Genetic Variance in Environmental Variance leads to non-linear Relationships between Traits with Application to Birth Weight and Survival in Piglets

***H.A. Mulder*[*], W.G. Hill[†], and E. F. Knol[‡].**

[*]Animal Breeding and Genomics Centre, Wageningen University, Wageningen, The Netherlands, [†]Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom, [‡]TOPIGS Research Center IPG, Beuningen, The Netherlands

ABSTRACT: There is abundant evidence for existence of genetic variance in environmental variance. Little is known about the relationships between environmental variance of one trait and levels of other traits, however. A genetic covariance between environmental variance of one trait and the level of another trait is expected to lead to non-linearity between them. Such non-linearity might occur for birth weight and survival of piglets, where animals of extreme weights have lower survival, and therefore uniformity in weight influences survival. The objectives were to derive this non-linear relationship analytically using multiple regression and apply it to piglet birth weight and piglet survival. Monte Carlo simulation was used to check the multiple regression equation derived. It explained most of the non-linearity between piglet birth weight and survival. The framework developed can help in developing optimal breeding strategies utilizing genetic variance in environmental variance.

Keywords: Environmental variance; Uniformity; Non-linear relationships; Pigs

Introduction

Subsequent to the publication by SanCristobal-Gaudy et al. (1998), there is growing evidence for the existence of genetic variance in environmental variance (V_E). Hill and Mulder (2010) reviewed the literature and concluded that the median genetic coefficient of variation for environmental variance is ~30%, indicating that the V_E could be increased/decreased by 30% if changed by one genetic standard deviation of V_E. Selection on V_E can be used to increase uniformity (Mulder et al. (2008)). For many traits, especially in pigs and poultry, it is beneficial to increase uniformity of animals by selection to increase ease of management and homogeneity of animal products. However, very little is known about the genetic correlation between V_E of one trait and the mean of other traits, i.e. existence of trade-offs. For within-litter variability of piglet birth weight there is, however, some evidence that increasing uniformity leads to higher survival of piglets (Kapell et al. (2011)). It is expected that piglet birth weight is an optimum trait, as both very small and very heavy piglets are less desirable for different reasons.

Genetic variation in V_E leads to genetic variation in the squared phenotypic deviation of that trait (Mulder et al. (2007)). If there is a genetic correlation between V_E of one trait and the mean of another, the phenotypic relationship between the two traits is expected to be non-linear. Non-linearity between them may have its consequences for estimation of genetic parameters and selection decisions (Sölkner and Fuerst-Waltl (1996)). The magnitude by which the genetic covariance between V_E of one trait and the mean of another trait can contribute to such non-linearity is unknown, however.

The objectives of this study were to derive this relationship analytically and to utilize it to investigate the non-linearity between piglet birth weight and piglet survival.

Materials and Methods

Quantitative genetic model: We investigated two traits for which the phenotypic deviations from the mean are the sum of additive genetic and environmental effects, with no dominance and epistasis (Falconer and Mackay (1996)). The first trait (P_1) is subject to additive genetic differences in V_E, so-called genetic heterogeneity of environmental variance (Mulder et al. (2007)), whereas the second trait (P_2) has homogeneous environmental variance. We assume the exponential model for V_E (SanCristobal-Gaudy (1998); see Hill and Mulder (2010), for an overview). The genetic models are:

$$P_1 = \mu_1 + A_{m,1} + E_1 = A_{m,1} + \chi\exp(\tfrac{(\ln(\sigma^2_{e_1})+A_{v,1}}{2}) \quad (1)$$

$$P_2 = \mu_2 + A_{m,2} + E_2 \quad (2)$$

where μ_1 and μ_2 are the population means for trait 1 and 2, $A_{m,1}$, $A_{v,1}$ and $A_{m,2}$ are respectively the breeding values for the mean trait value of trait 1, V_E of trait 1 and the mean of trait 2, χ is a standard normal deviate for the environmental effect of trait 1, $\sigma^2_{e_1}$ is the environmental variance when $A_{v,1} = 0$ and E_2 is the environmental effect for trait 2. It was assumed that the breeding values $A_{m,1}$, $A_{v,1}$ and $A_{m,2}$ are the sum of an infinite number of loci each with small additive genetic effects and follow a multivariate normal distribution:

$$\begin{bmatrix} A_{m,1} \\ A_{v,1} \\ A_{m,2} \end{bmatrix} \sim \mathrm{N}\left(\begin{matrix} 0 \\ 0, \\ 0 \end{matrix} \begin{bmatrix} \sigma^2_{A_{m,1}} & cov_{A_{m,1},A_{v,1}} & cov_{A_{m,1},A_{m,2}} \\ & \sigma^2_{A_{v,1}} & cov_{A_{v,1},A_{m,2}} \\ symmetric & & \sigma^2_{A_{m,2}} \end{bmatrix} \otimes \mathbf{A} \right),$$

where **A** is the additive genetic or numerator relationship matrix, $\sigma^2_{A_{m,1}}$, $\sigma^2_{A_{v,1}}$ and $\sigma^2_{A_{m,2}}$ are the additive genetic variances in $A_{m,1}$, $A_{v,1}$ and $A_{m,2}$, respectively, $cov_{A_{m,1},A_{v,1}}$, $cov_{A_{m,1},A_{m,2}}$ and $cov_{A_{v,1},A_{m,2}}$ are the additive genetic covariances between $A_{m,1}$ and $A_{v,1}$, between $A_{m,1}$ and $A_{m,2}$ and between $A_{v,1}$ and $A_{m,2}$, respectively. The environmental effects E_1 (=$\chi\exp(\tfrac{(\ln(\sigma^2_{e_1})+A_{v,1}}{2})$) and E_2 are bivariate normally distributed:

$\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} \sim N\left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{bmatrix} \sigma^2_{e_1}\exp(A_{v,1}) & cov_{e_1e_2} \\ symmetric & \sigma^2_{e_2} \end{bmatrix}\right)$, where $\sigma^2_{e_2}$ is the environmental variance of trait 2 and $cov_{e_1e_2} = r_e\sigma_{e_1}\exp(\frac{1}{2}A_v)\sigma_{e_2}$, where r_e is the environmental correlation.

Multiple regression: Multiple regression was used to derive the relationship between the two traits. Here we regressed P_2 on P_1 and P_1^2 to investigate the non-linear relationship between P_1 and P_2 (see Figure 1):

$$P_2 = b_1P_1 + b_2P_1^2 \quad (3)$$

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \mathbf{P^{-1}c} \quad (4)$$

where $\mathbf{P} = \begin{bmatrix} \text{var}(P_1) & \text{cov}(P_1, P_1^2) \\ \text{cov}(P_1, P_1^2) & \text{var}(P_1^2) \end{bmatrix}$ and $\mathbf{c} = \begin{bmatrix} \text{cov}(P_2, P_1) \\ \text{cov}(P_2, P_1^2) \end{bmatrix}$. The elements of **P** are:

$$\text{var}(P_1) = \sigma^2_{A_{m,1}} + \sigma^2_{e_1}\exp\left(0.5\sigma^2_{A_{v,1}}\right) \quad (5)$$

$$\text{cov}(P_1, P_1^2) = 3cov_{A_{m,1},A_{v,1}}\sigma^2_{e_1}\exp\left(0.5\sigma^2_{A_{v,1}}\right) \quad (6)$$

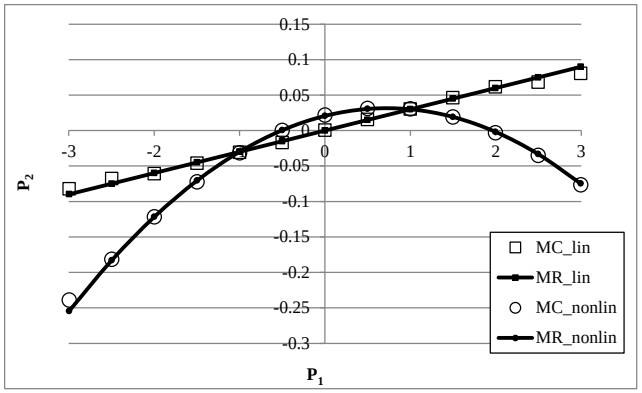
$$\text{var}(P_1^2) = 2\sigma^4_{A_{m,1}} + 4\left[\sigma^2_{A_{m,1}}\sigma^2_{e_1}\exp\left(0.5\sigma^2_{A_{v,1}}\right) + cov^2_{A_{m,1},A_{v,1}}\sigma^2_{e_1}\exp\left(0.5\sigma^2_{A_{v,1}}\right)\right] + 2cov^2_{A_{m,1},A_{v,1}}\sigma^2_{e_1}\exp\left(0.5\sigma^2_{A_{v,1}}\right) + 3\sigma^4_{e_1}\exp\left(2\sigma^2_{A_{v,1}}\right) - \sigma^4_{e_1}\exp\left(\sigma^2_{A_{v,1}}\right) \quad (7)$$

The elements of **c** are:

$$\text{cov}(P_2, P_1) = cov_{A_{m,1},A_{m,2}} + cov_{e_1e_2} \quad (8)$$

$$\text{cov}(P_2, P_1^2) = cov_{A_{v,1},A_{m,2}}\sigma^2_e\exp\left(0.5\sigma^2_{A_{v,1}}\right) + r_e\sigma_{e_1}\exp\left(0.25\sigma^2_{A_{v,1}}\right)\sigma_{e_2}cov_{A_{m,1},A_{v,1}} \quad (9)$$

Figure 1: The relationship between P_2 and P_1 either assuming no genetic correlation ($r_{A_{v,1},A_{m,2}} = 0$) between environmental variance in P_1 and the level of P_2 using Monte Carlo simulation (MC_lin) and multiple regression (MR_lin) or assuming a negative genetic correlation ($r_{A_{v,1},A_{m,2}} = -0.5$) between them (MC_nonlin, MR_nonlin) ($\sigma^2_{A_{m,1}} = \sigma^2_{A_{m,2}} = 0.3, \sigma^2_{A_{v,1}} = 0.05$, $\sigma^2_{e_1}\exp\left(0.5\sigma^2_{A_{v,1}}\right) = 0.7$, $\sigma^2_{e_2} = 0.7$, $r_{A_{m,1},A_{m,2}} = 0.1, r_{A_{m,1},A_{v,1}} = 0, r_e = 0$).



Monte Carlo simulation: We used Monte Carlo simulation to check the fit of the derived multiple regression equation 3. We simulated 1000 replicates according to the genetic models in equations 1 and 2.

Piglet data: We used 32,450 birth weight records from 2129 litters from 813 sows. Piglets were crossbred whereas the sows were purebred Landrace pigs from TOPIGS. Piglet survival rate was expressed as the proportion of piglets born alive per litter.

Statistical analysis: We analyzed piglet birth weight, its V_E and piglet survival rate with a double hierarchical generalized linear model in ASReml (Rönnegård et al. (2010); Felleki et al. (2012)). We assumed homogeneous residual variance for piglet survival rate and no residual correlation between the three traits. Fixed effects fitted were herd-year-season and parity of the sow for all three traits, and in addition sex of the piglet was fitted for piglet birth weight and V_E of piglet birth weight and sex ratio were fitted for piglet survival rate. The random effects included were maternal genetic effects for all three traits (trivariate normally distributed) and litter effects for birth weight and its V_E (bivariate normally distributed). Litter effects were not fitted for survival rate because survival rate was already at litter level. Models with permanent sow effects gave variance components at the boundaries. Therefore, permanent sow effects were excluded from the model.

The DHGLM was run for 12 ASReml runs to update the weights and the response variable for V_E when estimated parameters had converged (sum of relative squared deviations between current and previous run was smaller than $1.0 * 10^{-4}$). Subsequently, these estimates of parameters were used to predict the regression of piglet survival rate on birth weight and were compared to the averages of survival rate for each of 20 bins of equal numbers of piglet birth weight records. Because the statistical model for piglet weight and its V_E included litter effects, the litter variances and covariances were added to the genetic variances and covariances in equations 5, 6 and 7.

Results and Discussion

Monte Carlo simulation: Figure 1 shows that the relationship between P_1 and P_2 is curvilinear when the genetic correlation $r_{A_{v,1},A_{m,2}}$between V_E of P_1 and the mean of P_2 is negative. This graph shows that equation 3 predicts the relationship between P_1 and P_2 (R^2>0.99) very accurately. When both $r_{A_{v,1},A_{m,2}}$ and r_e are zero, the relationship is linear as expected ($b_2 = 0$), otherwise it has a minimum if $b_2 > 0$ or a maximum if $b_2 < 0$. The position of the minimum/maximum can be derived from equation 3, positive (negative) values for example implying that it is above (below) the current population mean of P_1. The analytical expressions can be used to derive properties about the optima of traits provided non-linear relationships between traits are due to genetic variance in V_E.

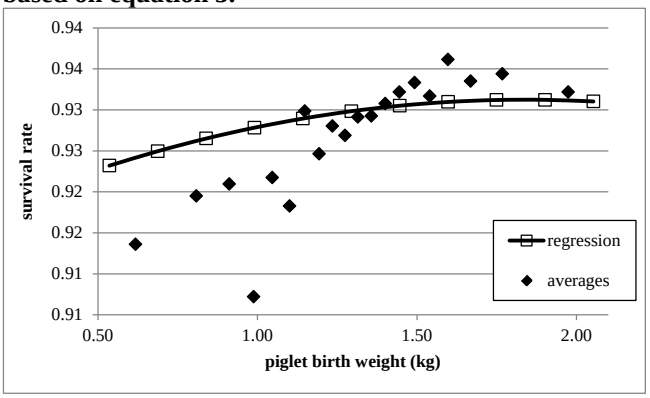
Table 1. Estimates of genetic variance for birth weight (BW), its environmental variance V_E (BWv) and survival rate (SU), and the genetic correlations between the three traits.

Parameter	Estimate	(se)
Genetic variance in birth weight (BW)	0.016	0.002
Genetic variance in V_E of birth weight (BWv)	0.053	0.008
Genetic variance in survival rate (SU)	0.001	0.0002
Genetic correlation BW, BWv	0.508	0.080
Genetic correlation BW, SU	0.116	0.097
Genetic correlation BWv, SU	-0.179	0.121

Piglet data: Table 1 shows the genetic variances and genetic correlations for piglet birth weight and survival rate. The genetic correlation between mean piglet birth weight and survival is seen to be slightly positive, whereas that between V_E of birth weight and survival rate is negative. This implies that higher birth weight and lower V_E leads to a genetic increase in survival rate. In other words, more uniform litters have higher survival rates. Figure 2 shows the regression using equation 3 (including the litter effects) in comparison to the average survival rates per bin of individual birth weights. The regression based on the genetic parameters, partly describes the non-linear relationship between survival rate and birth weight. The relatively poor fit on the right side of the curve is probably due to the assumption that the residual correlation between piglet birth weight and survival rate is zero. Assuming a small residual correlation, e.g. 0.02, already improves the fit from R^2=0.70 to R^2=0.72, while a linear regression line has R^2=0.68. In accordance with Sölkner and Fuerst-Waltl (1996), environmental or residual correlations can greatly dilute the non-linearity of the genetic relationship so that the phenotypic relationship is still almost linear. From a biological point of view, the non-linear relationship between piglet birth weight and survival rate can be explained by the fact that extremely small piglets decrease survival rate, as also do very large piglets but to a lesser extent. Van der Lende and De Jager (1991) reported that the optimal birth weight was 1.67kg, similar to that for this study, i.e. 1.84 kg.

Implications: The example for pigs shows there is a non-linear relationship between traits caused by a genetic covariance between V_E of birth weight and survival rate. There may also be such non-linear relationships for other pairs of traits, especially for those with an optimum, which exhibit genetic variance in V_E. Equation 3 can also be extended to prediction of response to selection and to determining optimum trait values. The existence of genetic and phenotypic optima may require frequent updating of genetic parameters because the population mean is being changed by selection.

Figure 2: The phenotypic relationship between survival rate at litter level and individual piglet birth weight (kg) based on averages per bin and regression based on equation 3.



Conclusion

Non-linear relationships between traits can occur when V_E of one trait is genetically correlated to the mean of others. An analytical expression was derived that can be used to explore the optimum values of traits or to predict response to selection. The theory was illustrated with the relationship between piglet birth weight and survival rate, which was shown to be non-linear. The multiple regression framework developed can help in developing optimal breeding strategies utilizing genetic variance in V_E.

Acknowledgement

The first author acknowledges financial support from the Dutch Ministry of Economic Affairs, Agriculture, and Innovation (Public-private partnership "Breed4Food" code KB-12-006.03-004-ASG-LR and KB-12-006.03-005-ASG-LR).

Literature Cited

Falconer, D. S. and Mackay, T. F. C. (1996). Introduction to quantitative genetics. 4th ed. Pearson Education Limited, Essex, United Kingdom.

Felleki, M. Lee, D., Lee, Y., et al. (2012). Genet. Res. 94:307-317.

Hill, W. G., and Mulder, H. A. (2010). Genet. Res. 92:381-395.

Kapell, D. N. R. G., Ashworth, C. J., Knap, P. W. (2011). Livest. Sci. 135: 215-224.

Mulder, H. A., Bijma, P., and Hill, W.G. (2007). Genetics 175: 1895-1910.

Mulder, H. A., Bijma, P., and Hill, W.G. (2008). Genet. Sel. Evol. 40: 37-59.

Rönnegård, L., Felleki, M., Fikse, F., et al. (2010). Genet. Sel. Evol. 42:8.

SanCristobal-Gaudy, M., Bodin, L., Elsen, J.M., et al. (2001). Genet. Sel. Evol. 33 : 249-271.

Sölkner, J. and Fuerst-Waltl, B. (1996). Interbull bull. 12:226-233.

Van der Lende, T., and De Jager, D. (1991). Livest. Prod. Sci. 28:73-84.