



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

Gabrielle Dubuc-Messier, Samuel Caro, Charles Perrier, Kees van Oers, Denis Reale, A. Charmantier

► To cite this version:

Gabrielle Dubuc-Messier, Samuel Caro, Charles Perrier, Kees van Oers, Denis Reale, et al.. Gene flow does not prevent personality and morphological differentiation between two blue tit populations. *Journal of Evolutionary Biology*, 2018, 31 (8), pp.1127-1137. 10.1111/jeb.13291 . hal-02326707

HAL Id: hal-02326707

<https://hal.science/hal-02326707>

Submitted on 22 Oct 2019

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

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



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

Journal:	<i>Journal of Evolutionary Biology</i>
Manuscript ID	JEB-2018-00170
Manuscript Type:	Research Papers
Keywords:	Cyanistes caeruleus, genetic divergence, local adaptation, personality, Phenotypic plasticity, Qst - Fst

Montreal, 8 april 2018

Dear editor,

Please find enclosed a manuscript entitled “Gene flow does not prevent personality and morphological differentiation between two blue tit populations” that we wish to submit for publication in your journal.

We believe our work will be of interest to the broad readership of your journal because it challenges the general idea that gene flow limits populations genetic divergence and the importance of plasticity versus genetic effects in shaping population divergence for personality traits. We report on a common garden experiment exploring the genetic basis of phenotypic differences observed in the wild for personality, physiological and morphological traits between two blue tit (*Cyanistes caeruleus*) populations inhabiting contrasting habitats separated by a small spatial scale and connected by gene flow. We raised nestlings originating from the two habitats in aviaries for up to five years and then compared their adult phenotypes. Our results revealed differences similar to those found in the wild, suggesting a genetic divergence for all traits. In addition, Q_{st} - F_{st} comparisons revealed that the observed quantitative genetic divergence is likely the result of contrasting selection pressures rather than of neutral processes. Our study is one of the first to report Q_{st} - F_{st} comparisons for personality traits and suggests that genetic divergence is possible at a small spatial scale for behavioural and physiological traits. Such small scale evolution of animal personality and physiology has rarely been reported and shows that population genetic divergence is possible at a small spatial scale for traits generally considered less prone to genetic divergence.

We would like to thank you and the reviewers for providing constructive comments on the first version of the manuscript that we think have greatly improved our work. Please find attached a revised version of our manuscript and a detailed answer to all comments.

We agree for the dataset to be shared on Dryad after the paper is published. The manuscript is not under consideration for publication in another journal. All persons entitled to authorship have been named and have approved the submission of this version of the manuscript. The manuscript is 5978 words.

We hope you will consider for publication in your journal this revised version of our manuscript and we look forward to your assessment.

Best regards,

The authors

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.

c. Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), PO Box 50,
6700 AB, Wageningen, The Netherlands.

* co-last authors

Running head

Divergence between blue tit populations

Corresponding author

Gabrielle Dubuc-Messier: dubuc-messier.gabrielle@courrier.ugam.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 CEFÉ 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-OREME 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.

Dear Dr Klingenberg,

Thank you for your feedback regarding our manuscript # JEB-2017-00630. We are very thankful for the new comments you provided as well as those from the two reviewers. We outline below how all these comments have been taken into account to provide a revised version of our manuscript that will now, we hope, be accepted for publication in Journal of Evolutionary Biology.

Comments from Editor

In my own reading of your paper, I have come across the act that you used the method of random skewers for comparing P matrices. This method has been criticised for lack of power and other statistical properties. Very recently, such a critique has been published by Jim Rohlf. Please have a look at this paper and think how the criticism applies to your analysis. I am not prescribing a particular course of action to you, but I would like you think about it as part of the further revisions.

>> Since the Q_{st} , P_{st} and F_{st} comparison is more robust and provides stand-alone results that do not necessitate any complementary analysis, we have decided to remove the matrix comparison done with the random skewer method that has been recently criticized. This does not change any conclusion from the manuscript, but it should remove any doubt regarding random skewers.

+++++

Comments from reviewer 1

I have only two comments:

1. - The first one refer to the use of the term phenotypic plasticity (or plastic response) as synonym of environmental factors affecting phenotypes. I think it may confound readers and, whenever possible, I suggest the used of environmental factors throughout the manuscript. This is because phenotypic plastic response to environmental condition may have a genetic component (see for instance Charmantier et al. (2008), Science) and, thus, phenotypic plastic responses may be due not only to environmental factors, but to genetic factors. Remember that the reaction norms may have strong genetic components, but you use plasticity as solely reflecting environmental effects. Traditional terms used in quantitative genetics are genetic, environmental and maternal effects and, thus, would strongly recommend the used of "environmental factors" instead of plasticity.

>> We partly disagree here with the reviewer's opinion that our use of the term "plasticity" brings confusion in all its use. We now clarify that by plasticity we refer to "the adjustment of individual phenotypes in response to environmental factors" (definition provided in line 30). Note that plasticity does not result in a genetic change, however variance in plasticity can have a genetic origin when plasticity is heritable and lead to evolution of different plasticity across populations. Although this is a side issue to our study, we now explain this in lines 65 to 68. Finally, when addressing the issue of the Q_{st}/P_{st} comparison, we have followed the reviewer's advice and changed plasticity to environmental factors (lines 119 and 454).

2. - The second comment refers to the scarce information provided for used mixed model. You should explain how mixed models accounted for the structure of the data set (random blocks) (see for instance Schielzeth, H. & Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. Behav Ecol 20: 416-420). Do you used random

intercept, random slope, or random intercept-and-slope models? Was identity of rearing brood nested within identity brood of origin? In the case of bird identity, this random factor should be nested within the interaction between brood of origin and rearing brood (or, in the case of a completely hierarchized structure of random factors, at the lower level). Remember that a correct definition of hierarchized random blocks is essential for proper estimations of genetic and environmental factor explaining phenotypic variation.

>> Thank you for pointing out this imprecision. We used random intercepts for bird identity, brood of rearing and brood of origin. We now clarified this (L249 to 253).

Comments from reviewer 2

Comments on your written responses:

1. Re. my comment 12: I meant that care takers may not have fed blind with respect to origin. While I don't really think there are box/nest/location, date or care taker effects, some readers may disagree, so perhaps it is best to mention these possibilities and state why you discard it. I think it is always a strength of a paper if it presents (and if space permits) discusses its potential weaknesses, as long as it doesn't distract.

>> The populations that we study here differ in many traits, one of them being phenology (i.e. timing of reproduction). Because of this timing difference, raising the chick at the same time is impossible. We have made every effort possible for keeping the birds of the two populations into the same conditions, but for sure there are many little, sometimes unidentified, differences, that were impossible to control. Temperature was certainly a bit different (later in the season for the Fango birds), as well as humidity, photoperiod, they travelled in a different boat, etc. among which are the rearing conditions (caretakers, etc). Thus, it is impossible to pinpoint and inventory the exact differences that might have been of relevance to the birds. We have therefore clarified this timing difference between the populations (see L151 to 157), but we believe it is not possible to discuss it much further since we do not know if there were any relevant difference (and which ones) that the birds could have cued on despite our efforts to homogenize their environments and experiences.

2. Re. comment 41: you don't address the concern of multiple testing, so no mention of the need for correction for a large number of tests, which inflate the probability to find significant results. So Bonferroni correction or something similar might be called for.

>> Bonferroni corrections have been criticized for being too conservative (Moran 2003). However, applying a Bonferroni correction to our results would lead to the same conclusions as we found a p -value < 0.01 for all study traits except body mass. If the editor deems it useful, we can add this information in the results section but prefer not to.

*Moran, M. D. (2003). Arguments for rejecting the sequential Bonferroni in ecological studies. *Oikos*, 100(2), 403-405.*

And instead of Wright, I meant Fisher's method (https://en.wikipedia.org/wiki/Fisher%27s_method), which to my opinion can also be used as a within-study meta-analysis. You repeatedly test the same hypothesis, just changing the response variable. So you could apply Fisher's method to the p -values obtained for the habitat of origin tests.

>> We now use this test. We provide this information and the results of this test in lines 262 to 264 and 306 to 308 ($p\text{-value} = 3.675 \times 10^{-8}$).

3. Re. comment 46: with evergreen habitat being atypical, I mean not the habitat itself but in the evolutionary history of the Blue tit, assuming that the species is largely adapted to deciduous trees and that colonisation of evergreen forest is relatively recent, or that gene flow will be mostly from deciduous to evergreen populations (given the relative population sizes of Blue tits in each of the habitats at a large, European?, scale). Then, even if Blue tits in evergreen habitats are exposed to specific selection pressures and have actually evolved in that direction, that doesn't mean that they are now adapted – there may still be an evolutionary lag. In fact, if you found directional selection on traits *within* a habitat, this would confirm local maladaptation. Whether or not size-corrected body mass is a good measure of overall condition (whatever that is) in ad libitum fed, captive birds, is debatable, but I'm thinking of effects that are more general for maladaptation, such as maternal stress hormones or lack of micronutrients deposited in eggs affecting overall development (body, physiology, brain, behaviour). Again, I don't know if this is likely, but if you think it is not, then explaining this in the paper might be the best way forward.

>> We agree that ongoing selection can be a sign of maladaptation. We also agree that we cannot completely exclude that early environmental effects such as maternal effects occurring before and soon after the chicks were sampled from their nest were at least partly responsible for the observed patterns. We discussed this possibility in length in lines 358 to 377.

Comments on the new version:

4. L 36: Edelaar & Bolnick 2012 TREE is also a useful citation in this section

>> We added this citation (L36)

5. L 53: and/or (you don't need both)

>> We modified the sentence (L51)

6. L 54: also Richardson et al. 2014 TREE microgeographic divergence

>> We added this citation (L52).

7. L 67: given that you (correctly) scale you question to the dispersal distance, whether a species is highly mobile or not is now unimportant, so remove that part after the comma.

>> We changed the sentence (L65).

8. L 110: I like the layered approach to the issue of divergence in this paper, with three different analyses/data sets.

>> Thank you for the positive comment.

9. L 110: in this section you need to introduce the issue of Pst and Qst, explain what Pst is (not so commonly known), and why/that you will assume that the Pst of lab-reared birds should approximate Qst to a greater degree than the Pst of wild birds. (see L 106-107, where you do mention that P approximates G), if there is plasticity in the wild.

>> We added some details and explanations on this topic in L103 to L122.

10. L 144: do you actually know this? (I assume the birds were ringed)

>> We meant they were assigned randomly to a cage (L148).

11. L 145: fed ad libitum (saturation)?

>> ad libitum means that there was always food and water available for the birds. We never force-fed any bird, neither were they food-restricted. This term is commonly used in the literature.

12. L 153: tarsus at which age?

>> Adult (> 1 year of age), we now provide this information (L164).

13. L 188: blind with respect to what?

>> We now provide this information (L199).

14. L 211: what percentage of SNPs did you remove for being potentially under selection? I suppose this must be well under 1%. Nonetheless, many of these outliers will be false positives, i.e. highly divergent SNPs that actually are neutral and therefore should not be filtered out. So you are caught between a rock and a hard place. Perhaps, to obtain a conservative (upper) estimate for Fst, you could also report what value you obtain without this filtering step, so just overall genomic divergence.

>> We removed 0.7% of the total number of SNPs. We now provide this information (L223). Including or excluding SNPs putatively under divergent selection (Fst of 0.004 and 0.006 respectively) provided high Qst/Fst ratio in both cases. Excluding loci putatively under selection therefore had little effect on the general conclusion, (which is expected in a context of putatively highly polygenic traits and very low genetic differentiation among populations, resulting in few outliers SNPs of large effects and high differentiation).

15. L 245: check use of singular/plural

>> Done (L257)

16. L 265-266: so move this introduction up, to line 110

>> Done (L114 to 117).

17. L 278: habitat of

>> Done (L279-280)

18. L 279: and individual in the case of body mass? So basically, all random effects variances plus residual variance?

>> *Yes, we now provide this information (L280-281)*

19. L 280: this is the Pst for wild birds? Why not include the random effect of brood?

>> *These birds were measured once adult, for many of them we did not know the brood of origin.*

And I guess observer is random here, because many people were involved?

>> *Yes, L288*

But did you then include this observer variance component into the within-population variance? (For Pst you don't specify which variances were used to calculate the within-pop variance component).

>> *Yes, we now provide this information (L287 to 289)*

This to me would not seem correct, since observer effects are not within-population genetic variation (as mentioned before in my comment 34 of the previous revision). The same might actually be true for rearing brood effect: this could be indirect genetic effects, but could also be purely or mostly environmental. And the same issue for the residual variance: to what extent is this genetic variance, or just unexplained environmental noise around the genotypic value? I think you now assume it is fully genetic. (Note that I'm partly disagreeing with myself, re. comment 34 last review).

What I'm missing in this section is an explanation/justification of which variables are or are not included in the between- and within-population variance components, and methodological/theoretical citations to back this up. And then the effects of the assumptions and decisions on results made should be discussed later on.

>> *We rewrote partly this paragraph and hope that it is now clearer (L275 to 289). For Qst calculation, following the reviewer's comments, we decided to calculate σ_B as the variance attributable to habitat of origin and σ_W as the residual variance (or for body mass as the sum of the variance attributable to the residual and to the individual identity). We did not include any broods effects in these models anymore because the variance attributable to the brood is also attributable to the population of origin in our case. These changes did not affect our conclusions. We changed the results in Table 2 and S5.*

In our opinion, the observer effect should not be included in the between population variance for Pst calculation but could be considered as residual variance in the context of our analysis. This is why we calculated σ_B as the variance attributable to habitat of origin and σ_W as the sum of the variance attributable to the observer and to the residual variance and the individual effect (L287 to 289).

20. L 285: I think you mean slightly informative priors? Slightly UNinformative means very informative.

>> Yes, we changed the sentence (L292)

21. L 286: I'm surprised you stay with a burn-in of 500 iterations, when running a model of 10 million iterations. Normally the burn-in is more like a third or half of the total iterations. With a thinning of 200, you now discard only the first 2 effective samples, and keep the remaining 50,000 samples. So basically, you have no burn-in period. I propose a burn-in of at least 1 million iterations, unless you know that your chains converge very quickly (in which case you don't need 10 million iterations, which looks like overkill anyway).

>> We compared models with different burn-in period, iterations and nu and found no important difference between these models and really small autocorrelation (L291 to 297).

22. L 295: maybe this information goes better with line 278?

>> Yes, we agree, we changed the location of the sentence to L283-284.

23. L 331: I still don't get this result: you state that the two matrices are more similar to each other than two random matrices. This is very unsurprising. And irrelevant. What we want to know is if they are dissimilar, as you claim you will test in L 108. But you don't do that. I mentioned this before. I think this analysis is not relevant at all – look for an analysis that will tell us if the matrices are identical or not (e.g. if your correlation of 0.9 is different from 1.0, not if it is different from 0.0, but probably random skewers is not the thing to do).

>> This comment is in line with the Editor concern regarding the random skewer method. As explained in our response to the Editor, we have removed this comparison.

24. L 338: extent

>> We removed this part of the sentence

25. L 343: suggests

>> Modified (L343)

26. L 349: replace cannot be for is not

>> Modified (L349)

27. L 351: indicates

>> Modified (L351)

28. L 353: remove genetically

>> Removed

29. L 354: replace cannot for does not

>> *Modified (L355)*

30. L 355: replace could be for are

>> *Modified (L356)*

31. L 360: remove genetic

>> *Removed*

32. L 367: mention here also the significant effects you found for rearing brood, suggesting the acting of such early effects

>> *We discuss the rearing brood effects in lines 433 to 443.*

33. L 387: or that you removed high Fst neutral ones

>> *We now mention this possibility (L387) but we choose to keep it in parenthesis since this issue is not important and not discussed further.*

34. L 388: but you should not use microsatellites for Qst-Fst comparisons, see several papers on this by Jost, Edelaar and other authors, in Molecular Ecology (partly cited in Leinonen et al. 2013).

>> *Yes, this is why we did not use microsatellites in this study, the microsatellite study is mentioned to discuss the Fst level.*

35. L 390: this statement needs to be re-evaluated after checking what SNP filtering does, and whether the Qst calculations are changing based on my comments above.

>> *This statement still holds after re-evaluation of the consequences of the filtering.*

36. L 397: paper by Dingemanse et al. on predator-presence related population divergence in stickleback personalities also comes to mind (J Anim Ecol?)

>> *We added this reference (L398)*

37. L 400: studies

>> *Modified (L419)*

38. L 404: better refer to the Qst values in Table 2?

>> *Modified (L404)*

39. L 413: I agree, but also mention/discuss if you would expect the traits you used to be correlated in your variance covariance matrix (which they are hardly)

>> *As explained in our response to the Editor, we have removed the comparison between the covariance matrices from the manuscript.*

40. L 428: again, you tested if they are similar, but you should test if they are different

>> *This section was removed.*

41. L 435: associated with

>> *Modified (L431)*

42. L 445: remove (2 to 12 days) since this is irrelevant – statistical power is the issue, not the time frame.

>> *We agree, we changed the location of the parenthesis in the sentence (L441).*

43. L 447: this is the same as your first argument – effectively not sensitive relative to other environmental effects

>> *We removed this sentence*

44. L 459: again, this may need to be revised depending on any Q_{st} and P_{st} recalculations. It is kind of strange that divergence is not reduced under a common environment, assuming that any plasticity in the wild would tend to operate in the direction of the divergent selection between habitats. I think this is also the common observation (often even no remaining divergence in a common environment).

>> *We meant that the environmental effects might not be very high in the wild otherwise we would have found a P_{st} significantly different from the Q_{st} . See our response to previous comments about the filtering.*

45. L 463: remove phenotypic

>> *Done.*

46. L 464: replace L 465-466 by: but that genomically are diverged much less.

>> *We prefer the original sentence.*

47. L 466: replace past study on by past results for

>> *Done (L461).*

48. L 470: scales (relative to dispersal ability) for

>> *Done (L464).*

Table S6: for HR and tarsus length, the mean is higher than the credible interval, so this needs to be fixed

>> *We corrected this information (Table S5).*

Table 1: as far as I’m concerned (and you, as you don’t test its significance), you can remove the info for the intercepts).

Time of day: based on d.f. you fitted a linear effect – have you checked if the effect is actually non-linear?As you have enough data, perhaps fit time as categorical (by hour for example). Same for age. This might change (improve?) your Qst estimates, see L 338.

>> *We have only fitted here a linear effects.*

+++++

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, $Q_{st} - F_{st}$

23 **Introduction**

24 Understanding the evolutionary causes of phenotypic divergence among populations is an
25 important aspect of the study of diversity. Environmental heterogeneity can have a major role in
26 generating phenotypic divergence among populations (Wang & Bradburd 2014). Spatial variation
27 in selection pressures resulting from such environmental heterogeneity can lead to genotype by
28 environment interactions for fitness and produce phenotypic and genetic divergence between
29 populations that can lead to local adaptations (Kawecki & Ebert 2004; Wang & Bradburd 2014).
30 Spatial heterogeneity in ecological conditions can also favour the evolution of phenotypic
31 plasticity, i.e. the adjustment of individual phenotypes in response to environmental factors
32 (Pigliucci 2005) and cause phenotypic divergence of populations in the absence of genetic
33 divergence or local adaptation (Sultan & Spencer 2002; Réale *et al.* 2003; Pigliucci 2005).
34 Phenotypic divergence of populations can also be produced by non-random dispersal of
35 individuals between habitat types (Wang & Bradburd 2014). Importantly, phenotypic divergence
36 of populations does not necessarily involve an adaptive process since phenotypic plasticity and
37 non-random dispersal can be non-adaptive (Edelaar & Bolnick 2012; Fitzpatrick 2012; Wang &
38 Bradburd 2014) and can occur in the same or in the opposite direction to genetic divergence
39 (Fitzpatrick 2012). In addition, strong founder effects or genetic drift can also lead to phenotypic
40 and genetic divergence of populations (Slatkin 1987). Establishing the relative importance of
41 environmental versus genetic effects involved in the phenotypic divergence of populations
42 provides fundamental information about the origin of intra-specific diversity in the wild. In
43 addition, determining if this divergence is adaptive or the result of neutral processes is essential
44 because it gives important indications about the eco-evolutionary dynamics of traits and their
45 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 than 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 (Kluen & Brommer 2013; Mutzel *et al.* 2013; Dubuc-Messier *et al.* 2017). For details regarding the phenotyping of wild birds used in the P_{st} 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 CEFÉ-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 CEFÉ-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.

210 Molecular markers and F_{st} calculation

211 For logistical reasons, we were not able to perform a molecular analysis on the birds used in the
212 common garden experiment. As an alternative, we used a dataset, published by Szulkin *et al.*
213 (2016) of wild birds from these two populations (i.e. deciduous, $n = 49$; evergreen, $n = 83$
214 individuals) and genotyped at several thousand SNP using RAD-sequencing. We retained loci
215 genotyped over at least 75% of the individuals. To avoid bias during filtering and in the F_{st}
216 estimates, we pruned highly related individuals from the dataset to keep only individuals linked
217 with values of kinship lower than 0.05 (coefficient of Loiselle; Loiselle *et al.* 1995; Cheverud
218 1996) computed in Genodive 2.27 (Meirmans & Van Tienderen 2004). In order to retain loci
219 more likely to be informative, we applied a 5% MAF threshold (*Minor Allele Frequency*, using
220 vcfTools 0.1.11; Danecek *et al.* 2011). We pruned the dataset for SNPs that deviated from Hardy-
221 Weinberg-Equilibrium in at least one of the two populations (p -value < 0.05) using vcfTools
222 0.1.11. We retained only the first SNP of each 100 bp locus. To obtain a set of SNPs more likely
223 to be neutral, we filtered out SNPs potentially under divergent selection between the two habitats
224 (p -value < 0.015 ; 0.7 % of total SNPs removed). This was done with a Bayescan 2.0 test (Foll &
225 Gaggiotti 2008; 5 000 pilot iterations, 50 000 burnin, prior odds of 100). Average F_{st} and 95%
226 confidence intervals were estimated using the R-package hierfstat 0.04-22 (Goudet 2005). The
227 final dataset contained 69 individuals (32 and 37 individuals in the deciduous and evergreen
228 habitats, respectively) genotyped at 5407 SNPs.

229

230 Statistical analysis

231 Genetic divergence between habitats of origin

232 We tested for a genetic difference between the two habitats for each trait with univariate linear
233 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 origin 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.* 2011). 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).

Q_{st} , P_{st} and F_{st} 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 Q_{st} 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 Q_{st} as:

$$Q_{st} = \sigma_B / (\sigma_B + 2 * \sigma_W)$$

Where σ_B is the between-habitat phenotypic variance and σ_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 σ_B as the variance attributable to the habitat

of origin and σ_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 σ_B as the variance attributable to habitat of origin and σ_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 $Q_{st} / \text{mean } F_{st}$ 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.

Q_{st} , P_{st} and F_{st} comparison

We found a significant but small genetic differentiation between the two populations [mean F_{st} over all loci = 0.004 (95% CrI: 0.003; 0.005), p-value < 0.001]. Q_{st} was higher than F_{st} with non-overlapping intervals for all traits. The ratio between the Q_{st} and F_{st} was significantly greater than one for all traits. Credibility intervals for Q_{st} and P_{st} 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 than 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 M€erila 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 M€erila 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.

444

445 Genetic and environmental effects are not mutually exclusive

446 Genetic divergence does not preclude a plastic response to ecological conditions specific to each

447 habitat. For example, in the wild, the phenotypic difference in male heart rate between habitats

448 was not significant (Dubuc-Messier *et al.* 2017), but using the common garden experiment we

449 found here a significant difference in male heart rate. It is thus possible that plastic responses of

450 heart rate to habitat specific ecological conditions in the wild may have hidden the genetic

451 divergence (Conover & Schultz 1995). In addition, the important temporal variation in mean

452 handling aggression in the wild shown by Dubuc-Messier *et al.* (2017) in each population,

453 suggests that individuals can partly adjust their personality phenotype for this trait depending on

454 the current local conditions. However, for all traits, the P_{st} between wild birds was not

455 statistically different from their Q_{st} , suggesting that environmental effects in the wild might not

456 result in stronger or weaker differentiation compared to the genetic differentiation.

457

458 Conclusion

459 Our study suggests a genetic divergence for personality, physiological and morphological traits

460 between two blue tit populations that occupy different habitats but that are separated by small

461 spatial distances compared to the dispersal ability of the species and connected by gene flow. The

462 present study and past results for this system suggest that these differences are likely due to

463 different selection pressures and may represent local adaptations. These results thus emphasize

464 the role of environmental heterogeneity for intra-specific phenotypic diversity and suggest that

465 genetic population divergence is possible at small spatial scales (relative to their dispersal ability)

466 for behavioural traits.

467

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

- Atwell, J.W., Cardoso, G.C., Whittaker, D.J., Campbell-Nelson, S., Robertson, K.W. & Ketterson, E.D. 2012. Boldness behavior and stress physiology in a novel urban environment suggest rapid correlated evolutionary adaptation. *Behav. Ecol* **23**: 960–969. doi:10.1093/beheco/ars059.
- Bates, D., Mächler, M., Bolker, B.M., & Walker, S.C. 2014. Fitting linear mixed-effects models using lme4. available at <https://cran.r-project.org/web/packages/lme4/vignettes>.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**: 1–48.
- Bell, A.M. 2005. Behavioural differences between individuals and two populations of Stickleback (*Gasterosteus Aculeatus*). *J. Evol. Biol.* **18**: 464–473.

490 Bertrand, J. A., Delahaie, B., Bourgeois, Y. X., Duval, T., García-Jiménez, R., Cornuault,
491 J., Pujol, B., Thébaud, C., Milá, B. (2016). The role of selection and historical factors in
492 driving population differentiation along an elevational gradient in an island bird. *J. Evol.*
493 *Biol.* 29: 824-836.

494

495 Blondel, J., Perret, P., Anstett, M.-C. & Thebaud, C. 2002. Evolution of sexual size dimorphism
496 in birds: test of hypotheses using Blue tits in contrasted mediterranean habitats. *J. Evol.*
497 *Biol.* **15**: 440–450.

498

499 Charmantier, A., Doutrelant, C., Dubuc-Messier, G., Fargevieille, A., & Szulkin, M. 2016.
500 Mediterranean Blue tits as a case study of local adaptation. *Evol. Appl.* **9**: 135–152.

501

502 Cheverud, J.M. 1996. Developmental integration and the evolution of pleiotropy. *Amer. Zool.*
503 **36**: 44–50.

504

505 Cheverud, J.M., Rutledge, J.J. & Atchley, W.R. 1983. Quantitative genetics of development:
506 genetic correlations among age-specific trait values and the evolution of ontogeny.
507 *Evolution* **37**: 895.

508

509 Cheverud, J.M., & Marroig, G. 2007. Comparing covariances matrices: random skewers method
510 compared to the common principal components model. *J. Genet. Mol. Biol.* **30**: 461-469.

511

512 Class, B., Klueck, E., & Brommer, J.E. 2014. Evolutionary quantitative genetics of behavioral
513 responses to handling in a wild passerine. *Ecol. Evol.* **4**: 427–440.

- 514
- 515 Conover, D.O., & Schultz, E.T. 1995. Phenotypic similarity and evolutionary significance of
516 contergradient variation. *Trends Ecol. Evol.* **10**: 248–252.
- 517
- 518 Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E.
519 *et al.* 2011. The variant call format and VCFtools. *Bioinformatics* **27**: 2156–2158.
- 520
- 521 Dewey, M. 2017. metap: meta-analysis of significance values. R package version 8.0.
- 522
- 523 Dingemanse, N. J., Wright, J., Kazem, A. J., Thomas, D. K., Hickling, R. Dawnay, N. 2007.
524 Behavioural syndromes differ predictably between 12 populations of three-spined
525 stickleback. *J. Animal Ecol.* **76**: 1128–1138.
- 526
- 527 Drent, P.J., van Oers, K. & van Noordwijk, A.J. 2003 Realized heritability of personalities in the
528 Great Tit (*Parus Major*). *Proc. R. Soc. B* **270**: 45–51.
- 529
- 530 Dubuc-Messier, G., Réale, D., Perret, P. & Charmantier, A. 2017. Environmental heterogeneity
531 and population differences in Blue tits personality traits. *Behav. Ecol.* **28**: 448–459.
- 532
- 533 Edelaar, P & Bolnick, DI. 2012. Non-random gene flow: an underappreciated force in evolution
534 and ecology. *Trends Ecol. Evol.* **27**: 659–665.
- 535
- 536 Fitzpatrick, B.M. 2012 Underappreciated consequences of phenotypic plasticity for ecological
537 speciation. *Int. J. Ecol.* **2012**: 1–12.

- 538
- 539 Foll, M. & Gaggiotti, O. 2008 A genome-scan method to identify selected loci appropriate for
540 both dominant and codominant markers: A bayesian perspective'. *Genetics* **180**: 977–993.
- 541
- 542 Goudet, J. 2005. HIERFSTAT, a package for R to compute and test hierarchical F-Statistics'.
543 *Mol. Ecol. Notes* **5**:184–86.
- 544
- 545 Green J A. 2011. The heart rate method for estimating metabolic rate: review and
546 recommendations. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* **158**:287–304.
- 547
- 548 Grosbois, V., Henry, P.-Y., Blondel, J., Perret, P., Lebreton, J.-D., Thomas, D.W. *et al.* 2006.
549 Climate impacts on mediterranean Blue Tit survival: an investigation across seasons and
550 spatial scales. *Glob. Change Biol.* **12**: 2235–2249.
- 551
- 552 Hadfield, J. 2010. MCMC methods for multi-response generalized linear mixed models: The
553 MCMCglmm R Package. *J. Stat. Softw.* **33**:1–22.
- 554
- 555 Herczeg, G., Ab Ghani, N.I. & Merilä, J. 2013. Evolution of Stickleback feeding behaviour:
556 genetics of population divergence at different ontogenetic stages. *J. Evol. Biol.* **26**: 955–
557 962.
- 558
- 559 Jacquin, L., Reader, S.M., Boniface, A., Mateluna, J. Patalas, I., Pérez-Jvostov, F *et al.* 2016.
560 Parallel and nonparallel behavioural evolution in response to parasitism and predation in
561 Trinidadian guppies. *J. Evol. Biol.* **29**: 1406–1422.

- 562
- 563 Kawecki, T.J. & Ebert, D. 2004 Conceptual issues in local adaptation. *Ecol. Lett.* **7**: 1225–1241.
- 564
- 565 Kluen, E., & Brommer, J.E. 2013. Context-specific repeatability of personality traits in a wild
- 566 bird: A reaction-norm perspective. *Behav. Ecol.* **24**: 650–658.
- 567
- 568 Koolhaas, J.M., Korte, S.M., De Boer, S.F., Van Der Vegt, B.J., Van Reenen, C.G., Hopster, H.,
- 569 D *et al.* 1999. Coping styles in animals: Current status in behavior and stress-physiology.
- 570 *Neurosci. Biobehav. Rev.* **23**: 925–935.
- 571
- 572 Kruuk, L.E.B., & Hadfield, J.D. 2007. How to separate genetic and environmental causes of
- 573 similarity between relatives: separating genetic and environmental effects. *J. Evol. Biol.*
- 574 **20**:1890–1903.
- 575
- 576 Lambrechts, M.M., Blondel, J., Hurtez-Bousses, S., Maistre, M. & Perret, P. 1997. Adaptive
- 577 inter-population differences in Blue tit life-history traits on Corsica. *Evol. Ecol.* **11**: 599–
- 578 612.
- 579
- 580 Lambrechts, M. M., Perret, P., Maistre, M., and Blondel, J. 1999. Do experiments with captive
- 581 non-domesticated animals make sense without population field studies ? A case study
- 582 with blue tits' breeding time. *Proc. R. Soc. B* **266**:1311-13.
- 583
- 584 Laurila, A., S. Karttunen, S. & J. Merila. 2002. Adaptive phenotypic plasticity and genetics of
- 585 larval life histories in two *Rana temporaria* populations. *Evolution* **56**:617-627.

- 586
- 587 Leinonen, T., McCairns, R.J.S., O'Hara, R.B. & Merilä, J. 2013. QST–FST comparisons:
588 evolutionary and ecological insights from genomic heterogeneity. *Nat. Rev. Genet.* **14**:
589 179–190.
- 590
- 591 Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* **17**: 183–
592 189.
- 593
- 594 Loiselle, B.A., Sork, V.L., Nason, J. & Graham, C. 1995. Spatial genetic structure of a tropical
595 understory shrub, *Psychotria Officinalis* (Rubiaceae). *Am. J. Bot.* **82**: 1420.
- 596
- 597 Meirmans, P.G., & Van Tienderen, P.H. 2004. Genotype and Genodive: Two programs for the
598 analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* **4**: 792–794.
- 599
- 600 Miranda, A.C., Schielzeth, H., Sonntag, T. & Partecke, J. 2013. Urbanization and its effects on
601 personality traits: A result of microevolution or phenotypic plasticity? *Glob. Change Biol.*
602 **19**: 2634–2644.
- 603
- 604 Mutzel, A., Dingemanse, N.J., Araya-Ajoy, Y.G. & Kempenaers, B. 2013. Parental provisioning
605 behaviour plays a key role in linking personality with reproductive success. *Proc. R. Soc.*
606 **B 280**: 20131019–20131019.
- 607 Petelle, M.B., Martin, J.G.A. & Blumstein, D.T. 2015. Heritability and genetic correlations of
608 personality traits in a wild population of Yellow-Bellied Marmots (*Marmota Flaviventris*).
609 *J. Evol. Biol.* **28**: 1840–1848.

- Pigliucci, M. 2005 Evolution of phenotypic plasticity: Where are we going now? *Trends Ecol. Evol.* **20**: 481–86.
- Plummer, M., Best, N., Cowles, K., & Vines, K. 2006. CODA: Convergence diagnosis and output analysis for MCMC. *R News* 6: 7-11.
- Porlier, M., Garant, D., Bourgault, P., Perret, P., Blondel, J. & Charmantier A. 2012. Habitat-linked population genetic differentiation in the Blue tit *Cyanistes caeruleus*. *Journal of Heredity* **103**: 781-791.
- R Core Team. 2017 R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL [Http://www.R-Project.org/](http://www.R-Project.org/). Accessed January 15. <https://www.r-project.org/>.
- Réale, D., Berteaux, D.A., McAdam, A.G. & Boutin, S. 2003 Lifetime selection on heritable life-history traits in a natural population of red squirrels. *Evolution* **57**: 2416–2423.
- Réale, D., Garant, D., Humphries, M.M., Bergeron, P., Careau, V. & Montiglio, P.-O. 2010. personality and the emergence of the pace-of-life syndrome concept at the population level. *Proc. R. Soc. B* **365**: 4051–4063.
- Réale, D., Reader, S.M., Sol, D., McDougall, P.T. & Dingemanse, N.J. 2007. Integrating animal temperament within ecology and evolution. *Biol. Rev.* 82, 291–318.

- 635 Reparaz, L.B., van Oers, K., Naguib, M., Doutrelant, C., Visser, M.E. & Caro, S.P. 2014. Mate
636 preference of female Blue tits varies with experimental photoperiod. *PLoS ONE* **9**:
637 e92527.
- 638
- 639 Revell, L.J. 2007. The G matrix under fluctuating correlational mutation and selection. *Evolution*
640 **61**: 1857–1872.
- 641
- 642 Revell, L.J. 2012. Phytools: An R package for phylogenetic comparative biology (and other
643 things): Phytools: R Package. *Methods Ecol. Evol* **3**: 217–23.
- 644
- 645 Richardson, J. L., Urban, M. C., Bolnick, D. I., & Skelly, D. K. 2014. Microgeographic
646 adaptation and the spatial scale of evolution. *Trends Ecol. Evol.* **29**:165-176.
- 647
- 648 Roff, D.A., & Fairbairn, D.J. 2012. A test of the hypothesis that correlational selection generates
649 genetic correlations. *Evolution* **66**: 2953–60.
- 650
- 651 Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science* **236**:
652 787–792.
- 653
- 654 Smith, B.R. & Blumstein, D.T. 2008. Fitness consequences of personality: A meta-analysis.
655 *Behav. Ecol.* **19**: 448–455.
- 656
- 657 Sultan, S.E. & Spencer, H.G. 2002. Metapopulation structure favors plasticity over local
658 adaptation. *Am. Nat.* **160**: 71–283.

- 659
- 660 Szulkin, M., Gagnaire, P.-A., Bierne, N. & Charmantier, A. 2016. Population genomic footprints
661 of fine-scale differentiation between habitats in mediterranean Blue tits. *Mol. Ecol.* **25**:
662 542–558.
- 663
- 664 Taylor, R.W., Boon, A.K., Dantzer, B., Réale, D., Humphries, M.M., Boutin, S., Gorrell, J.C.,
665 Coltman, D.W. & McAdam, A.G. (2012) Low heritabilities, but genetic and maternal
666 correlations between red squirrel behaviours: G-matrix of red squirrel personality traits. *J*
667 *Evol. Biol.* **25**: 614–624.
- 668
- 669 Teplitsky, C., Tarka, M., Møller, A.P., Nakagawa, S., Balbontín, J., Burke, T.A., Doutrelant, C.
670 *et al.* 2014. Assessing multivariate constraints to evolution across ten long-term avian
671 studies. *PLoS ONE* **9**: e90444.
- 672
- 673 van Oers, K., Kohn, G.M., Hinde, C.A. & Naguib, M. 2015. Parental food provisioning is related
674 to nestling stress response in wild Great tit nestlings: implications for the development of
675 personality. *Front. Zool.* **12**: 1–10.
- 676
- 677 van Oers, K., & Sinn, D.L. 2011. Toward a basis for the phenotypic gambit: Advances in the
678 evolutionary genetics of animal personality. *From Genes to Animal Behavior* (eds M.
679 Inoue-Murayama, S. Kawamura, and A. Weiss). pp. 165–183. Tokyo: Springer Japan.
- 680
- 681 Wang, I.J., & Bradburd, G.S. 2014. Isolation by environment. *Mol. Ecol.* **23**: 5649–5662.
- 682

683 Weir, B. S. & Cockerham, C. C. 1984. Estimating F_{ST} statistics for the analysis of population
684 structure. *Evolution* **38**:1358-1370.

685

686 Wilson, A.J., Pemberton, J.M., Pilkington, J.G., Clutton-Brock, T.H., Coltman, D.W. & Kruuk,
687 L.E.B. 2007. Quantitative genetics of growth and cryptic evolution of body size in an
688 island population. *Evol. Ecol.* **21**: 337–356.

689

690 Wright, S. 1949. The genetical structure of populations. *Ann. Hum. Genet.* **15**: 323–354.

691

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 $1.5 * \text{the interquartile range}$, dots represent extreme data. All differences are significant (see Table 1 for details).

700 Table 1. Final models describing the phenotype of blue tits originating from two distinct
701 populations and habitats (deciduous and evergreen) in Corsica (France) and reared in a common
702 garden.

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

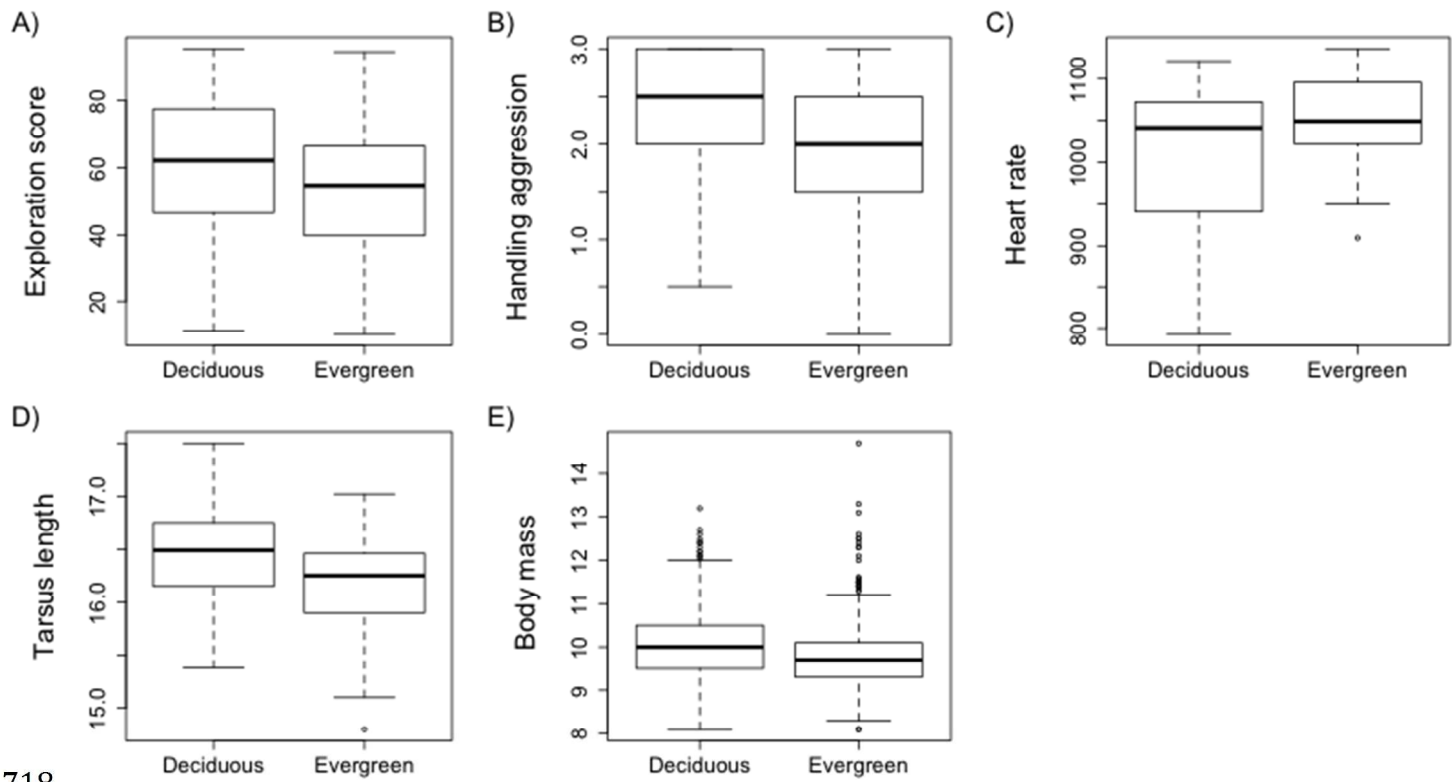
703 The deciduous habitat, females, and cohort 2010 were set as references in models. Estimates are
704 from a model with the brood of rearing and brood of origin in random effect (and individuals
705 identity for body mass), variance estimates are shown in Table S4. L-ratio and p-values are from
706 the comparison of a full model and a model without the variable of interest. The p-values and L-
707 ratio associated with each parameter in initial models before selection are presented in Table S3.
708

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.

Traits	Q_{st} (95% CrI)	P_{st} (95% CrI)	Q_{st} / F_{st} ratio (95% CrI)
Exploration score	0.084 (0.029; 0.804)	0.063 (0.018; 0.727)	20.982 (7.266; 201.065)
Handling aggression	0.129 (0.034; 0.832)	0.045 (0.011; 0.692)	32.309 (8.525; 208.025)
Heart rate during manual restraint (HR)	0.101 (0.033; 0.846)	0.032 (0.007; 0.562)	25.320 (8.244; 211.475)
Body mass	0.069 (0.018; 0.736)	0.095 (0.030; 0.773)	17.144 (4.541; 183.998)
Tarsus length	0.197 (0.050; 0.872)	0.212 (0.048; 0.864)	49.368 (12.455; 217.881)
<i>Mean F_{st}</i>	<i>0.004</i> (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 Q_{st} .

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.

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

¹ mean maximal frass mg/m² per day in each population (sampled between 2011 and 2015 during the breeding period using 0.25m² 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.

Score	Wings spread	Tail feathers spread	Bird strikes fingers
0	No	No	No
1	No	No	Yes, but only if provoked
2	No	Yes	Yes, spontaneously
3	Yes	Yes	Yes, spontaneously

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.

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

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.

	<i>Brood of origin</i>				<i>Rearing broods</i>				<i>Residuals</i>
Traits	Variance (95% CI)	L-ratio	d.f.	p-value	Variance (95% CI)	L-ratio	d.f.	p-value	Variance (95% CI)
Exploration score	0.05 (0.00; 0.15)	0.76	1	0.38	0 (0.00; 0.001)	0	1	1	0.71 (0.56; 0.89)
Handling aggression	0.01 (0.00; 0.13)	0.002	1	0.97	0.08 (0.00; 0.24)	1.52	1	0.22	0.70 (0.54; 0.94)
HR	0.05 (0.00; 0.30)	0	1	1.00	0.25 (0.00; 0.70)	3.35	1	0.07	0.66 (0.42; 1.02)
Tarsus length	0.09 (0.00; 0.29)	1.69	1	0.19	0.14 (0.01; 0.37)	4.74	1	0.03	0.44 (0.34; 0.57)
Body mass	0.07 (0.01; 0.14)	4.10	1	0.04	0 (0.00; 0.001)	0	1	1.00	0.02 (0.22; 0.26)

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.

Traits	Between habitat variance (95% CrI)
Exploration score	1700 (11.88; 2464)
Handling aggression	2.909 (0.021; 4.470)
Heart rate during manual restraint (HR)	34 971 (245.5; 57 937)
Body mass	1.564 (0.015; 3.041)
Tarsus length	0.142 (0.114; 0.181)

References

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(1): 135–52.

Dubuc-Messier G, Réale D, Perret P, Charmantier A. 2016. Environmental heterogeneity and population differences in blue tits personality traits. *Behav Ecol.* arw148.

Dubuc-Messier G, Charmantier A, Doutrelant C, Perret S, Pradel R, Choquet R, Réale D. In prep. Environmental heterogeneity and differential local selection patterns on personlaity traits in blue tits populations.

Zandt H, Strijkstra A, Blondel J, van Balen H. 1990. Food in two mediterranean blue tit populations: Do differences in caterpillar availability explain differences in timing of the breeding season? In Blondel J, Gosler A, Lebreton JD, McCleery R, editors. *Population Biology of Passerine Birds: An Integrated Approach*. Berlin: Springer-Verlag. Pages 145–155.