Received: 7 October 2021 | Accepted: 8 March 2023

DOI: 10.1111/2041-210X.14098

Methods in Ecology and Evolution

RESEARCH ARTICLE

Estimation of additive genetic variance when there are gene–environment correlations: Pitfalls, solutions and unexplored questions

Gabriel Munar-Delgado[1] | Yimen G. Araya-Ajoy[2] | Pim Edelaar[3,4]

[1]Department of Biodiversity, Ecology and Evolution. Faculty of Biological Sciences, Complutense University of Madrid, 28040, Madrid, Spain

[2]Center for Biodiversity Dynamics, Norwegian University of Science and Technology, Trondheim, 7012, Norway

[3]Department of Molecular Biology and Biochemical Engineering, University Pablo de Olavide, Seville, 41013, Spain

[4]Swedish Collegium for Advanced Study, Uppsala, 752 38, Sweden

Correspondence
Gabriel Munar-Delgado
Email: gmunar@ucm.es

Pim Edelaar
Email: edelaar@upo.es

Funding information
Norges Forskningsråd, Grant/Award Number: 325826 and PID 302619; Spanish Ministry of Science and Innovation, Grant/Award Number: CGL2016-79483-P and PID2019-108971GB-I00; Spanish Ministry of Universities, Grant/Award Number: FPU17/06268; Swedish Collegium for Advanced Study

Handling Editor: Michael Morrissey

Abstract

1. Estimating the genetic variation underpinning a trait is crucial to understanding and predicting its evolution. A key statistical tool to estimate this variation is the *animal model*. Typically, the environment is modelled as an external variable independent of the organism, affecting the focal phenotypic trait via phenotypic plasticity. We studied what happens if the environment is not independent of the organism because it chooses or adjusts its environment, potentially creating non-zero genotype–environment correlations.
2. We simulated a set of biological scenarios assuming the presence or absence of a genetic basis for a focal phenotypic trait and/or the focal environment (treated as an extended phenotype), as well as phenotypic plasticity (the effect of the environment on the phenotypic trait) and/or 'environmental plasticity' (the effect of the phenotypic trait on the local environment). We then estimated the additive genetic variance of the phenotypic trait and/or the environment by applying five animal models which differed in which variables were fitted as the dependent variable and which covariates were included.
3. We show that animal models can estimate the additive genetic variance of the local environment (i.e. the extended phenotype) and can detect environmental plasticity. We show that when the focal environment has a genetic basis, the additive genetic variance of a phenotypic trait increases if there is phenotypic plasticity. We also show that phenotypic plasticity can be mistakenly inferred to exist when it is actually absent and instead environmental plasticity is present. When the causal relationship between the phenotype and the environment is misunderstood, it can lead to severe misinterpretation of the genetic parameters, including finding 'phantom' genetic variation for traits that, in reality, have none. We also demonstrate how using bivariate models can partly alleviate these issues. Finally, we provide the mathematical equations describing the expected estimated values.

Munar-Delgado and Araya-Ajoy authors contributed equally to this work

4. This study highlights that not taking gene–environment correlations into account can lead to erroneous interpretations of additive genetic variation and phenotypic plasticity estimates. If we aim to understand and predict how organisms adapt to environmental change, we need a better understanding of the mechanisms that may lead to gene–environment correlations.

KEYWORDS
additive genetic variance, animal model, bivariate model, environmental plasticity, extended phenotype, gene–environment correlation, gene–environment covariance, phenotypic plasticity

1 | INTRODUCTION

One of the biggest current challenges in evolutionary biology is understanding how populations adapt to their environment and predicting if they will be able to cope with the pace of anthropogenically induced environmental change. It is thus essential to understand the genetic and phenotypic changes that allow populations to cope with environmental challenges (Chevin et al., 2010). Within this context, the field of quantitative genetics studies the additive genetic variance of traits. This is a key determinant of their heritability and evolutionary potential and, thus, of the ability of populations to adapt to their environment (Falconer & Mackay, 1996). One of the most widely used statistical tools for estimating the additive genetic variance of a trait is a type of mixed model called the 'animal model' (Kruuk, 2004; Lynch & Walsh, 1998).

Various studies have highlighted possible pitfalls when using animal models, and the errors associated with their application and biological interpretation (de Villemereuil et al., 2018; Kruuk, 2004; Kruuk et al., 2008; Postma & Charmantier, 2007; Wilson et al., 2010). For example, failing to account for maternal effects may cause additive genetic variance (and thus heritability) to be overestimated due to the genetic covariance between siblings as generated by the shared environment (e.g. the care of their mother; Wilson et al., 2010).

Another problem with the animal model is that the assumed model structure might not reflect the actual biology of the system (Westneat et al., 2020), and this has received far less attention. Here we specifically focus on the assumption that genes and environment are uncorrelated (Lynch & Walsh, 1998) (Figure 1a). If not, the gene–environment covariance changes the estimated phenotypic variance:

$$V_P = V_G + V_E + 2\mathrm{Cov}[G,E], \quad (1)$$

where V_P is the phenotypic trait variance, V_G is its additive genetic variance, V_E its environmental variance and $Cov[G,E]$ is the gene–environment covariance. That this covariance is zero (i.e. that genotypes are randomly distributed across environments) may be true in captivity or other controlled conditions, but not necessarily so in natural populations. Although this problem has long been recognized (e.g. Falconer, 1960), it seems to be systematically ignored (e.g. it is not mentioned in Charmantier et al., 2014).

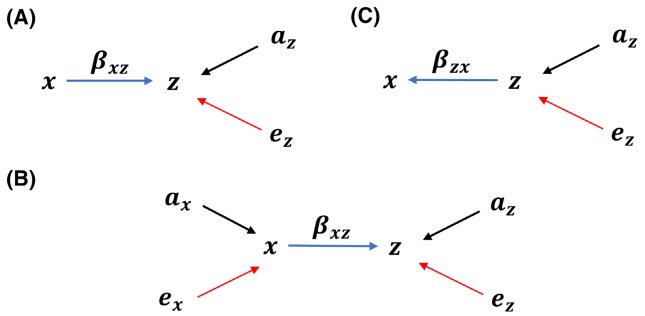

FIGURE 1 (a) Typical quantitative genetic partitioning of focal phenotypic trait z into direct additive genetic (a_z), environmental (x) and residual (e_z) components. A known environment may affect the phenotypic trait via phenotypic plasticity (β_{xz}), but this effect is independent of the organism's genes. (b) Compared to a), when an organism has a genetic preference (a_x) to choose or adjust its local environment (x), the genes of the organism and its environment are no longer independent. In this scenario, the genes influencing choice or adjustment of the local environment also indirectly influence the phenotypic trait through phenotypic plasticity (β_{xz}). (c) When the phenotypic trait (z) of an organism has a genetic component (a_z) and affects the choice or adjustment of the local environment (x) via 'environmental plasticity' (β_{zx}), the genes underpinning the expression of the phenotypic trait indirectly affect the choice or adjustment of the environment. Thus, the genes of the organism and its environment are no longer independent, causing a genetic correlation between the organism and its environment.

A common approach to deal with phenotype–environment relationships is to fit the environment as a covariate. However, this approach implicitly assumes that their covariance is environmental instead of genetic. This assumption may be incorrect in natural populations for two reasons. First, there might be a genetic basis to aspects of the environment that we include in the model (Figure 1b). This could occur when the organism has a (genetic) preference to occur in a specific type of environment ('selection (choice) of the environment' cf. Edelaar & Bolnick, 2019, e.g. habitat choice), or when it has a (genetic) inclination to change its environment to a different state ('adjustment of the environment' cf. Edelaar & Bolnick, 2019, for example habitat construction). Although it has been shown that genetic variation can affect and determine an individual's

environment (e.g. Dawkins, 1982; Jaenike & Holt, 1991; Weber et al., 2013), this possibility is rarely explored using animal models (see Järvinen et al., 2017; Gervais et al., 2020, 2022 for exceptions).

A second case that violates the assumption that the organism and the environment are independent occurs when a phenotypic trait influences how the local environment is adjusted or which environment is chosen (e.g. phenotype-dependent matching habitat choice; Edelaar et al., 2008; Figure 1c). This is an effect of the phenotypic trait on the focal environment, the reverse effect of phenotypic plasticity (where the environment affects the phenotypic trait). For lack of an established term (as far as we know), we will call this 'environmental plasticity' as a logical equivalent of phenotypic plasticity. Environmental plasticity covers the phenotype dependency of both choice and adjustment of the local environment that an individual experiences. Note that for habitat choice/selection, for an outside observer the environment does not undergo any transformation. However, the observing researcher is irrelevant: the local environment that the individual organism experiences does indeed change, and this is what matters. In this way, if an organism's focal phenotypic trait harbours genetic variation and affects the choice or adjustment of its local environment (i.e. there is environmental plasticity), then a genetic covariance between the trait and the environment is expected. Assuming the inverse causal relationship of an effect of the focal environment on the phenotypic trait (i.e. assuming phenotypic plasticity when there is environmental plasticity) could cause wrong inferences of the animal model estimates. As an example, imagine we are studying the heritability or plasticity of behavioural boldness of a breeding wild bird population. We might be tempted to add nest distance from the closest road as a fixed effect to control for the influence of human disturbance on boldness (i.e. phenotypic plasticity). Alternatively, boldness could affect nest distance choice (i.e. environmental plasticity, with less bold individuals preferring to breed further away from roads) instead of the other way around (Holtmann et al., 2017). For this scenario, treating boldness as a response variable and nest distance as the independent variable would not reflect the true causal relationship. This could result in misinterpretation of the estimates provided by the animal model due to the misspecification of the causal structure of the model.

When studying gene–environment correlations, it is important to emphasize that for any hypothesized relationship between phenotypes and environment, the estimated genetic variance can be underpinned by 'direct' genetic variance affecting a trait or environment versus 'indirect' genetic variance. We therefore use the term 'direct additive genetic variance' to refer to the variance caused by alleles 'directly' affecting a trait (path a_z to z in Figure 1a), in the sense that the causally intermediate traits are not measured or of interest. Within a path analysis context, this has been referred to as exogenous variance because factors outside the causal pathway cause it (e.g. de Villemereuil et al., 2018). In the path diagram depicted in Figure 1a, the direct variance can also be thought of as the expected genetic variance on the trait conditional on all individuals having the same focal environment. In contrast, we use the term 'indirect additive genetic variance' to acknowledge that alleles underpinning traits affecting the environment may cause indirect genetic variance on other traits, because of the indirect effects of alleles on a phenotype through the environment (path a_x to x to z in Figure 1b). In other words, variance that is caused by a plastic response to variation in a phenotype or environment with a genetic underpinning. Finally, we use the term 'total additive genetic variance' to refer to the sum of both direct and indirect genetic variance.

When the local environment varies depending on the individual's genotype, it may be seen as an extended phenotype, that is the expression of genes in traits outside of what is typically considered the organism (Dawkins, 1982; Edelaar & Bolnick, 2019). Studying environmental variables as extended phenotypes allows linking the study of gene–environment correlations to previous treatments on selection on causally covarying traits (e.g. Morrissey, 2014). Contrary to Dawkins's definition, here we use extended phenotype without implying that the choice or adjustment of the focal environment is expected to result in a change in fitness. (see Supporting Information I for further discussion on treating environmental variables as extended phenotypes.)

In this paper, we combine data simulations with mathematical derivations to describe the consequences for the animal model estimates and their interpretation when the focal environment depends on the genes of our study organism. We simulated data assuming 12 different biological scenarios and then analysed these data with a set of 5 animal models with different causal structures. The fitted animal models varied in which trait was the focal trait and the covariate (including bivariate models). We then show how the additive genetic variance estimated by the different animal models ($\hat{V}_a$) fitted for the different biological scenarios can either reflect a trait's *direct* additive genetic variance, its *indirect* additive genetic variance, the sum of these two (i.e. its *total* additive genetic variance), or a biased estimate which is not consistent with anything of the above. What the animal model provides us with depends on the existence of gene–environment correlation and the appropriateness of the model structure for each biological scenario (see Table I for the expected estimated values for each model for each scenario).

2 | MATERIALS AND METHODS

2.1 | General simulation design

We developed a simulation procedure in *R* 4.0.2 (R Core Team, 2020) to study the impact of gene–environment correlations on the estimates of plasticity and additive genetic variance. This simulation generates an individual phenotypic trait (z) which has a genetic underpinning summarized by a breeding value (α_z), which responds plastically (β_{xz}) to a focal environmental variable (x) and which is affected by unknown effects summarized in a residual value (ε_z),

$$z = \alpha_z + \beta_{xz}x + \varepsilon_z. \quad (2)$$

In a similar way, the focal environment variable may have a genetic underpinning summarized by a breeding value (α_x), it can also be affected by the organism's phenotypic trait (z) proportional to the coefficient β_{zx}, and it is also affected by unknown residual effects (ϵ_x),

$$x = \alpha_x + \beta_{zx} z + \epsilon_x. \quad (3)$$

The unknown residual effects on the phenotypic trait (ϵ_z) and the environment (ϵ_x) are assumed to be a realization of a normal distribution with a mean of zero and variance $\sigma^2_{\epsilon_z}$ and $\sigma^2_{\epsilon_x}$, respectively.

The breeding values for each individual are the sum of the effects of n loci (in linkage disequilibrium) influencing trait expression. Each locus has two alleles coded as 0 and 1, and for simplicity, we assume that their effects on the trait are additive. Also, for simplicity, we assume that allele frequencies for all loci are 0.5. Therefore, the expected direct additive genetic variance in a trait ($\sigma^2_{\alpha_z}$) or environment ($\sigma^2_{\alpha_x}$) is equal to $0.5^2 n$, where n can differ between the phenotypic trait and the focal environment.

We start by simulating the genotypes of a founder population of 750 individuals; then sexes are randomly assigned ensuring a sex ratio of 1. Individuals mate randomly, there is no natural selection, and there are no overlapping generations. Alleles follow Mendelian segregation, without mutation. We simulate 10 non-overlapping generations and each pair produces two individuals, ensuring a constant population size across generations. The genotype of each individual is then used to calculate its breeding value. Finally, pedigrees are built based on the parent–offspring relationships (see Supplementary Material for more details on the simulation procedure).

2.2 | Biological scenarios

We simulated 12 different biological scenarios (Table 1) that involved the presence or absence of the following factors: direct additive genetic variance for the phenotypic trait ($\sigma^2_{\alpha_z} = 25$), direct additive genetic variance for the focal environment ($\sigma^2_{\alpha_x} = 12.5$), phenotypic plasticity ($\beta_{xz} = 0.7$; i.e. the effect of the focal environment on the phenotypic trait, Figure 1a,b) and environmental plasticity ($\beta_{zx} = 0.5$; i.e. the effect of the phenotypic trait on the focal environment, Figure 1c). A fixed variance in z and x due to unknown residual effects was always simulated (120 and 70 respectively). These values are arbitrary, but we intentionally used values to make the simulation results clear. Every scenario was simulated 100 times, resulting in 100 datasets that differed due to stochastic variation. The simulation results are only valid for the specific choices of each simulation, and serve mostly as examples. Therefore, we generalize the simulation results by providing the analytical formulas for the expected estimated values by the different animal models for each scenario.

Scenarios 1–4 correspond to populations with a phenotypic trait with a direct genetic basis and/or phenotypic plasticity. These scenarios were used to check the simulation procedures since they are the classical scenarios usually assumed when using animal models. In scenarios 5–8, we simulated a direct genetic basis for the focal environment. In scenarios 7 and 8, we also simulated phenotypic plasticity. In these hypothetical scenarios, the focal phenotypic trait is affected by an environmental variable with a direct genetic basis. For example, when the water depth at which a deep-sea fish forages has a direct genetic basis (Gaither et al., 2018) and this depth (focal environment) affects its body mass (phenotypic trait). For scenarios 9–12, we simulated environmental plasticity instead of phenotypic plasticity. In other words, in these scenarios, we simulated that the focal environment was affected by the focal phenotypic trait. One such example is shown in Camacho et al. (2020), where ground-perching grasshoppers of a specific colour (focal phenotypic trait) choose a substrate of a colour (environmental variable) similar to their own to increase crypsis. This does not change the actual colour of any of the available substrates, but it does change the colour of the environment that each individual experiences, which is our focal trait. For simplicity, we did not simulate scenarios where phenotypic and environmental plasticity are simultaneously present, as this leads to feedback and possible order effects, although it appears to be possible in nature (Boyle & Start, 2020; Lowe & Addis, 2019).

2.3 | Statistical analyses: Animal models

We fitted the animal models using the R package ASReml-R (Butler, 2020). See Table 2 for parameter descriptions.

We fitted five different animal models. Models 1 and 2 correspond to the typical structures usually used to estimate the genetic parameters of the phenotypic trait of interest (z). Model 1 is the simplest case:

$$z = u_z + a_z + e_z. \quad (4)$$

where u_z is the population mean, a_z are the breeding values with variance $\widehat{V}_{a_z}$ and e_z is the residual term with variance $\widehat{V}_{e_z}$.

Simulated parameter	Scenario											
	1	2	3	4	5	6	7	8	9	10	11	12
$\sigma^2_{\alpha_z} = 25$	–	–	✓	✓	–	✓	–	✓	–	–	✓	✓
$\sigma^2_{\alpha_x} = 12.5$	–	–	–	–	✓	✓	✓	✓	–	✓	–	✓
$\beta_{xz} = 0.7$	–	✓	–	✓	–	–	✓	✓	–	–	–	–
$\beta_{zx} = 0.5$	–	–	–	–	–	–	–	–	✓	✓	✓	✓

TABLE 1 Parameters used for simulating in each scenario. '✓' indicate when a particular parameter was involved using the given values, where $\sigma^2_{\alpha_z}$ is direct additive genetic variance for the focal phenotypic trait, $\sigma^2_{\alpha_x}$ is direct additive genetic variance for the focal environment, β_{xz} is phenotypic plasticity and β_{zx} is environmental plasticity.

TABLE 2 Notation and description of each of the parameters that were simulated. To distinguish estimated parameters for (co)variances and types of plasticity from simulated ones, we denote parameter estimates using a hat symbol (e.g. $\hat{\beta}_{xz}$). Note that breeding values and residual values are mean centred for Equations 4–8, representing the statistical analyses but not for Equations 2 and 3, describing the simulation process. We thus represent them with different symbols. We also refer to $\hat{V}_a$ to the estimates of additive genetic variance to highlight that this may or may not represent any of the values used for the simulation.

Parameters	
Simulated	**Description**
z	Focal phenotypic trait
x	Focal environment/extended phenotype
a_z	Breeding value for the phenotype
e_z	Effect of residual variables on the phenotypic trait
a_x	Breeding value for the focal environment
e_x	Effect of residual variables on the focal environment
$\sigma^2_{a_z}$	Direct additive genetic variance of the phenotypic trait
$\sigma^2_{a_x}$	Direct genetic variance of the focal environment
β_{xz}	Strength of the effect of the focal environment on the phenotypic trait (i.e. phenotypic plasticity)
β_{zx}	Strength of the effect of the phenotypic trait on the focal environment (i.e. environmental plasticity)

Model 2 has a similar model structure but includes the focal environment of each individual (x) as a fixed effect and the effect of such an environment on the individuals' phenotypic trait ($\hat{\beta}_{xz}$):

$$z = u_z + \hat{\beta}_{xz}x + a_z + e_z. \quad (5)$$

Model 3 is similar to model 1, but now the environment is the trait of interest (the dependent variable), potentially with its own genetic basis:

$$x = u_x + a_x + e_x. \quad (6)$$

where x is the experienced focal environment variable of each individual, u_x is the population mean, a_x are the breeding values with variance $\hat{V}_{a_x}$ and e_x represent the residual effects with variance $\hat{V}_{e_x}$.

Following the logic and structure of model 2, in model 4 we fit the phenotypic trait as a fixed effect to estimate and control for its influence on the focal environment variable (i.e. environmental plasticity):

$$x = u_x + \hat{\beta}_{zx}z + a_x + e_x. \quad (7)$$

Finally, model 5 is the bivariate model that is fitted to estimate the additive genetic covariance $Cov[a_z, a_x]$ between the phenotypic and the environmental variable:

$$\begin{bmatrix} z \\ x \end{bmatrix} = u + a + e. \quad (8)$$

$$\begin{bmatrix} a_z \\ a_x \end{bmatrix} \sim \text{MVN}(0, \mathbf{G}): \begin{bmatrix} \hat{V}_{a_z} & \widehat{\text{Cov}}[a_z, a_x] \\ \widehat{\text{Cov}}[a_x, a_z] & \hat{V}_{a_x} \end{bmatrix}. \quad (9)$$

where $\widehat{\text{Cov}}[a_z, a_x]$ represents the estimated additive genetic covariance between the phenotypic and the focal environment, and $\hat{V}_{a_z}$ and $\hat{V}_{a_x}$ represent the estimated additive genetic variance in the phenotypic trait and focal environment, respectively. MVN(0, **G**) represents a multivariate normal distribution where G is the genetic variance–covariance matrix.

For bivariate models, starting values for the variances in the additive genetic (**G**) and the residual (R) covariance matrices were based on the output from the univariate models (Wilson et al., 2010) to improve model convergence. For simplicity, initial values were set to 0.1 when univariate models detected additive genetic variance and set to 0.0000001 when the estimated additive genetic variance was close to 0. Covariance starting values were always 0.0000001.

3 | RESULTS

In general, applying different animal models (i.e. assuming different causal structures) results in the estimation of different parameters (total, direct or indirect additive genetic variance, or a biased estimate that does not correspond with any potential parameter of interest) depending on the specific characteristics of the different biological scenarios. Table I in the Supporting Information provides a summary for all possible model-scenario combinations.

3.1 | Analysing the standard scenarios with classical animal model structures

When models 1 and 2 were applied to standard scenarios 1–4 where a direct genetic basis for the phenotypic trait and/or phenotypic plasticity was simulated, the expected value for the estimated genetic variance is equal to the simulated direct additive genetic variance of the phenotypic trait ($\mathbb{E}\left[\hat{V}_{a_z}\right] = \sigma^2_{a_z}$; Supporting Information results, Table S2). Note that we refer to the expected value of the estimates ($\mathbb{E}$) because each estimate will vary around the value used as input for the simulations because of finite sample size.

3.2 | What happens if the focal environment has a genetic basis, and it is fitted as a dependent variable in an animal model?

When the presence or absence of a direct genetic basis for the focal environment and/or environmental plasticity were simulated (scenarios 5, 9 and 10) and we applied models with the focal environment as the dependent variable and the phenotypic trait as a covariate (models 3 and 4), the expected value for the estimated genetic variance is equal to the simulated direct additive genetic variance of the focal environment ($\mathbb{E}\left[\hat{V}_{a_x}\right] = \sigma^2_{a_x}$; Figure S2). The estimated strength of environmental plasticity also corresponded to the simulated value ($\mathbb{E}\left[\hat{\beta}_{zx}\right] = \beta_{zx}$; Figure S3, Table S3).

3.3 | How are genetic variance estimates affected when the focal environment has a genetic basis and the phenotypic trait responds plastically to it?

When we simulated a genetic basis for the phenotypic trait and the focal environment but no plasticity (scenarios 5 and 6) and then applied models with the phenotypic trait as the dependent variable, without or with the focal environment as a covariate (models 1 and 2), the expected value for the estimated genetic variance is equal to the simulated direct genetic variance ($\mathbb{E}\left[\hat{V}_{a_z}\right] = \sigma^2_{a_z}$). The estimated phenotypic plasticity also matched the simulated value ($\mathbb{E}\left[\hat{\beta}_{xz}\right] = \beta_{xz}$; Figures 2 and 3; Table S4). Symmetrically, the same happened in scenarios 3 and 6 for the focal environment estimates when we applied models 3 and 4 (Supporting Information Results).

In contrast, when we simulated phenotypic plasticity alongside a genetic basis for the focal environment (scenarios 7 and 8) and applied the model with the phenotypic trait as dependent variable and no covariates (model 1), the expected value for the estimated genetic variance is equal to the sum of the simulated direct genetic effects and indirect genetic effects ($\mathbb{E}\left[\hat{V}_{a_z}\right] = \sigma^2_{a_z} + \beta^2_{xz}\sigma^2_{a_x}$; Figure 2; Table S4). On the other hand, when the focal environment was added as a covariate (model 3), the indirect genetic effects of the environment were statistically removed. The expected value for the estimated genetic variance is then equal to the simulated direct genetic variance ($\mathbb{E}\left[\hat{V}_{a_z}\right] = \sigma^2_{a_z}$) (Figure 2, Table S3). For these scenarios, phenotypic plasticity was correctly estimated (Figure 3, Table S3).

Symmetrically, the same was true for the genetic variance of the focal environment and environmental plasticity estimates in scenarios 11 and 12 when models 3 and 4 were applied (Supporting Information Results).

3.4 | What happens if there is environmental plasticity, yet the classical animal model structures are applied?

When applying the model with the phenotypic trait as the dependent variable and no environmental covariate (model 1) to scenarios where environmental plasticity was simulated (9, 10, 11 and 12), the expected value for the estimated genetic variance is equal to the simulated direct genetic variance ($\mathbb{E}\left[\hat{V}_{a_z}\right] = \sigma^2_{a_z}$; Figure 4, Table S5). However, when we added the focal environment as a covariate (model 2) and thus applied a model assuming the wrong causal relationship (assuming phenotypic plasticity when there is environmental plasticity), phenotypic plasticity was estimated to be present when it was absent (Figure 3, Table S5). Moreover, due to this misspecification, the additive genetic variance was misestimated (see Section 4; Figure 4, scenarios 10–12).

Symmetrically, the same happened for the scenarios 2, 4, 7 and 8 when model 4 was applied (see Supporting Information for Results).

3.5 | What if bivariate models are applied?

When applying the bivariate model (model 5) to all scenarios, estimates for the genetic variances of the phenotypic trait and the focal environment matched estimates by models 1 ($\hat{V}_{a_z}$) and 3 ($\hat{V}_{a_x}$) (i.e. univariate models without a covariate; Supporting Information Results). Moreover, the genetic covariance between the phenotypic trait and the local environment was correctly estimated when it was simulated to exist (phenotypic plasticity together with a genetic basis for the focal environment, scenarios 7 and 8; or environmental plasticity together with a genetic basis for the focal phenotypic trait, scenarios 11 and 12; Figure 5).

4 | DISCUSSION

We compared the results of different animal model structures applied to simulated data reflecting different biological scenarios. Our results show (i) how the genetic basis of the focal environment can be estimated using animal models, (ii) how animal models can estimate not only phenotypic plasticity but also environmental plasticity if the correct model structure is fitted, (iii) how plasticity can increase the additive genetic variance of the focal trait and how the additive genetic estimates provided by the animal model can be potentially misinterpreted, (iv) how fitting the wrong causal structure can result in wrong inferences about the additive genetic variance and type of plasticity and (v) how bivariate models can detect a genetic covariance between the phenotype and the focal environment and may help differentiating between alternative scenarios.

4.1 | The genetic basis of the focal environment can be estimated with animal models

The additive genetic variance of the focal environment can be estimated by fitting it as a dependent variable in an animal model (models 3 and 4). The estimated genetic variance of the focal environment should reflect genetic variation for an individual's preference and ability to choose or adjust its environment (Akcali & Porter, 2017; Edelaar & Bolnick, 2019). This modelling approach allows studying

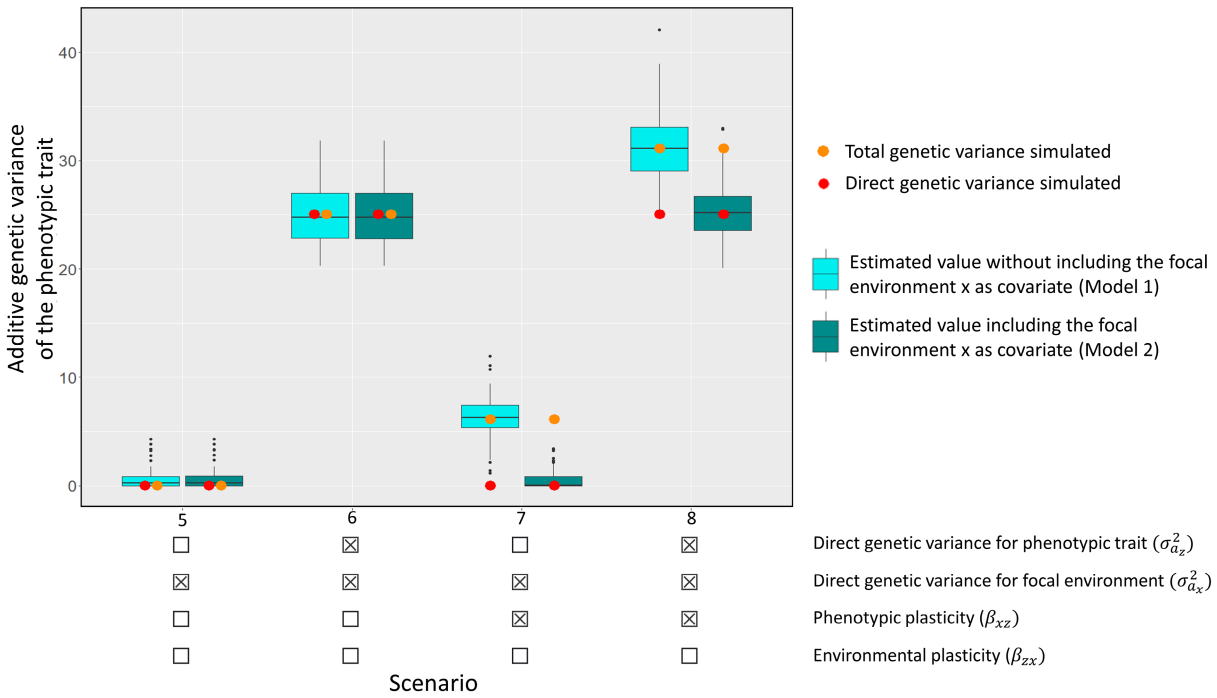

FIGURE 2 Estimated additive genetic variance of the phenotypic trait for models 1 (phenotypic trait as dependent variable) and 2 (with the focal environment as a covariate) for the simulated values in scenarios 5–8. The box plots illustrate the distribution of estimates of the 100 simulations for each scenario (the bottom and the top of the boxes are the first and third quartiles, the middle band is the median, its whiskers extend from the box to highest and lowest points within 1.5 times the interquartile range. Outliers are represented with black dots. Red dots are the simulated direct genetic variances for the focal phenotypic trait. Orange dots are the simulated total genetic variance (direct + indirect; see Section 4). Crossed squares (☒) indicate if non-zero direct genetic variance for the phenotypic trait, direct genetic variance for the focal environment, phenotypic plasticity and/or environmental plasticity were simulated.

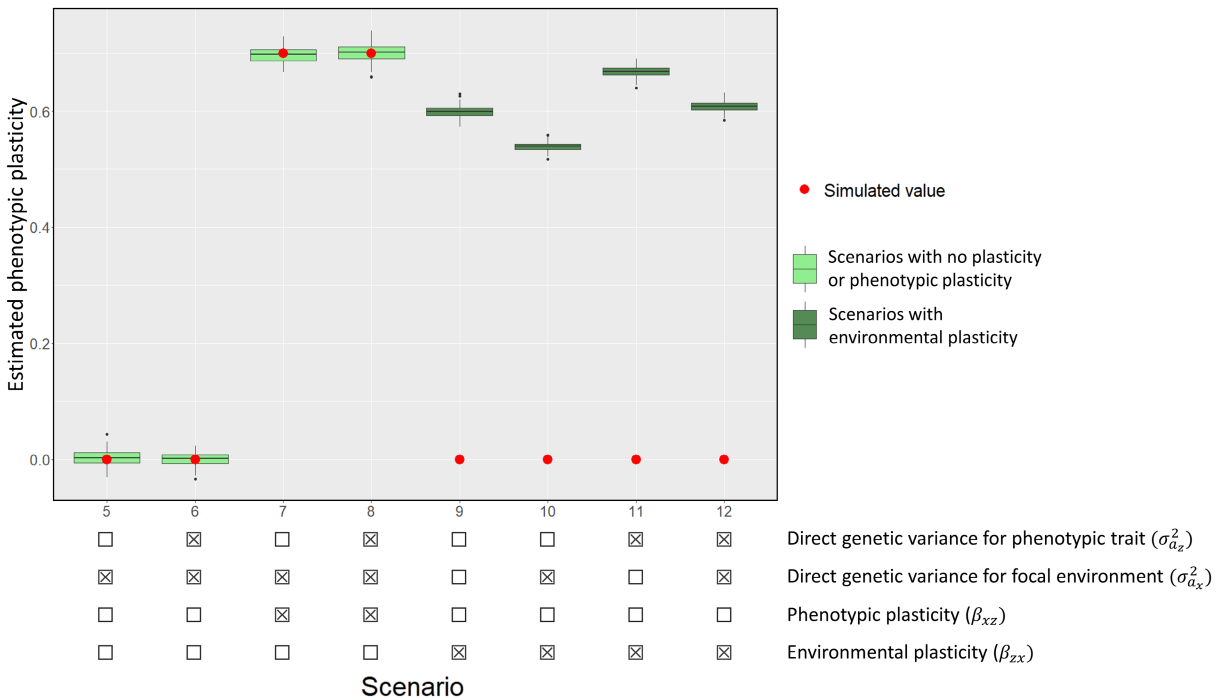

FIGURE 3 Estimated effects of the focal environment on the phenotypic trait (i.e. strength of phenotypic plasticity), for model 2 (focal environment as covariate). See Figure 2 for box plot description and legend explanation.

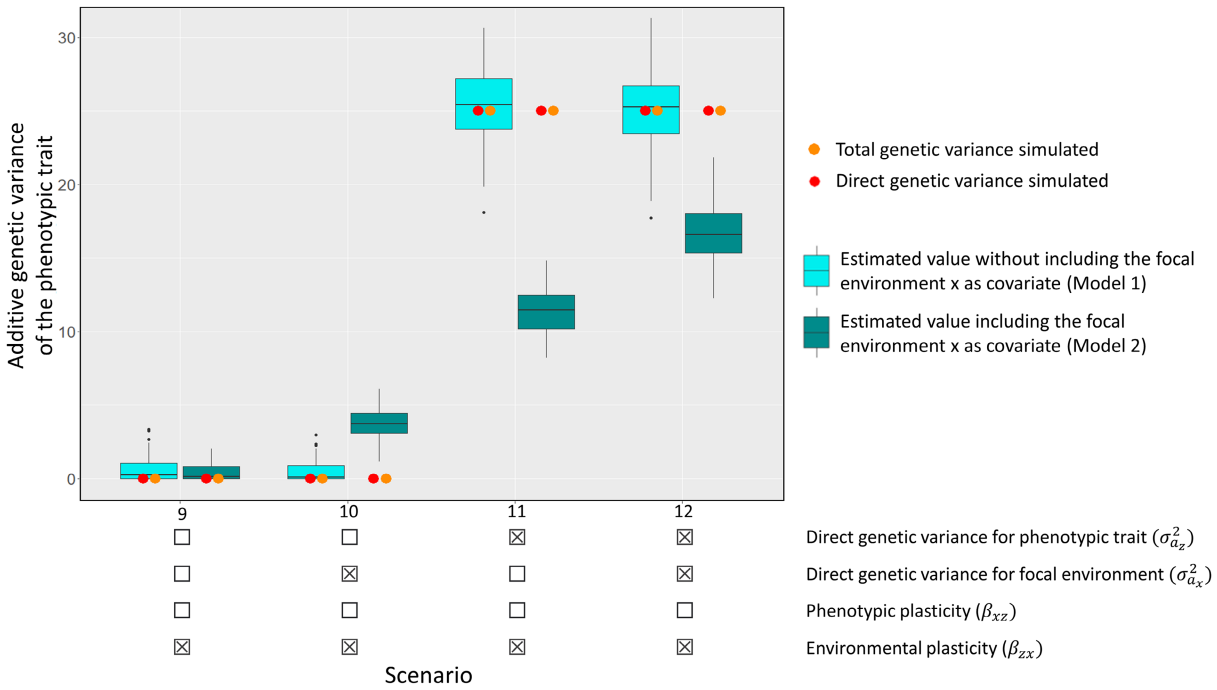

FIGURE 4 Distribution of the estimated values of additive genetic variance of the phenotypic trait for models 1 (phenotypic trait as dependent variable) and 2 (focal environment as covariate) for the simulated values in scenarios 9–12 (see Figure 2 for box plot description and legend explanation).

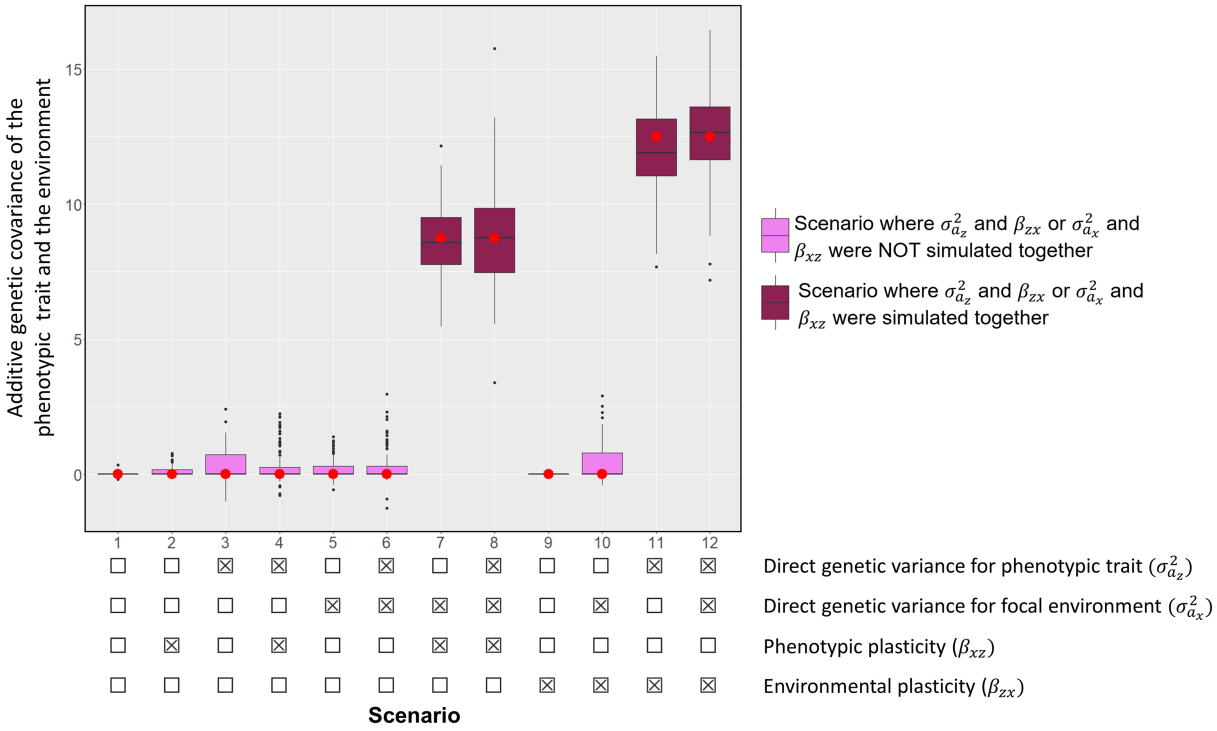

FIGURE 5 Estimated values of additive genetic covariance between the phenotypic trait and the focal environment (model 5, bivariate model) for the simulated values in all (12) scenarios. Red dots are the simulated additive genetic covariance between the focal phenotypic trait and the focal environment (see Figure 2 for box plot description and legend explanation).

the 'heritability' of environmental variables potentially chosen or adjusted by individuals in wild populations (i.e. extended phenotypes) using pedigree or genetic relatedness information. This approach would also be useful whenever the trait underpinning the choice or adjustment of the environment is unknown or cannot be measured directly. As discussed further below, what the estimate provided by the animal model means (direct, indirect or total additive genetic variance of the focal environment) will depend on the model structure.

Social scientists have already recognized that many environments are heritable, since humans select, modify and create environments using behaviours with a genetic basis (Plomin et al., 2016; Saltz, 2019). However, this view has hardly been adopted for wild non-human populations (see Møller, 2006; Weber et al., 2013 for exceptions), and animal models have almost never been applied for this purpose (Gervais et al., 2020, 2022; Järvinen et al., 2017; are the only exceptions we know of). We hope that researchers will recognize the potential of studying the heritable variation of the environment experienced by individuals by treating it as an extended phenotype. In this way, the focal environment is no longer exclusively modelled as an external ecological context imposing selective pressures, but as one that may be chosen or adjusted by the organism (Edelaar & Bolnick, 2019), and therefore as an integral part of an organism's adaptive potential, with its own genetic basis and evolutionary dynamics. The approach we outline here provides a powerful tool to quantify the heritability of the environment. As more heritability estimates of the environment accumulate, comparative analysis can provide insights on which types of environments have no, low or high heritability, improving our ability to predict evolutionary responses to environmental change.

4.2 | Animal models can estimate environmental plasticity if the right model structure is fitted

Treating the focal environment as a dependent variable, a phenotypic trait can be fitted as a fixed effect to estimate its impact on the focal environment. This could be expanded to any type of regression that allows estimating the reaction norm of the focal environment, that is as a function of the phenotypic trait. The advantage of using animal models in this context is that they take into account the non-independence caused by relatedness among individuals.

In the last decade, research into phenotype-responsive choice of the environment in the form of the so-called matching habitat choice (Edelaar et al., 2008) has gained relevance and consolidation (Camacho et al., 2020; Lowe & Addis, 2019), but there are still many open questions. The animal model may be a valuable addition to the toolbox for further study. In contrast, the effect of variation in individual phenotypes on adjustment of the environment (e.g. niche construction) appears to have been virtually ignored in the scientific literature (Edelaar & Bolnick, 2019). The animal model structure we propose might be able to shed some light on this too. As is the case for phenotypic plasticity (Pigliucci, 2005), genetic variation in environmental plasticity is necessary for it to evolve. Thus, estimating genetic variation in environmental plasticity, or in other words G(ene) by P(henotype) interaction is another interesting avenue for future research.

4.3 | Animal models detect increased genetic variation via pleiotropy/indirect genetic effects

We found that the estimated genetic variance of the phenotypic trait is larger when the phenotypic trait itself is affected by a focal environment that harbours genetic variance (Figure 2, scenarios 7 and 8). Here, the genes underpinning the chosen or adjusted focal environment have an indirect pleiotropic effect on the plastic phenotypic trait (Figure 1b). The estimates of the genetic variance correspond to the sum of the simulated direct genetic effects (the direct genetic basis for the phenotypic trait, $\sigma^2_{a_z}$) and indirect genetic effects (the direct genetic basis for the focal environment, $\sigma^2_{a_x}$, proportional to the square of the strength of phenotypic plasticity, β^2_{xz}). This form of pleiotropy is sometimes called environmental pleiotropy (Paaby & Rockman, 2013; Saltz, 2019). If desired, these indirect effects can be controlled for and filtered out by fitting the focal environment as a fixed effect in the animal model (Figure 2, scenarios 7 and 8, model 2). Doing so allows estimating the direct genetic variance used in our simulations. Note that for real populations, the estimated genetic variance for a trait after controlling for a focal environmental variable would still be the sum of its additive direct genetic variance plus the indirect effects of other unmeasured environmental (and phenotypic) variables affecting the focal phenotypic trait. Therefore, the value and interpretation of the direct estimated additive direct genetic variance is contingent on the model structure. All of the above applies symmetrically to instances where a focal phenotypic trait with a genetic basis affects the focal environment (scenarios 11 and 12).

A similar issue has been addressed before, when focusing on the covariance between two different phenotypic traits (de Villemereuil et al., 2018), showing how the inclusion of a phenotypic trait as a covariate 'explains away' some of the additive genetic variance. Generalizing this to a focal environment as a second trait, empiricists should be aware that the environment is not always independent of the organism (i.e. can be an extended phenotype), and that including it in the model changes the interpretation of what is estimated.

4.4 | Consequences of fitting the wrong causal model

When including a focal environment with a genetic basis as a covariate in the animal model with the intention to disentangle direct and indirect genetic effects, the model structure assumes that the phenotype–environment covariance is due to a causal effect of the environment on the phenotypic trait (i.e. phenotypic plasticity).

However, this is problematic when in reality it is the focal phenotypic trait that influences the focal environment (i.e. there is environmental plasticity). In this case, the 'classical' structure of the animal model does not reflect the real causal structure of the biological system. When fitting this model, a 'false' phenotypic plasticity is estimated (Figure 3) as a function of:

$$\mathrm{E}\left[\hat{\beta}_{xz}\right] = \beta_{zx}\frac{\sigma_z^2}{\sigma_x^2 + \beta_{zx}^2\sigma_z^2} = \beta_{zx}\rho. \quad (10)$$

where β_{zx} is the simulated environmental plasticity, σ_z^2 is the variance of the phenotypic trait, σ_x^2 the variance of the focal environment and $\beta_{zx}^2\sigma_z^2$ is the indirect variance caused by environmental plasticity. The false estimate of the phenotypic plasticity is thus a function of the environmental plasticity β_{zx} and the coefficient ρ, representing the ratio between the phenotypic variance and the total variance in the environment.

Consequently, assuming the wrong causal structure may also result in the additive genetic variance to be misestimated (see Supporting Information for more details). In this scenario, the estimated additive genetic variance of the trait is a function of:

$$\mathrm{E}\left[\hat{V}_{a_z}\right] \approx \sigma_{a_z}^2 + \beta_{zx}^2\left[\sigma_{a_x}^2\rho^2 + \sigma_{a_z}^2\left(\beta_{zx}^2\rho^2 - 2\rho\right)\right]. \quad (11)$$

The second term of the right-hand side of this equation captures the bias caused by fitting the wrong causal structure. The estimated value is dependent on the additive direct genetic variance of both the phenotype and the environment, $\sigma_{a_z}^2$ and $\sigma_{a_x}^2$, the magnitude of environmental plasticity, β_{zx}, and the ratio between the phenotypic variance and the total variance in the environment, ρ (Figure 4, scenario 12). When there is no environmental plasticity ($\beta_{zx} = 0$), the estimated additive genetic variance of the phenotype is equal to the additive genetic variance of the focal trait (i.e. $\mathrm{E}\left[\hat{V}_{a_z}\right] = \sigma_{a_z}^2$). However, if environmental plasticity is not 0, even if there is no additive genetic variance for the phenotypic trait ($\sigma_{a_z}^2 = 0$), a 'phantom' additive genetic variance can be estimated when there is additive genetic variance for the focal environment ($\sigma_{a_x}^2 \neq 0$; Figure 4, scenarios 10 and 11). Furthermore, even in a scenario where there is no additive genetic variance in the environment ($\sigma_{a_x}^2 = 0$), if there is indeed some genetic variance in the phenotype ($\sigma_{a_z}^2 \neq 0$), the additive genetic variance in the phenotype will be misestimated as a function of the phenotypic and environmental variance ratio, and the strength of environmental plasticity.

These results highlight the importance of correctly identifying the causal relationship between local environments and phenotypic traits before fitting an environmental covariate in an animal model. Researchers are sometimes tempted to add many environmental variables hypothesized to be affecting the phenotypic trait, because they are believed to cause an overestimation of the genetic variance when related individuals share the same environments (Wilson et al., 2010), or to estimate conditional heritabilities (de Villemereuil et al., 2018). However, this could lead to a wrong interpretation of the results if the animal model does not reflect the actual causal structure underpinning phenotypic expression. Previous studies including environmental covariates have almost invariably not taken alternative biological scenarios into consideration. This could have resulted in wrong inferences about the genetic architecture and evolutionary potential of phenotypic traits. Following up the example in the introduction, if boldness is affecting nest site selection (and not the other way around), then treating boldness as a response variable and nest distance to the closest road as a covariate would result in detecting a false influence of the nest distance on boldness (i.e. inferring a false phenotypic plasticity), likely biasing the estimates of genetic variance for boldness. Importantly, these problems are not unique to phenotype–environment relationships, but to any wrongly inferred casual phenotype–phenotype relationship fitted in an animal model.

4.5 | Role of bivariate animal models

Bivariate animal models allow estimating the total genetic variance of both the phenotypic trait and the focal environment. Moreover, they can estimate the genetic covariance $\mathrm{Cov}\left[a_z, a_x\right]$ that can arise when the focal environment has a genetic basis and affects the phenotypic trait via phenotypic plasticity, or when the phenotypic trait has a genetic basis and affects the local environment via environmental plasticity (Figure 5). From the estimates of the variance–covariance matrix, it is possible to obtain an estimate of the strength of phenotypic and environmental plasticity as:

$$\hat{\beta}_{xz} = \frac{\widehat{\mathrm{Cov}}\left[a_z, a_x\right] + \widehat{\mathrm{Cov}}\left[e_z, e_x\right]}{\hat{V}_{a_x} + \hat{V}_{e_x}}. \quad (12)$$

$$\hat{\beta}_{zx} = \frac{\widehat{\mathrm{Cov}}\left[a_z, a_x\right] + \widehat{\mathrm{Cov}}\left[e_z, e_x\right]}{\hat{V}_{a_z} + \hat{V}_{e_z}}. \quad (13)$$

However, it would be necessary to know which type of plasticity (genetic or environmental) is acting in the studied population. Furthermore, the sign of the genetic and residual covariance should be the same if there is only one process underpinning the relationships between trait and environment.

It is also possible to calculate the direct genetic variance of the phenotypic trait after controlling for the indirect effects of the environmental variable and vice versa, for instance, in our simulation:

$$\sigma_{a_z}^2 = \mathrm{E}\left[\hat{V}_{a_z} - \frac{\widehat{\mathrm{Cov}}\left[a_z, a_x\right]}{\hat{V}_{a_x}}\right] = \mathrm{E}\left[\hat{V}_{a_z} - \hat{\beta}_{xz}^2\hat{V}_{a_x}\right]. \quad (14)$$

$$\sigma_{a_x}^2 = \mathrm{E}\left[\hat{V}_{a_x} - \frac{\widehat{\mathrm{Cov}}\left[a_z, a_x\right]}{\hat{V}_{a_z}}\right] = \mathrm{E}\left[\hat{V}_{a_x} - \hat{\beta}_{zx}^2\hat{V}_{a_z}\right]. \quad (15)$$

However, these calculations will be correct only if we know whether environmental or phenotypic plasticity is causing the genetic covariance between the phenotypic trait and focal environment, and thus depend on knowing the correct causal structure.

Thus, bivariate animal models with the focal phenotypic trait and the focal environment as response variables could be applied as a first step to disentangle alternative biological scenarios. First, it could indicate to what extent the phenotypic trait and the focal environment have a genetic variance. Second, if a genetic covariance is detected, there could be some type of plasticity influencing the total genetic variance of one of those traits. Biological insight or additional (experimental) datasets might then help to clarify the causal relationships, that is whether there is phenotypic plasticity, environmental plasticity or other types of non-random assortment of genotypes in their environment (see below).

4.6 | Additional sources of genotype–environment covariance

Genotype–environment covariance may arise due to biological processes other than plasticity. One is divergent natural selection. When phenotypic traits (genotypes) are divergently selected across local environments, a gene–environment correlation is generated and a genetic covariance for the phenotypic and environmental traits is detected by the animal model. Relatives are more similar phenotypically and therefore more likely to occur (survive and reproduce) in the same habitats. Thus, a genetic variance for the environment could also be detected. Therefore, the animal model could detect non-existing phenotypic or environmental plasticity, whereas in reality the covariate does not influence the *development* of the focal trait, but its continued *presence* (via differential survival or reproduction).

Other scenarios where a genetic covariance could emerge are those where plastic habitat choice is present. That occurs when a preference for an environment is induced by an environmental cue during ontogeny, that is imprinting or learned habitat choice (Akcali & Porter, 2017; sometimes also called social learning, e.g. Lillie et al., 2018). For example, in some species, individuals choose environments similar to those they experienced during their natal stage (the period between birth and independence from the parent). Therefore, they choose an environment similar to the one their parents chose (e.g. Nielsen et al., 2013). This causes parents and offspring to share local environments, which again could generate a gene–environment correlation. Something similar happens when there is a degree of philopatry, that is dispersal does not randomize relatives across environments (Ducros et al., 2020; Gervais et al., 2022).

Finally, genetic covariance between the phenotypic trait and the environment will also emerge when there are loci that directly affect both traits (i.e. there is direct pleiotropy).

4.7 | Other limitations

We acknowledge here that we only addressed simplified scenarios compared to those in real biological systems. First, the phenotypic trait and the focal environment are assumed to follow a Gaussian distribution and plasticity is linear. The effect of non-Gaussian traits and/or non-linear reaction norms on the estimation and potential misinterpretation of genetic parameters can be more complex and may need a specific treatment (see de Villemereuil et al., 2018; Morrissey, 2015). Second, for all models, we assumed that the expression of the phenotypic trait and the focal environment (extended phenotype) are independent of the frequency and density of other phenotypes, which are roughly constant across our simulations anyway. Finally, we just presented bivariate scenarios where only two traits are involved. Increasing the number of traits and environments would greatly increase the number of potential causal structures. Moreover, these causal structures can even change across different environments (e.g. Tonsor & Scheiner, 2007). Nevertheless, this paper provides some general conclusions that can be the starting point for further studies.

4.8 | Advice to empiricists: Estimating total, direct, indirect or uninterpretable additive genetic variance

Researchers should decide what kind of additive genetic variance they want to estimate (total, direct and/or indirect) before performing any kind of analysis, to avoid misinterpretation of the heritability of the focal trait. First, the total additive genetic variance for the focal trait can be estimated by not including any covariate with a genetic basis in the model or by performing a bivariate animal model. Second, the direct additive genetic variance of a focal trait, that is what remains after removing the indirect effects of a covarying trait (an extended phenotype, or any other correlated phenotypic trait for that matter), can be estimated by fitting the covariate as a fixed effect in the animal model or by performing a bivariate animal model (Equations 14 and 15). In the same way, by subtracting the estimate of direct additive genetic variance from the estimated total additive genetic variance, the indirect genetic effects of the covariate can be estimated.

However, relevant caution is needed in two additional steps prior to analysis. First, when fitting a focal environment as a fixed effect in the animal model. Both total or direct (total minus indirect genetic effects) genetic variance could be estimated depending on whether the environment has a genetic basis or not (e.g. when applying model 2 for scenarios 4 and 7, respectively). Thus, it is necessary to know whether the focal environment is heritable. Not being aware of this could lead to part of the genetic variance being wrongly interpreted to reflect non-heritable environmental effects (Gervais et al., 2022). Second, when choosing which causal structure is fitted when making inferences about the process underpinning the covariance between phenotypes and environments (or two phenotypic traits). Fitting the wrong

structure results in a misinterpretation or even miscalculation of the genetic parameters and which type of plasticity is acting (e.g. when applying model 2 to scenarios 10–12). Biological insight or additional measures and experiments might be needed. If a phenotype–environment correlation is detected in a study population, we should first rule out the possibility that the environmental variable may be (partially) genetically inherited (or non-genetically inherited for that matter). If that is not possible, we should proceed to test experimentally the existence of phenotypic or environmental plasticity by manipulating the environment or the phenotypic trait (see Camacho et al., 2020 for an example of testing for environmental plasticity). If we are uncertain about the correct causal structure, we should proceed with great caution when making inferences about the obtained estimates of genetic variance when fitting a heritable variable as fixed effect, and at the very least the assumptions and possible consequences underlying a chosen model structure should be discussed. If we ignore the potential mechanisms underpinning gene–environment correlations, we could arrive at misleading conclusions about the adaptive potential and expected evolutionary dynamics of the phenotypes of our study organisms and, therefore, on their ability to cope with environmental change.

AUTHOR CONTRIBUTIONS

Pim Edelaar conceived the idea; Yimen G. Araya-Ajoy and Pim Edelaar designed the general methodology; Yimen G. Araya-Ajoy designed the simulations; Gabriel Munar-Delgado implemented the simulations and data analysis, Gabriel Munar-Delgado and Yimen G. Araya-Ajoy derived mathematical approximations of the results; all authors interpreted the results; Gabriel Munar-Delgado led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We thank Anne Charmantier, Niels Dingemanse, Erik Postma and Jon Wright for discussion, and Michael Morrissey, Sam Scheiner and an anonymous reviewer for comments that improved the paper. This research was supported by the Spanish Ministry of Universities (FPU17/06268 to G.M.-D.), the Spanish Ministry of Science and Innovation (CGL2016-79483-P and PID2019-108971GB-I00 to P.E.), the Swedish Collegium for Advanced Study (P.E.) and the Norwegian Research Council (PID 302619 and 325826 to Y.G.A.-A.).

CONFLICT OF INTEREST STATEMENT

The authors are not aware of any conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14098.

DATA AVAILABILITY STATEMENT

Code to simulate the data is available from Zenodo repository at https://doi.org/10.5281/zenodo.7728542 (Munar-Delgado et al., 2023).

ORCID

Gabriel Munar-Delgado https://orcid.org/0000-0002-6481-4754
Pim Edelaar https://orcid.org/0000-0003-4649-6366

REFERENCES

Akcali, C. K., & Porter, C. K. (2017). Comment on Van Belleghem et al. 2016: Habitat choice mechanisms in speciation and other forms of diversification. *Evolution*, *71*(11), 2754–2761. https://doi.org/10.1111/EVO.13375

Boyle, J., & Start, D. (2020). Plasticity and habitat choice match colour to function in an ambush bug. *Functional Ecology*, *34*(4), 822–829. https://doi.org/10.1111/1365-2435.13528

Butler, D. (2020). Asreml: Fits the linear mixed model. R package version 4.1.0.130. www.vsni.co.uk

Camacho, C., Sanabria-Fernández, A., Baños-Villalba, A., & Edelaar, P. (2020). Experimental evidence that matching habitat choice drives local adaptation in a wild population. *Proceedings of the Royal Society B*, *287*(1927), 20200721. https://doi.org/10.1098/rspb.2020.0721

Charmantier, A., Garant, D., & Kruuk, L. E. (Eds.). (2014). *Quantitative genetics in the wild*. OUP Oxford.

Chevin, L. M., Lande, R., & Mace, G. M. (2010). Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. *PLoS Biology*, *8*(4), e1000357. https://doi.org/10.1371/JOURNAL.PBIO.1000357

Dawkins, R. (1982). *The extended phenotype*. Oxford University Press.

de Villemereuil, P., Morrissey, M. B., Nakagawa, S., & Schielzeth, H. (2018). Fixed-effect variance and the estimation of repeatabilities and heritabilities: Issues and solutions. *Journal of Evolutionary Biology*, *31*(4), 621–632. https://doi.org/10.1111/jeb.13232

Ducros, D., Morellet, N., Patin, R., Atmeh, K., Debeffe, L., Cargnelutti, B., Chaval, Y., Lourtet, B., Coulon, A., & Hewison, A. M. (2020). Beyond dispersal versus philopatry? Alternative behavioural tactics of juvenile roe deer in a heterogeneous landscape. *Oikos*, *129*(1), 81–92. https://doi.org/10.1111/oik.06793

Edelaar, P., & Bolnick, D. I. (2019). Appreciating the multiple processes increasing individual or population fitness. *Trends in Ecology & Evolution*, *34*(5), 435–446. https://doi.org/10.1016/j.tree.2019.02.001

Edelaar, P., Siepielski, A. M., & Clobert, J. (2008). Matching habitat choice causes directed gene flow: A neglected dimension in evolution and ecology. *Evolution*, *62*(10), 2462–2472. https://doi.org/10.1111/j.1558-5646.2008.00459.x

Falconer, D. S. (1960). *Introduction to quantitative genetics*. Oliver and Boyd.

Falconer, D. S., & Mackay, T. F. (1996). *Introduction to quantitative genetics* (4th ed.). Longman.

Gaither, M. R., Gkafas, G. A., de Jong, M., Sarigol, F., Neat, F., Regnier, T., Moore, D., Gröcke, D. R., Hall, N., Liu, X., Kenny, J., Lucaci, A., Hughes, M., Haldenby, S., & Hoelzel, A. R. (2018). Genomics of habitat choice and adaptive evolution in a deep-sea fish. *Nature Ecology & Evolution*, *2*(4), 680–687. https://doi.org/10.1038/s41559-018-0482-x

Gervais, L., Hewison, A. J. M., Morellet, N., Bernard, M., Merlet, J., Cargnelutti, B., Chaval, Y., Pujol, B., & Quéméré, E. (2020). Pedigree-free quantitative genetic approach provides evidence for heritability of movement tactics in wild roe deer. *Journal of Evolutionary Biology*, *33*(5), 595–607. https://doi.org/10.1111/JEB.13594

Gervais, L., Morellet, N., David, I., Hewison, A. J. M., Réale, D., Goulard, M., Chaval, Y., Lourtet, B., Cargnelutti, B., Merlet, J., Quéméré, E., & Pujol, B. (2022). Quantifying heritability and estimating evolutionary potential in the wild when individuals that share genes also share environments. *Journal of Animal Ecology*, *91*, 1239–1250. https://doi.org/10.1111/1365-2656.13677

Holtmann, B., Santos, E. S., Lara, C. E., & Nakagawa, S. (2017). Personality-matching habitat choice, rather than behavioural plasticity, is a likely driver of a phenotype–environment covariance. *Proceedings*

of the Royal Society B: Biological Sciences, *284*(1864), 20170943. https://doi.org/10.1098/rspb.2017.0943

Jaenike, J., & Holt, R. D. (1991). Genetic variation for habitat preference: Evidence and explanations. *The American Naturalist*, *137*, S67–S90. https://doi.org/10.1086/285140

Järvinen, P., Kluen, E., & Brommer, J. E. (2017). Low heritability of nest construction in a wild bird. *Biology Letters*, *13*(10), 20170246. https://doi.org/10.1098/rsbl.2017.0246

Kruuk, L. E. (2004). Estimating genetic parameters in natural populations using the "animal model". *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1446), 873–890. https://doi.org/10.1098/rstb.2003.1437

Kruuk, L. E. B., Slate, J., & Wilson, A. J. (2008). New answers for old questions: The evolutionary quantitative genetics of wild animal populations. *Annual Review of Ecology, Evolution, and Systematics*, *39*, 525–548. https://doi.org/10.1146/ANNUREV.ECOLSYS.39.110707.173542

Lillie, K. M., Gese, E. M., Atwood, T. C., & Sonsthagen, S. A. (2018). Development of on-shore behavior among polar bears (*Ursus maritimus*) in the southern Beaufort Sea: Inherited or learned? *Ecology and Evolution*, *8*(16), 7790–7799. https://doi.org/10.1002/ece3.4233

Lowe, W. H., & Addis, B. R. (2019). Matching habitat choice and plasticity contribute to phenotype–environment covariation in a stream salamander. *Ecology*, *100*(5), e02661. https://doi.org/10.1002/ecy.2661

Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits*. Sinauer Associates.

Møller, A. P. (2006). Rapid change in nest size of a bird related to change in a secondary sexual character. *Behavioral Ecology*, *17*(1), 108–116. https://doi.org/10.1093/beheco/arj003

Morrissey, M. B. (2014). Selection and evolution of causally covarying traits. *Evolution*, *68*(6), 1748–1761. https://doi.org/10.1111/evo.12385

Morrissey, M. B. (2015). Evolutionary quantitative genetics of nonlinear developmental systems. *Evolution*, *69*(8), 2050–2066. https://doi.org/10.1111/evo.12728

Munar-Delgado, G., Edelaar, P., & Araya-Ajoy, Y. (2023). Estimation of additive genetic variance when there are gene-environment correlations: Pitfalls, solutions and unexplored questions. In Methods in Ecology and Evolution (version 1). *Zenodo*. https://doi.org/10.5281/zenodo.7728542

Nielsen, S. E., Shafer, A. B. A., Boyce, M. S., & Stenhouse, G. B. (2013). Does learning or instinct shape habitat selection? *PLoS ONE*, *8*(1), e53721. https://doi.org/10.1371/JOURNAL.PONE.0053721

Paaby, A. B., & Rockman, M. V. (2013). The many faces of pleiotropy. *Trends in Genetics*, *29*(2), 66–73. https://doi.org/10.1016/j.tig.2012.10.010

Pigliucci, M. (2005). Evolution of phenotypic plasticity: Where are we going now? *Trends in Ecology & Evolution*, *20*(9), 481–486. https://doi.org/10.1016/j.tree.2005.06.001

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science*, *11*(1), 3–23. https://doi.org/10.1177/1745691615617439

Postma, E., & Charmantier, A. (2007). What "animal models" can and cannot tell ornithologists about the genetics of wild populations. *Journal of Ornithology*, *148*(2), 633–642. https://doi.org/10.1007/s10336-007-0191-8

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Saltz, J. B. (2019). Gene–environment correlation in humans: Lessons from psychology for quantitative genetics. *Journal of Heredity*, *110*(4), 455–466. https://doi.org/10.1093/jhered/esz027

Tonsor, S. J., & Scheiner, S. M. (2007). Plastic trait integration across a CO_2 gradient in Arabidopsis thaliana. *The American Naturalist*, *169*(5), E119–E140. https://doi.org/10.1086/513493

Weber, J. N., Peterson, B. K., & Hoekstra, H. E. (2013). Discrete genetic modules are responsible for complex burrow evolution in *Peromyscus* mice. *Nature*, *493*(7432), 402–405. https://doi.org/10.1038/nature11816

Westneat, D. F., Araya-Ajoy, Y. G., Allegue, H., Class, B., Dingemanse, N., Dochtermann, N. A., Garamszegi, L. Z., Martin, J. G. A., Nakagawa, S., Réale, D., & Schielzeth, H. (2020). Collision between biological process and statistical analysis revealed by mean centring. *Journal of Animal Ecology*, *89*(12), 2813–2824. https://doi.org/10.1111/1365-2656.13360

Wilson, A. J., Réale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E. B., & Nussey, D. H. (2010). An ecologist's guide to the animal model. *Journal of Animal Ecology*, *79*(1), 13–26. https://doi.org/10.1111/J.1365-2656.2009.01639.X

SUPPORTING INFORMATION

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

Supplementary Material I. About considering the extended phenotype as a focal trait.

Supplementary Material II. Supplementary Results.

Table S1. Expected values for the estimates of the additive genetic variance ($E\left[\hat{V}_{a_z}\right]$ and $E\left[\hat{V}_{a_x}\right]$) for each model-scenario combination.

Table S2. Difference between the mean estimated and simulated value of the direct additive genetic variance of the focal phenotypic trait and phenotypic plasticity with models 1 and 2 for scenarios 1–4.

Table S3. Difference between the mean estimated and simulated value of the direct additive genetic variance of the focal environment and environmental plasticity with models 3 and 4 for scenarios 5, 9 and 10.

Table S4. Difference between the mean estimated and simulated direct value for the additive genetic variance of the phenotypic trait and phenotypic plasticity with models 1 and 2 for scenario 5, 6, 7 and 8.

Table S5. Difference between the mean estimated and simulated direct value for the additive genetic variance of the phenotypic trait and phenotypic plasticity with models 1 and 2 for scenario 9, 10, 11 and 12.

Figure S1. Additive genetic variance estimates for the phenotypic trait for models 1 (phenotypic trait as dependent variable) and 2 (phenotypic trait as the dependent variable and focal environment as covariate) applied to the simulated data for scenarios 1 to 4.

Figure S2. Additive genetic variance estimates for the environment for models 3 (focal environment as dependent variable) and 4 (focal environment as the dependent variable and phenotypic trait as a covariate) applied to the simulated data for scenarios 3, 6, 11 and 12.

Figure S3. Additive genetic variance estimates for the environment for models 3 (focal environment as dependent variable) and 4 (focal environment as the dependent variable and phenotypic trait as a covariate) applied to the simulated data for scenarios 2, 4, 7 and 8.

Figure S4. Distribution of the estimated values of additive genetic variance of the phenotypic trait for model 1 (phenotypic trait as dependent variable) versus model 5 (phenotypic trait and focal environment bivariate model).

Figure S5. Additive genetic variance estimates for the focal environment for models 3 (focal environment as the dependent

variable) and 4 (phenotypic trait as a covariate) applied to the simulated data for scenarios 2, 4, 7 and 8.

Figure S6. Distribution of the estimated values of additive genetic variance of the phenotypic trait for models 1 (phenotypic trait as dependent variable) versus model 5 (phenotypic trait and focal environment bivariate model).

Supplementary Material IV. Supplementary equations and mathematical approximations.

Supplementary Material V. References.

How to cite this article: Munar-Delgado, G., Araya-Ajoy, Y. G., & Edelaar, P. (2023). Estimation of additive genetic variance when there are gene–environment correlations: Pitfalls, solutions and unexplored questions. *Methods in Ecology and Evolution*, *14*, 1245–1258. https://doi.org/10.1111/2041-210X.14098