

# General methods for evolutionary quantitative genetic inference from generalised mixed models

Pierre de Villemereuil<sup>\*</sup>, Holger Schielzeth<sup>†,‡</sup>, Shinichi Nakagawa<sup>§</sup>, and Michael Morrissey<sup>\*\*</sup>

<sup>\*</sup>Laboratoire d'Écologie Alpine, CNRS UMR 5553  
Université Joseph Fourier, BP53, 38041 Grenoble, Cedex 9, France  
`bonamy@horus.ens.fr`

<sup>†</sup>Department of Evolutionary Biology  
Bielefeld University, Bielefeld, Germany

<sup>‡</sup>Institute of Ecology  
Dornburger Str. 159, 07743 Jena  
`holger.schielzeth@uni-jena.de`

<sup>§</sup>Evolution and Ecology Research Centre  
University of New South Wales, Sydney, NSW 2052, Australia  
`s.nakagawa@unsw.edu.au`

<sup>\*\*</sup>School of Evolutionary Biology  
University of St Andrews, St Andrews, UK, KY16 9TH  
`michael.morrissey@st-andrews.ac.uk`

May 18, 2016

**Keywords:** quantitative genetics, generalised linear model, statistics, theory, evolution, additive genetic variance, G matrix

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

**Abstract**

Methods for inference and interpretation of evolutionary quantitative genetic parameters, and for prediction of the response to selection, are best developed for traits with normal distributions. Many traits of evolutionary interest, including many life history and behavioural traits, have inherently non-normal distributions. The generalised linear mixed model (GLMM) framework has become a widely used tool for estimating quantitative genetic parameters for non-normal traits. However, whereas GLMMs provide inference on a statistically-convenient latent scale, it is sometimes desirable to express quantitative genetic parameters on the scale upon which traits are expressed. The parameters of a fitted GLMMs, despite being on a latent scale, fully determine all quantities of potential interest on the scale on which traits are expressed. We provide expressions for deriving each of such quantities, including population means, phenotypic (co)variances, variance components including additive genetic (co)variances, and parameters such as heritability. We demonstrate that fixed effects have a strong impact on those parameters and show how to deal for this effect by averaging or integrating over fixed effects. The expressions require integration of quantities determined by the link function, over distributions of latent values. In general cases, the required integrals must be solved numerically, but efficient methods are available and we provide an implementation in an R package, QGGLMM. We show that known formulae for quantities such as heritability of traits with Binomial and Poisson distributions are special cases of our expressions. Additionally, we show how fitted GLMM can be incorporated into existing methods for predicting evolutionary trajectories. We demonstrate the accuracy of the resulting method for evolutionary prediction by simulation, and apply our approach to data from a wild pedigreed vertebrate population.

## 26 Introduction

27 Additive genetic variances and covariances of phenotypic traits determine the response to se-  
28 lection, and so are key determinants of the processes of adaptation in response to natural  
29 selection and of genetic improvement in response to artificial selection (Fisher, 1918; Falconer,  
30 1960; Lynch and Walsh, 1998; Walsh and Lynch, forthcoming). While the concept of additive  
31 genetic variance (Fisher, 1918; Falconer, 1960) is very general, being applicable to any type of  
32 character with any arbitrary distribution, including, for example, fitness (Fisher, 1930), tech-  
33 niques for estimating additive genetic variances and covariances are best developed for Gaussian  
34 traits (i.e., traits that follow a normal distribution; Henderson 1950; Lynch and Walsh 1998).  
35 Furthermore, quantitative genetic theory for predicting responses to selection are also best  
36 developed and established for Gaussian characters (Walsh and Lynch, forthcoming), but see  
37 Morrissey (2015). Consequently, although many characters of potential evolutionary interest  
38 are not Gaussian (e.g. survival or number of offspring), they are not well-handled by existing  
39 theory and methods. Comprehensive systems for estimating genetic parameters and predict-  
40 ing evolutionary trajectories of non-Gaussian traits will hence be very useful for quantitative  
41 genetic studies of adaptation.

42 For the analysis of Gaussian traits, linear mixed model-based (LMM) inferences of genetic  
43 parameters, using the ‘animal model’, have become common practice in animal and plant  
44 breeding (Thompson, 2008; Hill and Kirkpatrick, 2010), but also in evolutionary studies on  
45 wild populations (Kruuk, 2004; Wilson *et al.*, 2010). Recently, the use of generalised linear  
46 mixed models (GLMMs) to analyse non-Gaussian traits has been increasing (e.g. Milot *et al.*,  
47 2011; Wilson *et al.*, 2011; Morrissey *et al.*, 2012; de Villemereuil *et al.*, 2013; Ayers *et al.*,  
48 2013). Whereas LMM analysis directly estimates additive genetic parameters on the scale on  
49 which traits are expressed and selected, and upon which we may most naturally consider their  
50 evolution, this is not the case for GLMMs. In this paper, we offer a comprehensive description  
51 of the assumptions of GLMMs and their consequences in terms of quantitative genetics and a  
52 framework to infer quantitative genetic parameters from GLMMs output. This work applies and  
53 extends theory in Morrissey (2015), to handle the effects of (non-linear) relationships among the  
54 scale upon which inference is conducted in a GLMM and the scale of data, and to accommodate

55 the error structures that arise in GLMM analysis. These results generalise existing expressions  
56 for specific models (threshold model and Poisson with a log-link, Dempster and Lerner, 1950;  
57 Robertson, 1950; Foulley and Im, 1993). We show that fixed effects in GLMMs raise special  
58 complications and we offer some efficient approaches for dealing with this issue.

59 While it will undoubtedly be desirable to develop a comprehensive method for making data-  
60 scale inferences of quantitative genetic parameters with GLMMs, such an endeavour will not  
61 yield a system for predicting evolution in response to natural or artificial selection, even if a  
62 particular empirical system is very well served by the assumptions of a GLMM. This is because  
63 systems for evolutionary prediction, specifically the Breeder's equation (Lush, 1937; Fisher,  
64 1924) and the Lande equation (Lande, 1979; Lande and Arnold, 1983), assume that breeding  
65 values (and in most applications, phenotypes) are multivariate normal or make assumptions  
66 such as linearity of the parent-offspring regression, which are unlikely to hold for non-normal  
67 traits (Walsh and Lynch, forthcoming). Even if it is possible to estimate additive genetic vari-  
68 ances of traits on the scale upon which traits are expressed, we will show that these quantities  
69 will not strictly be usable for evolutionary prediction. However, we will see that the scale on  
70 which estimation is performed in a GLMM does, by definition, satisfy the assumptions of the  
71 Breeder's and Lande equations. Thus, for the purpose of predicting evolution, it may be useful  
72 to be able to express selection of non-Gaussian traits on this scale. Such an approach will yield  
73 a system for evolutionary prediction of characters that have been modelled with a GLMM,  
74 requiring no more assumptions than those that are already made in applying the statistical  
75 model.

76 The main results in this paper are arranged in four sections. First, we describe the GLMM  
77 framework: its relationship to the more general (Gaussian) LMM and especially to the Gaussian  
78 animal model (Henderson, 1973; Kruuk, 2004; Wilson *et al.*, 2010), how GLMMs can be usefully  
79 viewed as covering three scales and how some special interpretational challenges arise and are  
80 currently dealt with. Second, we propose a system for making inferences of quantitative genetic  
81 parameters on the scale upon which traits are expressed for arbitrary GLMMs. We show how  
82 to estimate genotypic and additive genetic variances and covariances on this scale, accounting  
83 for fixed effects as necessary. We lay out the formal theory underlying the system, apply it to  
84 an empirical dataset. The relationships between existing analytical formulae and our general

85 framework are also highlighted. Third, we illustrate the issues when inferring quantitative  
86 genetic parameters using a GLMM with an empirical example on Soay sheep (*Ovis aries*) and  
87 how our framework can help to overcome them. Fourth, we outline a system of evolutionary  
88 prediction for non-Gaussian traits that capitalises on the fact that the latent scale in a GLMM  
89 satisfies the assumptions of available equations for the prediction of evolution. We show in  
90 a simulation study that (i) evolutionary predictions using additive genetic variances on the  
91 observed data scale represent approximations, and can, in fact, give substantial errors, and  
92 (ii) making inferences via the latent scale provides unbiased predictions, insofar as a GLMM  
93 may provide a pragmatic model of variation in non-Gaussian traits. The framework introduced  
94 here (including both quantitative genetic parameters inference and evolutionary prediction) has  
95 been implemented in a package for the R software (R Core Team, 2015) available at <https://github.com/devillemereuil/qgglm>.

## 97 **The generalised linear mixed model framework**

### 98 **Linear mixed models for Gaussian traits**

99 For Gaussian traits, a linear mixed model allows various analyses of factors that contribute to  
100 the mean and variance of phenotype. In particular, a formulation of a linear mixed model called  
101 the ‘animal model’ provides a very general method for estimating additive genetic variances and  
102 covariances, given arbitrary pedigree data, and potentially accounting for a range of different  
103 types of confounding variables, such as environmental effects, measurement error or maternal  
104 effects. A general statement of an animal model analysis decomposing variation in a trait,  $\mathbf{z}$ ,  
105 into additive genetic and other components would be

$$\mathbf{z} = \mu + \mathbf{X}\mathbf{b} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_1\mathbf{u}_1 + \dots + \mathbf{Z}_k\mathbf{u}_k + \mathbf{e}, \quad (1)$$

106 where  $\mu$  is the model intercept,  $\mathbf{b}$  is a vector of fixed effects such as sex and age, relating  
107 potentially both continuous and categorical effects to observations via the fixed effects design  
108 matrix  $\mathbf{X}$ , just as in an ordinary linear model, and  $\mathbf{e}$  is the vector of normally-distributed  
109 residuals. An arbitrary number of random effects can be modelled, with design matrices  $\mathbf{Z}$ ,

110 where effects ( $\mathbf{a}$ ,  $\mathbf{u}_1 \dots \mathbf{u}_k$ ) are assumed to be drawn from normal distributions with variances  
111 to be estimated. The key feature of the animal model is that it includes individual additive  
112 genetic effects, or breeding values, conventionally denoted  $\mathbf{a}$ . These additive genetic effects  
113 and, critically, their variance, are estimable given relatedness data, which can be derived from  
114 pedigree data, or, more recently, from genomic estimates of relatedness (Sillanpää, 2011). The  
115 covariances of breeding values among individuals can be modelled according to

$$\mathbf{a} \sim N(\mathbf{0}, \mathbf{A}V_A), \quad (2)$$

116 where  $\mathbf{A}$  is the additive genetic relatedness matrix derived from the pedigree and  $V_A$  is the  
117 genetic additive variance.

### 118 **Common issues with non-Gaussian traits**

119 Many non-Gaussian traits, however, cannot be strictly additive on the scale on which they are  
120 expressed. Consider, for example, survival probability that is bounded at 0 and 1 so that effects  
121 like the substitution effect of one allele for another necessarily must be smaller when expressed  
122 in individuals that otherwise have expected values near zero or one. In such a scenario, it may  
123 be reasonable to assume that there exists an underlying scale, related to survival probability,  
124 upon which genetic and other effects are additive.

125 In addition to inherent non-additivity, many non-Gaussian traits will have complex patterns  
126 of variation. Over and above sources of variation that can be modelled with fixed and random  
127 effects, as in a LMM (e.g., using Eqs. 1 and 2), residual variation may include both inherently  
128 stochastic components, and components that correspond to un-modelled systematic differences  
129 among observations. In a LMM, such differences are not distinguished, but contribute to resid-  
130 ual variance. However, for many non-Gaussian traits it may be desirable to treat the former  
131 as arising from some known statistical distribution, such as the binomial or Poisson distribu-  
132 tion, and to deal with additional variation via a latent-scale residual (i.e. an overdispersion  
133 term). Separation of these two kinds of variation in residuals may be very generally useful in  
134 evolutionary quantitative genetic studies.

## 135 The scales of the generalised linear mixed model

136 Generalised linear mixed model (GLMM) analysis can be used for inference of quantitative  
137 genetic parameters, and provides pragmatic ways of dealing with inherent non-additivity and  
138 with complex sources of variation. The GLMM framework can be thought of as consisting of  
139 three distinct scales on which we can think of variation in a trait occurring (see Fig. 1). A *latent*  
140 *scale* is assumed (Fig. 1, top), on which effects on the propensity for expression of some trait  
141 are assumed to be additive. A function, called a ‘link function’ is applied that links expected  
142 values for a trait to the latent scale. For example, a trait that is expressed in counts, say,  
143 number of behaviours expressed in a unit time, is a strictly non-negative quantity. As depicted  
144 in Fig. 1, a strictly positive distribution of expected values may related to latent values ranging  
145 from  $-\infty$  to  $+\infty$  by a function such as the log link. Finally, a distribution function (e.g.  
146 Binomial, Poisson, etc.) is required to model the “noise” of observed values around expected  
147 value (Fig. 1, bottom). Different distributions are suitable for different traits. For example,  
148 with a count trait such as that depicted in Fig. 1, observed values may be modelled using the  
149 Poisson distribution, with expectations related to the latent scale via the log link function.

150 More formally, these three scales of the GLMM can be written:

$$\ell = \mu + \mathbf{X}\mathbf{b} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_1\mathbf{u}_1 + \dots + \mathbf{Z}_k\mathbf{u}_k + \mathbf{o}, \quad (3a)$$

151

$$\boldsymbol{\eta} = g^{-1}(\ell), \quad (3b)$$

152

$$\mathbf{z} \sim \mathcal{D}(\boldsymbol{\eta}, \boldsymbol{\theta}), \quad (3c)$$

153 where Eq. 3a is just as for a LMM (Eq. 1), except that it describes variation on the *latent*  
154 *scale*  $\ell$ , rather than the response directly. Note that we now refer to the “residual” (noted  $\mathbf{e}$  in  
155 Eq. 1) as “overdispersion” (denoted  $\mathbf{o}$ , with a variance denoted  $V_O$ ), since residuals (variation  
156 around expected values) are defined by the distribution function,  $\mathcal{D}$ , in this model. Just as for  
157 the LMM (Eq. 1), all random effects are assumed to follow normal distributions with variances  
158 to be estimated on the latent scale. Particularly, the variance of additive genetic effects  $\mathbf{a}$  is  
159 assumed to follow Eq. 2 on the latent scale.

160 Eq. 3b formalises the idea of the link function. Any link function has an associated inverse

161 link function,  $g^{-1}$ , which is often useful for converting specific latent values to expected values.  
162 The expected values  $\boldsymbol{\eta}$  constitute what we call the *expected data scale*. For example, where  
163 the log link function translates expected values to the latent scale, its inverse, the exponential  
164 function, translates latent values to expected values. Finally, Eq. 3c specifies the distribution  
165 by which the observations  $\mathbf{z}$  scatter around the expected values according to some distribution  
166 function, that may involve parameters (denoted  $\boldsymbol{\theta}$ ) other than the expectation. We call this  
167 the *observed data scale*. Some quantities of interest, such as the mean, are the same on the  
168 expected data scale and on the observed data scale. When parameters are equivalent on these  
169 two scales, we will refer to them together as the *data scales*.

170 The distinction we make between the expected and observed data scales is one of convenience  
171 as they are not different scales *per se*. However, this distinction allows for more biological  
172 subtlety when interpreting the output of a GLMM. The expected data scale can be thought  
173 of as the “intrinsic” value of individuals (shaped by both the genetic and the environment),  
174 but this intrinsic value can only be studied through random realisations. As we will see,  
175 because breeding values are intrinsic individual values, the additive genetic variance is the  
176 same for both scales, but, due to the added noise in observed data, the heritabilities are not.  
177 Upon which scale to calculate heritability depends on the underlying biological question. For  
178 example, individuals (given their juvenile growth and genetic value) might have an intrinsic  
179 annual reproductive success of 3.4, but can only produce a integer number of offspring each  
180 year (say 2, 3, 4 or 5): heritabilities of both intrinsic expectations and observed numbers can  
181 be computed, but their values and interpretations will differ.

## 182 **Current practices and issues to compute genetic quantitative parameters from** 183 **GLMM outputs**

184 Genetic variance components estimated in a generalised animal model are obtained on the  
185 latent scale. Hence, the “conventional” formula to compute heritability:

$$h_{\text{lat}}^2 = \frac{V_{A,\ell}}{V_{A,\ell} + V_{\text{RE}} + V_{\text{O}}}, \quad (4)$$

186 where  $V_{RE}$  is the summed variance of all random effects apart from the additive genetic variance,  
187 and  $V_O$  is the overdispersion variance, is the heritability on the latent scale, not on the observed  
188 data scale (Morrissey *et al.*, 2014). Here, and throughout this paper,  $V_{A,\ell}$  stands for the additive  
189 genetic variance on the latent scale. Although it might sometimes be sensible to measure the  
190 heritability of a trait on the latent scale (for example, in animal breeding, where selection might  
191 be based on latent breeding values), it is natural to seek inferences on the scale upon which the  
192 trait is expressed, and on which we may think of selection as acting. Some expressions exist  
193 by which various parameters can be obtained or approximated on the observed data scale. For  
194 example, various expressions for the intra-class correlation coefficients on the data scale exist  
195 (reviewed in Nakagawa and Schielzeth, 2010), but, contrary to LMM, heritabilities on the data  
196 scales within a GLMM framework cannot be considered as intra-class correlation coefficients.  
197 Exact analytical expressions exist for the additive genetic variance and heritability on the  
198 observed data scale for two specific and important families of GLMMs (i.e. combinations of  
199 link functions and distribution functions): for a binomial model with a probit link function (i.e.,  
200 the “threshold model,” Dempster and Lerner, 1950) and for a Poisson model with a logarithm  
201 link function (Foulley and Im, 1993). A general system for calculating genetic parameters on  
202 the expected and observed data scales for arbitrary GLMMs is currently lacking.

203 In addition to handling the relationship between observed data and the latent trait via  
204 the link and distribution functions, any system for expected and observed scale quantitative  
205 genetic inference with GLMMs will have to account for complex ways in which fixed effects  
206 can influence quantitative genetic parameters. It is currently appreciated that fixed effects  
207 in LMMs explain variance, and that variance associated with fixed effects can have a large  
208 influence on summary statistics such as repeatability (Nakagawa and Schielzeth, 2010) and  
209 heritability (Wilson, 2008). This principle holds for GLMMs as well, but fixed effects cause  
210 additional, important complications for interpreting GLMMs. While random and fixed effects  
211 are independent in a GLMM on the latent scale, the non-linearity of the link function renders  
212 them inter-related on the expected and observed scales. Consequently, and unlike in a LMM  
213 or in a GLMM on the latent scale, variance components on the observed scale in a GLMM  
214 depend on the fixed effects. Consider, for example, a GLMM with a log link function. Because  
215 the exponential is a convex function, the influence of fixed and random effects will create more

216 variance on the expected and observed data scales for larger values than for smaller values.

## 217 Quantitative genetic parameters in GLMMs

218 Although all examples and most equations in this article are presented in a univariate form, all  
219 our results are applicable to multivariate analysis, which is implemented in our software. Unless  
220 stated otherwise, the equations below assume that no fixed effect (apart from the intercept)  
221 were included in the GLMM model.

### 222 Phenotypic mean and variances

223 **Expected population mean** The expected mean phenotype  $\bar{z}$  on the data scale (i.e., applying  
224 to both the mean expected value and mean observed value) is given by

$$\bar{z} = \int g^{-1}(\ell) f_{\mathcal{N}}(\ell, \mu, V_{A,\ell} + V_{RE} + V_O) d\ell, \quad (5)$$

225 where  $f_{\mathcal{N}}(\ell, \mu, V_{A,\ell} + V_{RE} + V_O)$  is the probability density of a Normal distribution with mean  
226  $\mu$  and variance  $V_{A,\ell} + V_{RE} + V_O$  evaluated at  $\ell$ .

227 **Expected-scale phenotypic variance** Phenotypic variance on the expected data scale can be  
228 obtained analogously to the data scale population mean. Having obtained  $\bar{z}$ , the phenotypic  
229 variance is

$$V_{P,\text{exp}} = \int (g^{-1}(\ell) - \bar{z})^2 f_{\mathcal{N}}(\ell, \mu, V_{A,\ell} + V_{RE} + V_O) d\ell. \quad (6)$$

230 **Observed-scale phenotypic variance** Phenotypic variance of observed values is the sum of the  
231 variance in expected values and variance arising from the distribution function. Since these  
232 variances are independent by construction in a GLMM, they can be summed. This distribution  
233 variance is influenced by the latent trait value, but might also depend on additional distribution  
234 parameters included in  $\boldsymbol{\theta}$  (see Eq. 3c). Given a distribution-specific variance function  $v$ :

$$V_{P,\text{obs}} = V_{P,\text{exp}} + \int v(\ell, \boldsymbol{\theta}) f_{\mathcal{N}}(\ell, \mu, V_{A,\ell} + V_{RE} + V_O) d\ell. \quad (7)$$

235 **Genotypic variance on the data scales, arising from additive genetic variance on**  
236 **the latent scale**

237 Because the link function is non-linear, additive genetic variance on the latent scale is manifested  
238 as a combination of additive and non-additive variance on the data scales. Following Falconer  
239 (1960) the total genotypic variance on the data scale is the variance of genotypic values on  
240 that scale. Genotypic values are the expected data scale phenotypes, given latent scale genetic  
241 values. The expected phenotype of an individual with a given latent genetic value  $a$ , i.e., its  
242 genotypic value on the data scales  $E[z|a]$ , is given by

$$E[z|a] = \int g^{-1}(\ell) f_{\mathcal{N}}(\ell, \mu + a, V_{\text{RE}} + V_{\text{O}}) d\ell. \quad (8)$$

243 The total genotypic variances on the expected and observed data scales are the same, since  
244 genotypic values are expectations that do not change between the expected and observed scales.  
245 The total genotypic variance on both the expected and observed data scales is then

$$V(E[z|a]) = \int (E[z|a] - \bar{z})^2 f_{\mathcal{N}}(a, 0, V_{\text{A},\ell}) da. \quad (9)$$

246 This is the total genotypic variance on the data scale, arising from strictly additive genetic  
247 variance on the latent scale. If non-additive genetic effects are modelled on the latent scale,  
248 they would be included in the expectations and integrals in Eqs. 8 and 9.

249 **Additive genetic variance on the data scales**

250 The additive variance on the data scales is the variance of breeding values computed on the  
251 data scales. Following Robertson (1950; see also Fisher 1918), breeding values on the data  
252 scales, i.e.,  $a_{\text{exp}}$  and  $a_{\text{obs}}$ , are the part of the phenotype  $z$  that depends linearly on the latent  
253 breeding values. The breeding values on the data scale can then be defined as the predictions  
254 of a least-squares regression of the observed data on the latent breeding values,

$$a_{\text{obs}} = \hat{z}|a = m + ba, \quad (10)$$

255 where  $\hat{z}$  is the value of  $z$  predicted by the regression,  $a$  the latent breeding value and  $m$  and  $b$   
256 the parameters of the regression. Thus, we have  $V_{A,\text{obs}} = b^2 V_{A,\ell}$  and, from standard regression  
257 theory:

$$b = \frac{\text{cov}(z, a)}{V_{A,\ell}}. \quad (11)$$

258 Because of the independence between the expected values of  $z$  (i.e. the expected data scale  
259  $g^{-1}(\ell)$ ) and the distribution “noise” (see Eq. 7), we can obtain the result that  $\text{cov}(z, a) =$   
260  $\text{cov}(g^{-1}(\ell), a)$ , hence:

$$b = \frac{\text{cov}(g^{-1}(\ell), a)}{V_{A,\ell}}. \quad (12)$$

261 Stein’s (1973) lemma states that if  $X$  and  $Y$  are bivariate normally distributed random variables,  
262 then the covariance of  $Y$  and some function of  $X$ ,  $f(X)$ , is equal to the expected value of  $f'(X)$   
263 times the covariance between  $X$  and  $Y$ , so,

$$\text{cov}(g^{-1}(\ell), a) = E \left[ \frac{dg^{-1}(\ell)}{d\ell} \right] \text{cov}(\ell, a) = E \left[ \frac{dg^{-1}(\ell)}{d\ell} \right] V_{A,\ell}, \quad (13)$$

264 noting that the covariance of latent breeding values and latent values is the variance of latent  
265 breeding values. Finally, combining Eq. 12 with Eq. 13, we obtain:

$$b = E \left[ \frac{dg^{-1}(\ell)}{d\ell} \right]. \quad (14)$$

266 To avoid confusion with various uses of  $b$  as other forms of regression coefficients, and for  
267 consistency with Morrissey (2015), we denote the average derivative of expected value with  
268 respect to latent value as  $\Psi$ :

$$\Psi = E \left[ \frac{dg^{-1}(\ell)}{d\ell} \right] = \int \frac{dg^{-1}(\ell)}{d\ell} f_{\mathcal{N}}(\ell, \mu, V_{A,\ell} + V_{\text{RE}} + V_{\text{O}}) d\ell. \quad (15)$$

269 The additive genetic variance on the expected and observed scales are still the same and are  
270 given by

$$V_{A,\text{obs}} = V_{A,\text{exp}} = \Psi^2 V_{A,\ell}. \quad (16)$$

## 271 Including fixed effects in the inference

272 **General issues** Because of the non-linearity introduced by the link function in a GLMM, all  
273 quantitative genetic parameters are directly influenced by the presence of fixed effects. Hence,  
274 when fixed effects are included in the model, it will often be important to marginalise over them  
275 to compute accurate population parameters. There are different approaches to do so. We will  
276 first describe the simplest approach (i.e. directly based on GLMM assumptions).

277 **Averaging over predicted values** In a GLMM, no assumption is made about the distribution of  
278 covariates in the fixed effects. Given this, we can marginalise over fixed effects by averaging over  
279 predicted values (marginalised over the random effects, i.e.  $\mathbf{X}\hat{\mathbf{b}}$ , where  $\hat{\mathbf{b}}$  are the fixed effects  
280 estimates). Note that, doing so, we implicitly make the assumption that our sample is repre-  
281 sentative of the population of interest. Using this approach, we can compute the population  
282 mean in Eq. 5 as:

$$\bar{z} = \frac{1}{N} \sum_{i=1}^N \int g^{-1}(\ell) f_{\mathcal{N}}(\ell, \mu + \hat{\boldsymbol{\ell}}_i, V_{A,\ell} + V_{RE} + V_O) d\ell, \quad (17)$$

283 where  $N$  is the number of predicted latent values in  $\hat{\boldsymbol{\ell}} = \mathbf{X}\hat{\mathbf{b}}$ . Typically,  $\mathbf{X}$  will be the fixed  
284 effects design matrix used when fitting the generalised animal model (Eqs. 1, 2, and 3), and  
285  $N$  will be the number of data observations. Furthermore, this assumes that all fixed effects  
286 represent biologically relevant variation, rather than being corrections for the observation pro-  
287 cess or experimental condition. From this estimate of  $\bar{z}$ , we can compute the expected-scale  
288 phenotypic variance:

$$V_{P,\text{exp}} = \frac{1}{N} \sum_{i=1}^N \int (g^{-1}(\ell) - \bar{z})^2 f_{\mathcal{N}}(\ell, \mu + \hat{\boldsymbol{\ell}}_i, V_{A,\ell} + V_{RE} + V_O) d\ell. \quad (18)$$

289 Note that we are not averaging over variances computed for each predicted values, since the  
290 value of the mean  $\bar{z}$  is the same across the computation. Eqs. 7, 8, 9 and 15 are to be modified  
291 accordingly to compute all parameters, including  $\Psi$ . This approach has the advantages of being  
292 simple and making a direct use of the GLMM inference without further assumptions.

293 **Sampled covariates are not always representative of the population** The distribution of covariate  
294 values in  $\mathbf{X}$  may not be representative of the population being studied. In such cases, integration  
295 over available values of fixed effects may be inappropriate. For example, a population may be  
296 known (or assumed) to have an equal sex ratio, but one sex may be easier to catch, and  
297 therefore over-represented in any given dataset. In such a situation, incorporation of additional  
298 assumptions or data about the distribution of covariates (e.g., of sex ratio) may be useful.  
299 A first approach is to predict values according to a new set of covariates constructed to be  
300 representative of the population (e.g. with balanced sex ratio). Given these new predicted  
301 values, the above approach can readily be used to compute quantitative genetic parameters  
302 of interest. A drawback of this approach is that it requires one to create a finite sample of  
303 predicted values instead of a full distribution of the covariates. A second approach will require  
304 one to specify such a distribution for fixed covariates, here noted  $f_X(\mathbf{X})$ . In that case, Eq. 17  
305 can be modified as follows

$$\bar{z} = \iint g^{-1}(\ell) f_{\mathcal{N}}(\ell, \mu + \mathbf{X}\hat{\mathbf{b}}, V_{A,\ell} + V_{RE} + V_O) f_X(\mathbf{X}) d\mathbf{X} d\ell. \quad (19)$$

306 All relevant equations (Eqs. 6, 7, 8, 9 and 15) are to be modified accordingly. This approach is  
307 the most general one, but requires the ability to compute  $f_X(\mathbf{X})$ . Note that this distribution  
308 should also account for potential covariance between covariates.

### 309 **Summary statistics and multivariate extensions**

310 Eqs. 5 through 16 give the values of different parameters that are useful for deriving other  
311 evolutionary quantitative genetic parameters on the observed data scale. Hence, from them,  
312 other parameters can be computed. The narrow-sense heritability on the observed data scale  
313 can be written as

$$h_{\text{obs}}^2 = \frac{V_{A,\text{obs}}}{V_{P,\text{obs}}}. \quad (20)$$

314 Replacing  $V_{P,\text{obs}}$  by  $V_{P,\text{exp}}$  will lead to the heritability on the expected data scale  $h_{\text{exp}}^2$ :

$$h_{\text{exp}}^2 = \frac{V_{A,\text{exp}}}{V_{P,\text{exp}}}. \quad (21)$$

315 Recalling that  $V_{A,obs} = V_{A,exp}$ , but  $V_{P,obs} \neq V_{P,exp}$ , note that the two heritabilities above differ.  
316 Parameters such as additive genetic coefficient of variance and evolvability (Houle, 1992) can  
317 be just as easily derived. The coefficient of variation on the expected and observed data scales  
318 are identical and can be computed as

$$CV_{A,obs} = CV_{A,exp} = 100 \frac{\sqrt{V_{A,obs}}}{\bar{z}}, \quad (22)$$

319 and the evolvability on the expected and observed data scales will be

$$I_{A,obs} = I_{A,exp} = \frac{V_{A,obs}}{\bar{z}^2}. \quad (23)$$

320 The multivariate genetic basis of phenotypes, especially as summarised by the  $\mathbf{G}$  matrix,  
321 is also often of interest. For simplicity, all expressions considered to this point have been pre-  
322 sented in univariate form. However, every expression has a fairly simple multivariate extension.  
323 Multivariate phenotypes are typically analysed by multi-response GLMMs. For example, the  
324 vector of mean phenotypes in a multivariate analysis on the expected data scale is obtained by  
325 defining the link function to be a vector-valued function, returning a vector of expected values  
326 from a vector of values on the latent scale. The phenotypic variance is then obtained by inte-  
327 grating the vector-valued link function times the multivariate normal distribution total variance  
328 on the latent scale, as in Eq. 5 and Eq. 7. Integration over fixed effects for calculation of the  
329 multivariate mean is directly analogous to either of the extensions of Eq. 5 given in Eqs. 17  
330 or 19. Calculation of other parameters, such as multivariate genotypic values, additive-derived  
331 covariance matrices, and phenotypic covariance matrices, have directly equivalent multivariate  
332 versions as well. The additive genetic variance-covariance matrix (the  $\mathbf{G}$  matrix) on the ob-  
333 served scale is simply the multivariate extension of Eq. 16, i.e.,  $\mathbf{G}_{obs} = \mathbf{\Psi} \mathbf{G}_\ell \mathbf{\Psi}^T$ . Here,  $\mathbf{G}_\ell$  is the  
334 latent  $\mathbf{G}$  matrix and  $\mathbf{\Psi}$  is the average gradient matrix of the vector-valued link function, which  
335 is a diagonal matrix of  $\Psi$  values for each trait (simultaneously computed from a multivariate  
336 version of Eq. 15).

## 337 Relationships with existing analytical formulae

### 338 Binomial distribution and the threshold model

339 Heritabilities of binary traits have a long history of analysis with a threshold model (Wright,  
340 1934; Dempster and Lerner, 1950; Robertson, 1950), whereby an alternate trait category is  
341 expressed when a trait on a latent “liability scale” exceeds a threshold. Note that this liability  
342 scale is not the same as the latent scale hereby defined for GLMM (see Fig. S1 in Supplementary  
343 Information). However, it can be shown (see Supplementary Information, section A) that  
344 a GLMM with a binomial distribution and a probit link function is exactly equivalent to  
345 such a model, only with slightly differently defined scales. For threshold models, heritability  
346 can be computed on this liability scale by using adding a so-called “link variance”  $V_L$  to the  
347 denominator (see for example Nakagawa and Schielzeth, 2010; de Villemereuil *et al.*, 2013):

$$h_{\text{liab}}^2 = \frac{V_{A,\ell}}{V_{A,\ell} + V_{\text{RE}} + V_{\text{O}} + V_L}. \quad (24)$$

348 Because the probit link function is the inverse of the cumulative standard normal distribution  
349 function, the “link variance”  $V_L$  is equal to one in this case. One can think of the “link variance”  
350 as arising in this computation because of the reduction from three scales (in case of a GLMM)  
351 to two scales (liability and observed data in the case of a threshold model): the liability scale  
352 includes the link function.

353 When the heritability is computed using the threshold model, Dempster and Lerner (1950)  
354 and Robertson (1950) derived an exact analytical formula relating this estimate to the observed  
355 data scale:

$$h_{\text{obs}}^2 = \frac{t^2}{p(1-p)} h_{\text{liab}}^2, \quad (25)$$

356 where  $p$  is the probability of occurrence of the minor phenotype and  $t$  is the density of a  
357 standard normal distribution at the  $p$ th quantile (see also Roff, 1997). It can be shown (see  
358 SI, section A) that this formula is an exact analytical solution to Eqs. 5 to 21 in the case of  
359 a GLMM with binomial distribution and a probit link. When fixed effects are included in the  
360 model, it is still possible to use these formulae by integration over the marginalised predictions  
361 (see SI, section A). Note that Eq. 25 applies only to analyses conducted with a probit link, it

362 does not apply to a binomial model with a logit link function.

363 **Poisson distribution with a logarithm link**

364 For a log link function and a Poisson distribution, both the derivative of the inverse link function,  
365 and the variance of the distribution, are equal to the expected value. Consequently, analytical  
366 results are obtainable for a log/Poisson model for quantities such as broad- and narrow-sense  
367 heritabilities. Foulley and Im (1993) derived an analytical formula to compute narrow-sense  
368 heritability on the observed scale:

$$h_{\text{obs}}^2 = \frac{\lambda^2 V_{A,\ell}}{\lambda^2 [\exp(V_{A,\ell} + V_{\text{RE}} + V_{\text{O}}) - 1] + \lambda} = \frac{\lambda V_{A,\ell}}{\lambda [\exp(V_{A,\ell} + V_{\text{RE}} + V_{\text{O}}) - 1] + 1}, \quad (26)$$

369 where  $\lambda$  is the data scale phenotypic mean, computed analytically as:

$$\lambda = \exp\left(\mu + \frac{V_{A,\ell} + V_{\text{RE}} + V_{\text{O}}}{2}\right). \quad (27)$$

370 Again, it can be shown (see SI, section B) that these formulae are exact solutions to Eq. 5 to  
371 21 when assuming a Poisson distribution with a log link. The inclusion of fixed effects in the  
372 model make the expression slightly more complicated (see SI, section B). These results can also  
373 be extended to the Negative-Binomial distribution with log link with slight modifications of  
374 the analytical expressions (see SI, section B).

375 The component of the broad-sense heritability on the observed data scale that arises from  
376 additive genetic effects on the latent scale can be computed as an intra-class correlation coefficient (i.e. repeatability) for this kind of model (Foulley and Im, 1993; Nakagawa and Schielzeth,  
377 2010):

$$H_{\text{obs}}^2 = \frac{V(E[z|a])}{V_{\text{P,obs}}} = \frac{\lambda(\exp(V_{A,\ell}) - 1)}{\lambda [\exp(V_{A,\ell} + V_{\text{RE}} + V_{\text{O}}) - 1] + 1}. \quad (28)$$

379 If non-additive genetic component were fitted in the model (e.g. dominance variance), they  
380 should be added to  $V_{A,\ell}$  in Eq. 28 to constitute the total genotypic variance, and thus obtain  
381 the actual broad-sense heritability. Note that the Eqs. 28 and 26 converge together for small  
382 values of  $V_{A,\ell}$ .

## 383 **Example analysis: quantitative genetic parameters of a non-normal** 384 **character**

385 We modelled the first year survival of Soay sheep (*Ovis aries*) lambs on St Kilda, Outer He-  
386 brides, Scotland. The data are comprised of 3814 individuals born between 1985 and 2011,  
387 and that are known to either have died in their first year, defined operationally as having died  
388 before the first of April in the year following their birth, or were known to have survived be-  
389 yond their first year. Months of mortality for sheep of all ages are generally known from direct  
390 observation, and day of mortality is typically known. Furthermore, every lamb included in this  
391 analysis had a known sex and twin status (whether or not it had a twin), and a mother of a  
392 known age.

393 Pedigree information is available for the St Kilda Soay sheep study population. Maternal  
394 links are known from direct observation, with occasional inconsistencies corrected with genetic  
395 data. Paternal links are known from molecular data. Most paternity assignments are made  
396 with very high confidence, using a panel of 384 SNP markers, each with high minor allele  
397 frequencies, and spread evenly throughout the genome. Details of marker data and pedigree  
398 reconstruction are given in Bérénos *et al.* (2014). The pedigree information was pruned to  
399 include only phenotyped individuals and their ancestors. The pedigree used in our analyses  
400 thus included 4687 individuals with 4165 maternal links and 4054 paternal links.

401 We fitted a generalised linear mixed model of survival, with a *logit* link function and a  
402 binomial distribution function. We included fixed effects of individual's sex and twin status,  
403 and linear, quadratic, and cubic effects of maternal age ( $matAge_i$ ). Maternal age was mean-  
404 centred by subtracting the overall mean. We also included an interaction of sex and twin status,  
405 and an interaction of twin status with maternal age. We included random effects of breeding  
406 value (as for Eq. 2), maternal identity, and birth year. Because the overdispersion variance  $V_O$   
407 in a binomial GLMM is unobservable for binary data, we set its variance to one. The model was  
408 fitted in MCMCGLMM (Hadfield, 2010), with diffuse independent normal priors on all fixed  
409 effects, and parameter-expanded priors for the variances of all estimated random effects.

410 We identified important effects on individual survival probability, i.e., several fixed effects  
411 were substantial, and also, each of the additive genetic, maternal, and among-year random

412 effects explained appreciable variance (Table 1). The model intercept corresponds to the ex-  
413 pected value on the latent scale of a female singleton (i.e. not a twin) lamb with an average  
414 age (4.8 years) mother. Males have lower survival than females, and twins have lower survival  
415 than singletons. There were also substantial effects of maternal age, corresponding to a rapid  
416 increase in lamb survival with maternal age among relatively young mothers, and a negative  
417 curvature, such that the maximum survival probabilities occur among offspring of mothers aged  
418 6 or 7 years. The trajectory of maternal age effects in the cubic model are similar to those  
419 obtained when maternal age is fitted as a multi-level effect.

420 To illustrate the consequences of accounting for different fixed effects on expected and ob-  
421 served data scale inferences, we calculated several parameters under a series of different treat-  
422 ments of the latent scale parameters of the GLMM. We calculated the phenotypic mean, the  
423 additive genetic variance, the total variance of expected values, the total variance of observed  
424 values, and the heritability of survival on the expected and observed scales.

425 First, we calculated parameters using only the model intercept ( $\mu$  in Eq. 1 and 3a). This  
426 intercept, under default settings, is arbitrarily defined by the linear modelling software imple-  
427 mentation and is thus software-dependent. In the current case, due to the details of how the  
428 data were coded, the intercept is the latent scale prediction for female singletons with average  
429 aged (4.8 years) mothers. In an average year, singleton females with average aged mothers have  
430 a probability of survival of about 80%. The additive genetic variance  $V_{A,obs}$ , calculated with  
431 Eq. 16 is about 0.005, and corresponds to heritabilities on the expected and observed scales of  
432 0.115 and 0.042 (Table 2).

433 In contrast, if we wanted to calculate parameters using a different (but equally arbitrary)  
434 intercept, corresponding to twin males, we would obtain a mean survival rate of 0.32, an additive  
435 genetic variance that is twice as large, but similar heritabilities (Table 1). Note that we have  
436 not modelled any systematic differences in genetic parameters between females and males, or  
437 between singletons and twins. These differences in parameter estimates arise from the exact  
438 same estimated variance components on the latent scale, as a result of different fixed effects.

439 This first comparison has illustrated a major way in which the fixed effects in a GLMM  
440 influence inferences on the expected and observed data scales. For linear mixed models, it  
441 has been noted that variance in the response is explained by the fixed predictors, and that

442 this may inappropriately reduce the phenotypic variance and inflate heritability estimates for  
443 some purposes (Wilson, 2008). However, in the example so far, we have simply considered two  
444 different intercepts (i.e. no difference in explained variance): female singletons vs male twins,  
445 in both cases, assuming focal groups of individuals are all born to average aged mothers. Again  
446 these differences in phenotypic variances and heritabilities arise from differences in intercepts,  
447 not any differences in variance explained by fixed effects. All parameters on the expected and  
448 observed value scales are dependent on the intercept, including the mean, the additive genetic  
449 variance and the total variance generated from random effects. Heritability is modestly affected  
450 by the intercept, because additive genetic and total variances are similarly, but not identically,  
451 influenced by the model intercept.

452 Additive genetic effects are those arising from the average effect of alleles on phenotype,  
453 integrated over all background genetic and environmental circumstances in which alternate  
454 alleles might occur. Fixed effects, where they represent biologically-relevant variation, are  
455 part of this background. Following our framework (see Eq. 17), we can solve the issue of the  
456 influence of the intercept by integrating our calculation of  $\Psi$  and ultimately  $V_{A,obs}$  over all fixed  
457 effects. This approach has the advantage of being consistent for any chosen intercept, as the  
458 value obtained after integration will not depend on that intercept. Considering all fixed and  
459 random effects, quantitative genetic parameters on the expected and observed scales are given  
460 in table 2, third column. Note that additive genetic variance is not intermediate between the  
461 two extremes (concerning sex and twin status), that we previously considered. The calculation  
462 of  $V_{A,obs}$  now includes an average slope calculated over a wide range of the steep part of the  
463 inverse-link function (near 0 on the latent scale, and near 0.5 on the expected data scale), and so  
464 is relatively high. The observed total phenotypic variance  $V_{P,obs}$  is also quite high. The increase  
465 in  $V_{P,obs}$  has two causes. First the survival mean is closer to 0.5, so the random effects variance  
466 is now manifested as greater total variance on the expected and observed scales. Second, there  
467 is now variance arising from fixed effects that is included in the total variance.

468 Given this, which estimates should be reported or interpreted? We have seen that when  
469 fixed effects are included in a GLMM, the quantitative genetic parameters calculated without  
470 integration are sensitive to an arbitrary parameter: the intercept. Hence integration over fixed  
471 effects may often be the best strategy for obtaining parameters that are not arbitrary. In

472 the case of survival analysed here,  $h_{\text{obs}}^2$  is the heritability of realised survival, whereas  $h_{\text{exp}}^2$  is  
473 the heritability of “intrinsic” individual survival. Since realised survival is the one “visible”  
474 by natural selection,  $h_{\text{obs}}^2$  might be a more relevant evolutionary parameter. Nonetheless, we  
475 recommend that  $V_{\text{P,exp}}$  and  $V_{\text{P,obs}}$  are both reported.

## 476 **Evolutionary prediction**

477 Systems for predicting adaptive evolution in response to phenotypic selection assume that the  
478 distribution of breeding values is multivariate normal, and in most applications, that the joint  
479 distribution of phenotypes and breeding values is multivariate normal (Lande, 1979; Lande and  
480 Arnold, 1983; Morrissey, 2014; Walsh and Lynch, forthcoming). The distribution of breeding  
481 values is assumed to be normal on the latent scale in a GLMM analysis, and therefore the  
482 parent-offspring regression will also be normal on that scale, but not necessarily on the data  
483 scale. Consequently, evolutionary change predicted directly using data-scale parameters may  
484 be distorted. The Breeder’s and Lande equations may hold approximately, and may perhaps be  
485 useful. However, having taken up the non-trivial task of pursuing GLMM-based quantitative  
486 genetic analysis, the investigator has at their disposal inferences on the latent scale. On this  
487 scale, the assumptions required to predict the evolution of quantitative traits hold. In this  
488 section we will first demonstrate by simulation how application of the Breeder’s equation will  
489 generate biased predictions of evolution. We then proceed to an exposition of some statistical  
490 machinery that can be used to predict evolution on the latent scale (from which evolution on  
491 the expected and observed scale can subsequently be calculated, using Eq. 5), given inference  
492 of the function relating traits to fitness.

## 493 **Direct application of the Breeder’s and Lande equations on the data scale**

494 In order to explore the predictions of the Breeder’s equation applied at the level of observed  
495 phenotype, we conducted a simulation in which phenotypes were generated according to a  
496 Poisson GLMM (Eq. 3a to 3c, with a Poisson distribution function and a log link function), and  
497 then selected the largest observed count values (positive selection) with a range of proportions  
498 of selected individuals (from 5% to 95%, creating a range of selection differentials), a range

of latent-scale heritabilities (0.1, 0.3, 0.5 and 0.8, with a latent phenotypic variance fixed to 0.1), and a range of latent means  $\mu$  (from 0 to 3). We simulated 10,000 replicates of each scenario, each composed of a different array of 10,000 individuals. For each simulation, we simulated 10,000 offspring. For each offspring, a breeding value was simulated according to  $a_{\ell,i} \sim \mathcal{N}((a_{\ell,d} + a_{\ell,s})/2, V_{A,\ell}/2)$ , where  $a_{\ell,i}$  is the focal offspring's breeding value,  $a_{\ell,d}$  and  $a_{\ell,s}$  are the breeding values of simulated dams and sires and  $V_{A,\ell}/2$  represents the segregational variance assuming parents are not inbred. Dams and sires were chosen at random with replacement from among the pool of simulated selected individuals. For each scenario, we calculated the realised selection differential arising from the simulated truncation selection,  $S_{\text{obs}}$ , and the average evolutionary response across simulations,  $R_{\text{obs}}$ . For each scenario, we calculated the heritability on the observed scale using Eq. 20. If the Breeder's equation was strictly valid for a Poisson GLMM on the observed scale, the realised heritability  $R_{\text{obs}}/S_{\text{obs}}$  would be equal to the observed-scale heritability  $h_{\text{obs}}^2$ .

The correspondence between  $R_{\text{obs}}/S_{\text{obs}}$  and  $h_{\text{obs}}^2$  is approximate (Fig. 2), and strongly depends on the selection differential (controlled here by the proportion of selected individuals). Note that, although the results presented here depict a situation where the ratio  $R_{\text{obs}}/S_{\text{obs}}$  is very often larger than  $h_{\text{obs}}^2$ , this is not a general result (e.g. this is not the case when using negative instead of positive selection, data not shown). In particular, evolutionary predictions are poorest in absolute terms for large  $\mu$  and large (latent) heritabilities. However, because we were analysing simulation data, we could track the selection differential of latent value (by calculating the difference in its mean between simulated survivors and the mean simulated before selection). We can also calculate the mean latent breeding value after selection. Across all simulation scenarios, the ratio of the change in breeding value after selection, to the change in breeding value before selection was equal to the latent heritability (see Fig. 2), showing that evolutionary changes could be accurately predicted on the latent scale.

## Evolutionary change on the latent scale, and associated change on the expected and observed scales

In an analysis of real data, latent (breeding) values are, of course, not measured. However, given an estimate of the effect of traits on fitness, say via regression analysis, we can derive

528 the parameters necessary to predict evolution on the latent scale. The idea is thus to relate  
529 measured fitness on the observed data scale to the latent scale, compute the evolutionary  
530 response on the latent scale and finally compute the evolutionary response on the observed  
531 data scale.

532 To relate the measured fitness on the observed scale to the latent scale, we need to compute  
533 the expected fitness  $W_{\text{exp}}$  given latent trait value  $\ell$ , which is

$$W_{\text{exp}}(\ell) = \sum_k W_P(k)P(Z = k|\ell), \quad (29)$$

534 where  $W_P(k)$  is the measure of fitness for the  $k$ th data scale category (assuming the observed  
535 data scale is discrete as in most GLMMs). Population mean fitness, can then be calculated in  
536 an analogous way to Eq. 5:

$$\bar{W} = \int W_{\text{exp}}(\ell)f_{\mathcal{N}}(\ell, \mu, V_{A,\ell} + V_{\text{RE}} + V_{\text{O}})d\ell. \quad (30)$$

537 These expressions comprise the basic functions necessary to predict evolution. Given a fitted  
538 GLMM, and a given estimate of the fitness function  $W_P(k)$ , each of several approaches could  
539 give equivalent results. For simplicity, we proceed via application of the breeder's equation at  
540 the level of the latent scale.

541 The change in the mean genetic value of any character due to selection is equal to the  
542 covariance of breeding value with relative fitness (Robertson, 1966, 1968). Using Stein's (1973)  
543 lemma once more, this covariance can be obtained as the product of the additive genetic variance  
544 of latent values and the average derivative of expected fitness with respect to latent value, i.e.,  
545  $E \left[ \frac{dW_{\text{exp}}}{d\ell} \right]$ . Evolution on the latent scale can therefore be predicted by

$$\Delta\mu = V_A E \left[ \frac{dW_{\text{exp}}}{d\ell} \right] \frac{1}{\bar{W}}. \quad (31)$$

546 In the case of a multivariate analysis, note that the derivative above should be a vector of  
547 partial derivatives (partial first order derivative with respect to latent value for each trait) of  
548 fitness.

549 If fixed effects need to be considered, the approach can be modified in the same way as

550 integration over fixed effects is accomplished for calculating other quantities, i.e. the expression

$$\bar{W} = \frac{1}{N} \sum_{i=1}^N \int W_{\text{exp}}(\ell) f_{\mathcal{N}}(\ell, \mu + \hat{\ell}_i, V_{A,\ell} + V_{\text{RE}} + V_{\text{O}}) d\ell \quad (32)$$

551 would be used in calculations of mean fitness and the average derivative of expected fitness  
552 with respect to latent value.

553 Phenotypic change caused by changes in allele frequencies in response to selection is calcu-  
554 lated as

$$\Delta \bar{z} = \int g^{-1}(\ell) f_{\mathcal{N}}(\ell, \mu + \Delta\mu, V_{A,\ell} + V_{\text{RE}} + V_{\text{O}}) d\ell - \bar{z}. \quad (33)$$

555 Or, if fixed effects are included in the model:

$$\Delta \bar{z} = \frac{1}{N} \sum_{i=1}^N \int g^{-1}(\ell) f_{\mathcal{N}}(\ell, \mu + \hat{\ell}_i + \Delta\mu, V_{A,\ell} + V_{\text{RE}} + V_{\text{O}}) d\ell - \bar{z}. \quad (34)$$

556 Note that, in this second equation,  $\bar{z}$  must be computed as in Eq. 17 and that this equation  
557 assumes that the distribution of fixed effects for the offspring generation is the same as for the  
558 parental one.

559 Another derivation of the expected evolutionary response using the Price-Robertson identity  
560 (Robertson, 1966; Price, 1970) is given in the Supplementary Information (section C).

## 561 **The simulation study revisited**

562 Using the same replicates as in the simulation study above, we used Eqs. 29 to 34 to predict  
563 phenotypic evolution. This procedure provides greatly improved predictions of evolutionary  
564 change on the observed scale (Fig. 3, top row). However, somewhat less response to selection is  
565 observed than is predicted. This deviation occurs because, in addition to producing a perma-  
566 nent evolutionary response in the mean value on the latent scale, directional selection creates  
567 a transient reduction of additive genetic variance due to linkage disequilibrium. Because the  
568 link function is non-linear, this transient change in the variance on the latent scale generates  
569 a transient change in the mean on the expected and observed scales. Following several genera-  
570 tions of random mating, the evolutionary change on the observed scale would converge on the  
571 predicted values. We simulated such a generation at equilibrium by simply drawing breeding

572 values for the post-selection sample from a distribution with the same variance as in the parental  
573 generation. This procedure necessarily generated a strong match between predicted and simu-  
574 lated evolution (Fig. 3, bottom row). Additionally, the effects of transient reduction in genetic  
575 variance on the latent scale could be directly modelled, for example, using Bulmer's (1971)  
576 approximations for the transient dynamics of the genetic variance in response to selection.

## 577 Discussion

578 The general approach outlined here for quantitative genetic inference with GLMMs has several  
579 desirable features: *(i)* it is a general framework, which should work with any given GLMM  
580 and especially, any link and distribution function, *(ii)* provides mechanisms for rigorously han-  
581 dling fixed effects, which can be especially important in GLMMs, and *(iii)* it can be used for  
582 evolutionary prediction under standard GLMM assumptions about the genetic architecture of  
583 traits.

584 Currently, with the increasing applicability of GLMMs, investigators seem eager to convert  
585 to the observed data scale. It seems clear that conversions between scales are generally useful.  
586 However, it is of note that the underlying assumption when using GLMMs for evolutionary  
587 prediction is that predictions hold on the latent scale. Hence, some properties of heritabilities  
588 for additive Gaussian traits, particularly the manner in which they can be used to predict  
589 evolution, do not hold on the data scale for non-Gaussian traits, even when expressed on the  
590 data scale. Yet, given an estimate of a fitness function, no further assumptions are necessary to  
591 predict evolution on the data scale, via the latent scale (as with Eqs. 29, 31, and 33), over and  
592 above those that are made in the first place upon deciding to pursue GLMM-based quantitative  
593 genetic analysis. Hence we recommend using this framework to produce accurate predictions  
594 about evolutionary scenarios.

595 We have highlighted important ways in which fixed effects influence quantitative genetic in-  
596 ferences with GLMMs, and developed an approach for handling these complexities. In LMMs,  
597 the main consideration pertaining to fixed effects is that they explain variance, and some or all  
598 of this variance might be inappropriate to exclude from an assessment of  $V_P$  when calculating  
599 heritabilities (Wilson, 2008). This aspect of fixed effects is relevant to GLMMs, but further-

600 more, all parameters on the expected and observed scales, not just means, are influenced by  
601 fixed effects in GLMMs; this includes additive genetic and phenotypic variances. This fact  
602 necessitates particular care in interpreting GLMMs. Our work clearly demonstrates that con-  
603 sideration of fixed effects is essential, and the exact course of action needs to be considered  
604 on a case-by-case basis. Integrating over fixed effects would solve, in particular, the issue of  
605 intercept arbitrariness illustrated with the Soay sheep example. Yet cases may often arise where  
606 fixed effects are fitted, but where one would not want to integrate over them (e.g. because they  
607 represent experimental rather than natural variability). In such cases, it will be important to  
608 work with a biologically meaningful intercept, which can be achieved for example by centring  
609 covariates on relevant values (Schielzeth, 2010). Finally, note that this is not an all-or-none  
610 alternative: in some situations, it could be relevant to integrate over some fixed effects (e.g. of  
611 biological importance) while some other fixed effects (e.g. those of experimental origins) would  
612 be left aside.

613 One of the most difficult concepts in GLMMs seen as a non-linear developmental model  
614 (Morrissey, 2015) is that an irreducible noise is attached to the observed data. This is the  
615 reason why we believe that distinguishing between expected and observed data scale does have  
616 a biological meaning. Researchers using GLMMs need to realise that this kind of model can  
617 assume a large variance in the observed data with very little variance on the latent and expected  
618 data scales. For example, a Poisson/log GLMM with a latent mean  $\mu = 0$  and a total latent  
619 variance of 0.5 will result in observed data with a variance  $V_{P,obs} = 2.35$ . Less than half of this  
620 variance lies in the expected data scale ( $V_{P,exp} = 1.07$ ), the rest is residual Poisson variation.  
621 Our model hence assumes that more than half of the measured variance comes from totally  
622 random noise. Hence, even assuming that the whole latent variance is composed of additive  
623 genetic variance, the heritability will never reach a value above 0.5. Whether this random noise  
624 should be accounted for when computing heritabilities (i.e. whether we should compute  $h_{exp}^2$   
625 or  $h_{obs}^2$ ) depends on the evolutionary question under study. In many instances, it is likely that  
626 natural selection will act directly on realised values rather than their expectations, in which case  
627  $h_{obs}^2$  should be preferred. We recommend however, that, along with  $V_{A,obs}$ , all other variances  
628 (including  $V_O$ ,  $V_{P,exp}$  and  $V_{P,obs}$ ) are reported by researchers.

629 The expressions given here for quantitative genetic parameters on the expected and ob-

630 served data scales are exact, given the GLMM model assumptions, in two senses. First, they  
631 are not approximations, such as might be obtained by linear approximations (Browne *et al.*,  
632 2005). Second, they are expressions for the parameters of direct interest, rather than convenient  
633 substitutes. For example, the calculation (also suggested by Browne *et al.* 2005) of variance  
634 partition coefficients (i.e. intraclass correlations) on an underlying scale only provides a value  
635 of the broad-sense heritability (e.g. using the genotypic variance arising from additive genetic  
636 effects on the latent scale).

637 The framework developed here (including univariate and multivariate parameters computa-  
638 tion and evolutionary predictions on the observed data scale) is implemented in the R package  
639 QGGLMM available at <https://github.com/devillemereuil/qgglm>. Note that the package  
640 does not perform any GLMM inference but only implements the hereby introduced framework  
641 for analysis posterior to a GLMM inference. While the calculations we provide will often (i.e.  
642 when no analytical formula exists) be more computationally demanding than calculations on  
643 the latent scale, they will be direct ascertainments of specific parameters of interest, since the  
644 scale of evolutionary interest is likely to be the observed data scale, rather than the latent  
645 scale (unless some artificial selection is applied to predicted latent breeding values as in mod-  
646 ern animal breeding). Most applications should not be onerous. Computations of means and  
647 (additive genetic) variances took less than a second on a 1.7 GHz processor when using our  
648 R functions on the Soay sheep data set. Summation over fixed effects, and integration over  
649 1000 posterior samples of the fitted model took several minutes. When analytical expressions  
650 are available (e.g. for Poisson/log, Binomial/probit and Negative-Binomial/log, see the sup-  
651plementary information and R package documentation), these computations are considerably  
652 accelerated.

## 653 **Acknowledgements**

654 PdV was supported by a doctoral studentship from the French *Ministère de la Recherche et de*  
655 *l'Enseignement Supérieur*. HS was supported by an Emmy Noether fellowship from the German  
656 Research Foundation (DFG; SCHI 1188/1-1). SN is supported by a Future Fellowship, Aus-  
657 tralia (FT130100268). MBM is supported by a University Research Fellowship from the Royal

658 Society (London). The Soay sheep data were provided by Josephine Pemberton and Loeske  
659 Kruuk, and were collected primarily by Jill Pilkington and Andrew MacColl with the help of  
660 many volunteers. The collection of the Soay sheep data is supported by the National Trust  
661 for Scotland and QinetQ, with funding from NERC, the Royal Society, and the Leverhulme  
662 Trust. We thank Kerry Johnson, Paul Johnson, Alastair Wilson, Loeske Kruuk and Josephine  
663 Pemberton for valuable discussions and comments on this manuscript.

## 664 **References**

- 665 Ayers, D. R., R. J. Pereira, A. A. Boligon, F. F. Silva, F. S. Schenkel, V. M. Roso, and L. G.  
666 Albuquerque, 2013 Linear and Poisson models for genetic evaluation of tick resistance in  
667 cross-bred Hereford x Nellor cattle. *Journal of Animal Breeding and Genetics* 130: 417–424.
- 668 Browne, W. J., S. V. Subramanian, K. Jones, and H. Goldstein, 2005 Variance partitioning  
669 in multilevel logistic models that exhibit overdispersion. *Journal of the Royal Statistical*  
670 *Society* 168: 599–613.
- 671 Bulmer, M. G., 1971 The effect of selection on genetic variability. *The American Natural-*  
672 *ist* 105: 201–211.
- 673 Béréños, C., P. A. Ellis, J. G. Pilkington, and J. M. Pemberton, 2014 Estimating quantitative  
674 genetic parameters in wild populations: a comparison of pedigree and genomic approaches.  
675 *Molecular Ecology* 23: 3434–3451.
- 676 de Villemereuil, P., O. Gimenez, and B. Doligez, 2013 Comparing parent–offspring regression  
677 with frequentist and Bayesian animal models to estimate heritability in wild populations: a  
678 simulation study for Gaussian and binary traits. *Methods in Ecology and Evolution* 4(3):  
679 260–275.
- 680 Dempster, E. R. and I. M. Lerner, 1950 Heritability of Threshold Characters. *Genetics* 35(2):  
681 212–236.
- 682 Falconer, D. S., 1960 *Introduction to Quantitative Genetics*. Oliver and Boyd.

- 683 Fisher, R. A., 1918 The correlation between relatives on the supposition of Mendelian inheri-  
684 tance. *Transactions of the Royal Society of Edinburgh* 52: 399–433.
- 685 Fisher, R. A., 1924 The biometrical study of heredity. *Eugenics Review* 16: 189–210.
- 686 Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.
- 687 Foulley, J. L. and S. Im, 1993 A marginal quasi-likelihood approach to the analysis of Poisson  
688 variables with generalized linear mixed models. *Genetics Selection Evolution* 25(1): 101.
- 689 Hadfield, J., 2010 MCMC Methods for Multi-Response Generalized Linear Mixed Models:  
690 The MCMCglmm R Package. *Journal of Statistical Software* 33(2): 1–22.
- 691 Henderson, C. R., 1950 Estimation of genetic parameters. *Annals of Mathematical Statis-*  
692 *tics* 21: 309–310.
- 693 Henderson, C. R., 1973 *Proceedings of the Animal Breeding and Genetics Symposium in*  
694 *Honour of Dr. Jay L. Lush*, Chapter Sire evaluation and genetic trends. Published by the  
695 American Society of Animal Science and the American Dairy Science Association.
- 696 Hill, W. G. and M. Kirkpatrick, 2010 What animal breeding has taught us about evolution.  
697 *Annual Review of Ecology, Evolution, and Systematics* 41: 1–19.
- 698 Houle, D., 1992 Comparing evolvability and variability of quantitative traits. *Genetics* 130(1):  
699 195–204.
- 700 Kruuk, L. E. B., 2004 Estimating genetic parameters in natural populations using the ‘animal  
701 model’. *Philosophical Transactions of the Royal Society of London B* 359: 873–890.
- 702 Lande, R., 1979 Quantitative genetic analysis of multivariate evolution, applied to brain:body  
703 size allometry. *Evolution* 33: 402–416.
- 704 Lande, R. and S. J. Arnold, 1983 The measurement of selection on correlated characters.  
705 *Evolution* 37: 1210–1226.
- 706 Lush, J. L., 1937 *Animal breeding plans*. Ames, Iowa: Iowa State College Press.
- 707 Lynch, M. and B. Walsh, 1998 *Genetics and analysis of quantitative traits*. Sunderland, MA:  
708 Sinauer.

- 709 Milot, E., F. M. Mayer, D. H. Nussey, M. Boisvert, F. Pelletier, and D. Reale, 2011 Evidence for  
710 evolution in response to natural selection in a contemporary human population. *Proceedings*  
711 of the National Academy of Sciences 108(41): 17040–17045.
- 712 Morrissey, M. B., 2014 In search of the best methods for multivariate selection analysis.  
713 *Methods in Ecology and Evolution* 5: 1095–1109.
- 714 Morrissey, M. B., 2015 Evolutionary quantitative genetics of non-linear developmental systems.  
715 *Evolution* 69: 2050–2066.
- 716 Morrissey, M. B., P. de Villemereuil, B. Doligez, and O. Gimenez, 2014 Bayesian approaches to  
717 the quantitative genetic analysis of natural populations. In A. Charmantier, D. Garant, and  
718 L. E. Kruuk (Eds.), *Quantitative Genetics in the Wild*, pp. 228–253. Oxford (UK): Oxford  
719 University Press.
- 720 Morrissey, M. B., C. A. Walling, A. J. Wilson, J. M. Pemberton, T. H. Clutton-Brock, and  
721 L. E. B. Kruuk, 2012 Genetic Analysis of Life-History Constraint and Evolution in a Wild  
722 Ungulate Population. *The American Naturalist* 179(4): E97–E114.
- 723 Nakagawa, S. and H. Schielzeth, 2010 Repeatability for Gaussian and non Gaussian data: a  
724 practical guide for biologists. *Biological Reviews* 85(4): 935–956.
- 725 Price, G. R., 1970 Selection and covariance. *Nature* 227: 520–521.
- 726 R Core Team, 2015 *R: A Language and Environment for Statistical Computing*. Vienna,  
727 Austria: R Foundation for Statistical Computing.
- 728 Robertson, A., 1950 Heritability of threshold characters. *Genetics* 35(2): 212–236.
- 729 Robertson, A., 1966 A mathematical model of the culling process in dairy cattle. *Animal*  
730 *Production* 8: 95–108.
- 731 Robertson, A., 1968 *Population Biology and Evolution*, Chapter The spectrum of genetic  
732 variation, pp. 5–16. New York: Syracuse University Press.
- 733 Roff, D. A., 1997 *Evolutionary quantitative genetics*. New York (US): Chapman & Hall.

- 734 Schielzeth, H., 2010 Simple means to improve the interpretability of regression coefficients.  
735 *Methods in Ecology and Evolution* 1(2): 103–113.
- 736 Sillanpää, M. J., 2011 On statistical methods for estimating heritability in wild populations.  
737 *Molecular Ecology* 20(7): 1324–1332.
- 738 Stein, C. M., 1973 Estimation of the mean of a multivariate normal distribution. Proceedings  
739 of the Prague Symposium on Asymptotic Statistics 1: 345–381.
- 740 Thompson, R., 2008 Estimation of quantitative genetic parameters. Proceedings of the Royal  
741 Society B: Biological Sciences 275(1635): 679–686.
- 742 Walsh, B. and M. Lynch, forthcoming *Evolution and selection of quantitative traits*. Sinauer.
- 743 Wilson, A. J., 2008 Why  $h^2$  does not always equal  $VA/VP$ ? *Journal of Evolutionary Biol-*  
744 *ogy* 21(3): 647–650.
- 745 Wilson, A. J., M. B. Morrissey, M. J. Adams, C. A. Walling, F. E. Guinness, J. M. Pemberton,  
746 T. H. Clutton-Brock, and L. E. B. Kruuk, 2011 Indirect genetics effects and evolutionary  
747 constraint: an analysis of social dominance in red deer, *Cervus elaphus*. *Journal of Evolu-*  
748 *tionary Biology* 24(4): 772–783.
- 749 Wilson, A. J., D. Reale, M. N. Clements, M. B. Morrissey, C. A. W. E. Postma, L. E. B.  
750 Kruuk, and D. H. Nussey, 2010 An ecologist's guide to the animal model. *Journal of Animal*  
751 *Ecology* 79: 13–26.
- 752 Wright, S., 1934 An analysis of variability in number of digits in an inbred strain of Guinea  
753 pigs. *Genetics* 19(6): 506–536.

Table 1: Parameters from the GLMM-based quantitative genetic analysis of Soay sheep (*Ovis aries*) lamb first-year survival. All estimates are reported as posterior modes with 95% credible intervals. The intercept in this model is arbitrarily defined for female lambs without twins, born to average age (4.8 years) mothers.

| Parameter                    | Posterior mode with 95% CI |
|------------------------------|----------------------------|
| (a) Fixed effects            |                            |
| Intercept                    | 2.686 (1.631 – 3.299)      |
| Sex (male vs. female)        | -1.185 (-1.436 – -0.932)   |
| Twin (twin vs. singleton)    | -2.383 (-3.315 – -1.760)   |
| Maternal age, linear term    | 0.238 (0.092 – 0.384)      |
| Maternal age, quadratic term | -0.169 (-0.196 – -0.148)   |
| Maternal age, cubic term     | 0.014 (0.010 – 0.019)      |
| Sex-twin interaction         | 0.497 (0.016 – 1.016)      |
| Sex-maternal age interaction | -0.020 (-0.103 – 0.070)    |
| (b) Random effects           |                            |
| $V_{A,\ell}$                 | 0.882 (0.256 – 1.542)      |
| $V_{\text{mother}}$          | 0.467 (0.213 – 0.876)      |
| $V_{\text{year}}$            | 3.062 (1.814 – 5.635)      |

Table 2: Estimates of expected and observed scale phenotypic mean and variances, and additive genetic variance, for three different treatments of the fixed effects, as modelled on the linear scale with a GLMM, and reported in table 1. Additive genetic variance and heritability on the latent scales are also reported for comparison. Note that  $h_{\text{lat}}^2$  is slightly lower when averaging over fixed effects, since the variance they explain is then accounted for.

| Quantity            | Arbitrary intercept<br>(singleton female) | Arbitrary intercept<br>(twin male) | Averaging over all fixed effects |
|---------------------|---|------------------------------------|----------------------------------|
| $V_{A,\ell}$        | 0.915 (0.275 – 1.66)                      | 0.915 (0.275 – 1.66)               | 0.915 (0.275 – 1.66)             |
| $h_{\text{lat}}^2$  | 0.159 (0.056 – 0.27)                      | 0.159 (0.056 – 0.27)               | 0.117 (0.0417 – 0.194)           |
| $\bar{z}$           | 0.838 (0.721 – 0.887)                     | 0.352 (0.220 – 0.474)              | 0.446 (0.334 – 0.511)            |
| $V_{A, \text{obs}}$ | 0.006 (0.002 – 0.015)                     | 0.011 (0.005 – 0.023)              | 0.013 (0.005 – 0.021)            |
| $V_{P, \text{exp}}$ | 0.060 (0.034 – 0.095)                     | 0.098 (0.072 – 0.124)              | 0.123 (0.106 – 0.137)            |
| $V_{P, \text{obs}}$ | 0.136 (0.108 – 0.206)                     | 0.250 (0.183 – 0.250)              | 0.250 (0.227 – 0.250)            |
| $h_{\text{exp}}^2$  | 0.109 (0.043 – 0.201)                     | 0.122 (0.054 – 0.227)              | 0.102 (0.039 – 0.166)            |
| $h_{\text{obs}}^2$  | 0.047 (0.017 – 0.085)                     | 0.051 (0.022 – 0.101)              | 0.043 (0.020 – 0.087)            |

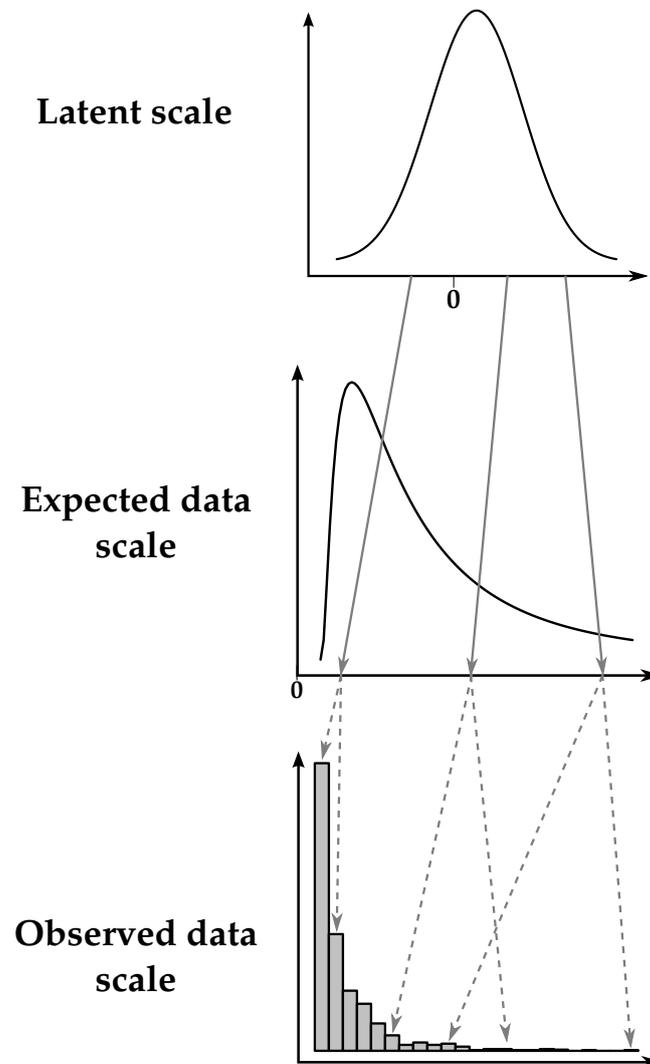


Figure 1: Example of the relationships between the three scales of the GLMM using a Poisson distribution and a logarithm link function. Deterministic relationships are denoted using grey plain arrows, whereas stochastic relationships are denoted using grey dashed arrows. Note that the latent scale is depicted as a simple Gaussian distribution for the sake of simplicity, whereas it is a mixture of Gaussian distributions in reality.

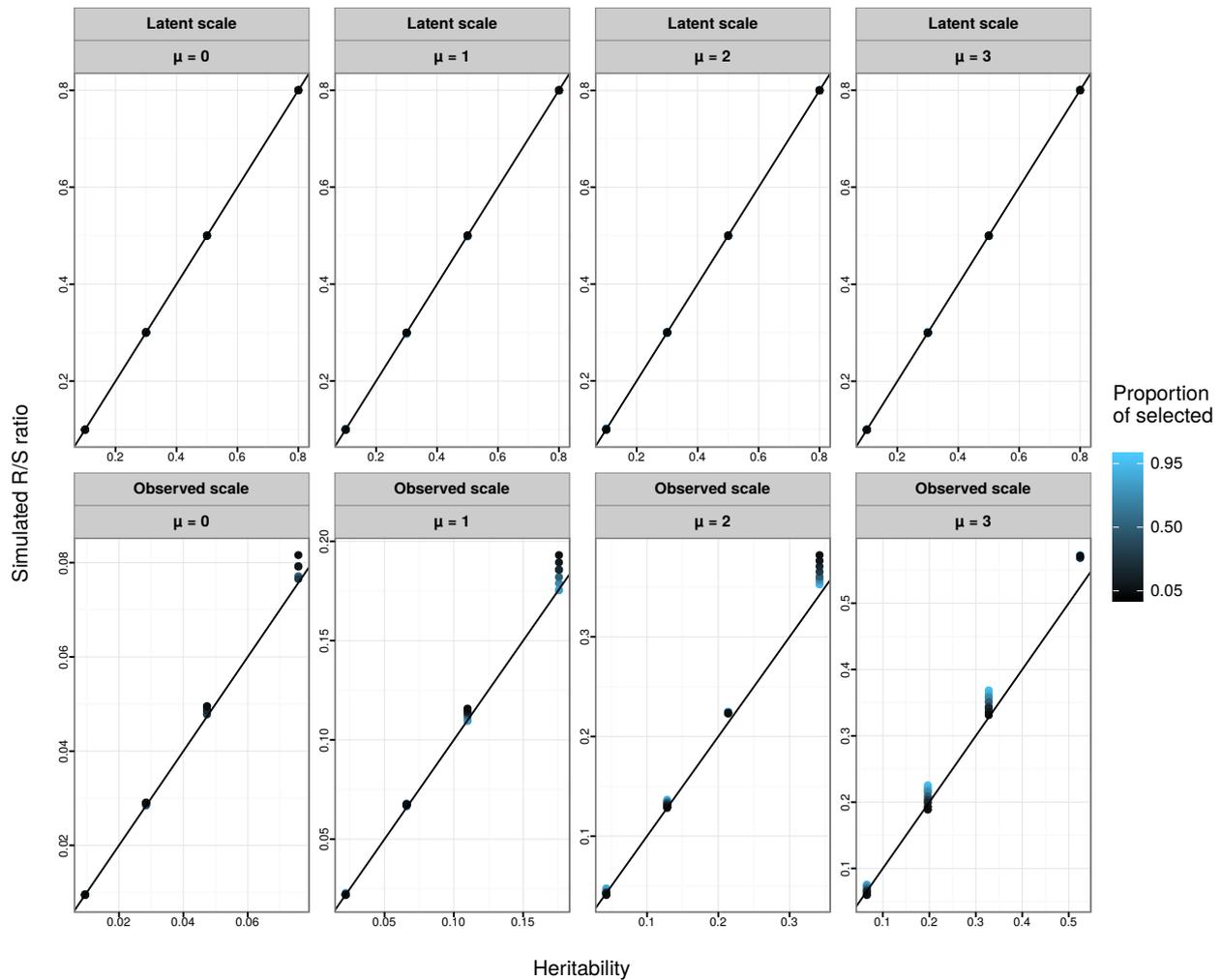


Figure 2: Simulated  $R/S$  (evolutionary response over selection differential, or the realised heritability) on the latent (upper panels) or observed data (lower panels) scales against the corresponding-scale heritabilities. Each data point is the average over 10,000 replicates of 10,000 individuals for various latent heritabilities  $h_{\text{lat}}^2$  (0.1, 0.3, 0.5, 0.8), latent population mean ( $\mu$  from 0 to 3, from left to right) and proportion of selected individuals (5%, 10%, 20%, 30%, 50%, 70%, 80%, 90%, 95%, varying from black to blue). The 1:1 line is plotted in black. The breeder's equation is predictive on the latent scale (upper panels), but approximate on the observed data scale (lower panels), because phenotypes and breeding values are not jointly multivariate normal on that scale.

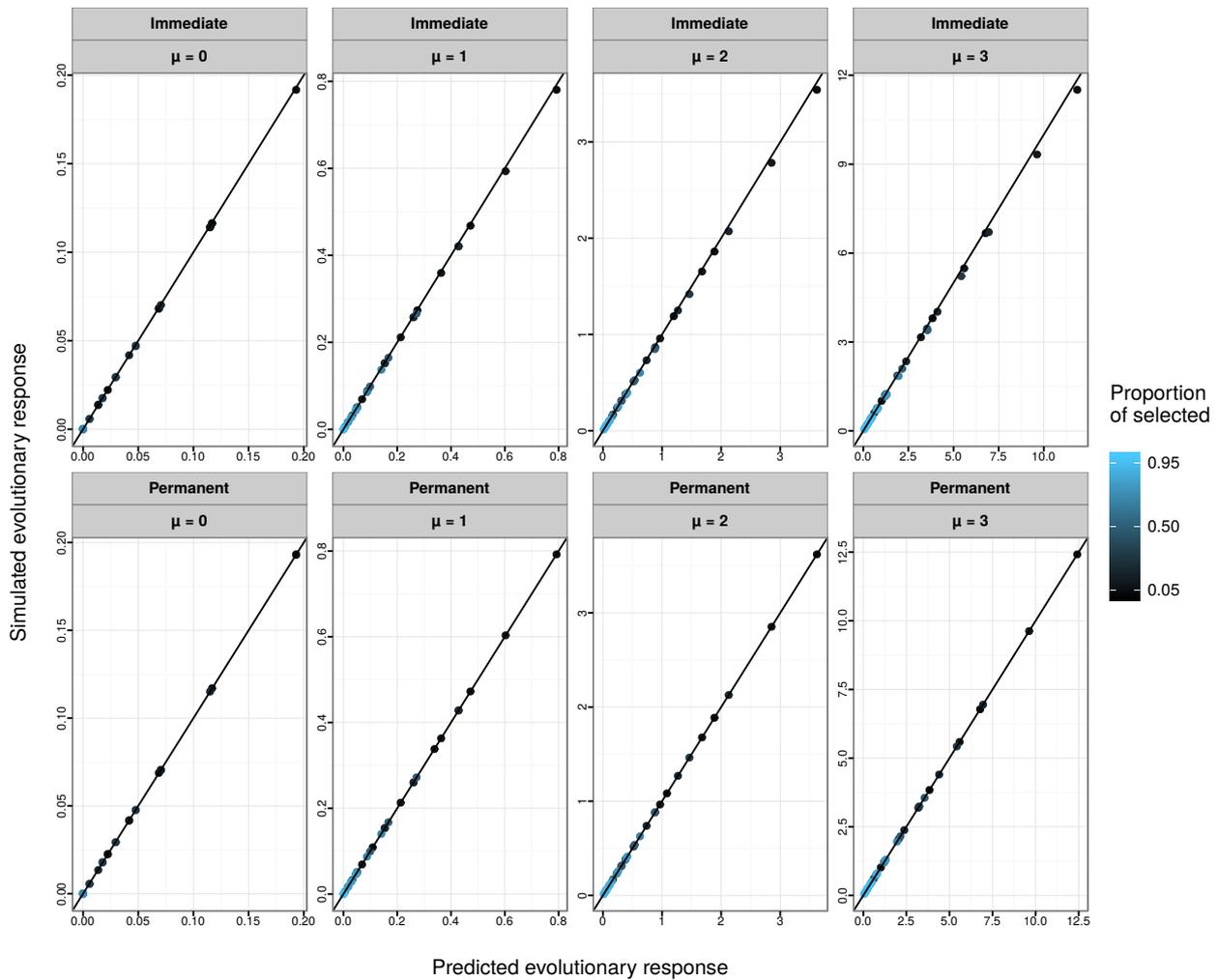


Figure 3: Predicted  $R_{\text{obs}}$  (phenotypic evolutionary response on the observed scale, see Eq. 34) against the simulated  $R_{\text{obs}}$ , via evolutionary predictions applied on the latent scale. Each data point is the average over 10,000 replicates of 10,000 individuals for various latent heritabilities  $h_{\text{lat}}^2$  (0.1, 0.3, 0.5, 0.8), latent population mean ( $\mu$  from 0 to 3) and proportion of selected individuals (5%, 10%, 20%, 30%, 50%, 70%, 80%, 90%, 95%, varying from black to blue). The 1:1 line is plotted in black. The upper panels (“Immediate”) show simulations for the response after a single generation, which include both a permanent and transient response to selection arising from linkage disequilibrium. The bottom panels (“permanent”) show simulation results modified to reflect only the permanent response to selection.