Title
Gene flow does not prevent personality and morphological differentiation between two blue tit populations.

Authors
Gabrielle Dubuc-Messier a,b, Samuel P. Caro a,c, Charles Perrier a, Kees van Oers c, Denis Réale b*, Anne Charmantier a,b*

Affiliations
a. Centre d’Écologie Fonctionnelle et Évolutive, Unité Mixte de Recherche CNRS 5175, 1919 Route de Mende, Montpellier Cedex 5, France.
b. Département des sciences biologiques, Université du Québec à Montréal, Case postale 8888, succursale Centre-ville, Montréal (Québec), H3C 3P8, Canada.
*co-last authors

Running head
Divergence between blue tit populations

Corresponding author
Gabrielle Dubuc-Messier: dubuc-messier.gabrielle@courrier.uqam.ca

Université du Québec à Montréal, Département des sciences biologiques, Case postale 8888, succursale Centre-ville, Montréal (Québec), H3C 3P8, Canada.

Tél: 514-987-3000 ext.2265
Fax: 514-987-4647

Acknowledgements

We thank the landowners for permission to work on their properties and all the blue tit crew for their work, specifically Marta Szulkin, Christophe de Franceschi, Philippe Perret, Virginie Demeyrier, Samuel Perret and Boris Delahaie. We thank Julie Morand-Ferron, Andrew Handry, Arnaud Bechet and Jean-Francois Giroux for their comments. We also thank Marylou Aaldering, Floor Petit, Franca Kropman and Sonja Schaper for taking good care of the birds in captivity. The behavioural tests run at the NIOO-KNAW were approved by the Animal Experimentation Committee of the Royal Dutch Academy of Sciences (DEC-KNAW; permit number CTE09–04 and NIOO11.09). The work performed in the field was approved by the prefectural office of Corsica and the Regional Direction of Environment (DIREN) committee (permit numbers 2009-0379 and 3467) and the tests run at the CEFE were approved by the Comité Institutionnel de Protection des Animaux (UQAM; CIPA-769-2015; 0413-R1-769-0414) in Canada. This project was funded by grants from the Agence Nationale de la Recherche (BioAdapt grant ANR-12-ADAP-0006-02-PEPS), the European Research Council (Starting grant ERC-2013-StG-337365-SHE) and the OSU-OOREME to AC, by the Natural Science and Engineering Research Council of Canada (Discovery grant) to DR and the Netherlands Organisation for Scientific Research
(NWO-VENI grant 863.09.011) to SPC. GDM received PhD fellowships from the Fonds de Recherche Québec Nature et Technologies and the Natural Science and Engineering Research Council of Canada. We declare no conflict of interests.
Abstract

Understanding the causes and consequences of population phenotypic divergence is a central goal in ecology and evolution. Phenotypic divergence among populations can result from genetic divergence, phenotypic plasticity or a combination of the two. However, few studies have deciphered these mechanisms for populations geographically close and connected by gene flow, especially in the case of personality traits. In this study, we used a common garden experiment to explore the genetic basis of the phenotypic divergence observed between two blue tit (Cyanistes caeruleus) populations inhabiting contrasting habitats separated by 25 km, for two personality traits (exploration speed and handling aggression), one physiological trait (heart rate during restraint) and two morphological traits (tarsus length and body mass). Blue tit nestlings were removed from their population and raised in a common garden for up to five years. We then compared adult phenotypes between the two populations, as well as trait-specific \(Q_{st}\) and \(F_{st}\). Our results revealed differences between populations similar to those found in the wild, suggesting a genetic divergence for all traits. \(Q_{st}\) - \(F_{st}\) comparisons revealed that the traits divergences likely result from dissimilar selection patterns rather than from genetic drift. Our study is one of the first to report a \(Q_{st}\) - \(F_{st}\) comparison for personality traits and adds to the growing body of evidence that population genetic divergence is possible at a small scale for a variety of traits including behavioural traits.

Keywords

Cyanistes caeruleus, genetic divergence, local adaptation, personality, plasticity, \(Q_{st}\) - \(F_{st}\)
Introduction

Understanding the evolutionary causes of phenotypic divergence among populations is an important aspect of the study of diversity. Environmental heterogeneity can have a major role in generating phenotypic divergence among populations (Wang & Bradburd 2014). Spatial variation in selection pressures resulting from such environmental heterogeneity can lead to genotype by environment interactions for fitness and produce phenotypic and genetic divergence between populations that can lead to local adaptations (Kawecki & Ebert 2004; Wang & Bradburd 2014).

Spatial heterogeneity in ecological conditions can also favour the evolution of phenotypic plasticity, i.e. the adjustment of individual phenotypes in response to environmental factors (Pigliucci 2005) and cause phenotypic divergence of populations in the absence of genetic divergence or local adaptation (Sultan & Spencer 2002; Réale et al. 2003; Pigliucci 2005). Phenotypic divergence of populations can also be produced by non-random dispersal of individuals between habitat types (Wang & Bradburd 2014). Importantly, phenotypic divergence of populations does not necessarily involve an adaptive process since phenotypic plasticity and non-random dispersal can be non-adaptive (Edelaar & Bolnick 2012; Fitzpatrick 2012; Wang & Bradburd 2014) and can occur in the same or in the opposite direction to genetic divergence (Fitzpatrick 2012). In addition, strong founder effects or genetic drift can also lead to phenotypic and genetic divergence of populations (Slatkin 1987). Establishing the relative importance of environmental versus genetic effects involved in the phenotypic divergence of populations provides fundamental information about the origin of intra-specific diversity in the wild. In addition, determining if this divergence is adaptive or the result of neutral processes is essential because it gives important indications about the eco-evolutionary dynamics of traits and their evolutionary trajectories.
Traditionally, it has been considered that the homogenizing effect of gene flow prevents genetic divergence of populations (Sultan & Spencer 2002; Lenormand 2002). Thus, most research on genetic divergence focused on populations separated by large spatial scales or by important landscape barriers to dispersal (Slatkin 1987; Lenormand 2002). Nevertheless, recent theoretical and empirical studies revealed that even in the presence of gene flow, phenotypic divergence can have a genetic origin when there is strong divergent selection and/or non-random dispersal (Richardson et al. 2014; Wang & Bradburd 2014). Despite growing interest for such isolation by environment, there is little empirical data on the mechanisms underlying the phenotypic divergence of populations separated by small geographic distances and connected by gene flow.

Behavioural traits have often been considered as highly plastic and thus less prone to genetic divergence. However, several studies are now showing that among-individual differences in behaviour can be repeatable (personality; Réale et al. 2007), heritable (van Oers & Sinn 2011), and related to fitness (Smith & Blumstein 2008) and could thus evolve in response to local conditions. In this context, an increasing number of studies have compared the personality phenotypes of individuals inhabiting contrasted ecological conditions (Bell 2005; Quinn et al. 2009; Atwell et al. 2012; Herczeg et al. 2013; Miranda et al. 2013; Karlsson et al. 2016; Jacquin et al. 2016). However, fewer studies have disentangled the role of plasticity from that of genetic effects in shaping phenotypic divergence between populations separated by distances that are within the dispersal ability of a species (Atwell et al. 2012; Miranda et al. 2013). Note that the plastic response to environmental factors can itself have a genetic basis, hence plasticity levels can differ across populations because plasticity can be heritable and evolve differently across populations (e.g. Laurila et al. 2002).
Previously, we have revealed a phenotypic divergence for personality and morphological traits between two wild populations of blue tits (Cyanistes caeruleus) living in contrasting habitats in a Mediterranean landscape (Charmantier et al. 2016; Dubuc-Messier et al. 2017). These populations occupy habitats and valleys dominated by either evergreen (holm oak, Quercus ilex) or deciduous oaks (downy oak, Quercus pubescens) yet are separated only by 25 km, which is within the typical dispersal range of the species (Tufto et al. 2005; Winkel & Frantzen 1991). The dominant tree species in each habitat and valley is suspected to have an important influence on blue tits’ ecological context that translates into phenotypic divergence between populations for numerous types of traits despite a spatial proximity and gene flow among them (Charmantier et al. 2016). For example, blue tits from the evergreen habitat have a higher adult survival probability, a lower body mass, a smaller tarsus length, a higher docility (lower handling aggression), and a slower exploration in a novel environment, compared to birds from the deciduous habitat (Table S1; Grosbois et al. 2006; Charmantier et al. 2016; Dubuc-Messier et al. 2017). In addition, past studies in this system revealed that small birds (mass and tarsus length) have a selective advantage in the evergreen habitat (Blondel et al. 2002; Teplitsky et al. 2014), suggesting that at least some of the observed phenotypic divergence between habitats could be adaptive.

In this study, we used a common garden experiment to assess whether the personality and morphological divergence between these two blue tit populations could have a genetic basis. We collected blue tit nestlings from the evergreen and deciduous habitats and raised them for up to five years in aviaries, subsequently comparing their personality, physiological and morphological phenotypes once adults. Previous experiments in aviaries on this system have found a genetic
divergence between these habitats for life-history traits (Lambrechts et al. 1997). Based on these results, we hypothesized that the phenotypic divergence found previously in the wild for personality and morphological traits would also reflect a genetic divergence. Therefore, we predicted that, following the common garden experiment, individuals originating from the evergreen habitat would show a slower exploration in the novel environment, a higher docility (lower handling aggression), a smaller tarsus and a lower body mass then individuals originating from the deciduous habitat. We also compared heart rate during manual restraint of birds originating from the two habitats, a physiological measure of stress reaction often used in personality studies (Koolhaas et al. 1999).

Second, we investigated whether the potential genetic divergence between these habitats could be attributed to different selection pressures or to genetic drift using a $Q_{st}$ - $F_{st}$ comparison approach (Leinonen et al. 2013). A trait $Q_{st}$ measures the amount of additive genetic variance among populations relative to the total genetic variance in the trait (Leinonen et al. 2013). $F_{st}$ is the equivalent measure for neutral molecular variance (Weir and Cockerham 1984) and can be used as a null expectation for the degree of population divergence due to genetic drift and gene flow. If $Q_{st} > F_{st}$, the trait divergence is higher than the neutral expectation and is likely the result of directional selection favouring local adaptation (Leinonen et al. 2013) rather than the result of drift. The two blue tit populations have very large effective population sizes (roughly estimated around 10,000 in each valley, Perrier et al., genomic work in progress) and have been found weakly genetically differentiated (Szulkin et al. 2016). Consequently, it is unlikely that any genetic divergence for these traits would be produced by genetic drift. We considered that, because birds were raised in a common garden, a phenotypic difference among individuals was a realistic approximation of an additive genetic effect. We thus used the phenotype of the common
garden birds to calculate the $Q_{st}$ and predicted that the $Q_{st}$ of each trait would significantly exceed
the $F_{st}$ estimated between both populations. In addition, in order to better understand the
importance of environmental factors in shaping the observed phenotypic differentiation in the
wild, we compared the genetic differentiation ($Q_{st}$) of birds from the common garden experiment
with the phenotypic differentiation of wild birds for the same traits ($P_{st}$; the amount of
phenotypic variance among wild populations relative to the total phenotypic variance in the trait).

**Materials and Methods**

The population located in the evergreen habitat (Evergreen-Pirio) is in the Corsican Fango valley
(42°34′N, 08°44′E; 200m elevation) and contains 205 nest-boxes distributed across two study
plots. The population located in the deciduous habitat (Deciduous-Muro) is in the Corsican
Regino valley (42°32′N, 08°55′E, 350 m elevation) and contains 110 nest-boxes distributed
across three study plots. A weekly to daily monitoring over the course of the breeding season
(from early April to the end of June) allowed the recording of exact laying dates and hatching
dates for all broods established in nest boxes.

Nestlings were collected for the common garden experiment at 7 to 12 days of age and were
brought to the Netherlands Institute of Ecology (NIOO-KNAW, Wageningen, Netherlands)
where they were hand raised under standardized conditions. We used 169 blue tits that were
collected in 2010 and 2011 in the deciduous habitat (2010: 42 birds, 7 broods; 2011: 39 birds, 6
broods) and in the evergreen habitat (2010: 44 birds, 10 broods; 2011: 44 birds, 8 broods). In
2010, before collecting chicks, broods were cross-fostered between nests for another experiment.
For this experiment, at 2 to 4 days old, half of the chicks from a given brood were exchanged with half of the chicks of another brood from the same population.

Once collected, all birds were transported by car and hand-fed from Corsica to the Netherlands, and were hand reared until independence as described in Reparaz et al. (2014). Briefly, all the chicks from a given habitat and year were kept in boxes divided into multiple compartments that were not isolated from one another, each compartment containing one nest of 3 to 5 nestlings, until fledgling. Chicks from adjacent nests could easily change compartment, meaning that chicks from different nests were quickly mixed. After fledgling, birds were housed in cages in groups of 2 to 4 birds, irrespective of their sex and nest of origin (assigned randomly). Up to that period, chicks were fed every half-hour, 14 hours per day (7:00 am - 9:00 pm), with a diet consisting of a mixture of curd cheese, ground beef heart, baby cereal, multivitamin solution and calcium carbonate, supplemented with wax moth larvae and bee larvae, until independence. Raising chicks from the different habitats at exactly the same time would have been ideal but was impossible because chicks in the Regino and the Fango valleys hatch one month apart. However, chicks from different nests and habitats could easily see and hear each other, as they were raised in the same rooms, and fledglings from the Regino valley were still present in the cages when the younger chicks from the Fango valley arrived in the laboratory. Caretakers were the same for birds of different origins.

At independence, about 35 days after hatching, birds were relocated to larger individual cages or aviaries. Food and water were provided ad libitum. In 2012 and 2015, birds were moved to the Centre d'Écologie Fonctionnelle et Évolutive (CEFE-CNRS; Montpellier, France), where they were kept in outdoor aviaries before being released back into their natal habitat in Corsica.
Morphological measurements were taken during the period at the NIOO-KNAW. Tarsus length was measured once (at > 1 year of age) but body mass was measured several times, always by the same person. We were interested in testing for a difference in adult body mass and thus kept in the analysis only the measures made at one year of age and older.

Behavioural and physiological trials

In total, 169 birds were tested for their exploration behaviour and, among these birds, 137 were tested for handling aggression and 57 for heart rate. All behavioural and physiological traits were measured once for each bird, which prevented us from reporting their repeatability. However, these behavioural and physiological traits have been shown to be repeatable in these two populations in the wild, with repeatability estimates ranging from 0.26 to 0.75 depending on the trait (see Dubuc-Messier et al. 2017 for details). In the present study, exploration behaviour was measured using a different protocol (see below) than the one used in the wild (Dubuc-Messier et al. 2017). Nevertheless, we are confident that the exploration behaviour measured here represents repeatable characteristics of the individuals because this measure has been shown to be repeatable in blue tits in several studies using different protocols (Klun & Brommer 2013; Mutzel et al. 2013; Dubuc-Messier et al. 2017). For details regarding the phenotyping of wild birds used in the $P_{sl}$ calculations, please refer to Dubuc-Messier et al. (2017).

Exploration behaviour

Exploration behaviour trials were done in fall 2011 in the Netherlands Institute of Ecology as described by Reparaz et al. (2014) and using a novel environment chamber slightly modified from Drent et al. (2003). The novel environment chamber consisted of a 4.0 x 2.4 x 2.5m room with five artificial trees. Individuals were placed in cages adjacent to the main chamber 30 to 120
minutes before the trials and introduced in the main chamber through a sliding door. For two
minutes, the observer counted the number of movements between trees and the number of small
jumps on a given tree / branch. Exploration scores was the sum of both and varied from 10 (a
very slow exploration pattern) to 92 (a very fast exploration pattern; Reparaz et al. 2014).

Handling aggression

Handling aggression was measured assessing the bird’s aggression towards a manipulator
(Dubuc-Messier et al. 2017). We used a score ranging from 0 to 3. A score of 0 was the lowest
aggression score (no reaction; high docility) and 3 the highest (see Table S2 for detailed protocol).
Handling aggression was recorded in 2012 and 2015 at the CEFE-CNRS (France). Birds from the
2010 cohort were tested for handling aggression in 2012 or 2015 (at 2 or 5 years of age), while
the entire cohort from 2011 was tested for handling aggression in 2015 (at 4 years of age).
Handling aggression score was assessed blindly with respect to habitat of origin in 2015 and was
assessed by two different observers, one in 2012 and one in 2015.

Heart rate during manual restraint

Heart rate was recorded in 2012 at the CEFE-CNRS (for the 2010 cohort only), as described by
Dubuc-Messier et al. (2017). Within a few minutes after capture, we recorded heart rate for 30
seconds using a digital recorder. We used the software Avisoft SASLab Pro version 5.1 to extract
the mean time interval (sec) between two heartbeats using approximately 100 consecutive
heartbeats per individual.
Molecular markers and \(F_{st}\) calculation

For logistical reasons, we were not able to perform a molecular analysis on the birds used in the common garden experiment. As an alternative, we used a dataset, published by Szulkin et al. (2016) of wild birds from these two populations (i.e. deciduous, \(n = 49\); evergreen, \(n = 83\) individuals) and genotyped at several thousand SNP using RAD-sequencing. We retained loci genotyped over at least 75% of the individuals. To avoid bias during filtering and in the \(F_{st}\) estimates, we pruned highly related individuals from the dataset to keep only individuals linked with values of kinship lower than 0.05 (coefficient of Loiselle; Loiselle et al. 1995; Cheverud 1996) computed in Genodive 2.27 (Meirmans & Van Tienderen 2004). In order to retain loci more likely to be informative, we applied a 5% MAF threshold (Minor Allele Frequency, using vcf tools 0.1.11; Danecek et al. 2011). We pruned the dataset for SNPs that deviated from Hardy-Weinberg-Equilibrium in at least one of the two populations (p-value < 0.05) using vcf tools 0.1.11. We retained only the first SNP of each 100 bp locus. To obtain a set of SNPs more likely to be neutral, we filtered out SNPs potentially under divergent selection between the two habitats (p-value < 0.015; 0.7 % of total SNPs removed). This was done with a Bayescan 2.0 test (Foll & Gaggiotti 2008; 5 000 pilot iterations, 50 000 burnin, prior odds of 100). Average \(F_{st}\) and 95% confidence intervals were estimated using the R-package hierfstat 0.04-22 (Goudet 2005). The final dataset contained 69 individuals (32 and 37 individuals in the deciduous and evergreen habitats, respectively) genotyped at 5407 SNPs.

Statistical analysis

Genetic divergence between habitats of origin

We tested for a genetic difference between the two habitats for each trait with univariate linear mixed-models using the phenotype of each bird as a response variable and habitat of origin, sex,
and their interaction as fixed effects. When we found a significant interaction between habitat of
group and sex, we ran a separate model for each sex. Specific confounding variables were added
as fixed effects for each particular trait. For exploration score, we added a cohort term as fixed
effect to test for any environmental effect early in life or during the hand-rearing period in
captivity. Novel environment tests were done on the two cohorts at the same time (in autumn
2011). Thus, at the time of the test, individuals born in 2010 were almost 1½ years old, while
individuals born in 2011 were 5 months old. Hence, in this model, the cohort term controlled for
the combined effect of cohort and age. For handling aggression score, we added cohort, time of
day (hour), and year of test (2012 or 2015) as fixed effects. For heart rate models, we added as
fixed effect mean individual adult body mass because heart rate is related to the metabolic rate
and both are positively related to body mass (Green et al. 11). Heart rate recordings were done
in 2012 on the 2010 cohort only. We therefore did not add a fixed effect for bird age, cohort or
year to avoid redundancy. We also added in heart rate models the time of day (hour) as a fixed
effect. For body mass, we added age as a continuous variable, cohort, and time of day (hour). For
tarsus length, we added cohort only as fixed effect (i.e. 2010 and 2011).

In all models, we used random intercepts for the brood of origin and rearing brood to account for
the non-independence of birds coming from the same brood or / and the effect of foster parents
for nestlings that have been cross-fostered prior to the captivity period. Because body mass was
measured several times for each bird, we also added a random intercept for bird identity for this
trait.

All response variables were Z-transformed prior to analyses. We tested the significance of the
fixed effects and selected a minimal models by LRT (log likelihood ratio test) in a stepwise
elimination procedure starting with a model that included all variables (Bates et al. 2014). We kept all the random effects in final minimal models. We present in Table S3 the L-ratios and p-values associated with all variables in initial models. Analyses were done with R (R Core Team 2017) using the function lmer of the package lme4 (Bates et al. 2015). Confidence intervals (95%) were generated with the function confint.merMod (lme4). We assumed a Gaussian distribution for all traits, which was confirmed after visual inspection of the residuals. We also evaluated the population of origin effect across all five traits using Fisher’s combined probability test run with the sumlog function of the R package metap (Dewey 2017).

Qst, Fst and Fst comparison

Because birds were raised in a common garden, we considered that a phenotypic difference among individuals was a realistic approximation of an additive genetic effect. For each trait, we thus calculated the Qst between the two habitats based on the phenotypes of birds from the common garden using a procedure similar to Bertrand et al. (2016) with univariate mixed models in a Bayesian framework. We calculated Qst as:

\[ Q_{st} = \frac{\sigma_B}{\sigma_B + 2\sigma_W} \]

Where \( \sigma_B \) is the between-habitat phenotypic variance and \( \sigma_W \) the within-habitat variance (or residual; Wright 1949). The two variance components were extracted from a univariate linear mixed model including habitat of origin (and identity of the bird for body mass) as random intercepts. We also included the fixed effects structure selected previously (minimal model) excluding the term habitat of origin. We calculated \( \sigma_B \) as the variance attributable to the habitat
of origin and $\sigma_W$ as the residual variance (or for body mass as the sum of the variance attributable to the residual and to the individual identity; Bertrand et al. 2016). We did not include any broods effects in these models because the variance attributable to the brood is also attributable to the population of origin. We present the between-habitat variance for each study trait extracted from the models used to calculate $Q_{st}$ in Table S5. We calculated $P_{st}$ as $Q_{st}$ but used as random intercepts habitat of origin, the identity of the bird and the observer identity (for handling aggression and heart rate) along with the significant fixed effects detailed in Dubuc-Messier et al. (2017). For $P_{st}$ calculation, we calculated $\sigma_B$ as the variance attributable to habitat of origin and $\sigma_W$ as the sum of the variance attributable to the observer, to the residual variance and the individual identity.

These models were performed with MCMCglmm package (Hadfield 2010) in R using slightly informative priors (i.e. $V = V_p / n$, $nu = 1$ or 0.5; $V_p$ is the total phenotypic variance of the trait and $n$ the number of random effects), 10 million iterations, a thinning of 200 and a burn-in phase of 500. Because the results of the models with different $nu$ were similar, we used the posterior distribution of models with $nu = 1$ in $Q_{st}$ and $P_{st}$ calculations. We assessed the presence of autocorrelation with the function autocorr (MCMCglmm package). All models showed an autocorrelation less than $10^{-4}$. We checked for model convergence with the function gewe.diag of the coda package (Plummer et al. 2006). For all traits, we calculated the ratio $Q_{st}$/mean $F_{st}$ for each sample of the posterior distribution and report the posterior mode of the ratio and its 95% credibility intervals (calculated using the HPDinterval function of the package lme4). We assumed that $Q_{st}$ differed significantly from $F_{st}$ when the credibility interval around the ratio did not include one.
Results

Divergence between habitats of origin

The Fisher combined probability test method on all studied traits indicated an overall significant effect of the habitat of origin (chi-squared: 54.647, df = 10 and p-value < 0.001). Below we present the results for each trait separately.

Behavioural and physiological traits

For birds in the common garden experiment, habitat of origin had a significant effect on the two behavioural traits: blue tits from the deciduous habitat were faster explorers and were more aggressive to the handler (Table 1; Fig. 1). Birds from the deciduous habitat had a lower heart rate than birds from the evergreen habitat (Table 1; Fig. 1). We found a trend for an interaction between habitat of origin and sex for heart rate (L-ratio = 3.360, d.f. = 1, p-value = 0.067): evergreen males had a higher heart rate than deciduous males [estimate = 1.24 (95% CI: 0.31; 2.17), L-ratio = 6.260, d.f. = 1, p-value = 0.010] but there was no habitat of origin effect for females (L-ratio = 2.150, d.f. = 1, p-value = 0.142). There was no interaction between sex and habitat of origin for the two behavioural traits, but there was a difference in exploration score between sexes (Table 1).

Morphological traits

Habitat of origin also had a significant effect on the two morphological traits: deciduous birds were heavier and had a longer tarsus than evergreen birds (Table 1; Fig. 1). We did not find any interaction between habitat of origin and sex for these two traits (tarsus length: L-ratio = 0.226,
d.f. = 1, p-value = 0.634; body mass: L-ratio = 0.155, d.f. = 1, p-value = 0.694). Among-
individual differences in body mass were significant and represented 45% of the total variance of
the trait [variance = 0.34 (95% CI: 0.26; 0.46), L-ratio = 421.95, p-value < 0.001].

Brood effects

Differences among broods of origin explained a significant portion (78%) of the total phenotypic
variance in body mass, but not for the other traits (Table S4). Differences among rearing broods
explained a significant portion of the total variance in tarsus length (22%) and a marginally
significant portion of total variance in heart rate (30%, p-value = 0.07) but not for the other traits.

Qst, Pst and Fst comparison

We found a significant but small genetic differentiation between the two populations [mean Fst
over all loci = 0.004 (95% CrI: 0.003; 0.005), p-value < 0.001]. Qst was higher than Fst with non-
overlapping intervals for all traits. The ratio between the Qst and Fst was significantly greater than
one for all traits. Credibility intervals for Qst and Pst overlapped for all traits (Table 2).

Discussion

Our common garden experiment suggests a genetic divergence in personality, physiological and
morphological traits between two blue tit populations inhabiting contrasted habitats separated by
a small spatial distance in regards to the species dispersal capacity. Adult blue tits originating
from the evergreen habitat displayed slower exploration behaviour, lower handling aggression
(higher docility), faster heart rate, lower body mass and shorter tarsus compared to birds from the
deciduous habitat (Table 1; Fig. 1). These differences are similar to the ones measured in the wild
suggesting that plasticity alone is not responsible for the observed phenotypic divergence in the wild (Charmantier et al. 2016; Dubuc-Messier et al. 2017). In addition, we found a significant F$_{st}$ between the two populations, but its low value (0.004) indicates current or past gene flow, in concordance with previous findings (Szulkin et al. 2016). The Q$_{st}$ - F$_{st}$ comparisons revealed that blue tits from these populations are more differentiated for personality, physiological and morphological traits then they are at the genome-wide level (Table 2). These results suggest that genetic drift alone does not explain the observed divergence between the two populations and that differences in selection regimes are responsible for this divergence.

The divergence we describe in personality, physiological and morphological traits is likely to be mainly of genetic origin, since birds from both habitats were raised in identical conditions from their first week of life to up to five years. In addition, the divergence found in this study for adult body size is consistent with previous studies that have found divergent selection between the two populations for morphological traits (Blondel et al. 2002; Teplitsky et al. 2014) and moderate to high heritability for these traits (0.29 to 0.51; Teplitsky et al. 2014). However, we cannot completely exclude that early environmental effects such as non genetic inheritance, occurring before the chicks were sampled from their nest were at least partly responsible for the observed patterns (Kruuk & Hadfield 2007; Räsänen et al. 2007; Bonduriansky & Day 2009; Bouwhuis et al. 2010; van Oers et al. 2015). Such early environmental effects might be particularly important for tarsus length, which is usually fixed at fifteen days of age for this species. However, for behavioural traits, such strong environmental effects lasting for up to five years are unlikely, since very few studies have reported long-term consequences of early environmental conditions for the studied traits (Taylor et al. 2012; Petelle et al. 2015) and because maternal effects are known to decrease during ontogeny (Cheverud et al. 1983; Wilson et al. 2007). One way to
control for very early environmental effects would be to allow the birds to breed in captivity and
compare the phenotypes in the offspring generation. However, this type of experiment presents
significant challenges that have so far prevented their feasibility in our study system. In
particular, while it is possible to maintain blue tits in aviaries for short time experiments (Reparaz
et al. 2014) it is difficult to make them breed in captivity (Lambrechts et al. 1999).

Some studies have raised concerns regarding $Q_{st}$ and $F_{st}$ estimation and their comparison
(Leinonen et al. 2013). In particular between-population variance and thus $Q_{st}$ estimation may be
imprecise when a small number of populations are compared like it is the case in our study
(O’Hara and Mërila 2005; Leinonen et al. 2013). However, simulations have shown that a small
number of populations results in a downward bias in $Q_{st}$ estimation when $Q_{st}$ is high (O’Hara and
Mërila 2005). Another important concern is whether genetic markers involved in $F_{st}$ estimation
are truly neutral (Leinonen et al. 2013). In this study, we used an $F_{st}$ calculated from markers that
included the whole genome. Although we filtered SNPs under potential divergent selection, it is
possible that we included potentially non-neutral regions (or that we removed some neutral ones).
However, using microsatellites, Porlier et al. (2012) have found a lower $F_{st}$ (0.001) between the
same populations during a similar time period (year 2009). Hence, although $Q_{st}$ and $F_{st}$
comparison have some limitations, these limitations should most probably have limited our
capacity to detect significant $Q_{st}$ - $F_{st}$ differences rather than reveal false differences.

Environmental heterogeneity, divergent selection and local adaptations

The importance of environmental heterogeneity and gene flow for phenotypic divergence has
mainly been studied for life history and morphological traits and much less for behavioural traits.
Indeed, few studies have disentangled so far the role of plasticity from that of genetic differences
in shaping the phenotypic divergence of populations for behavioural traits (Bell 2005; Dingemanse et al. 2007; Herczeg et al. 2013; Karlsson et al. 2016; Jacquin et al. 2016) and even fewer for highly mobile avian species (Atwell et al. 2012; Miranda et al. 2013). In addition, to our knowledge, no studies has until now reported $Q_{st}$ - $F_{st}$ comparisons involving personality traits. This shortage of study is probably due to the fact that personality traits are often considered plastic and thus less prone to genetic divergence and local adaptations than morphological traits. Yet, the results of our study suggest a genetic divergence for personality traits and that this divergence could be as strong as for morphological traits (Table 1 and 2).

Past studies in this system and on personality variation suggest that the genetic divergence found here could be the result of the coevolution of multiple types of traits in response to the ecological context of each habitat. Indeed, an increasing number of studies are suggesting that life-history and personality traits could have co-evolved to form a pace-of-life syndrome (Réale et al. 2010). For example, empirical and theoretical studies are suggesting that high investment in early reproduction at a cost of reduced residual reproductive value (either via survival or future reproduction) should be associated with boldness, fast exploration, and high aggressiveness (Réale et al. 2010; Wolf et al. 2007). Our results on this system are consistent with the pace-of-life syndrome hypothesis. Blue tits from the deciduous habitat, which are more aggressive and faster explorers, have a shorter lifespan and a lower residual reproductive value, but larger clutch sizes than birds from the evergreen habitat (Grosbois et al. 2006; Charmantier et al. 2016; Dubuc-Messier et al. 2017; Table S1). Our results suggest that these divergences for personality traits are genetic and the $Q_{st}$ - $F_{st}$ comparisons revealed that they are likely the result of divergent selection pressures rather than drift. In addition, studies on other blue tit or great tit (Parus major) populations have found that the personality phenotype is heritable and related to fitness (van Oers
& Sinn 2011; Class et al. 2014). Therefore, taken together, our results suggest that the personality phenotypes of birds living in these habitats could have evolved and be implicated in blue tit adaptation to local ecological conditions prevailing in each habitat.

Brood effects

We did not find any significant brood-of-origin effect for handling aggression, exploration score, heart rate, and tarsus length. Since all these traits except heart rate have been shown to be heritable in previous studies on blue tits (van Oers & Sinn 2011; Class et al. 2014; Teplitsky et al. 2014), the absence of heritable variance in our analysis is most probably explained by the relatively small number of broods. Estimating heritability was not the goal of this study, we only wanted to control for dependence issues associated with the use of sibs.

The partial cross-fostering manipulation before the common garden experiment revealed a significant rearing brood effect for tarsus length. This result suggests that the rearing environment between 2 days to 12 days old can have a significant impact on this morphological trait. Contrarily to the other traits that are more labile, tarsus length generally stabilises at fifteen days of age in blue tits. We were, therefore, able to capture the early environmental effect for this trait by measuring the adult phenotype. We found a marginally significant brood of rearing effect for heart rate but not for other traits. There may be several reasons for such results. First, these traits may not be sensitive to the rearing environment. Second, it is possible that - as for brood of origin – these traits are slightly sensitive to early environmental effect (2 to 12 days) but that we lack power to detect it.
Genetic and environmental effects are not mutually exclusive

Genetic divergence does not preclude a plastic response to ecological conditions specific to each habitat. For example, in the wild, the phenotypic difference in male heart rate between habitats was not significant (Dubuc-Messier et al. 2017), but using the common garden experiment we found here a significant difference in male heart rate. It is thus possible that plastic responses of heart rate to habitat specific ecological conditions in the wild may have hidden the genetic divergence (Conover & Schultz 1995). In addition, the important temporal variation in mean handling aggression in the wild shown by Dubuc-Messier et al. (2017) in each population, suggests that individuals can partly adjust their personality phenotype for this trait depending on the current local conditions. However, for all traits, the $P_{st}$ between wild birds was not statistically different from their $Q_{st}$, suggesting that environmental effects in the wild might not result in stronger or weaker differentiation compared to the genetic differentiation.

Conclusion

Our study suggests a genetic divergence for personality, physiological and morphological traits between two blue tit populations that occupy different habitats but that are separated by small spatial distances compared to the dispersal ability of the species and connected by gene flow. The present study and past results for this system suggest that these differences are likely due to different selection pressures and may represent local adaptations. These results thus emphasize the role of environmental heterogeneity for intra-specific phenotypic diversity and suggest that genetic population divergence is possible at small spatial scales (relative to their dispersal ability) for behavioural traits.
Acknowledgements

We thank the landowners for permission to work on their properties and all the blue tit crew for their work. We declare no conflict of interests.

Data accessibility

The dataset will be shared on dryad upon publication.

References


538  Foll, M. & Gaggiotti, O. 2008 A genome-scan method to identify selected loci appropriate for
541  Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-Statistics’.
544  Green J A. 2011. The heart rate method for estimating metabolic rate: review and
Climate impacts on mediterranean Blue Tit survival: an investigation across seasons and
551  Hadfield, J. 2010. MCMC methods for multi-response generalized linear mixed models: The
554  Herczeg, G., Ab Ghani, N.I. & Merilä, J. 2013. Evolution of Stickleback feeding behaviour:
genetics of population divergence at different ontogenetic stages. J. Evol. Biol. 26: 955–
962.
Parallel and nonparallel behavioural evolution in response to parasitism and predation in


Figure 1. Mean phenotypes of blue tits originating from two distinct populations and habitats (deciduous and evergreen) in Corsica (France) and reared in a common garden. A) exploration score, B) handling aggression score, C) heart rate during manual restraint (heart beats/min.), D) tarsus length (mm) and E) adult body mass (g). Boxplots on raw data, the boxes represent the first and the third quartile, the lines represent the median, the ends of the whiskers represent the minimum data in the 1.5 * the interquartile range, dots represent extreme data. All differences are significant (see Table 1 for details).
Table 1. Final models describing the phenotype of blue tits originating from two distinct populations and habitats (deciduous and evergreen) in Corsica (France) and reared in a common garden.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Terms</th>
<th>Estimates</th>
<th>95% CI</th>
<th>L-ratio</th>
<th>d.f.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploration score</td>
<td>Intercept</td>
<td>-0.32</td>
<td>-0.62; -0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>-0.48</td>
<td>-0.78; -0.19</td>
<td>9.70</td>
<td>1</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>0.26</td>
<td>0.004; 0.52</td>
<td>3.97</td>
<td>1</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>Cohort</td>
<td>0.88</td>
<td>0.59; 1.17</td>
<td>23.91</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Handling aggression</td>
<td>Intercept</td>
<td>0.45</td>
<td>0.18; 0.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>-0.82</td>
<td>-1.18; -0.46</td>
<td>14.96</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Heart rate during restraint (HR)</td>
<td>Intercept</td>
<td>-0.57</td>
<td>-1.06; -0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>0.98</td>
<td>0.35; 1.62</td>
<td>8.17</td>
<td>1</td>
<td>0.004</td>
</tr>
<tr>
<td>Body mass</td>
<td>Intercept</td>
<td>-1.07</td>
<td>-1.40; -0.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>-0.33</td>
<td>-0.63; -0.03</td>
<td>4.46</td>
<td>1</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>-0.56</td>
<td>-0.77; -0.35</td>
<td>25.08</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>0.27</td>
<td>0.21; 0.33</td>
<td>74.23</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Time of day</td>
<td>0.09</td>
<td>0.07; 0.11</td>
<td>75.50</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Tarsus length</td>
<td>Intercept</td>
<td>-0.25</td>
<td>-0.58; 0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>-0.60</td>
<td>-1.00; -0.19</td>
<td>7.74</td>
<td>1</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>1.04</td>
<td>0.81; 1.28</td>
<td>61.46</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

The deciduous habitat, females, and cohort 2010 were set as references in models. Estimates are from a model with the brood of rearing and brood of origin in random effect (and individuals identity for body mass), variance estimates are shown in Table S4. L-ratio and p-values are from the comparison of a full model and a model without the variable of interest. The p-values and L-ratio associated with each parameter in initial models before selection are presented in Table S3.
Table 2. $Q_{st}$ and $P_{st}$ values (posterior mode) for each trait (and 95% credible interval (CrI)), mean $F_{st}$ and $Q_{st} / F_{st}$ ratio [posterior mode and associated 95% CrI] between two blue tits populations originating from distinct populations and habitats (deciduous or evergreen) in Corsica (France) and reared in a common garden.

<table>
<thead>
<tr>
<th>Traits</th>
<th>$Q_{st}$ (95% CrI)</th>
<th>$P_{st}$ (95% CrI)</th>
<th>$Q_{st} / F_{st}$ ratio (95% CrI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploration score</td>
<td>0.084 (0.029; 0.804)</td>
<td>0.063 (0.018; 0.727)</td>
<td>20.982 (7.266; 201.065)</td>
</tr>
<tr>
<td>Handling aggression</td>
<td>0.129 (0.034; 0.832)</td>
<td>0.045</td>
<td>32.309 (8.525; 208.025)</td>
</tr>
<tr>
<td>Heart rate during manual restraint (HR)</td>
<td>0.101 (0.033; 0.846)</td>
<td>0.032</td>
<td>25.320 (8.244; 211.475)</td>
</tr>
<tr>
<td>Body mass</td>
<td>0.069 (0.018; 0.736)</td>
<td>0.095</td>
<td>17.144 (4.541; 183.998)</td>
</tr>
<tr>
<td>Tarsus length</td>
<td>0.197 (0.050; 0.872)</td>
<td>0.212</td>
<td>49.368 (12.455; 217.881)</td>
</tr>
</tbody>
</table>

$Mean F_{st}$: 0.004 (0.003; 0.005)

$Q_{st}$ have been calculated from the phenotype of birds raised in a common garden and $P_{st}$ from the phenotype of wild birds.
Figure 1.
Online supporting information

Table S1. Caterpillar abundance, life-history, morphological and personality phenotypes (mean (n)) of the two Corsican blue tit populations (France) in the wild.

Table S2. Blue tit handling aggression scale.

Table S3. L-Ratio, degree of freedom and p-values associated with each parameter in initial models describing the phenotype of blue tits originating from two distinct habitats (deciduous and evergreen) in Corsica (France) and reared in a common garden.

Table S4. Variance components, L-ratio and p-values for studied traits in two blue tits populations in Corsica (France) reared in a common garden.

Table S5. Between-habitat variance (posterior mean and 95% CrI) for each study trait extracted from the models used to calculate Qst.
Online Supporting Information

Average caterpillar abundance, life-history, morphological and personality phenotypes

Table S1. Caterpillar abundance, life-history, morphological and personality phenotypes (mean ($n$)) of the two Corsican blue tit populations (France) in the wild.

<table>
<thead>
<tr>
<th>Populations</th>
<th>Deciduous</th>
<th>Evergreen</th>
</tr>
</thead>
<tbody>
<tr>
<td>First year of monitoring</td>
<td>1993</td>
<td>1976</td>
</tr>
<tr>
<td>Caterpillar abundance(^1)</td>
<td>762.87</td>
<td>87.10</td>
</tr>
<tr>
<td>Annual adult survival probability(^2)</td>
<td>0.391 (6)</td>
<td>0.574 (14)</td>
</tr>
<tr>
<td>Date of first egg laying (1 = March 1(^{st}))(^3)</td>
<td>38.56 (1233)</td>
<td>70.08 (1920)</td>
</tr>
<tr>
<td>Male body mass (g)(^3)</td>
<td>9.82 (1032)</td>
<td>9.37 (1607)</td>
</tr>
<tr>
<td>Female body mass (g)(^3)</td>
<td>9.66 (1153)</td>
<td>9.23 (1616)</td>
</tr>
<tr>
<td>Male tarsus length (mm)(^3)</td>
<td>16.52 (578)</td>
<td>16.27 (789)</td>
</tr>
<tr>
<td>Female tarsus length (mm)(^3)</td>
<td>16.05 (614)</td>
<td>15.84 (798)</td>
</tr>
<tr>
<td>Clutch size(^3)</td>
<td>8.50 (1235)</td>
<td>6.61 (1913)</td>
</tr>
<tr>
<td>Number of fledglings(^3)</td>
<td>6.60 (1092)</td>
<td>4.15 (1273)</td>
</tr>
<tr>
<td>Mean exploration speed (cm/s) ± s.d.(^4)</td>
<td>13.52 ± 8.39 (176)</td>
<td>10.37 ± 7.49 (117)</td>
</tr>
<tr>
<td>Mean handling aggression score ± s.d.(^4)</td>
<td>1.69 ± 0.95 (703)</td>
<td>1.49 ± 0.99 (549)</td>
</tr>
<tr>
<td>Mean heart rate during manual restraint ± s.d.(^4)</td>
<td>963.30 ± 87.80 (159)</td>
<td>976.24 ± 86.99 (91)</td>
</tr>
</tbody>
</table>

\(^1\) mean maximal frass mg/m\(^2\) per day in each population (sampled between 2011 and 2015 during the breeding period using 0.25m\(^2\) trays placed under the forest canopy and collected twice a week, see Zandt et al. 1990 for details about the sampling procedure); \(^2\) Dubuc-Messier et al. In prep; \(^3\) Charmantier et al. 2016 (collected between the first year of monitoring and 2014); \(^4\) Dubuc-Messier et al. 2016).
Handling aggression scores

The test was done within two minutes after capture and prior to any other manipulation. The handler held the bird with one hand and placed the bird’s legs between his forefinger and his thumb to let the bird free to move its tails and wings. The handler pointed the forefinger of his other hand at a spot about 2 to 3 cm in front of the bird’s beak and noted if the bird struck at his finger, and the position of its wings and tail. After two seconds in this position, the handler moved his forefinger towards the bird’s beak two or three times and recorded its reaction.

Table S2. Blue tit handling aggression scale.

<table>
<thead>
<tr>
<th>Score</th>
<th>Wings spread</th>
<th>Tail feathers spread</th>
<th>Bird strikes fingers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>No</td>
<td>No</td>
<td>Yes, but only if provoked</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>Yes</td>
<td>Yes, spontaneously</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes, spontaneously</td>
</tr>
</tbody>
</table>

When the bird displayed one reaction specific to one score and another reaction specific to another score, it received an average score between the two. For example, a bird that struck without any provocation (score 2) but did not have its wings and tail feathers spread (score 1) would be scored as 1.5.
Initial models

Table S3. L-Ratio, degree of freedom and p-values associated with each parameter in initial models describing the phenotype of blue tits originating from two distinct habitats (deciduous and evergreen) in Corsica (France) and reared in a common garden.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Terms</th>
<th>L-ratio</th>
<th>d.f.</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploration score</td>
<td>Cohort</td>
<td>23.912</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Sex * Habitat of origin</td>
<td>1.104</td>
<td>1</td>
<td>0.293</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>3.970</td>
<td>1</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>9.697</td>
<td>1</td>
<td>0.002</td>
</tr>
<tr>
<td>Handling aggression</td>
<td>Time of day</td>
<td>0.258</td>
<td>1</td>
<td>0.612</td>
</tr>
<tr>
<td></td>
<td>Cohort</td>
<td>0.052</td>
<td>1</td>
<td>0.819</td>
</tr>
<tr>
<td></td>
<td>Year of trial</td>
<td>0.001</td>
<td>1</td>
<td>0.973</td>
</tr>
<tr>
<td></td>
<td>Sex * Habitat of origin</td>
<td>0.615</td>
<td>1</td>
<td>0.432</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>0.092</td>
<td>1</td>
<td>0.761</td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>20.592</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Heart rate during restraint</td>
<td>Mean body mass</td>
<td>0.256</td>
<td>1</td>
<td>0.873</td>
</tr>
<tr>
<td></td>
<td>Sex * Habitat of origin</td>
<td>3.3601</td>
<td>1</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>1.9081</td>
<td>1</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>9.012</td>
<td>1</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Time of day</td>
<td>0.449</td>
<td>1</td>
<td>0.502</td>
</tr>
<tr>
<td>Body mass</td>
<td>Time of day</td>
<td>75.500</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>74.230</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Cohort</td>
<td>0.014</td>
<td>1</td>
<td>0.905</td>
</tr>
<tr>
<td></td>
<td>Sex * Habitat of origin</td>
<td>0.155</td>
<td>1</td>
<td>0.694</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>25.080</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>4.460</td>
<td>1</td>
<td>0.034</td>
</tr>
<tr>
<td>Tarsus length</td>
<td>Cohort</td>
<td>0.350</td>
<td>1</td>
<td>0.554</td>
</tr>
<tr>
<td></td>
<td>Sex * Habitat of origin</td>
<td>0.226</td>
<td>1</td>
<td>0.634</td>
</tr>
<tr>
<td></td>
<td>Sex</td>
<td>25.070</td>
<td>1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td></td>
<td>Habitat of origin</td>
<td>4.457</td>
<td>1</td>
<td>0.034</td>
</tr>
</tbody>
</table>

The brood of rearing and brood of origin identity are fitted as random effect in all models (and individuals identity for body mass), variance estimates are shown in Table S3. L-ratio and p-values are from the comparison of a full model and a model without the variable of interest.
Variance components

Table S4. Variance components (brood of origin, brood of rearing, and residuals), L-ratio, and p-values for studied traits in two blue tits populations in Corsica (France) reared in a common garden.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Brood of origin</th>
<th>Rearing broods</th>
<th>Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Variance (95% CI)</td>
<td>L-ratio</td>
<td>d.f.</td>
</tr>
<tr>
<td>Exploration score</td>
<td>0.05 (0.00; 0.15)</td>
<td>0.76</td>
<td>1</td>
</tr>
<tr>
<td>Handling aggression</td>
<td>0.01 (0.00; 0.13)</td>
<td>0.002</td>
<td>1</td>
</tr>
<tr>
<td>HR</td>
<td>0.05 (0.00; 0.30)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Tarsus length</td>
<td>0.09 (0.00; 0.29)</td>
<td>1.69</td>
<td>1</td>
</tr>
<tr>
<td>Body mass</td>
<td><strong>0.07 (0.01; 0.14)</strong></td>
<td>4.10</td>
<td>1</td>
</tr>
</tbody>
</table>

L-ratio and p-values are from the comparison of a full model and a model without the variable of interest. Bold indicates significant variance components.
Table S5. Between-habitat variance (posterior mean and 95% CrI) for each study trait extracted from the models used to calculate Qst.

<table>
<thead>
<tr>
<th>Traits</th>
<th>Between habitat variance (95% CrI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploration score</td>
<td>1700 (11.88; 2464)</td>
</tr>
<tr>
<td>Handling aggression</td>
<td>2.909 (0.021; 4.470)</td>
</tr>
<tr>
<td>Heart rate during manual restraint (HR)</td>
<td>34.971 (245.5; 57937)</td>
</tr>
<tr>
<td>Body mass</td>
<td>1.564 (0.015; 3.041)</td>
</tr>
<tr>
<td>Tarsus length</td>
<td>0.142 (0.114; 0.181)</td>
</tr>
</tbody>
</table>

References


