THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The relative importance of plasticity versus genetic differentiation in explaining between population differences; a metaanalysis

OPEN ACCESS

LETTER

The relative importance of plasticity versus genetic differentiation in explaining between population differences; a meta-analysis

Megan A. Stamp and Jarrod D. Hadfield iD

Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

*Correspondence: E-mail: j.hadfield@ed.ac.uk

Abstract
Both plasticity and genetic differentiation can contribute to phenotypic differences between populations. Using data on non-fitness traits from reciprocal transplant studies, we show that approximately 60% of traits exhibit co-gradient variation whereby genetic differences and plasticity-induced differences between populations are the same sign. In these cases, plasticity is about twice as important as genetic differentiation in explaining phenotypic divergence. In contrast to fitness traits, the amount of genotype by environment interaction is small. Of the 40% of traits that exhibit counter-gradient variation the majority seem to be hyperplastic whereby non-native individuals express phenotypes that exceed those of native individuals. In about 20% of cases plasticity causes non-native phenotypes to diverge from the native phenotype to a greater extent than if plasticity was absent, consistent with maladaptive plasticity. The degree to which genetic differentiation versus plasticity can explain phenotypic divergence varies a lot between species, but our proxies for motility and migration explain little of this variation.

Keywords
Counter-gradient variation, Gene-flow, local adaptation, phenotypic plasticity.

Ecology Letters (2020) **23**: 1432–1441

INTRODUCTION

When environmental conditions vary in space, individuals of the same species often differ in phenotype in a way that increases their fitness in the local environment (Hereford, 2009). These phenotypic differences arise through two different mechanisms: phenotypic plasticity, in which phenotypic expression is a direct response to the environment without genetic change (Pigliucci 2001), and local adaptation, wherein phenotypic differences are determined by genetic differences (Kawecki & Ebert 2004). Although many studies implicitly assume phenotypic differences between populations are mainly, if not completely, genetic (Brommer 2011), in reality the relative importance of these two processes in driving spatial differentiation in phenotype is currently unclear.

Reciprocal transplant studies have been widely used to estimate the contribution of plasticity and genetic differentiation to spatial phenotypic divergence (Turesson 1922), and to our knowledge four studies have synthesised their findings. Leimu & Fischer (2008) and Hereford (2009) both conducted meta-analyses that focused on traits positively associated with fitness, and found strong evidence for local adaptation. Using studies on plants, Leimu & Fischer (2008) found that the average performance of a native population is 0.16 within-population standard deviations greater than the performance of a non-native population, and using studies of both plants and animals, Hereford (2009) found a 45% increase in performance of native individuals. Such home versus away comparisons, when averaged over all possible reciprocal transplants,

measure the genotype by environment interaction for fitness (Blanquart *et al.,* 2013). Presumably differences in performance also exist because of the main effects of genotype and environment, but the magnitude of these differences were not characterised. In the context of fitness this is understandable as environmental differences in fitness probably reflect between-site differences in habitat quality rather than an active plastic response on the part of the organisms (Blanquart *et al.,* 2013).

For traits other than fitness, while genotype by environment interactions may exist, it is also meaningful to consider environmentally induced variation in local optima, and therefore genetic differentiation and phenotypic plasticity in those traits in response to divergent selection. Palacio-López *et al.* (2015) synthesised data from reciprocal transplant studies of non-fitness traits in plants. For those traits they classified as plastic (the 48% of traits that exhibited a> 53% change in phenotype when home versus away), 49.4% exhibited 'perfectly' adaptive plasticity, 19.5% 'partially' adaptive plasticity and 31% mal-adaptive plasticity (See Fig. 1). The adaptive plasticity categories represent situations where plasticity causes trait values to be closer to their putative optima compared to the situation where plasticity is absent (Ghalambor *et al.,* 2007). In contrast, maladaptive plasticity were those traits in which the plastic response was more than twice the difference in putative optima ('too-steep'; 9.8%) or opposite in sign ('wrong-sign'; 21.2%). While the classification system of Palacio-López *et al.* (2015) has merits, it suffers from the implicit assumption that the observed phenotypic divergence between populations is equal to the difference in their respective optima. An alternative but overlapping classification system is to distinguish co-gradient variation, where plasticity causes differences in trait
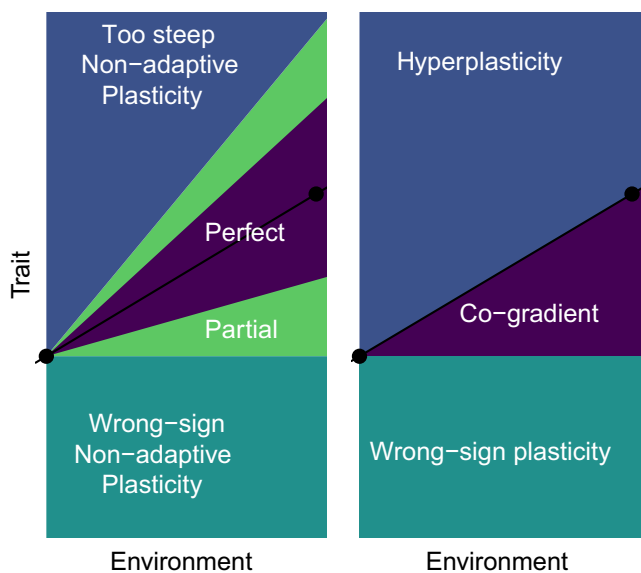
Figure 1 The classification scheme of Palacio-López *et al.* (2015) (left) and a classification scheme based on co/counter gradient variation (right). In both cases the black dots represent the mean phenotypes of two populations raised in their home environments such that their difference is the in situ divergence ($P_A E_A - P_B E_B$, where $P_i E_j$ refers to the average phenotype of individuals from population i raised in the environment of population of j: see Fig. 2). The solid black line represents the scenario where 100% of the divergence is due to plasticity with no genetic differentiation, and in this case the difference in mean phenotype between the same population assessed in the two environments (either $P_A E_B - P_A E_A$ or $P_B E_B - P_B E_A$) would track this line. Coloured regions denote plasticity induced phenotypes ($P_A E_B - P_A E_A$) that fall into the different classes, which for perfect adaptive plasticity must lie in the region 47–153% of the phenotypic divergence, and between 0–47% or 153–200% for partial adaptive plasticity.

value in the range 0-100% of the phenotypic divergence, from counter-gradient variation in which the plastic response is either greater than the phenotypic divergence (hyperplasticity) or opposite in sign (wrong-sign plasticity) (Levins 1968; Conover & Schultz 1995). Surprisingly, an informal review suggests that for traits in which spatial differentiation in phenotype can be attributed to both genetic and plastic responses, 84% show counter-gradient variation (Conover *et al.*, 2009). The related question of whether changes in mean phenotype within a population over time are due to plasticity or genetic adaptation was also touched on in Conover *et al.* (2009) but several more focussed taxonomic reviews have appeared subsequently (Gienapp (2008), Boutin & Lane (2014); Charmantier & Gienapp (2014); Franks *et al.* (2014); Stoks *et al.* (2014); Schilthuizen & Kellermann (2014) summarised in Merilä & Hendry (2014)). The consensus from these papers is that the contribution of plasticity outweighs that of genetic adaptation, although this conclusion is largely based on a failure to reject the null hypothesis that all phenotypic change is due to plasticity.

The conditions under which plasticity or genetic differentiation are favoured has been explored extensively in a theoretical context. When there is no cost to plasticity and the environmental cue is perfect, all spatial differentiation is

predicted to arise from a direct plastic response to spatial variation in the environment (Via & Lande 1985). However, as the cost of plasticity increases (van Tienderen 1997) or cue reliability decreases (Gavrilets & Scheiner 1993) the plastic response is reduced and genetic differences contribute to spatial differentiation (de Jong 1999; Tufto 2000). As the scale of environmental variation increases relative to the scale of dispersal, genetic differences start to play an increasingly dominant role (Hadfield 2016). Although not well developed theoretically (but see Edelaar *et al.* 2017), it has also been suggested that species without active motility (such as plants) do not have the capacity to move to environments in which they are suited and are therefore exposed to a greater range of environmental variation which requires a plastic response (Bradshaw 1972; Huey *et al.* 2002).

Low gene-flow between environments, either through active habitat choice or low migration, combined with high costs of plasticity and low cue reliability are therefore expected to favour genetic differentiation over plasticity as a cause of spatial differentiation. Given the difficulty of measuring the cost and accuracy of plasticity, empirical work has mainly concentrated on the explanatory power of gene-flow. Meta-analyses have been used to show that the *absolute* strength of plasticity is greater in plants than in animals (Acasuso-Rivero *et al.* 2019), which while consistent with active habitat choice reducing the need for plasticity, could also be because the continually growing modular structure of plants is more developmentally labile (De Kroon *et al.* 2005). In contrast, the *absolute* magnitude of genetic differentiation does not appear to increase with increasing geographic distance (Leimu & Fischer 2008) which is inconsistent with reduced gene-flow promoting genetic differentiation. Outside of a meta-analytic approach, Jacob *et al.* (2017) used an elegant experimental evolution approach testing the effects of motility and migration. As predicted, they showed in their ciliate microcosms that active habitat choice increases the amount of local adaptation, but surprisingly, reducing gene-flow by reducing the rate of random migration had little effect.

While previous syntheses give some insight into the relative contributions of plasticity and genetic differentiation to phenotypic divergence in space (Conover *et al.* 2009; Palacio-López *et al.* 2015) and time (Merilä & Hendry 2014), they suffer from several limitations. First, the use of informal, extreme or arbitrary inclusion criteria makes it hard to know whether their findings are general. Second, the unnecessary (and arbitrary, in the case of perfect/partial adaptive plasticity) discretisation of a continuous metric makes it hard to judge in quantitative terms the relative importance of plasticity versus genetic differentiation. Third, the reliance on significance testing may mean that observed patterns simply reflect statistical power rather than biological effect, and finally, without correcting for measurement error the number of studies falling into rarer classes, such as maladaptive plasticity, are likely to be inflated. In this paper, we collate data on non-fitness traits from reciprocal transplant studies and avoid the above issues using meta-analytic techniques to determine the average relative strength of genetic differentiation versus plasticity. In addition, we quantify the degree to which the strength of the two processes varies over species and traits and test leading
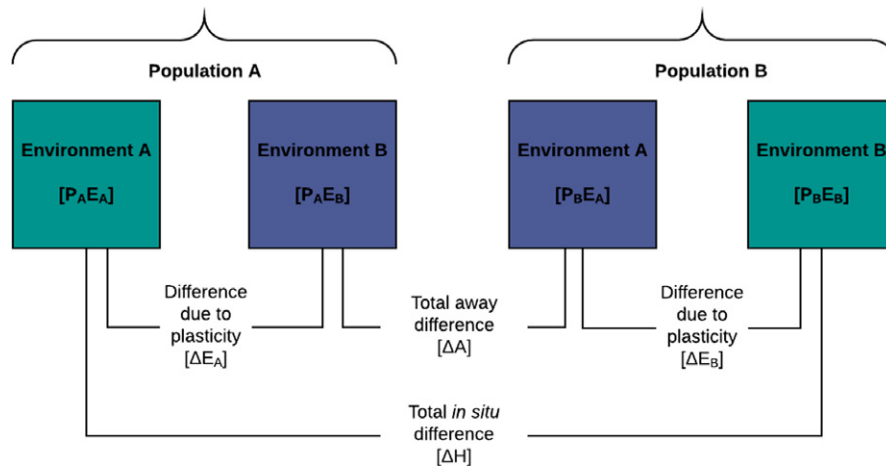
Figure 2 The different populations involved in a reciprocal transplant experiment and the way in which they can be used to determine the plastic component of phenotypic differences between populations. The green boxes indicate populations in their home environment, and orange boxes are populations that have been transplanted. The phenotypic difference observed between the individuals of the same population in different environments is identified as the difference due to plasticity.

hypotheses about what factors might promote plasticity over genetic differentiation.

METHODS

Literature search

Data were collected using the search term 'reciprocal transplant experiment' on the ISI Web of Science database on the 26th January 2018 and 22nd June 2018. Reciprocal transplant experiments are those in which two populations are assayed in their own and each other's environment (Fig. 2) to test whether phenotypic differences between populations are due to genetic differences or a plastic response to environmental variation. Two reciprocal transplant studies were also collected using the search term 'common garden experiment', as initially common garden studies were being screened as well. There proved to be a lack of suitable common garden studies, however, so this type of study was excluded from the analysis.

In total, 682 studies were screened from 1981 to 2018, which comprises the total number of studies returned by the search. Studies were chosen for inclusion in the meta-analysis based on the following criteria: a phenotypic trait measurement for each of the four treatment groups in the reciprocal transplant (Fig. 2) was reported; standard errors for these measurements could be extracted or calculated; distance between the studied populations could be determined; the populations involved in the reciprocal transplant were of the same species; and each phenotypic measurement corresponded to one population. Studies were excluded if they used lab populations or replicated natural conditions in a laboratory or greenhouse, or if the reciprocal transplant did not take place at the site where individuals were collected. For 218 studies it was determined from the abstract that the inclusion criteria were not met, and of the remaining 464 studies a full reading was required. For those studies, 375 did not meet the inclusion criteria (a summary of the reasons are given in the Supplementary materials Data S1) and 87 studies were selected for inclusion in this

meta-analysis. Of the 87 species, only three species had been subject to independent reciprocal transplant experiments.

Phenotypic means and their standard errors were either extracted from the text or tables, calculated from publicly-available raw data, or extracted from graphs using Web Plot Digitizer (Rohatgi 2012). For studies where phenotypes were measured over multiple time periods, the last time period was used in each study for consistency, except for cases where no standard error, or no measurement was reported for one or more groups at the final time point, in which case the previous time point was used. In studies that performed reciprocal transplants with more than two populations, only the first two listed populations were used to avoid issues with non-independence during analysis. All traits were used unless they were calculated from the same information (i.e. leaf width and leaf area), in which case the first listed trait was used. In cases where standard deviation and sample size were reported, these were used to calculate the standard error if it was not reported. Fitness traits – traits that are inextricably tied to fitness (i.e. measures of survival and/or fecundity) – were excluded. In total 200 traits were included, and there was little evidence of any publication bias (see Supplementary materials Data S1).

Effect size

The effect size extracted from these studies is the component of the in situ phenotypic divergence between populations that can be explained by plasticity, as opposed to genetic differentiation. This is obtained from the four phenotypic measurements collected from each study (Fig. 2), by calculating the plastic component of the phenotypic difference (ΔE) and dividing it by the total in-situ phenotypic divergence (ΔH – the difference when in their *home* environments) to give the plasticity metric (PL). The plastic component can be determined as follows. Individuals from Population A in Environment A ($P_A E_A$) and from Population A in Environment B ($P_A E_B$) are from a common genetic background but experience different environmental conditions, and so the

difference in phenotype can be ascribed to plasticity ($\Delta E_A = P_A E_A - P_A E_B$). Likewise, the difference in phenotype between individuals from Population B in Environment A ($P_B E_A$) and from Population B in Environment B ($P_B E_B$) can be ascribed to plasticity (ΔE_B). We take the average of these as the plastic component of the phenotypic difference ($\Delta E = (\Delta E_A + \Delta E_B)/2$). If there is no genotype by environment interaction, such that the reaction norms of the two populations only differ in intercept and not slope, we expect $\Delta E_A = \Delta E_B$. The in situ phenotypic divergence (ΔH) is simply the difference in phenotype between the two populations in their home environments; the difference between Population A in Environment A ($P_A E_A$) and Population B in Environment B ($P_B E_B$) (Box 1). The plasticity metric (PL) is then $\Delta E/\Delta H$ and lies between zero and one if there is co-gradient variation, but may be negative if there is wrong-sign plasticity, or greater than one if there is hyperplasticity. It should be noted that $1-PL$ can be interpreted as the component of the in-situ phenotypic divergence between populations that can be explained by gene divergence.

Our PL metric is similar to that used by Palacio-López *et al.* (2015) but differs in an important way from *absolute* measures of plasticity (Acasuso-Rivero *et al.* 2019) and genetic differentiation (Leimu & Fischer 2008); absolute measures face the risk of confounding the capacity to respond to an environmental difference with the magnitude of the environmental difference itself. For example, if researchers are better able to identify and manipulate environmental variables that are important to plants than they are for animals, then plasticity-induced changes in phenotype would be larger in plants (Acasuso-Rivero *et al.* 2019) even if plastic responses are comparable to those in animals. Using the notation of Chevin *et al.* (2010) to make this point clearer; if we assume the reaction norm b and the environmental sensitivity of selection B are both linear functions of the environment E, it is hard to tell whether the differences between plants and animals in their plasticity-induced response to contrasting environments ($|b[E_A - E_B]|$), is driven by differences in $|b|$ or $|E_A - E_B|$. Likewise, if environmental differences increase with geographic distance it is hard to ascertain whether greater genetic differentiation between distant sites (Leimu & Fischer 2008) is due to low-gene flow facilitating a response to divergent selection or whether the strength of divergent selection on breeding value $((B-b)[E_A - E_B])$ is itself greater. Since our PL metric is a *relative* measure of plasticity versus genetic differentiation the magnitude of the underlying environmental difference ($E_A - E_B$) should be largely controlled for when assessing the ability to be plastic.

Moderators

All plants (48 species, 118 traits) were classified as sessile, and animals (39 species, 82 traits) were classified as sessile (17 species) or motile (22 species) depending on whether their adult forms are anchored to a substrate. As proxies for the amount of gene flow, the distance between study populations was also recorded and plants were categorised into whether they were wind (10), water (15) or animal (23) pollinated. Traits were also classified as either being morphological (115), physiological (34), growth (25), timing (17) or behavioural (6) following

Hansen *et al.* (2011). Three sex-allocation traits were left uncategorised. Morphological traits mainly include measures of organismal size, such as height, biomass and number of structures such as leaves or branches. Physiological traits refer to various traits related to metabolism, macromolecule content, and other biochemical processes. Growth traits refer to changes in a quantity (usually a morphological trait) over time. Timing traits (originally classified as life-history traits in Hansen *et al.* (2011)) include measures of phenology and the timing or duration of life-history stages.

Statistical analyses

If the sampling errors around the four assay means are independent and normally distributed (as would be predicted from large-sample theory) then the sampling distribution of the PL metric can be derived (Marsaglia 1965, 2006). In general the distribution is heavier-tailed than the normal, can be asymmetric and bimodal and the median (the mean is undefined) may not coincide with the true value. We employ three strategies to overcome these issues which are discussed at length in the Supplementary materials Data S1:

The *normal* model. If the sampling distribution of $\widehat{\Delta H}$ does not have much density close to zero (either because the in situ phenotypic divergence is large, or because it is precisely measured) the sampling distribution of $\widehat{PL}$ can be approximated by a normal (Marsaglia 2006). We therefore employ meta-analysis using the delta method to obtain an approximate standard error for PL given the standard errors of the four means (using the msm package in R (Jackson 2011). Between-observation effects not due to measurement error (i.e. random-effect meta-analysis) and species effects were fitted as random. This model was also refitted with log-distance between populations and trait type as moderators together with one of plant/animal, mobile/sessile or pollination mode (plants only). The models were fitted in MCMCglmm (Hadfield 2010) using default (flat) priors.

The *normal* model failed to capture the leptokurtosis in $\widehat{PL}$ so we also conducted a bivariate meta-analysis of $\widehat{\Delta H}$ and $\widehat{\Delta A}$ (the estimated difference in the two populations phenotypes when *away* in each other's environment; $P_B E_A - P_A E_B$). From the joint distribution of ΔH and ΔA (after accounting for measurement error) we can obtain the distribution of PL using results in Marsaglia (2006) since it can be obtained as $(\Delta H + \Delta A)/(2\Delta H)$. However, even after accounting for variation in measurement error, the distribution of ΔH and ΔA were far from normal. This is most likely due to different traits been measured on different scales, and so we adopted two strategies.

The *ratio-t* model. First we assumed independent Gaussian sampling errors around the true values of ΔH and ΔA, and pairs of these true values were assumed to come from a common (over pairs of values) bivariate normal distribution after being subject to a rescaling:

$$\hat{\Delta}_{ij} = (\mu_j + u_{ij})s_i + e_{ij}$$

where i indexes pairs of values, j indexes ΔH or ΔA, μ are fixed intercepts and u and e are observation level random effects with the variance of e fixed at the statistics' sampling

variance. The squared scales of measurement (s^2) were assumed to come from an inverse gamma distribution with the scale parameter being equal to the shape parameter. This forces $E[1/s^2]$ to be one (since it is confounded with the variance of u) but has a free parameter determining the spread of scales. The resulting compound distribution is a bivariate generalised t-distribution.

The *ratio-sd-scaled* model. Second, we standardised ΔH and ΔA by the (weighted) average standard deviation of trait values (from $P_A E_A$ and $P_B E_B$ for ΔH, and from $P_A E_B$ and $P_B E_A$ for ΔA) and assumed the sampling errors come from a scaled non-central t-distribution (Hedges 1981; Camilli *et al.* 2010). The true values of ΔH and ΔA were then assumed to come from a common bivariate normal distribution as in the *ratio-t* model. For the *ratio-sd-scaled* model 70 out of 200 observations had to be discarded because the standard deviations were not available.

A difficulty with these two joint models of ΔH and ΔA is that the signs of the two variables are arbitrary (depending on whether population A is compared to B or vice versa) suggesting that a model in which the moderators and species effects determine their absolute values would make sense. However, the ratio of their signs is not arbitrary. This precludes taking absolute values, and in the absence of a solution to this problem we fitted a model without moderators and species effects. The *ratio-t* model was fitted in STAN using diffuse normal priors on the fixed effects, diffuse half-Cauchy priors on the standard deviations (Gelman 2006) and inverse-gamma shape parameter, and an LKJ prior (Lewandowski *et al.* 2009) with a shape parameter of one on the correlation matrix. The *ratio-sd-scaled* was fitted in MCMCglmm with flat priors.

The numerator of the *PL* metric is the average plastic response of the two populations (ΔE_A and ΔE_B) and so only captures the main effects of plasticity and not any GxE interaction. While we expect GxE interaction to be a dominant source of variation for fitness traits (which we exclude) the contribution of GxE interaction to non-fitness traits is less clear (see Discussion). In order to test whether the slopes of the reaction norms differed between populations (i.e. GxE interaction) we also fitted an identical model to the *ratio-sd-scaled* model but for ΔE_A and ΔE_B (rather than ΔH and ΔA). We used the *ratio-sd-scaled* model rather than the *ratio-t* model because differences in scale across traits might cause ΔE_A and ΔE_B to be strongly correlated, and we felt that the *ratio-sd-scaled* model would suffer from this issue less.

RESULTS

Although the quantitative inferences varied over model types, and all models suffered inadequacies of some form in terms of model fit, the general conclusions were reasonably consistent. The posterior means (and 95% credible intervals) for the median value of *PL* in the three models were *normal*) 0.703 [0.623–0.787] (without moderators), *ratio-t*) 0.750 [0.694–0.870] and *ratio-sd-scaled*) 0.677 [0.615–0.752] indicating that in co-gradient cases plasticity on average explains more than two-thirds of the between-population divergence. However, there was substantial variation around this expectation and the probability of a trait showing hyper-plasticity was *normal*)

0.159 [0.110–0.275] (without moderators), *ratio-t*) 0.314 [0.238–0.381] and *ratio-sd-scaled*) 0.244 [0.199–0.310] or showing wrong-sign plasticity *normal*) 0.011 [0.002–0.032] *ratio-t*) 0.117 [0.094–0.152] and *ratio-sd-scaled*) 0.142 [0.108–0.171]. The probability of a trait showing too-steep maladaptive plasticity (i.e. $PL > 2$) was essentially zero in the *normal* model but 0.079 [0.060 - 0.102] and 0.068 [0.057–0.093] for the *ratio-t* and *ratio-sd-scaled* models, respectively. The probabilities of counter-gradient variation (i.e. either hyper or negative-plasticity) are *normal*) 0.219 [0.120–0.290] *ratio-t*) 0.437 [0.346–0.516] and *ratio-sd-scaled*) 0.354 [0.310–0.457] and the probabilities of maladaptive plasticity (i.e. either hyper (wrong-sign) plasticity or too-steep plasticity) are *normal*) 0.011 [0.001–0.032] *ratio-t*) 0.209 [0.156–0.250] and *ratio-sd-scaled*) 0.215 [0.166–0.261].

Best estimates (using the posterior means of the relevant parameters) for the distribution of *PL* are plotted in Fig. 3. As can be seen, the distributions of *PL* under the three models have similar central tendencies, but the *normal* model differs substantially from the other models in terms of tail behaviour, and therefore the probabilities of hyper, negative or too-steep plasticity reported earlier. We believe the *ratio* models provide more robust estimates of these probabilities given they allow for thick-tailed distributions that are a characteristic of ratio distributions. Fig. 4 plots the joint distribution of *PL* and phenotypic divergence for each data point using either the raw data or the posterior mean estimates from the *ratio-t* and *ratio-sd-scaled* models.

For the *normal* model with moderator variables fitted, contrary to expectation the contribution of plasticity to population divergence was estimated to be smaller in plants than animals by an amount −0.113 [−0.300 to 0.051], but this was
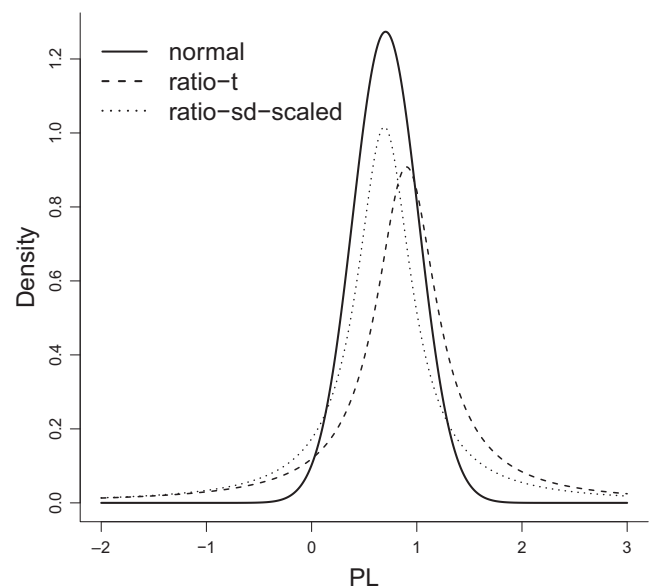

Figure 3 Inferred distribution for the ratio of plastic response to in-situ divergence (*PL*) using the posterior mode parameter values from the three models. Note the *normal* model assumes the distribution of *PL* is normal, but the *ratio-t* and *ratio-sd-scaled* models assume the component parts of *PL* are normal such that the distribution of *PL* is that given by Marsaglia (1965).
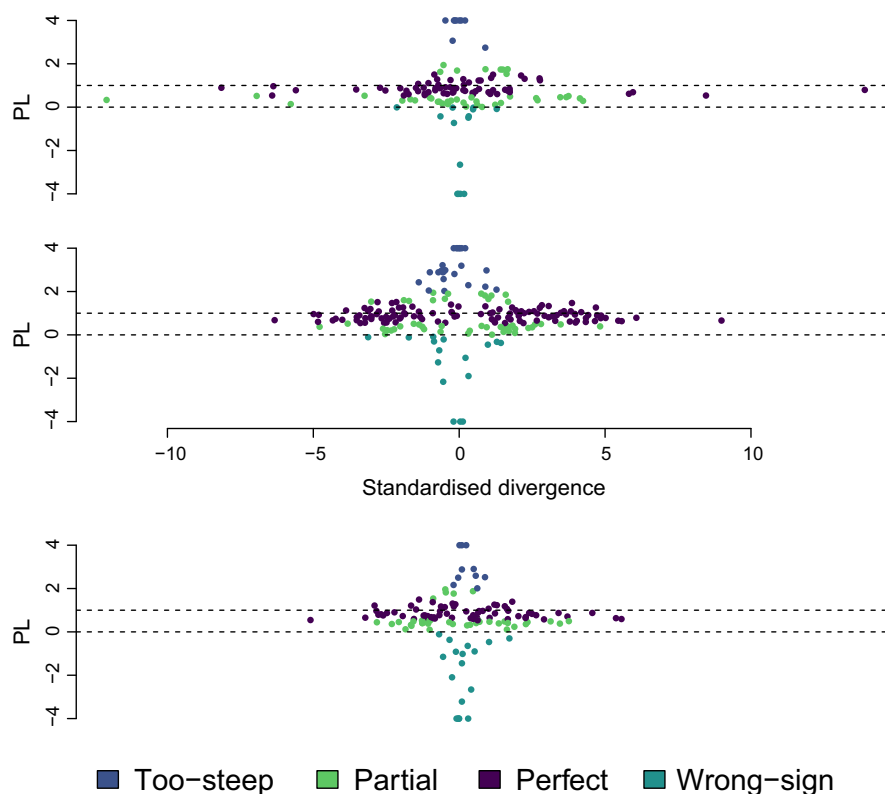
Figure 4 Joint distribution of the ratio of plastic response to in-situ divergence (*PL*) against the standardised in situ phenotypic divergence ($\Delta H = P_A E_A - P_B E_B$ where $P_i E_j$ refers to the average phenotype of individuals from population *i* raised in the environment of population of *j*) with points coloured by the classification system of Palacio-López *et al.* (2015) and the dashed lines separating (from top to bottom) wrong-sign plasticity, co-gradient variation and hyperplasticity. The top figure represents the raw data (i.e. $\widehat{PL}$) and phenotypic divergence is scaled by the average phenotype across the two environments $(P_A E_A + P_B E_B)/2$. The middle and bottom figures represent an MCMC draw from the meta-analytic estimates of *PL* and phenotypic divergence from the *ratio-t* and *ratio-sd-scaled* models respectively. In all plots, any points for which *PL* > 4 are plotted at *PL* = 4 and any points for which *PL* < −4 are plotted at *PL* = −4. In the three plots 14, 10, and 8 points have been subject to truncation, respectively.

far from significant (pMCMC = 0.172). However, lumping sessile animals with plants and contrasting them with motile animals (Huey *et al.* 2002) provided greater support for plasticity being weaker in sessile organisms (−0.258 [−0.476 to −0.057], pMCMC = 0.020). Using the model in which motile/sessile was fitted, there was little evidence that different trait types differed in the degree to which plasticity contributed to divergence (the *P*-value from a Wald test using the posterior means and covariances was 0.356), and the point estimates of the difference between trait types (with morphology as the reference level) were small for growth (−0.054 [−0.237 to 0.116]), physiological (−0.097 [−0.247 to 0.066]) and timing (−0.029 [−0.203 to 0.158]) traits, but larger for behavioural traits (−0.377 [−0.772 to 0.023]) for which there was substantial uncertainty. The effect of (log) distance between populations was small and non-significant (0.016 [−0.011 to 0.045]) as were the effects of pollination mode (Wald test, *P* = 0.200), with wind and water pollination (compared to animal pollination) coefficients being 0.107 [−0.157 to 0.356] and −0.208 [−0.483 to 0.112], respectively. The standard deviation of the species effects was substantial (0.265 [0.197 to 0.343]) and more than half the variation in *PL* (after controlling for the fixed effects) was explained by species (0.744 [0.539 to 0.900]).

The correlation between ΔE_A and ΔE_B across traits, after accounting for measurement error, was 0.872 [0.798 to 0.912]

indicating similar reaction norms and therefore weak genotype by environment interaction.

DISCUSSION

To the best of our knowledge, the relative importance of plasticity versus genetic differentiation at explaining phenotypic divergence has not previously been assessed in a fully quantitative manner. Our results suggest that there is substantial between-species and between-trait variation in the degree to which plasticity causes phenotypic divergence, but on the whole plasticity is the dominant cause. For traits that exhibit co-gradient variation, where plasticity-induced differentiation and genetic differentiation have the same sign (Conover & Schultz 1995), plasticity is approximately twice as important as genetic differentiation. Plasticity itself was only weakly genetically differentiated between populations (i.e little G by E interaction) but there was strong evidence that counter-gradient variation, where plasticity-induced differentiation and genetic differentiation have opposing signs (Conover & Schultz 1995), is moderately common.

In a similar literature-based study of plasticity in plants, Palacio-López *et al.* (2015) discretised a related metric to ours in order to make qualitative assessments. In particular, they assessed the relative frequency of adaptive to maladaptive

plasticity (Ghalambor *et al.* 2007) and found that maladaptive plasticity existed in a third of cases. Our best estimates of the prevalence of maladaptive plasticity are considerably lower than this because we control for the sampling errors that tend to result in estimates that fall into extreme categories. In addition, our analyses also suggest that many cases of maladaptive plasticity are most likely associated with very low phenotypic divergence such that the absolute strength of maladaptation may be weak (Fig. 3). In an informal qualitative review of spatial and temporal differentiation in plants and animals, Conover *et al.* (2009) suggested that more than three-quarters of traits exhibit counter-gradient variation, where plastic and genetic differentiation between populations differ in sign (Conover & Schultz 1995). Here we show that spatial counter-gradient variation is considerably rarer than this and is mainly caused by hyper-plasticity whereby the plasticity-induced response is the same sign as phenotypic divergence but greater in magnitude. Again, our more conservative estimates are expected since informal literature reviews compound the problems of ignoring sampling errors with selection bias and the conflation of effect size with statistical significance (Palmer 2000). The temporal studies included in Conover *et al.* (2009) are a case in point; the three studies included in the review (Merila *et al.* 2001; Garant *et al.* 2004; Wilson *et al.* 2007) were the only reports of putative temporal counter-gradient variation among examples dominated by co-gradient variation, and in all cases the exceptionally small effect sizes were overlooked in favour of statistical significance (that turned out to be erroneous, Hadfield *et al.* 2010).

Although we believe the importance of counter-gradient variation has been over-stated, we do acknowledge that our analyses suggest that it should exist with moderate frequency. This is surprising since co-gradient variation is the expected outcome from most theoretical models of spatially varying selection (Gavrilets & Scheiner 1993; de Jong 1999; Tufto 2000). Verbal models of why counter-gradient variation arises often invoke adaptive evolutionary changes that are required to counteract sub-optimal plastic responses induced by novel environments (Conover & Schultz 1995). Confirming this, counter-gradient variation in gene expression has been shown to repeatedly evolve when populations are exposed to new experimental environments (Ghalambor *et al.* 2015; Huang & Agrawal 2016). Over longer time-scales however, such sub-optimal plastic responses are expected to disappear, and so under this view the moderate prevalence of counter-gradient variation that we find suggests that populations are often in novel environments in which plastic responses have yet to evolve to their optimal values. Two ideas may be put forward against this viewpoint. First, counter-gradient patterns can also arise at equilibrium through adaptive plastic responses that have evolved to cope with both temporal and spatial environmental variation (King & Hadfield 2019) and there is little empirical work to gauge whether this is likely. Second, in our data, counter-gradient variation, like maladaptive plasticity, is often associated with low phenotypic divergence, and while low phenotypic divergence could be driven by genetic and plastic responses that are large in magnitude but opposite in sign, it seems more likely that low phenotypic divergence is also associated with low genetic and plastic divergence. As

phenotypic divergence approaches zero, our metric will tend to extreme values of hyper- or negative-plasticity, even if the absolute strength of plasticity is weak, and such patterns might simply be driven by drift as opposed to genetic compensation (Grether 2005) opposing strong maladaptive plasticity.

Our conclusion that plasticity plays a more important role than genetic differentiation in determining spatial divergence is in agreement with the qualitative conclusions drawn from studies that look at phenotypic differentiation in time (Gienapp *et al.* 2006; Merilä & Hendry 2014), but determining whether they are quantitatively similar will require a formal meta-analysis of temporal patterns. However, a number of methodological differences between quantifying spatial and temporal patterns would need to be considered. First, reciprocal transplant studies are a relatively clean way to separate plasticity from genetic differentiation, whereas the non-experimental model-based approaches for detecting temporal changes in genetic value can be biased by model assumptions (Hadfield *et al.* 2011). In contrast to this, the choice of time points to sample within a population is often made blindly with respect to environmental conditions, whereas the choice of populations is rarely random in reciprocal transplant studies. In particular, studies which choose to reciprocally transplant at small spatial scales seem to choose populations that are from contrasting environments (Galloway & Fenster 2000). In these cases plasticity is predicted to play a more dominant role compared to populations which had been chosen at random with respect to distance and/or environment (Hadfield 2016). This would inflate our estimates of the importance of plasticity, but the lack of relationship between the amount of plasticity and distance between populations [see below] suggests that the bias may not be large.

Different traits in the same species exhibited similar levels of plasticity, which at face value suggests there are species-level characteristics that promote plasticity over genetic differentiation. Our main predictor of whether a species should exhibit plasticity over genetic differentiation was whether the species was sessile or motile, with the expectation that sessile species should be more plastic because they cannot move to environments in which they are adapted (Bradshaw 1972). However, we found that if anything phenotypic divergence in sessile organisms has a reduced contribution from plasticity. This result is opposite to what we expect and appears to contradict a previous meta-analyses showing that plants have greater absolute levels of plasticity (Acasuso-Rivero *et al.* 2019). One explanation for this pattern is that absolute levels of plasticity are higher in sessile organisms, but the spatial scale of dispersal may be lower resulting in an even greater increase in the absolute amount of genetic differentiation (Slatkin 1978; Hadfield 2016). We tried to test more generally whether low rates of gene-flow facilitate genetic differentiation but although the distance between reciprocal transplant populations was positively related to the amount of phenotypic divergence explained by genetic differentiation, the relationship was weak and far from significant. This result should be taken with caution, however, because different species are likely to have very different dispersal distances and the correlation between distance and gene-flow might be quite weak. Compounding this problem, if researchers choose populations

to reciprocally transplant based on the scale of dispersal for that species (for example if populations at a distance of 1km are chosen for a low-dispersal species, but populations at a distance of 100km are chosen for a high-dispersal species), our proxy may then only be very weakly correlated with gene-flow, and at the limit may be uninformative when researchers can perfectly calibrate the distance between populations with the scale of dispersal. Within-species studies should suffer from this issue less, and indeed several such studies have found evidence of local-adaptation scaling with distance (e.g. Galloway & Fenster 2000; Joshi *et al.* 2001) despite a between-species meta-analysis failing to find such a pattern (Leimu & Fischer 2008). For plants we used pollination mode as an additional proxy for gene flow with the expectation that because wind-pollinated plants have increased gene flow compared to animal-pollinated plants (Hamrick *et al.* (1979), but see Friedman & Barrett (2009)) plasticity should play a more dominant role in any phenotypic divergence. As with distance, the point estimate was consistent with expectation but far from significant. Alternative proxies of gene flow using genetic marker information may prove more suitable (but see Bohonak (1999), Whitlock & Mccauley (1999)) but unfortunately only seven of the studies in our meta-analysis reported F_{st} values for their populations.

Although our predictors explained very little of the substantial between-species variation, it should be borne in mind that in the vast majority of species measurements of multiple traits came from a single paper and therefore a single pair of populations. It is therefore possible that some unknown fraction of the between-species variation in our plasticity metric is due to the particular pair of populations within each species that were transplanted. Moreover papers often focus on a non-random subset of traits and so it is possible that some of the observed species variation may also be due to variation across trait-types in their propensity to be plastic. However, our broad categorisation of traits into morphological, behavioural, physiological, growth and timing traits failed to find substantial differences. Previous meta-analyses and syntheses have found life-history traits to have lower heritabilities than other trait types (Postma 2014; Mittell *et al.* 2015) or morphological (Mousseau & Roff 1987), and specifically size traits (Hansen *et al.* 2011) to have higher heritabilities. The fact that we do not see these patterns recapitulated at the between-population level casts doubt on the utility of substituting phenotypic measures of relative divergence (P_{st}; Leinonen *et al.* (2006)) for genetic measures (Q_{st}; Wright (1951), Spitze (1993)) since the validity of this substitution assumes genetic variation has the same proportional contribution to both within and between population variation (Brommer 2011). However, given there are relatively few non-morphological traits in our analyses (and life-history traits were omitted) we urge caution in accepting our null result without further investigation.

For the non-fitness traits we analysed, we found a strong correlation between the plastic responses of each paired population, and so very little evidence for strong GxE interactions. This is in direct contrast to meta-analyses of fitness traits that have found substantial local-adaptation (Hereford 2009) since metrics of local-adaptation, when averaged over comparisons,

equal the difference in the plastic response of fitness among populations (i.e. $\Delta E_A - \Delta E_B$) and therefore represent GxE interaction (Blanquart *et al.* 2013). This suggests that the genetic determination of traits may be relatively insensitive to the environmental context, and that local adaptation is primarily driven by trait-fitness relationships that vary over environments. A similar pattern has been found in the sexual antagonism literature where the cross-sex genetic correlation is close to one for non-fitness traits, but reduced for fitness components (Poissant *et al.* 2010).

In conclusion, we show that plasticity plays a dominant role in explaining between-population phenotypic divergence, and that it usually acts in the same direction as genetic differentiation, consistent with it being adaptive. Nevertheless, substantial variation exists, and in a large minority of cases plasticity can act in opposition to genetic differentiation, and in a small minority of cases may even be maladaptive.

AUTHORSHIP

JDH conceived the study, MS collected the data, MS and JDH analysed the data and wrote the manuscript.

DATA AVAILABILITY STATEMENT

Data available from the Figshare Repository: https://doi.org/10.6084/m9.figshare.12311630.

REFERENCES

Acasuso-Rivero, C., Murren, C.J., Schlichting, C.D. & Steiner, U.K. (2019). Adaptive phenotypic plasticity for life-history and less fitness-related traits. *Proc. R. Soc. B*, 286, 20190653.

Blanquart, F., Kaltz, O., Nuismer, S.L. & Gandon, S. (2013). A practical guide to measuring local adaptation. *Ecol. Lett.*, 16, 1195–1205.

Bohonak, A.J. (1999). Dispersal, gene flow, and population structure. *Q. Rev. Biol.*, 74, 21–45.

Boutin, S. & Lane, J.E. (2014). Climate change and mammals: evolutionary versus plastic responses. *Evol. Appl.*, 7, 29–41.

Bradshaw, A. (1972). Some of the evolutionary consequences of being a plant. *Evol. Biol.*, 5, 25–47.

Brommer, J. (2011). Whither p_{ST}? the approximation of q_{ST} by p_{ST} in evolutionary and conservation biology. *J. Evol. Biol.*, 24, 1160–1168.

Camilli, G., de la Torre, J. & Chiu, C.Y. (2010). A noncentral t regression model for meta-analysis. *Journal of Educational and Behavioral Statistics*, 35, 125–153.

Charmantier, A. & Gienapp, P. (2014). Climate change and timing of avian breeding and migration: evolutionary versus plastic changes. *Evol. Appl.*, 7, 15–28.

Chevin, L.M., Lande, R. & Mace, G.M. (2010). Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. *PLoS Biol.*, 8, e1000357.

Conover, D.O., Duffy, T.A. & Hice, L.A. (2009). The covariance between genetic and environmental influences across ecological gradients. *Ann. N. Y. Acad. Sci.*, 1168, 100–129.

Conover, D.O. & Schultz, E.T. (1995). Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends Ecol. Evol.*, 10, 248–252.

De Kroon, H., Huber, H., Stuefer, J.F. & Van Groenendael, J.M. (2005). A modular concept of phenotypic plasticity in plants. *New Phytol.*, 166, 73–82.

Edelaar, P., Jovani, R. & Gomez-Mestre, I. (2017). Should I change or should I go? phenotypic plasticity and matching habitat choice in the adaptation to environmental heterogeneity. *Am. Nat.*, 190, 506–520.

Franks, S.J., Weber, J.J. & Aitken, S.N. (2014). Evolutionary and plastic responses to climate change in terrestrial plant populations. *Evol. Appl.*, 7, 123–139.

Friedman, J. & Barrett, S.C. (2009). Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants. *Ann. Bot.*, 103, 1515–1527.

Galloway, L.F. & Fenster, C.B. (2000). Population differentiation in an annual legume: local adaptation. *Evolution*, 54, 1173–1181.

Garant, D., Kruuk, L.E.B., McCleery, R.H. & Sheldon, B.C. (2004). Evolution in a changing environment: A case study with great tit fledgling mass. *Am. Nat.*, 164, E115–E129.

Gavrilets, S. & Scheiner, S.M. (1993). The genetics of phenotypic plasticity V. evolution of reaction norm shape. *J. Evol. Biol.*, 6, 31–48.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.*, 1, 515–533.

Ghalambor, C.K., Hoke, K.L., Ruell, E.W., Fischer, E.K., Reznick, D.N. & Hughes, K.A. (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature*, 525, 372.

Ghalambor, C.K., McKay, J.K., Carroll, S.P. & Reznick, D.N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Funct. Ecol.*, 21, 394–407.

Gienapp, P., Postma, E. & Visser, M.E. (2006). Why breeding time has not responded to selection for earlier breeding in a songbird population. *Evolution*, 60, 2381–2388.

Gienapp, P., Teplitsky, C., Alho, J.S., Mills, J.A. & Merila, J. (2008). Climate change and evolution: disentangling environmental and genetic responses. *Mol. Ecol.*, 17, 167–178.

Grether, G.F. (2005). Environmental change, phenotypic plasticity, and genetic compensation. *Am. Nat.*, 166, E115–E123.

Hadfield, J.D. (2010). MCMC methods for multi-response generalised linear mixed models: The MCMCglmm R package. *J. Stat. Softw.*, 33, 1–22.

Hadfield, J.D. (2016). The spatial scale of local adaptation in a stochastic environment. *Ecol. Lett.*, 19, 780–788.

Hadfield, J.D., Wilson, A.J., Garant, D., Sheldon, B.C. & Kruuk, L.E.B. (2010). The misuse of BLUP in ecology and evolution. *Am. Nat.*, 175, 116–125.

Hadfield, J.D., Wilson, A.J. & Kruuk, L.E.B. (2011). Cryptic evolution: does environmental deterioration have a genetic basis? *Genetics*, 187, 1099–1113.

Hamrick, J., Linhart, Y. & Mitton, J. (1979). Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. *Annu. Rev. Ecol. Syst.*, 10, 173–200.

Hansen, T.F., Pelabon, C. & Houle, D. (2011). Heritability is not evolvability. *Evol. Biol.*, 38, 258–277.

Hedges, L.V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *J. Edu. Stat.*, 6, 107–128.

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.*, 173, 579–588.

Huang, Y. & Agrawal, A.F. (2016). Experimental evolution of gene expression and plasticity in alternative selective regimes. *PLoS Genet.*, 12, e1006336.

Huey, R.B., Carlson, M., Crozier, L., Frazier, M., Hamilton, H., Harley, C. et al. (2002). Plants versus animals: do they deal with stress in different ways? *Integr. Comp. Biol.*, 42, 415–423.

Jackson, C.H. (2011). Multi-state models for panel data: The msm package for R. *J. Stat. Softw.*, 38, 1–28.

Jacob, S., Legrand, D., Chaine, A.S., Bonte, D., Schtickzelle, N., Huet, M. & et al. (2017). Gene flow favours local adaptation under habitat choice in ciliate microcosms. *Nature Ecol. Evol.*, 1, 1407.

de Jong, G. (1999). Unpredictable selection in a structured population leads to local genetic differentiation in evolved reaction norms. *J. Evol. Biol.*, 12, 839–851.

Joshi, J., Schmid, B., Caldeira, M., Dimitrakopoulos, P., Good, J., Harris, R. et al. (2001). Local adaptation enhances performance of common plant species. *Ecol. Lett.*, 4, 536–544.

Kawecki, T.J. & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecol. Lett.*, 7, 1225–1241.

King, J.G. & Hadfield, J.D. (2019). The evolution of phenotypic plasticity when environments fluctuate in time and space. *Evolution letters*, 3, 15–27.

Leimu, R. & Fischer, M. (2008). A meta-analysis of local adaptation in plants. *PLoS One*, 3, e4010.

Leinonen, T., Cano, J.M., Mäkinen, H. & Merilä, J. (2006). Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *J. Evol. Biol.*, 19, 1803–1812.

Levins, R. (1968). *Evolution in Changing Environments: Some Theoretical Explorations*. Princeton University Press, Princeton, NJ.

Lewandowski, D., Kurowicka, D. & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *J. Multivar. Anal*, 100, 1989–2001.

Marsaglia, G. (1965). Ratios of normal variables and ratios of sums of uniform variables. *J. Am. Stat. Assoc.*, 60, 193–204.

Marsaglia, G. (2006). Ratios of normal variables. *J. Stat. Softw.*, 16, 1–10.

Merilä, J. & Hendry, A.P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evol. Appl.*, 7, 1–14.

Merila, J., Kruuk, L.E.B. & Sheldon, B.C. (2001). Cryptic evolution in a wild bird population. *Nature*, 412, 76–79.

Mittell, E.A., Nakagawa, S. & Hadfield, J.D. (2015). Are molecular markers useful predictors of adaptive potential? *Ecol. Lett.*, 18, 772–778.

Mousseau, T.A. & Roff, D.A. (1987). Natural selection and the heritability of fitness components. *Heredity*, 59, 181–197.

Palacio-López, K., Beckage, B., Scheiner, S. & Molofsky, J. (2015). The ubiquity of phenotypic plasticity in plants: a synthesis. *Ecol. Evol.*, 5, 3389–3400.

Palmer, A.R. (2000). Quasireplication and the contract of error: Lessons from sex ratios, heritabilities and fluctuating asymmetry. *Annu. Rev. Ecol. Syst.*, 31, 441–480.

Pigliucci, M. (2001). *Phenotypic Plasticity: Beyond Nature and Nurture*. John Hopkins Univeristy Press, Baltimore, MD.

Poissant, J., Wilson, A.J. & Coltman, D.W. (2010). Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. *Evolution*, 64, 97–107.

Postma, E. (2014). Four decades of estimating heritabilities in wild vertebrate populations: improved methods, more data, better estimates. In: *Quantitative Genetics in the Wild* (edsCharmantier, A., Garant, D., & Kruuk, L.E.B.). Oxford University Press, Oxford, pp. 16–33.

Rohatgi, A. (2012). WebPlotDigitizer. Version: 4.2. Available at: https://automeris.io/WebPlotDigitizer

Schilthuizen, M. & Kellermann, V. (2014). Contemporary climate change and terrestrial invertebrates: evolutionary versus plastic changes. *Evol. Appl.*, 7, 56–67.

Slatkin, M. (1978). Spatial patterns in distributions of polygenic characters. *J. Theor. Biol.*, 70, 213–228.

Spitze, K. (1993). Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics*, 135, 367–374.

Stoks, R., Geerts, A.N. & De Meester, L. (2014). Evolutionary and plastic responses of freshwater invertebrates to climate change: realized patterns and future potential. *Evol. Appl.*, 7, 42–55.

van Tienderen, P.H. (1997). Generalists, specialists, and the evolution of phenotypic plasticity in sympatric populations of distinct species. *Evolution*, 51, 1372–1380.

Tufto, J. (2000). The evolution of plasticity and nonplastic spatial and temporal adaptations in the presence of imperfect environmental cues. *Am. Nat.*, 156, 121–130.

Turesson, G. (1922). The genotypical response of the plant species to the habitat. *Hereditas*, 3, 211–350.

Via, S. & Lande, R. (1985). Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution*, 39, 505–522.

Whitlock, M.C. & Mccauley, D.E. (1999). Indirect measures of gene flow and migration: $f_{ST} \neq 1/(4nm+1)$. *Heredity*, 82, 117–125.

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

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, 15, 323–354.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.