

# Response to mass selection when the genotype by environment interaction is modelled as a linear reaction norm

Rebecka KOLMODIN<sup>a\*</sup>, Piter BIJMA<sup>b</sup>

<sup>a</sup> Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences,  
P.O. Box 7023, 750 07 Uppsala, Sweden

<sup>b</sup> Animal Breeding and Genetics Group, Department of Animal Sciences,  
Wageningen University, P.O. Box 338, 6700 AH Wageningen, The Netherlands

(Received 16 June 2003; accepted 19 February 2004)

**Abstract** – A breeding goal accounting for the effects of genotype by environment interaction ( $G \times E$ ) has to define not only traits but also the environment in which those traits are to be improved. The aim of this study was to predict the selection response in the coefficients of a linear reaction norm, and response in average phenotypic value in any environment, when mass selection is applied to a trait where  $G \times E$  is modelled as a linear reaction norm. The optimum environment in which to test the selection candidates for a given breeding objective was derived. Optimisation of the selection environment can be used as a means to either maximise genetic progress in a certain response environment, to keep the change in environmental sensitivity at a desired rate, or to reduce the proportion of animals performing below an acceptance level. The results showed that the optimum selection environment is not always equal to the environment in which the response is to be realised, but depends on the degree of  $G \times E$  (determined by the ratio of variances in slope and level of a linear reaction norm), the correlation between level and slope, and the heritability of the trait.

**mass selection / selection response / reaction norm / genotype by environment interaction**

## 1. INTRODUCTION

Genotype by environment interaction ( $G \times E$ ) is becoming increasingly important due to the globalisation of animal breeding. With  $G \times E$ , the phenotypic expression of a trait in different environments, such as countries, climatic zones or production systems, is genetically not the same trait. In such cases, the breeding goal should define not only the traits but also the environment in which those traits are to be improved.

\* Corresponding author: Rebecka.Kolmodin@hgen.slu.se

The reaction norm model, where the phenotype is described as a continuous function of an environmental variable [14] is useful for studying  $G \times E$ , especially when phenotypes change gradually over an environmental scale [8], *e.g.*, production level or a climatic variable. In the range of environments normally encountered by a population of domestic animals, it is often reasonable to assume that reaction norms are linear functions of the environment, as has been found for milk production traits and fertility in dairy cattle [2, 11].

In animal breeding, a substantial research effort has been devoted to data analysis using reaction norm models or statistically similar random regression models [2, 7, 11], whereas a limited effort has been made for the optimisation of breeding programmes for those situations. Kirkpatrick and Bataillon [10] have derived equations for the maximisation of selection response in the phenotypic value in a specified environment to mass selection on a trait affected by  $G \times E$ . Their approach was to derive optimum index weights for observations recorded in different environments, modelling a covariance function without any assumptions of the shape of the reaction norm. Another approach is to model  $G \times E$  as a linear reaction norm. The advantage of this model is that selection response can be predicted not only in the phenotypic expression in any environment, but also in the environmental sensitivity of the trait (robustness or responsiveness to changes in the environment, the slope of a linear reaction norm). Selection response in reaction norm coefficients has previously been described in terms of a selection gradient, expressing the covariance between the coefficients and fitness [8, 9].

The objective of this study was to describe selection response of a trait affected by  $G \times E$  in terms of the selection index theory. Equations were derived for the prediction of genetic change in reaction norm coefficients depending on the environment in which the animals were tested. The response in the average phenotypic value in any environment can be calculated knowing the genetic change in reaction norm coefficients. With mass selection and  $G \times E$ , the variables available for the optimisation of selection response are selection intensity and the selection environment, which affect the accuracy of selection. The focus of this study was to derive prediction equations to find the optimum environment in which to test the selection candidates for a given breeding objective.

## 2. METHODS AND RESULTS

The selection index theory was combined with a reaction norm model and the Bulmer effect. Equations will be derived for the prediction of genetic change in reaction norm parameters at equilibrium genetic parameters.

From these equations, other equations will be derived in order to find the optimum selection environment for (i) maximising genetic progress in a certain environment, (ii) keeping the change in environmental sensitivity at a desired rate, or (iii) reducing the proportion of animals performing below an acceptable level. The equations were the main results of this study. The results are illustrated in connection to their derivation in the methods and results section. The implications of the results will be discussed in the discussion section.

### 2.1. The linear reaction norm model

The model was a linear reaction norm function for a single trait. In the following, the intercept and linear coefficient of a linear reaction norm will be referred to as level and slope. The phenotype of an individual  $j$  in an environment  $k$  was modelled as

$$y_{jk} = b_0 + b_1 x_k + a_{0j} + a_{1j} x_k + e_{0j} + e_{1j} x_k \quad (1)$$

where  $y_{jk}$  is the phenotypic value,  $b_0$  is the population average level (intercept) in the average environment,  $b_1$  is the population average slope,  $x_k$  is the effect of environment  $k$  on the phenotype,  $a_{0j}$  and  $a_{1j}$  are the true breeding values for level and slope, respectively, and  $e_{0j}$  and  $e_{1j}$  are the environmental (residual) effects on level and slope, respectively. The term  $b_0 + b_1 x_k$  represent the population average reaction norm. The intercept  $b_0$  is positioned in the average environment so that  $E(x) = 0$ , and breeding and environmental values are expressed as deviations from the average reaction norm. As is common in animal breeding, covariances between residuals of different individuals and covariances between breeding values and residuals were assumed to be zero. If there are reasons to assume that this is not true, the model could be extended by the inclusion of an effect of a common environment, *e.g.*, for individuals of the same litter.

The slope of a linear reaction norm is a measure of sensitivity towards environmental change, which can be treated as a trait of the animal [8]. Genetic variation for the trait environmental sensitivity results in  $G \times E$  and a genetic correlation  $<1$  between phenotypic values of another trait measured in two different environments.

Note that both the genetic and the environmental effects are assumed to be linear functions of the environmental value. Consequently, also genetic and environmental variances change with the environment. The phenotypic variance

in environment  $k$ ,  $\sigma_{y_k}^2$ , *i.e.*, the variance of equation (1), was

$$\begin{aligned}\sigma_{y_k}^2 &= \sigma_{a_0}^2 + 2x_k\sigma_{a_0a_1} + x_k^2\sigma_{a_1}^2 + \sigma_{e_0}^2 + 2x_k\sigma_{e_0e_1} + x_k^2\sigma_{e_1}^2 \\ &= \mathbf{x}'_k\mathbf{G}\mathbf{x}_k + \mathbf{x}'_k\mathbf{E}\mathbf{x}_k = \mathbf{x}'_k\mathbf{P}\mathbf{x}_k\end{aligned}\quad (2)$$

where  $\mathbf{x}'_k$  is a row vector  $[1 \ x_k]$  of the environment  $k$  and  $\mathbf{G}$ ,  $\mathbf{E}$  and  $\mathbf{P}$  are the genetic, environmental and phenotypic (co)variance matrices, respectively, of the reaction norm parameters.

$$\mathbf{G} = \begin{bmatrix} \sigma_{a_0}^2 & \sigma_{a_0a_1} \\ \sigma_{a_0a_1} & \sigma_{a_1}^2 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} \sigma_{e_0}^2 & \sigma_{e_0e_1} \\ \sigma_{e_0e_1} & \sigma_{e_1}^2 \end{bmatrix} \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} \sigma_{p_0}^2 & \sigma_{p_0p_1} \\ \sigma_{p_0p_1} & \sigma_{p_1}^2 \end{bmatrix}$$

where  $\sigma_{a_0}^2$  and  $\sigma_{e_0}^2$  are the genetic and environmental variances of level,  $\sigma_{a_1}^2$  and  $\sigma_{e_1}^2$  are the genetic and environmental variances of slope, and  $\sigma_{a_0a_1}$  and  $\sigma_{e_0e_1}$  are the genetic and environmental covariances between level and slope, and  $\mathbf{P} = \mathbf{G} + \mathbf{E}$ . The heritability in environment  $k$  becomes  $x'_k\mathbf{G}x_k/x'_k\mathbf{P}x_k$ .

## 2.2. Population parameters

To illustrate the theoretical results that will be derived in the following sections, a Fortran 90 deterministic simulation programme was written. The infinitesimal model was assumed. Input values for the simulation were genetic and environmental parameters for the base population. The base population total phenotypic variance was set to 1.0 in the environment of the intercept of the reaction norm. Due to the variance of the slope, the genetic, environmental and, consequently, also the phenotypic total variance changed with the environment. The genetic and phenotypic total variances in three environments are shown in Table I. When there was no correlation between level and slope, the variance increased symmetrically with increasing distance from the environment of the intercept. With a non zero correlation between level and slope, the variance changed asymmetrically at the two sides of the environment of the intercept.

To illustrate varying degrees of  $\mathbf{G} \times \mathbf{E}$ , the genetic correlation between the trait expressed in the environment of the intercept and an environment deviating 1 SD was set to 0.95, 0.80, and 0.60, corresponding to ratios of the genetic variances of slope and level of 0.11, 0.56, and 1.77, respectively, assuming the genetic and environmental correlations between level and slope are zero (Tab. I). Two other values,  $-0.4$  and  $0.4$ , were also studied for the correlation between level and slope. The correlation between level and slope had little

**Table I.** The correlation between the level and slope ( $r$  (level, slope)), genetic and environmental correlations assumed equal), variance ratio between slope and level, genetic correlation between the expression of a trait in the environment of the intercept and an environment deviating 1 SD from the environment of the intercept ( $r_g$  ( $x = 0, x = 1$ )), and the base population total genetic variance ( $\sigma_G^2$ , assuming  $h^2 = 0.5$  in the environment of the intercept) and total phenotypic variance ( $\sigma_P^2$ ) in the environment of the intercept ( $x = 0$ ) and an environment deviating 1 SD from the environment of the intercept ( $x = -1$  and  $x = 1$ ) assuming  $h^2$  was constant over the range of environments.

$r$ (level, slope)	-0.4			0			0.4		
Variance ratio	0.11	0.56	1.77	0.11	0.56	1.77	0.11	0.56	1.77
$r_g$ ( $x = 0, x = 1$ )	0.94	0.76	0.55	0.95	0.8	0.6	0.97	0.84	0.65
$\sigma_G^2 x = -1$	0.82	1.05	1.65	0.56	0.78	1.39	0.29	0.51	1.12
$\sigma_G^2 x = 0$	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$\sigma_G^2 x = 1$	0.29	0.51	1.12	0.56	0.78	1.39	0.82	1.05	1.65
$\sigma_P^2 x = -1$	1.38	1.83	3.04	1.11	1.56	2.77	0.84	1.29	2.50
$\sigma_P^2 x = 0$	1	1	1	1	1	1	1	1	1
$\sigma_P^2 x = 1$	0.84	1.29	2.50	1.11	1.56	2.77	1.38	1.83	3.04

influence on the genetic correlation between the trait expressed in the environment of the intercept and an environment deviating 1 SD when the variance in the slope was small (a small variance ratio, little  $G \times E$ ), but a larger influence with a larger variance in the slope (larger variance ratio, more  $G \times E$ ) (Tab. I). The genetic and environmental correlations between level and slope were assumed to be equal.

Two values of base population heritability in the environment of the intercept were studied: 0.2 and 0.5. For the illustration of selection response in level and slope and in Table I, the genetic and environmental variance ratios of slope and level were set to be equal. Then the genetic and environmental variances were affected proportionally by the environmental effect and heritability was constant over the environmental range. For the illustration of selection response of the total phenotypic value in a specified environment and of the optimum selection environment, the environmental variance of slope was increased so that heritability decreased with an increasing distance from the intercept, being 5% lower at 1 SD from the intercept.

The Bulmer effect was accounted for by an iterative reduction of the genetic parameters until equilibrium was established (App. A). The equilibrium values were used for calculations of selection response, accuracy of selection, and optimum selection environment. Mass selection with 10% of the males and 10% of the females selected was assumed. The selection response is expressed in phenotypic SD-units of the trait expressed in the environment of the intercept.

## 2.3. Genetic change

### 2.3.1. Genetic change in reaction norm parameters

The genetic change in the reaction norm parameters, level and slope, when selecting on the phenotypic value  $y_k$  is a function of the selection environment,  $k$ . From the regression of the breeding values for level and slope on the phenotypic selection differential in environment  $k$ , it follows that

$$\begin{bmatrix} \Delta a_0 & \Delta a_1 \end{bmatrix} = \mathbf{x}'_k \mathbf{G} i / \sigma_{y_k} \quad (3)$$

where  $\Delta a_0$  and  $\Delta a_1$  are the selection responses in level and slope, respectively,  $i$  is the selection intensity, and  $\sigma_{y_k} = \sqrt{\mathbf{x}'_k \mathbf{P} \mathbf{x}_k}$  is the phenotypic standard deviation in environment  $k$ .

The accuracy of selection is the correlation between the selection criterion (the phenotypic value) and the reaction norm parameter

$$r_{y_k, a_0} = (\sigma_{a_0} + x_k r_{a_0, a_1} \sigma_{a_1}) / \sigma_{y_k} \quad (4)$$

and

$$r_{y_k,a_1} = (r_{a_0,a_1}\sigma_{a_0} + x_k\sigma_{a_1})/\sigma_{y_k} \quad (5)$$

for level and slope, respectively, and where  $r_{a_0,a_1}$  is the genetic correlation between level and slope. The accuracies can be substituted into equation (3) to give  $\Delta a_0 = ir_{y_k,a_0}\sigma_{a_0}$  and  $\Delta a_1 = ir_{y_k,a_1}\sigma_{a_1}$ , which agrees with the classical  $\Delta G = ir_{IA}\sigma_A$  where  $r_{IA}$  is the correlation between the index and the true breeding value and  $\sigma_A$  is the genetic standard deviation [5].

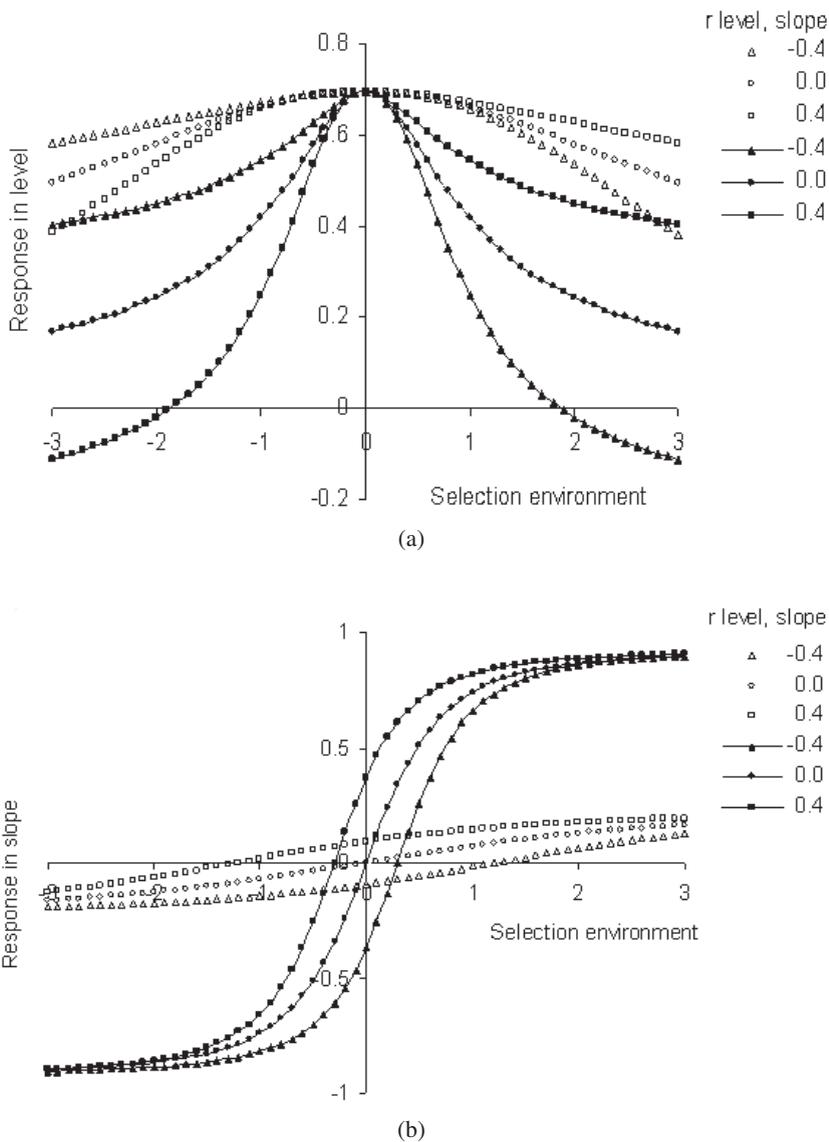
The selection responses in level and slope over a range of environments are shown in Figure 1 for a heritability that is constant over the environmental scale. As expected, the selection environment had a larger effect on the selection response with a higher variance in slope, *i.e.*, a the higher degree of  $G \times E$ . The correlation between level and slope affected the shape and location on the environmental axis of the response curve. The maximum selection response in level was achieved with selection in the average environment. When the correlation between level and slope was positive, the response in level was higher with a selection environment that was better than the average environment than in a selection environment that was worse, and *vice versa* with a negative correlation. Selection response in slope was always the highest with a selection environment that was better than average and with a positive correlation between level and slope. With the lower level of heritability (0.2) the response in both level and slope was smaller and less affected by the environmental value, but the shape of the response curve was the same as with the higher level of heritability (0.5) (not shown). The intermediate level of  $G \times E$  (variance ratio 0.56, not shown) yielded results that were intermediate between the high and low level of  $G \times E$  shown in Figure 1.

### 2.3.2. Genetic change in other environments

The genetic change in a defined environment,  $l$ , is a function of both the environment of selection,  $k$ , and the environment,  $l$ , where the results of selection are expressed. When multiplying equation (3) by  $\mathbf{x}_l$  we get

$$\Delta G_l = \mathbf{x}'_k \mathbf{G} \mathbf{x}_l / \sigma_{y_k} \quad (6)$$

where  $\Delta G_l$  is the genetic change in environment  $l$ ,  $\mathbf{x}_l$  is a column vector  $\begin{bmatrix} 1 \\ x_l \end{bmatrix}$ , and  $\mathbf{x}'_k \mathbf{G} \mathbf{x}_l$  is the covariance between the selection criterion  $y_k$  and the genetic merit in environment  $l$ .



**Figure 1.** Selection response in (a) level and (b) slope of a linear reaction norm as a function of the selection environment, degree of genotype by environment interaction (variance ratio between slope and level, 0.11, open symbols, or 1.77, filled symbols), and correlation between level and slope ( $-0.4$ ,  $0.0$ , or  $0.4$ ). Heritability was constant  $0.5$  over the environmental range. Selection response is expressed in phenotypic SD units per generation and the selection environment as deviation in environmental SD units from the average environment.

The accuracy of selection,  $r_{kl}$ , is the correlation between the selection criterion  $y_k$  and the genetic merit in environment  $l$ ,  $a_{0j} + a_{1j}x_l$ .

$$r_{kl} = \frac{\mathbf{x}'_k \mathbf{G} \mathbf{x}_l}{\sigma_{y_k} \sigma_{A_l}} \quad (7)$$

where  $\sigma_{A_l} = \sqrt{\mathbf{x}'_l \mathbf{G} \mathbf{x}_l}$  is the standard deviation of the genetic merit in environment  $l$ . Note that combining equations (6) and (7) gives  $\Delta G_l = ir_{kl} \sigma_{A_l}$ , as expected.

The genetic correlation,  $r_{g_{1,2}}$ , between the genetic merit in two environments 1 and 2, is

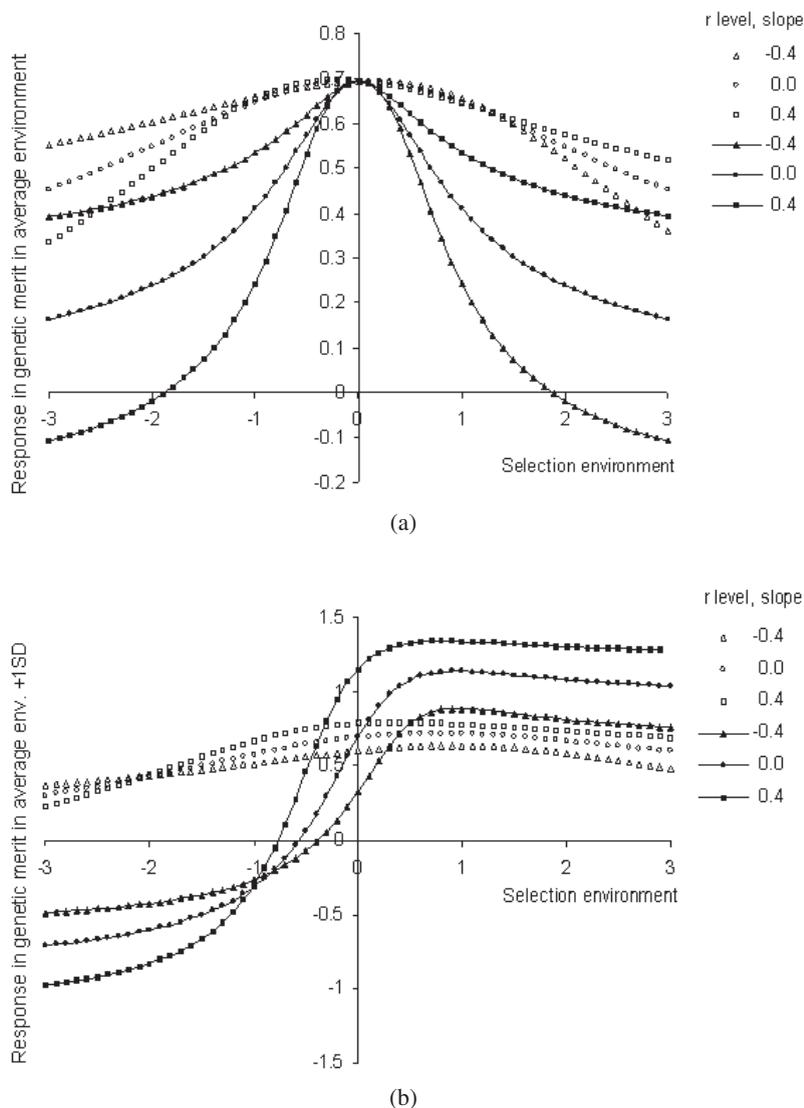
$$r_{g_{12}} = \frac{\mathbf{x}'_1 \mathbf{G} \mathbf{x}_2}{\sqrt{\mathbf{x}'_1 \mathbf{G} \mathbf{x}_1} \sqrt{\mathbf{x}'_2 \mathbf{G} \mathbf{x}_2}} \quad (8)$$

where  $\mathbf{x}'_1 \mathbf{G} \mathbf{x}_2$  is the covariance between the genetic merit in environment 1 and 2, and  $\sqrt{\mathbf{x}'_1 \mathbf{G} \mathbf{x}_1}$  and  $\sqrt{\mathbf{x}'_2 \mathbf{G} \mathbf{x}_2}$  are the standard deviations of the true breeding values in environments 1 and 2, respectively.

The selection response in environment  $l$  as a function of the selection environment  $k$  is illustrated in Figure 2 for a heritability of 0.5 in the environment of the intercept and 5% lower 1 SD from the intercept. Figure 2 shows that maximum gain was achieved when the selection environment was close to the response environment, but not necessarily equal. The correlation between level and slope affected the shape of the response curve. Heritability (0.2 or 0.5 in the environment of the intercept) affected the magnitude of response (not shown). As for the response in level and slope, the selection environment had a larger effect on the selection response when there was a higher variance in slope / and a higher  $G \times E$  (environmental effect with variance ratio  $1.77 > 0.56 > 0.11$ , variance ratio 0.56 not shown). This sensitivity of the response to the selection environment was asymmetric when level and slope were correlated, *i.e.*, the cost of a sub-optimal selection environment depended on which side of the optimum environment was the actual selection environment. With a constant heritability the maximum genetic gain was achieved when the selection environment was equal to the response environment (not shown).

### 2.3.3. Maximising genetic progress

A breeding goal that defines a goal trait and the environment in which the trait is to be improved can be expressed as the expected phenotypic value,  $y_{jl}$ , in environment  $l$  where the animals are expected to perform in the future,



**Figure 2.** Selection response in genetic merit in (a) the average environment and (b) an environment deviating +1 SD unit from average as a function of the selection environment, degree of genotype by environment interaction (variance ratio between slope and level, 0.11, open symbols, or 1.77, filled symbols), and correlation between level and slope ( $-0.4$ ,  $0.0$ , or  $0.4$ ). Heritability was  $0.5$  in the average environment and  $5\%$  lower  $1$  SD unit from average. Selection response is expressed in phenotypic SD units per generation and the selection environment as deviation in environmental SD units from the average environment.

$E(y_{jl}) = a_{0j} + a_{1j}x_l$ . The selection criterion is the phenotypic value,  $y_{jk}$ , in the selection environment  $k$ . The optimum selection environment,  $k_{opt}$ , that gives the highest  $\Delta G_l$  is found by solving  $\partial \Delta G_l / \partial x_k = 0$ . The derivation is shown in Appendix B.

For the general case,

$$x_{k_{opt}} = \mathbf{x}'_l \mathbf{G} \begin{bmatrix} \sigma_{p_0 p_1} \\ -\sigma_{p_0}^2 \end{bmatrix} \left/ \mathbf{x}'_l \mathbf{G} \begin{bmatrix} -\sigma_{p_1}^2 \\ \sigma_{p_0 p_1} \end{bmatrix} \right. \quad (9)$$

When there is neither a genetic nor an environmental correlation between level and slope the expression simplifies to  $x_{k_{opt}} = \frac{\sigma_{p_0}^2 \sigma_{a_1}^2 x_l}{\sigma_{a_0}^2 \sigma_{p_1}^2} = \frac{h_1^2}{h_0^2} x_l$ , when  $h_0^2$  is defined as  $\sigma_{a_0}^2 / \sigma_{p_0}^2$ , and  $h_1^2$  as  $\sigma_{a_1}^2 / \sigma_{p_1}^2$ . When the genetic and environmental variance ratios are equal so that heritability is constant over the environmental range  $x_{k_{opt}}$  simplifies to  $x_{k_{opt}} = x_l$ . Note that the latter result agrees with the classical expression for correlated response,  $CR_l = ih_k r_{g_{k,l}} \sigma_{A_l}$  [6]. At a constant heritability, the response in trait expression in environment  $l$ ,  $CR_l$ , to selection on trait expression in environment  $k$  is maximised when the genetic correlation between trait expression in the two environments,  $r_{g_{k,l}}$  is unity, *i.e.*, the environments are equal.

Equation (9) shows that the optimum selection environment is a function of the equilibrium genetic parameters. The equilibrium genetic parameters depend on the efficiency of selection which, in turn, depends on the selection environment. Therefore the optimum selection environment depends on the environment in which the equilibrium genetic parameters have been obtained. To find the equilibrium optimum selection environment, an optimum selection environment was first calculated using equation (9) and base population parameters. Then the equilibrium genetic parameters in this environment were calculated and used to calculate a new optimum environment. This procedure was iterated until the optimum environment did not change between rounds. In our simulation, equilibrium was reached after 1–5 iterations with a convergence criterion of less than 0.1% change in an optimum environment between rounds.

The equilibrium optimum selection environment for different response environments is shown in Figure 3a. The optimum selection environment was not equal to the response environment and their relationship was not precisely linear. With more G×E (higher variance ratio) and/or a flatter heritability function (2% difference between heritability in average and deviating environment, not shown) the optimum selection environment was closer to the response environment, than with the lower variance ratio and/or a more curved heritability function.

To investigate the need for an iterative solution of  $x_{k_{opt}}$ ,  $x_{k_{opt}}$  was calculated also using base generation parameters, and the results were compared to the solution using equilibrium parameters. Figure 3b shows that the optimum selection environment calculated from base population parameters was close to the equilibrium optimum selection environment. With a higher variance in slope, the difference between the optimum selection environment calculated using the base population or equilibrium parameters was even smaller than that shown in Figure 3b. Thus the calculation of  $x_{k_{opt}}$  directly from base generation parameters is probably sufficient for practical applications.

### 2.3.4. Keeping environmental sensitivity constant

Selection for high phenotypic value in combination with a continuous improvement of the environment and in the presence of  $G \times E$  has been shown to increase the environmental sensitivity of a population [12]. In modern animal production systems, genetic improvement and improvements of feeding and management occur simultaneously. A restriction of the genetic change in environmental sensitivity can be included in the breeding objective to avoid that the animals become increasingly sensitive to, *e.g.*, feed quality or disease.

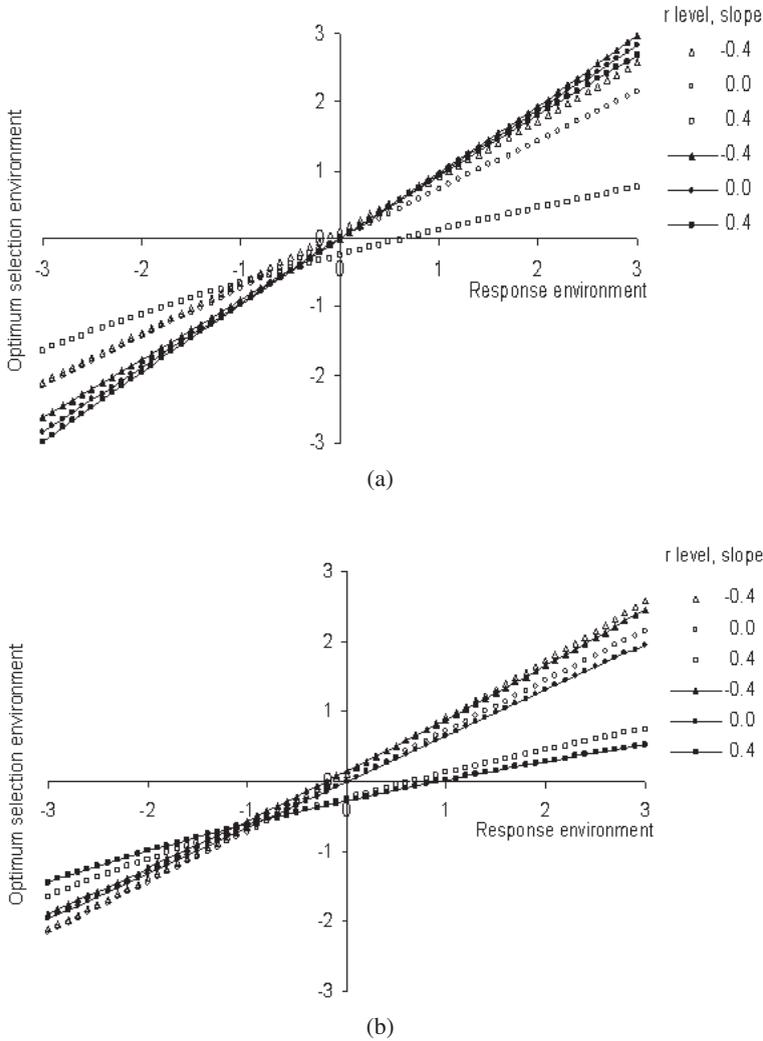
Since the genetic change in slope is a function of the selection environment,  $k$ , it is possible to find a selection environment that gives no change in slope. This environment can be found numerically from a plot of  $\Delta a_1$  over  $x_k$  (Fig. 1b). It can also be found analytically as the environment where there is no covariance between the selection criterion and the true breeding value for the slope. From equation (5) we get  $\sigma_{y_{jk}a_1} = \sigma_{a_0a_1} + x_k\sigma_{a_1}^2$ . Equating this to zero gives  $x_{k_{\Delta a_1=0}} = -\sigma_{a_0a_1}/\sigma_{a_1}^2$  or, rearranged,  $x_{k_{\Delta a_1=0}} = -r_{a_0,a_1}\sigma_{a_0}/\sigma_{a_1}$ . In words, the selection environment that gives no change in slope is “minus the correlation between level and slope, scaled by the ratio of the genetic standard deviations of level and slope”. Note that, if there is a positive correlation between level and slope,  $x_{k_{\Delta a_1=0}}$  must be negative, *i.e.*, worse than the average environment.

### 2.3.5. Desired gains in level and slope

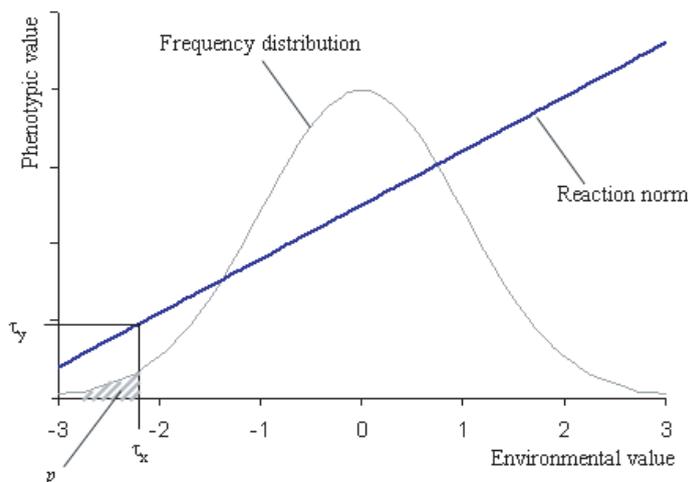
Keeping the slope constant, as in 2.3.4., is a special case of desired gains for level and/or slope. For the general case, a desired ratio,  $\delta$ , between the change in slope and level can be specified. From equation (3), it follows that

$$\delta = \frac{\Delta a_1}{\Delta a_0} = \frac{\sigma_{a_0a_1} + x_k\sigma_{a_1}^2}{\sigma_{a_0}^2 + x_k\sigma_{a_0a_1}}$$

Solving for  $x_{k_\delta}$ , the selection environment where this



**Figure 3.** (a) The equilibrium optimum selection environment for a given response environment as a function of the degree of genotype by environment interaction (variance ratio between slope and level, 0.11, open symbols, or 1.77, filled symbols), and correlation between level and slope ( $-0.4$ ,  $0.0$ , or  $0.4$ ). Heritability was  $0.5$  in the average environment and  $5\%$  lower  $1$  SD unit from average. The response environment is expressed as deviation in environmental SD units from the average environment. (b) The equilibrium optimum selection environment for a given response environment (open symbols) and the optimum selection environment calculated from base population parameters (filled symbols). The variance ratio was  $0.11$  and heritability, correlation between level and slope, and environmental values as in (a).



**Figure 4.** Illustration of the proportion,  $p$ , of animals performing below a defined acceptable phenotypic level,  $\tau_y$ , and the environmental truncation point,  $\tau_x$ , below which the phenotypic value falls below the threshold.

ratio is achieved, gives  $x_{k_\delta} = (\sigma_{a_0 a_1} - \delta \sigma_{a_0}^2) / (\delta \sigma_{a_0 a_1} - \sigma_{a_1}^2)$ . If no change in slope is desired, then  $\delta = 0$  and  $x_{k_\delta} = -\sigma_{a_0 a_1} / \sigma_{a_1}^2$ , as shown in the previous section.

### 2.3.6. Reducing the proportion of animals performing below an acceptable level

The main interest might not be in all cases to increase the level or to reduce the slope. For traits related to animal welfare, it may be more relevant to avoid poor animal performance. Besides the suffering of the animal and potential costs and worries for the farmer, poor animal performance may reduce consumer acceptance of the production system. Therefore, an objective may be to reduce the proportion,  $p$ , of animals that perform below an acceptable phenotypic level, say  $\tau_y$  (Fig. 4).

Consider the population average reaction norm,  $E(y_k) = b_0 + b_1 x_k$ . The animals in an environment  $x_k < \tau_x = (\tau_y - b_0) / b_1$  are expected to perform below the acceptance level  $\tau_y$ , where  $\tau_x$  is the environmental truncation point, below which  $y_{k|\tau_x}$  falls below the phenotypic threshold  $\tau_y = b_0 + b_1 \tau_x$ . The breeding objective,  $H$ , is to reduce the proportion  $p$ , thus  $H = -p$ , which is a non-linear function of level and slope. The selection criterion is the phenotypic value in the selection environment  $k$ . The problem is to find the selection environment

that yields the largest reduction of  $p$ . The first step is to find the optimum weights,  $v_0$  and  $v_1$ , for level and slope. The breeding objective can be approximated by a linear function of level and slope,  $H \approx v_0 a_0 + v_1 a_1$ . Values for  $v_0$  and  $v_1$  are obtained by taking the partial derivatives of  $-p$  with respect to level and slope, and calculating the values of the derivatives at the current population mean ( $b_0$  and  $b_1$ ). The relationship between reaction norm parameters and  $p$  arises solely *via*  $\tau_x$ , so that partial derivatives can be obtained *via*  $\tau_x$

$$v_{0/1} = \frac{\partial(-p)}{\partial\tau_x} \frac{\partial\tau_x}{\partial b_{0/1}}. \quad (10)$$

The partial derivatives are  $\partial(-p)/\partial\tau_x = -\varphi(\tau_x)$ ,  $\partial\tau_x/\partial b_0 = -1/b_1$ , and  $\partial\tau_x/\partial b_1 = (b_0 - \tau_y)/b_1^2$ , where  $\varphi$  is the density function of the distribution of the environmental values. The resulting weights are  $v_0 = \varphi(\tau_x)/b_1$  and  $v_1 = \varphi(\tau_x)(\tau_y - b_0)/b_1^2$ . Since the scaling of the weights does not affect the selection response,  $v_0$  and  $v_1$  may be divided by  $\varphi(\tau_x)/b_1$  to give  $v_0 = 1$  and  $v_1 = (\tau_y - b_0)/b_1 = \tau_x$ . Thus, the linearised breeding objective is  $H = a_0 + a_1 \tau_x$ . Note that this breeding goal corresponds to genetic gain in environment  $\tau_x$ , which shows that maximising the reduction of  $p$  is the same as maximising the selection response in environment  $\tau_x$ . Consequently, the environment of selection that maximises  $-\Delta p$  can be found by substituting  $x_l$  by  $\tau_x$  in equation (9).

The result that minimising  $p$  is the same as maximising the genetic change in environment  $\tau_x$  can be derived also in a more intuitive manner as follows: The intercept of a reaction norm can be defined in any environment without changing the reaction norm itself (see the discussion section). If the intercept is defined at the environmental truncation point  $\tau_x$ , instead of in the average environment, the rotation of the reaction norm (change in slope) around this intercept does not affect  $\tau_x$ . For small changes in level and slope, the shift of the truncation point  $\tau_x$ , and hence, the change in  $p$  is determined entirely by the change in level/phenotypic value in the environment  $\tau_x$ . This argument extends directly to higher-order reaction norms, meaning that also for quadratic and cubic reaction norms, maximising  $-\Delta p$  is identical to maximising  $\Delta G$  in the environment  $\tau_x$ . Equation (9) will need extension to higher-order reaction norms, though.

### 3. DISCUSSION

The described theory is a tool for understanding the effects of  $G \times E$  on selection response as a function of the environment of selection and the environment where selection response is to be realised. In the case of mass selection,

the derived equations can be used to solve the problem stated by Falconer [4] whether to select under the conditions in which the breed is required to live, or if it would be better to select “under some other conditions, for example under conditions more favourable for the expression of the desired character”. The old thesis that “the character required is best selected for under environmental conditions which favour its fullest expression” [6] may suggest that one should select in the environment, which maximises heritability. Equation (9), however, shows that the optimum selection environment depends on, besides the environment where selection response is to be realised, three factors: the degree of  $G \times E$  (determined by the ratio of variances in slope and level), the correlation between level and slope, and the heritability of the trait. The results show that the optimum selection environment is neither always equal to the environment in which selection response is to be realised, nor to the environment where heritability is the highest (Fig. 2).

The numerical illustrations show the relations between genetic and environmental parameters and selection response. The ratio between the variances in slope and level determines the magnitude of the environmental effect on selection response. With a high variation in slope, *i.e.*, a high degree of  $G \times E$ , the environmental value has a large effect on selection response. For prediction of selection response for a few generations of selection, which is often of interest in animal breeding, the prediction equations derived in this paper are relevant even when the conditions of the infinitesimal model do not hold exactly.

The equations presented here enable the breeder to quantify the value of data recording in a particular environment, and also to quantify the losses due to data recording in sub-optimal environments. The small difference between the optimum selection environment calculated using a base population or equilibrium parameters suggests that the optimum environment can, for convenience, be calculated from base population parameters without losing much in genetic gain.

It is important to realise that with a reaction norm the correlation between level and slope depends on the location of the intercept on the environmental scale [13]. Consider the phenotypic reaction norm  $y_{jk} = b_0 + b_1 x_k + p_{0j} + p_{1j} x_k$ , where  $p_{0j}$  and  $p_{1j}$  are the phenotypic values of level and slope of individual  $j$ . A non-zero correlation between  $p_{0j}$  and  $p_{1j}$  can be removed by redefining the intercept in such a way that the phenotype  $y_{jk}$  at the new intercept has no covariance with  $p_{1j}$ . The environment where this occurs is found by solving  $\sigma_{y_{jk}, p_{1j}} = \sigma_{p_{0j} + p_{1j} x_k, p_{1j}} = 0$ , which gives  $x_k = -\sigma_{p_0 p_1} / \sigma_{p_1}^2$ . If the original intercept is at  $x_k = 0$ , then the environment where level and slope are not correlated is at  $x_k = -\sigma_{p_0 p_1} / \sigma_{p_1}^2$ . Hence, the new value of the intercept is

$b_0^* = b_0 - b_1\sigma_{p_0p_1}/\sigma_{p_1}^2$ , and the slope is the same as for the original reaction norm. Thus, to remove a correlation between level and slope, a positive correlation requires shifting the intercept an amount  $\sigma_{p_0p_1}/\sigma_{p_1}^2$  downwards on the environmental scale, whereas a negative correlation requires shifting the intercept the same amount upwards on the environmental scale. A zero correlation can be achieved simultaneously on the genetic and environmental levels only when  $r_{a_0a_1} = r_{e_0e_1}$  and heritability is constant over the environmental scale.

Due to the dependency of the correlation between level and slope on the location of the intercept on the environmental scale, estimates of the correlation between level and slope should be interpreted with caution. To facilitate the comparison of results from different reaction norm studies, authors should report where the intercept was defined. We agree with the recommendation of van Tienderen and Koelewijn [13] to define the intercept in the average environment.

For species where the breeding population comprises almost the entire population, such as dairy cattle, the selection environment cannot simply be chosen. With progeny testing, the accuracy is generally rather high, so that the optimum environment for data recording will depend less on heritability and more on the genetic correlation with the breeding goal. Equation (9) does not apply to progeny testing. Instead one can argue that the optimum environment to record progeny information will be close to the environment where the population has to perform in the future. Then, when progeny testing of sires is combined with a nucleus herd for selection of cows, a near optimum solution would be to have the nucleus animals in an environment as indicated by equation (9), and the progeny in an environment similar to the expected future environment of production animals.

A problem arises if the environment that would give the maximum selection response is not an ethically acceptable environment for testing animals. Testing animals in the optimum environment for improving disease resistance, or to reduce the proportion of animals performing very poor, might be comparable to doing a challenge test. A major benefit in animal health and farmer economy in coming generations could justify such a test, though for practical as well as ethical reasons it should be done only with a small part of the population.

In order to be more applicable to practical animal breeding, the theory presented here needs to be extended to include multiple trait evaluations, more complicated selection schemes, such as selection on BLUP procedure breeding values, and non-linear reaction norm functions. Here we have assumed that animals are in a single environment, whereas a distribution of animals across environments is more realistic. Such a model would be useful for searching for

an optimum distribution of selection candidates over different environments, to optimise the use of testing capacity.

## ACKNOWLEDGEMENTS

This work was initiated during Rebecka Kolmodin's stay at Wageningen University from October to December 2002, financed by a Marie Curie Fellowship provided by the European Commission. We thank Birgitta Danell, Hossein Jorjani, Erling Strandberg, and two anonymous reviewers for valuable comments on earlier versions of this manuscript.

## REFERENCES

- [1] Bulmer M.G., The effect of selection on genetic variability, *Am. Nat.* 105 (1971) 201–211.
- [2] Calus M.P.L., Groen A.F., de Jong G., Genotype  $\times$  environment interaction for protein yield in Dutch dairy cattle as quantified by different models, *J. Dairy Sci.* 85 (2002) 3115–3123.
- [3] Cochran W.G., Improvement by means of selection, in: *Proceedings of the 2nd Berkeley Symposium on Mathematics, Statistics and Probability*, Neyman. J (Ed.), University of California Press, Berkeley (1951) pp. 449–470.
- [4] Falconer D.S., The problem of environment and selection, *Am. Nat.* 86 (1952) 293–298.
- [5] Falconer D.S., Mackay T.F.C., *Introduction to quantitative genetics*, 4th edn., Longman Group, Essex, 1996.
- [6] Hammond J., Animal breeding in relation to nutrition and environmental conditions, *Biol. Rev.* 22 (1947) 195–213.
- [7] Hayes B.J., Carrick M., Bowman P., Goddard M., Genotype  $\times$  environment interaction for milk production of daughters of Australian dairy sires from test-day records, *J. Dairy Sci.* 86 (2003) 3736–3744.
- [8] de Jong G., Phenotypic plasticity as a product of selection in a variable environment, *Am. Nat.* 145 (1995) 493–512.
- [9] de Jong G., Bijma P., Selection and phenotypic plasticity in evolutionary biology and animal breeding, *Livest. Prod. Sci.* 78 (2002) 195–214.
- [10] Kirkpatrick M., Bataillon T., Artificial selection on phenotypically plastic traits, *Genet. Res. Camb.* 74 (1999) 265–270.
- [11] Kolmodin R., Strandberg E., Madsen P., Jensen J., Jorjani H., Genotype by environment interaction in Nordic dairy cattle studied using reaction norms, *Acta Agric. Scand., Section A, Anim. Sci.* 52 (2002) 11–24.
- [12] Kolmodin R., Strandberg E., Jorjani H., Danell B., Selection in presence of a genotype by environment interaction: response in environmental sensitivity, *Anim. Sci.* 76 (2003) 375–385.

- [13] van Tienderen P.H., Koelewijn H.P., Selection on reaction norms, genetic correlations and constraints, *Genet. Res. Camb.* 64 (1994) 115–125.
- [14] Woltereck R., Weitere experimentelle Untersuchungen über Artveränderung, speziell über das Wesen quantitativer Artunterschiede bei Daphniden. *Verhandlungen der Deutschen zoologischen Gesellschaft* (1909) 110-172, in: Lynch M., Walsh J.B. (Eds.), *Genetics and analysis of quantitative traits*, Sinauer Associates, Inc. Publishers, Sunderland, MA, USA, 1998.

### APPENDIX A: EQUILIBRIUM GENETIC PARAMETERS

Applying the results of Bulmer [1] the equilibrium genetic parameters were found using the following equations iteratively:

$$\sigma_{a_{0/1}}^2(t+1) = \frac{1}{4}\sigma_{a_{0/1}}^2(t)^* (1 - k_m r_{y_i, a_{0/1}}^2) + \frac{1}{4}\sigma_{a_{0/1}}^2(t)^* (1 - k_f r_{y_i, a_{0/1}}^2) + \frac{1}{2}\sigma_{a_{0/1}}^2(t=0) \quad (11)$$

where  $\sigma_{a_{0/1}}^2(t)$  is the genetic variance in level/slope in generation  $t$ ,  $k_{m/f} = i_{m/f}(i_{m/f} - d_{m/f})$ , where  $i_{m/f}$  is the selection intensity of males/females, respectively, and  $d_{m/f}$  is the standardised deviation of the truncation point from the mean for males/females, respectively,  $r_{y_i, a_{0/1}}$  is the correlation between the phenotype and the true breeding value of level or slope (accuracy, Eqs. (4) and (5)), and  $\frac{1}{2}\sigma_{a_{0/1}}^2(t=0)$  is the Mendelian sampling variance. Note that since the accuracies depend on the selection environment, so do the equilibrium genetic parameters.

The covariance between level and slope,  $\sigma_{a_0 a_1}$ , in generation  $t+1$  is:

$$\sigma_{a_0 a_1}(t+1) = \frac{1}{4}\sigma_{m, a_0 a_1}(t) + \frac{1}{4}\sigma_{f, a_0 a_1}(t) + \frac{1}{2}\sigma_{a_0 a_1}(t=0) \quad (12)$$

where  $\sigma_{m/f, a_0 a_1}$  is the covariance contributed from the males/females, respectively. From Cochran [3] it follows that

$$\sigma_{m/f, a_0 a_1}(t+1) = \sigma_{a_0 a_1}(t) - k_{m/f} \left[ \sigma_{a_0}^2(t) + x_i \sigma_{a_0 a_1}(t) \right] \left[ \sigma_{a_0 a_1}(t) + x_i \sigma_{a_1}^2(t) \right] / \sigma_{y|x_i}^2 \quad (13)$$

where the terms between the square brackets represent the covariances between the phenotype and the true breeding value for level and slope, respectively.

## APPENDIX B: MAXIMISING $\Delta G$

The selection environment that maximises genetic progress in the goal environment is found by solving  $\partial\Delta G/\partial x_i = 0$ . Start from equation (6), in scalar notation  $\Delta G = [\sigma_{a_0}^2 + \sigma_{a_0a_1}(x_k + x_l) + \sigma_{a_1}^2 x_k x_l]i/\sigma_{y_k}$ . The selection intensity is independent of the selection environment and can therefore be ignored. Since the selection environment that maximises  $\Delta G$  is the same environment that maximises  $\Delta G^2$ ,  $\Delta G^2$  can replace  $\Delta G$  to make the calculations easier. To simplify the notation let  $\sigma_{a_0}^2 = a$ ,  $\sigma_{a_1}^2 = b$ ,  $\sigma_{a_0a_1} = c$ ,  $\sigma_{e_0}^2 = d$ ,  $\sigma_{e_1}^2 = e$ ,  $\sigma_{e_0e_1} = f$ ,  $\sigma_{p_0}^2 = g$ ,  $\sigma_{p_1}^2 = h$ ,  $\sigma_{p_0p_1} = j$ ,  $x_l = l$ , and  $x_k = x$ . Then

$$\Delta G^2 = \frac{a^2 + 2acl + c^2l^2 + 2a(bl + c)x + 2cl(bl + c)x + (bl + c)^2x^2}{g + 2jx + hx^2} = \frac{u}{v}.$$

Now take the derivative with respect to  $x$ ,  $\frac{\partial u}{\partial x} = \frac{v\frac{\partial u}{\partial x} - u\frac{\partial v}{\partial x}}{v^2} = 0$ . To find the solution we only need to consider the numerator,  $v\frac{\partial u}{\partial x} - u\frac{\partial v}{\partial x} = 0$ . The partial derivatives needed are  $\partial u/\partial x = 2(c + bl)^2x + 2(c + bl)(a + cl)$  and  $\partial v/\partial x = 2hx + 2j$ .

To solve the resulting second-order equation  $v\partial u/\partial x - u\partial v/\partial x = x^2[j(c + bl)^2 - h(c + bl)(a + cl)] + x[g(c + bl)^2 - h(a + cl)^2] + g(c + bl)(a + cl) - j(a + cl)^2 = 0$  we factorise it into

$$v\partial u/\partial x - u\partial v/\partial x = [(c + bl)x + (a + cl)][(j(c + bl) - h(a + cl))x + g(c + bl) - j(a + cl)].$$

The valid solution comes from equating the second term between square brackets to zero, giving  $x = [j(a + cl) - g(c + bl)]/[j(c + bl) - h(a + cl)]$ . Expressed in original terms, the solution is  $x_{k_{opt}} = \frac{\sigma_{p_0p_1}(\sigma_{a_0}^2 + \sigma_{a_0a_1}x_l) - \sigma_{p_0}^2(\sigma_{a_0a_1} + \sigma_{a_1}^2x_l)}{\sigma_{p_0p_1}(\sigma_{a_0a_1} + \sigma_{a_1}^2x_l) - \sigma_{p_1}^2(\sigma_{a_0}^2 + \sigma_{a_0a_1}x_l)}$ , or in matrix notation  $x_{k_{opt}} = \mathbf{x}'_l \mathbf{G} \begin{bmatrix} \sigma_{p_0p_1} \\ -\sigma_{p_0}^2 \end{bmatrix} / \mathbf{x}'_l \mathbf{G} \begin{bmatrix} -\sigma_{p_1}^2 \\ \sigma_{p_0p_1} \end{bmatrix}$ . That this extreme was indeed the maximum of the function was verified by plotting  $\Delta G$  over  $x_l$ .