**: 1516–1527.  
Ashman T-L (2003). Constraints on the evolution of males and sexual dimorphism: field estimates of genetic architecture of reproductive traits in three populations of gynodioecious *Fragaria virginiana*. *Evolution* **57**: 2012–2025.  
Basser PJ, Pajevic S (2007). Spectral decomposition of a 4th-order covariance tensor: applications to diffusion tensor MRI. *Signal Process* **87**: 220–236.  
Berner D, Adams DC, Grandchamp A-C, Hendry AP (2008). Natural selection drives patterns of lake-stream divergence in stickleback foraging morphology. *J Evol Biol* **21**: 1653–1665.  
Björklund M, Husby A, Gustafsson L (2013). Rapid and unpredictable changes of the G-matrix in a natural bird population over 25 years. *J Evol Biol* **26**: 1–13.  
Blondel J, Thomas DW, Charmantier A, Perret P, Bourgault P, Lambrechts MM (2006). A thirty-year study of phenotypic and genetic variation of blue tits in Mediterranean habitat mosaics. *Bioscience* **56**: 661–673.  
Bolstad GH, Cassara JA, Márquez E, Hansen TF, van der Linde K, Houle D *et al.* (2015). Complex constraints on allometry revealed by artificial selection on the wing of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **112**: 13284–13289.  
Both C, Artemyev AV, Blaauw B, Cowie RJ, Dekhuijzen AJ, Eeva T *et al.* (2004). Large-scale geographical variation confirms that climate change causes birds to lay earlier. *Proc R Soc Lond B Biol Sci* **271**: 1657–1662.  
Both C, Visser ME (2005). The effect of climate change on the correlation between avian life-history traits. *Glob Change Biol* **11**: 1606–1613.  
Cano JM, Laurila A, Palo J, Merilä J (2004). Population differentiation in G matrix structure due to natural selection in *Rana temporaria*. *Evolution* **58**: 2013–2020.  
Careau V, Wolak ME, Carter PA, Garland T (2015). Evolution of the additive genetic variance-covariance matrix under continuous directional selection on a complex behavioural phenotype. *Proc R Soc B* **282**: 20151119.  
Charmantier A, Blondel J (2003). A contrast in extra-pair paternity levels on mainland and island populations of Mediterranean blue tits. *Ethology* **109**: 351–363.  
Charmantier A, Blondel J, Perret P, Lambrechts MM (2004a). Do extra-pair paternities provide genetic benefits for female blue tits *Parus caeruleus*? *J Avian Biol* **35**: 524–532.  
Charmantier A, Doutrelant C, Dubuc-Messier G, Fargevieille A, Szulkin M (2016). Mediterranean blue tits as a case study of local adaptation. *Evol Appl* **9**: 135–152.  
Charmantier A, Garant D (2005). Environmental quality and evolutionary potential: lessons from wild populations. *Proc R Soc Lond B Biol Sci* **272**: 1415–1425.  
Charmantier A, Garant D, Kruuk LEB (eds) (2014). *Quantitative Genetics in the Wild*. Oxford University Press.  
Charmantier A, Gienapp P (2014). Climate change and timing of avian breeding and migration: evolutionary versus plastic changes. *Evol Appl* **7**: 15–28.  
Charmantier A, Kruuk LEB, Blondel J, Lambrechts MM (2004b). Testing for microevolution in body size in three blue tit populations. *J Evol Biol* **17**: 732–743.  
Charmantier A, Réale D (2005). How do misassigned paternities affect the estimation of heritability in the wild? *Mol Ecol* **14**: 2839–2850.  
Cheverud JM (1988). A comparison of genetic and phenotypic correlations. *Evolution* **42**: 958–968.  
Cresswell W, McCleery R (2003). How great tits maintain synchronization of their hatch date with food supply in response to long-term variability in temperature. *J Anim Ecol* **72**: 356–366.  
De Lathauwer L, De Moor B, Vandewalle J (2004). Computation of the canonical decomposition by means of simultaneous generalized schur decomposition. *SIAM J Matrix Anal Appl* **26**: 295–327.  
Doroszuk A, Wojewodzic MW, Gort G, Kammenga JE (2008). Rapid divergence of genetic variance-covariance matrix within a natural population. *Am Nat* **171**: 291–304.  
Dunn P (2004). Breeding dates and reproductive performance. In: *Birds and Climate Change* Vol. 35, Academic Press, pp 69–87.  
Falconer DS, Mackay TFC (1996). *Introduction to Quantitative Genetics*. 4th edn(16. print.) Pearson: Prentice Hall: Harlow.  
Firth JA, Hadfield JD, Santure AW, Slate J, Sheldon BC (2015). The influence of nonrandom extra-pair paternity on heritability estimates derived from wild pedigrees. *Evolution* **69**: 1336–1344.  
Fisher RA (1930). *The Genetical Theory Of Natural Selection*. Clarendon Press.  
Garant D, Hadfield JD, Kruuk LEB, Sheldon BC (2008). Stability of genetic variance and covariance for reproductive characters in the face of climate change in a wild bird population. *Mol Ecol* **17**: 179–188.  
Garant D, Kruuk LEB, McCleery RH, Sheldon BC (2007). The effects of environmental heterogeneity on multivariate selection on reproductive traits in female great tits. *Evolution* **61**: 1546–1559.  
Gienapp P, Teplitsky C, Alho JS, Mills JA, Merilä J (2008). Climate change and evolution: disentangling environmental and genetic responses. *Mol Ecol* **17**: 167–178.  
Goodenough AE, Elliot SL, Maitland DP, Hart AG (2009). Variation in the relationship between lay date and clutch size in three cavity-nesting woodland passerines. *Acta Ornithol* **44**: 27–36.  
Griswold CK, Logsdon B, Gomulkiewicz R (2007). Neutral evolution of multiple quantitative characters: a genealogical approach. *Genetics* **176**: 455–466.  
Guillaume F, Whitlock MC (2007). Effects of migration on the genetic covariance matrix. *Evolution* **61**: 2398–2409.  
Hadfield JD (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R Package. *J Stat Softw* **33**: 1–22.

- Hadfield JD, Heap EA, Bayer F, Mittell EA, Crouch NMA (2013). Intraclutch differences in egg characteristics mitigate the consequences of age-related hierarchies in a wild passerine. *Evolution* **67**: 2688–2700.
- Hadfield JD, Nutall A, Osorio D, Owens IPF (2007). Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *J Evol Biol* **20**: 549–557.
- Haftorn S (1988). Incubating female passerines do not let the egg temperature fall below the 'physiological zero temperature' during their absences from the nest. *Ornis Scand Scand J Ornithol* **19**: 97–110.
- Hansen TF, Houle D (2008). Measuring and comparing evolvability and constraint in multivariate characters. *J Evol Biol* **21**: 1201–1219.
- Hansen TF, Pélabon C, Houle D (2011). Heritability is not evolvability. *Evol Biol* **38**: 258–277.
- Hine E, Chenoweth SF, Rundle HD, Blows MW (2009). Characterizing the evolution of genetic variance using genetic covariance tensors. *Philos Trans R Soc B Biol Sci* **364**: 1567–1578.
- Houle D (1992). Comparing evolvability and variability of quantitative traits. *Genetics* **130**: 195–204.
- Houle D, Pélabon C, Wagner GP, Hansen TF (2011). Measurement and meaning in biology. *Q Rev Biol* **86**: 3–34.
- Husby A, Gustafsson L, Qvarnström A (2012). Low genetic variance in the duration of the incubation period in a collared flycatcher (*Ficedula albicollis*) population. *Am Nat* **179**: 132–136.
- Husby A, Nussey DH, Visser ME, Wilson AJ, Sheldon BC, Kruuk LEB (2010). Contrasting patterns of phenotypic plasticity in reproductive traits in two great tit (*Parus major*) populations. *Evolution* **64**: 2221–2237.
- Johansson F, Lind MI, Ingvarsson PK, Bokma F (2011). Evolution of the G-matrix in life history traits in the common frog during a recent colonisation of an island system. *Evol Ecol* **26**: 863–878.
- Jones AG, Arnold SJ, Bürger R (2007). The mutation matrix and the evolution of evolvability. *Evolution* **61**: 727–745.
- Jones AG, Arnold SJ, Bürger R, Houle D (2003). Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* **57**: 1747–1760.
- Klomp H (1970). The determination of clutch-size in birds. A review. *Ardea* **58**: 1–124.
- Kluen E, de Heij ME, Brommer JE (2011). Adjusting the timing of hatching to changing environmental conditions has fitness costs in blue tits. *Behav Ecol Sociobiol* **65**: 2091–2103.
- Konttiainen P, Pietiäinen H, Karell P, Pihlaja T, Brommer JE (2010). Hatching asynchrony is an individual property of female Ural owls which improves nestling survival. *Behav Ecol* **21**: 722–729.
- Kruuk LEB (2004). Estimating genetic parameters in natural populations using the 'animal model'. *Philos Trans R Soc B Biol Sci* **359**: 873–890.
- Kruuk LEB, Clutton-Brock TH, Slate J, Pemberton JM, Brotherstone S, Guinness FE (2000). Heritability of fitness in a wild mammal population. *Proc Natl Acad Sci USA* **97**: 698–703.
- Kruuk LEB, Slate J, Wilson AJ (2008). New answers for old questions: the evolutionary quantitative genetics of wild animal populations. *Annu Rev Ecol Syst* **39**: 525–548.
- Kvist L, Viiri K, Dias PC, Rytkönen S, Orell M (2004). Glacial history and colonization of Europe by the blue tit *Parus caeruleus*. *J Avian Biol* **35**: 352–359.
- Lande R (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33**: 402–416.
- Lynch M, Walsh B (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer.
- McCleery RH, Pettifor RA, Armbruster P, Meyer K, Sheldon BC, Perrins CM *et al.* (2004). Components of variance underlying fitness in a natural population of the great tit *Parus major*. *Am Nat* **164**: E62–E72.
- McGuigan K (2006). Studying phenotypic evolution using multivariate quantitative genetics. *Mol Ecol* **15**: 883–896.
- Merilä J, Sheldon BC, Kruuk LE (2001). Explaining stasis: microevolutionary studies in natural populations. *Genetica* **112–113**: 199–222.
- Morrissey MB, de Villemereuil P, Doligez B, Gimenez O (2014). Bayesian approaches to the quantitative genetic analysis of natural populations. In: Charmantier A, Garant D, Kruuk LEB (eds). *Quantitative Genetics in the Wild*. Oxford University Press: Oxford, UK, pp 228–253.
- Morrissey MB, Walling CA, Wilson AJ, Pemberton JM, Clutton-Brock TH, Kruuk LEB *et al.* (2012). Genetic analysis of life-history constraint and evolution in a wild ungulate population. *Am Nat* **179**: E97–E114.
- Morrissey MB, Wilson AJ (2010). pedantics: an R package for pedigree-based genetic simulation and pedigree manipulation, characterization and viewing. *Mol Ecol Resour* **10**: 711–719.
- Ovaskainen O, Cano JM, Merilä J (2008). A Bayesian framework for comparative quantitative genetics. *Proc R Soc Lond B Biol Sci* **275**: 669–678.
- Palumbi SR (2001). Humans as the world's greatest evolutionary force. *Science* **293**: 1786–1790.
- Perrins CM (1970). The timing of birds' breeding seasons. *Ibis* **112**: 242–255.
- Phillips PC, Whitlock MC, Fowler K (2001). Inbreeding changes the shape of the genetic covariance matrix in *Drosophila melanogaster*. *Genetics* **158**: 1137–1145.
- Porlier M, Charmantier A, Bourgault P, Perret P, Blondel J, Garant D (2012a). Variation in phenotypic plasticity and selection patterns in blue tit breeding time: between- and within-population comparisons. *J Anim Ecol* **81**: 1041–1051.
- Porlier M, Garant D, Perret P, Charmantier A (2012b). Habitat-linked population genetic differentiation in the blue tit *Cyanistes caeruleus*. *J Hered* **103**: 781–791.
- Postma E (2014). Four decades of estimating heritabilities in wild vertebrate populations: improved methods, more data, better estimates? In: Charmantier A, Garant D, Kruuk LEB (eds). *Quantitative Genetics in the Wild*. Oxford University Press: Oxford, UK, pp 16–33.
- Puentes A, Granath G, Ågren J (2016). Similarity in G matrix structure among natural populations of *Arabidopsis lyrata*. *Evolution* **70**: 2370–2386.
- Revell LJ (2007). The G matrix under fluctuating correlational mutation and selection. *Evolution* **61**: 1857–1872.
- Roff DA (1996). The evolution of genetic correlations: an analysis of patterns. *Evolution* **50**: 1392–1403.
- Roff D (2000). The evolution of the G matrix: selection or drift? *Heredity* **84**: 135–142.
- Roff DA, Fairbairn DJ (2012). A test of the hypothesis that correlational selection generates genetic correlations. *Evolution* **66**: 2953–2960.
- Roff DA, Mousseau T, Møller AP, de Lope F, Saino N (2004). Geographic variation in the G matrices of wild populations of the barn swallow. *Heredity* **93**: 8–14.
- Roff DA, Prokko JM, Krams I, Rantala MJ (2012). There is more than one way to skin a G matrix: comparing G matrices. *J Evol Biol* **25**: 1113–1126.
- Santure AW, Poissant J, De Cauwer I, van Oers K, Robinson MR, Quinn JL *et al.* (2015). Replicated analysis of the genetic architecture of quantitative traits in two wild great tit populations. *Mol Ecol* **24**: 6148–6162.
- Sheldon BC, Kruuk LEB, Merilä J (2003). Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution* **57**: 406–420.
- Siepielski AM, Gotanda KM, Morrissey MB, Diamond SE, DiBattista JD, Carlson SM (2013). The spatial patterns of directional phenotypic selection. *Ecol Lett* **16**: 1382–1392.
- Sniegula S, Golab MJ, Drobnik SM, Johansson F (2016). Seasonal time constraints reduce genetic variation in life-history traits along a latitudinal gradient. *J Anim Ecol* **85**: 187–198.
- Stephan SJ, Phillips PC, Houle D (2002). Comparative quantitative genetics: evolution of the G matrix. *Trends Ecol Evol* **17**: 320–327.
- Szulkin M, Gagnaire P-A, Bierre N, Charmantier A (2016). Population genomic footprints of fine-scale differentiation between habitats in Mediterranean blue tits. *Mol Ecol* **25**: 542–558.
- Teplitsky C, Mills JA, Yarrall JW, Merilä J (2009). Heritability of fitness components in a wild bird population. *Evolution* **63**: 716–726.
- Teplitsky C, Robinson MR, Merilä J (2014a). Evolutionary potential and constraints in wild populations. In: Charmantier A, Garant D, Kruuk LEB (eds). *Quantitative Genetics in the Wild*. Oxford University Press: Oxford, UK, pp 190–208.
- Teplitsky C, Tarka M, Møller AP, Nakagawa S, Balbontin J, Burke TA *et al.* (2014b). Assessing multivariate constraints to evolution across ten long-term avian studies. *PLoS One* **9**: e90444.
- Turelli M (1988). Phenotypic evolution, constant covariances, and the maintenance of additive variance. *Evolution* **42**: 1342–1347.
- Vedder O (2012). Individual birds advance offspring hatching in response to increased temperature after the start of laying. *Oecologia* **170**: 619–628.
- Visser ME (2008). Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proc R Soc Lond B Biol Sci* **275**: 649–659.
- Visser ME, Holleman LJM, Gienapp P (2006). Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia* **147**: 164–172.
- Walsh B, Blows MW (2009). Abundant genetic variation+strong selection=multivariate genetic constraints: a geometric view of adaptation. *Annu Rev Ecol Syst* **40**: 41–59.
- Willis JH, Coyne JA, Kirkpatrick M (1991). Can one predict the evolution of quantitative characters without genetics? *Evolution* **45**: 441–444.
- Winkler DW, Allen PE (1996). The seasonal decline in tree swallow clutch size: physiological constraint or strategic adjustment? *Ecology* **77**: 922–932.
- Wood CW, Brodie ED (2015). Environmental effects on the structure of the G-matrix. *Evolution* **69**: 2927–2940.

Supplementary Information accompanies this paper on Heredity website (<http://www.nature.com/hdy>)