2B-101: QUANTIFYING VARIATION IN ENVIRONMENTS WITHIN AND ACROSS HERDS

Report prepared for the

Co-operative Research Centre for High Integrity Australian Pork

By

Li LI and Susanne HERMESCH

Animal Genetics and Breeding Unit (AGBU)

University of New England

Armidale NSW 2351

December 2012

An Australian Government Initiative

Established and supported under the Australian Government's Cooperative Research Centres

Executive Summary

Animals have the ability to respond to differences in the environment, which his called environmental sensitivity. Different genotypes may not respond in the same way to diverse environments leading to genotype by environment interactions. These interactions have implications for breeding programs if animals are selected in one specific environment and progeny are expected to perform well in a wide range of environments. Breeding programs have focused on improving mean performance of pig genotypes and so far have paid little attention to incorporating environmental sensitivity in selection decisions.

A number of factors such as air quality, housing, feed quality and health status define specific environmental conditions a pig may experience during the growth phase affecting its performance. Information about individual environmental factors affecting performance is usually not available on farm and other measures are required to define the quality of the environment a pig has experienced. Pigs raised together in one group or at a similar time are exposed to the same overall environmental conditions. This aspect may be used to define the quality of an environment based on the mean performance of a group of pigs.

Performance records from over 265,000 pigs recorded in nine herds over a ten-year period were used for the analyses. There were genetic links between herds because a proportion of sires had progeny recorded across herds. Methodology and models were developed to quantify variation in environments within and across herds and to evaluate whether there were any breed by environment interactions or sire by environment interactions.

Variation in environments

The variation in environmental conditions was quantified through predictors of the mean performance of a group of pigs that were born in the same herd during a one-month period of time (contemporary group, CG). Alternative models were used to predict the mean performance of each CG. The distribution of mean performance of CGs showed a spread of about 150 g/day for growth rate and a spread of about 5 mm for backfat across herds and years independent of the model used. This considerable spread in mean performance of CG was observed in high-health herds with good overall husbandry and management conditions showing that even in these overall good conditions on farm, considerable environmental variation existed between individual groups of pigs.

A further important aspect is whether this variation in environments observed across herds is also apparent within herds and within years because pigs may be selected within-herds only at a specific point in time. A similar spread of mean performance within herds and within years was found for growth rate and backfat.

Breed by environment interactions

Breeds differed in their responsiveness to variation in environmental conditions with Large White being the most environmentally sensitive or least robust breed for growth rate and backfat. This breed was the leanest breed in comparison to Landrace and Duroc. The most robust breed was Duroc, which had a similar growth rate as Large White but was characterised by a higher backfat in comparison to the other breeds. Overall, these results support the hypothesis that leaner genotypes tend to be less robust and less able to perform consistently across a range of environmental conditions.

Sire by environment interactions

Extensive analyses were performed using random regression models to evaluate sire by environment interactions. For growth rate, sires differed in the response of their progeny to variation in the environment, which was less apparent for backfat. The implications of data structure in regard to the number of progeny per sire and the range in environmental variation per sire for the ability to estimate genetic differences in environmental sensitivity between sires were outlined. This information is useful for pig breeding companies to setup strategic recording procedures for genetic improvement of both, productivity and robustness.

Genetic parameters across environmental trajectory

Heritability estimates differed across the environmental trajectory for growth rate showing slightly higher estimates for inferior environments despite their higher residual and observed variances. The additive genetic variance across the trajectory was based on a function involving the environmental variable and estimates of the intercept and slope of sires. These estimates from the reaction norm model led to higher additive genetic variance for the lower range of the environmental trajectory.

Implications and recommendations

These analyses demonstrated a considerable spread in environments in high-health farms with good husbandry practices. The unadjusted mean performance of a group of pigs showed a similar distribution as the more complex environmental descriptors. This simple environmental descriptor can easily be derived from standard performance records collected on farms to quantify variation in environmental conditions within herds over time. Further research will now focus on developing more precise methodology to measure variation in environmental conditions on farm making use of information about seasonal effects, air quality, disease incidence and overall performance of pigs based on multiple traits. Multiple environmental descriptors were significant for growth rate, highlighting the need to develop an overall environmental index that combines multiple environmental factors.

Sires need to be represented across a wide range of environments. Avenues to increase the spread of progeny of sires across environmental conditions should be explored with breeders in order to improve data structures for future genetic analyses of sire by environment interactions.

Further, it was shown that residual variance was lower in superior environments in particular for growth rate. Therefore, the environmental descriptor may be extended by taking variation between animals within the same environmental into account. It is a further criterion that may capture environmental conditions better and will become available as more and more automatic weighing devices of individual pigs are installed on farms. Precision pork production will be developed further and measures of variation will be used to optimize pork production beyond current on-farm practices.

The methodology developed in this study to estimate sire by environment interactions can be used to estimate breeding values (EBVs) for the intercept and slope of sires. Firstly, these EBVs should be used to select sires with consistently superior performance across environments. This is a long-term breeding goal. Secondly, the variation in EBVs for the intercept and slope can be used to select sires whose progeny are best suited for a specific environment. This information should be used by breeders and producers to better match sires to the environments their progeny are likely to encounter.

Table of Contents

1. Introduction

Selection strategies and manipulation of pig husbandry practices aim to improve performance of pigs. However, the strong focus on productivity has had undesired effects on a number of physiological parameters with the consequence that highly productive pigs have increased difficulty in coping with environmental challenges (Prunier et al. 2010). Environmental constraints are due to multiple factors including pathogen and disease load, air quality, climatic conditions, stocking density and even some characteristics of its direct contemporaries housed in the same group. Usually these environmental parameters are not known on farm, although some of the effects of environmental stressors have been quantified in research projects as reviewed by Black et al. (2001).

Environmental sensitivity is the ability of an organism to change its phenotype in response to the environment. Further, genotype by environment interactions exist when genotypes differ in their response to environmental stimuli. In the past, a trait like growth rate has been defined as a different trait in each environment, which may have been defined by differences in feeding level (Hermesch, 2004). This approach is only feasible when there are few and distinct environments, which can be defined by feeding level for example, and it is not useful when many factors contribute to variation in environmental conditions.

Random regression models so called reaction norm (RN) models have been used to express the performance of a genotype as a function of the environment (Kolmodin et al. 2002). These models require definition of environmental constraints on a continuous scale and first studies are emerging that have investigated the response of sow genotypes in regard to their feed intake during lactation to differences in temperature (Lewis et al. 2010; Bergsma et al., 2011). However, often the environment is not defined via specific parameters and the mean performance of animals raised in the same environment is taken as a proxy for the sum of all environmental factors affecting performance of animals (e.g. Calus and Veerkamp, 2003; Hammami et al., 2009).

Reliable estimates of the parameters of reaction norm models can only be obtained when there is sufficient variation in the independent parameter, a descriptor of the environment, and when progeny of sires are distributed across environments to ensure that sires have a good representation across a wide range of environments. Therefore, existing data sets should be characterised to evaluate their suitability for genetic analyses of genotype by environment interaction using reaction norm models.

2. Methodology

2.1. Review of literature

Genotype by environment interaction

Both animals and plants have the ability to respond to changes in their environment, which is called environmental sensitivity or phenotypic plasticity (Bradshaw, 1965). Different genotypes do not respond in the same way to different environments leading to genotype by environment interactions (GxE). There are two types of GxE:

scaling effects or re-ranking for individuals. As different scaling effects for different traits can also results I n re-ranking of genotypes (Strandberg, 2006), both are important for animal breeding and selection. Two traditional methods to analysis GxE are the interaction-term model and the multiple-trait model. Both methods assume that distinct environments exist. When the production environment can be described as a continuous gradient, reaction norm (RN) models are a viable alternative. Numerous studies on GxE using RN models have reported results for dairy cattle (Kolmodin et al., 2002), beef (Mattar et al., 2011; Pegolo et al., 2011), sheep (Pollott and Greeff, 2004), swine (Hermesch et al., 2006; Knap and Su, 2008) and plants (Wu and Ying, 2001). This report reviews some important aspects of using RN models in the analyses of GxE.

Screening and editing procedures of data

Only records meeting certain criteria were used in analyses. For example, Kolmodin et al. (2002) used a number of inclusion criteria (at least 60 kg protein production, at least 15 days open, calving age from 18 to 48 months, at least one herd-mate in the same herd-year (contemporary group), known international identification of sire of the cow and at least 50 daughters per sire) in their Nordic dairy cattle study on 305 days kg protein production and days open in first lactation. Knap and Su (2008) used inclusion criteria of at least 30 records per herd-year-season (HYS) class, 50 litters per sire of the sows, five records per sire x farm x month class and two records per farm x month classes per sire in a study of litter size in pigs. Hermesch et al. (2006) restricted data to a minimum of ten progeny per sire and a live weight at start of test between 60 and 90 kg in the study of genotype by feeding level interaction in pigs.

Definitions of contemporary groups (CG)

The contemporary group (CG) size is an important factor for the calculation of environmental value. Different CG definitions were used across studies in terms of data size and its features. Some studies have used herd-year (HY) as formation of CG (Kolmodin et al., 2002). Herd was used as CG in some dairy cattle study (Fikse et al., 2003; Strandberg et al., 2009). Pollott and Greeff (2004) used flock, year, sex, age class and paddock as CG in studies on Merino sheep. Herd-year-season (HYS) as CG was used to investigate pig litter size (Knap and Su, 2008) and dairy cattle in New Zealand (Bryant et al., 2007).

Different definitions of CG will result in different data structures regarding to number of sire and genetic links across CGs. A simulation study by Calus et al. (2004) has shown that non-random use of sires, poor genetic links between herds (CGs) and small herd (CG) size had a large impact on the estimated covariance functions, estimated breeding values and calculated environmental parameters. Poor genetic connectedness and more diverse herd (CG) composition resulted in an over-estimate of genetic correlations between different expressions of a trait across environments and resulted in more biased breeding values. In the end, the author recommended the strategy of combining large numbers of animals per herd (CG) as the best possible solution. However, increasing CG size also means mixture of different environments and will disguise the real environmental driver for the animal performance, so a balance between these two factors is needed (Strandberg, 2006).

Definition of environmental descriptors (ED)

The environment is quantified on a continuous scale in RN models instead of being classified into groups in the traditional GxE methods. The environment is not only the space surrounding the animal; it also includes temperature, floor space, air, nutrition, feeding, vaccination etc. As the phenotype is the result of all environmental factors, it is common to use phenotypic average of CG as an environmental descriptor (ED). The average performance of a CG is either calculated as observed mean, by least squares means (Bryant et al., 2007; Kolmodin et al., 2002; Strandberg et al., 2009). Strandberg (2006) discussed the advantage and potential problems of these definitions of ED. This definition of ED is a simple but useful descriptor of a herd environment. The potential problems of this definition of ED includes mixture of genetic components of individuals in the environmental variable, especially for the non-random distribution of genotypes across herds no matter whether the individual's own records were included or excluded in the calculation of average ED. Another potential problem is that it linearizes the average RN, which may re-rank environments. Alternatively, some researchers inferred environmental values simultaneously with the other parameters in the model using Bayesian MCMC methodology (Knap and Su, 2008; Su et al., 2006).

Besides the phenotypic mean as ED, some other EDs have been assessed for studying RN in GxE. Fifteen variables including herd size, within-herd standard deviation, climatic information (temperature and rainfall) were used, among of which nine EDs were found to be significant in a study involving Guernsey cows (Fikse et al., 2003). Summer heat load index and altitude were used in a New Zealand dairy study (Bryant et al., 2007). An intensity index which combined climatic information with phenotypes, and a production index were used in UK dairy cattle study (Strandberg et al., 2009), in which both indexes were developed by principle component analysis and canonical correlation analysis (Haskell et al., 2007). Three different feeding levels were used as ED in a pig study (Hermesch et al., 2006). Percent North American Holstein genes was used and found to be a significant variable for the lactation milk yield in dairy (Zwald et al., 2003).

In order to minimize the link between dependent variable (y) and independent variable (environmental variable, x), some studies excluded the individual's records when ED for the individual was calculated.

It was quite common to standardize the ED before being used in the GxE estimations. Some studies standardized ED to range from -10 to 10 (Fikse et al., 2003). Others standardized ED to be mean zero and standard deviation unity (Strandberg et al., 2009).

Heterogeneous residual variances

Heterogeneous residual variances were expected for some groups of environments (Calus et al., 2005; Knap and Su, 2008). Heterogeneous residual variances increased linearly with increasing values of the mean for herd environments for both protein production and days open. Heritability estimates obtained fitting heterogeneous residual variances changed less between environments than those estimated using a homogeneous residual variance (Kolmodin et al., 2002). To overcome this problem, a few studies standardized traits to a mean of zero and a standard deviation of one within each CG to account for heterogeneous variances between environmental

groups (Hermesch et al., 2006; Pollott and Greeff, 2004). Other studies fitted the heterogeneous error in the model (Cardoso and Tempelman, 2012; Knap and Su, 2008; Mattar et al., 2011). However, the model fitting heterogeneous residual variances was not better than a model applying a homogeneous residual variance in the GxE research on long-yearling weight in beef cattle (Mattar et al., 2011).

Statistical models and genetic parameter estimation

The statistical RN model is:

$$y_{ml} = \mu + \sum_{i=1}^{n} \beta_i X_{ml}^i + a_m + \sum_{i=1}^{n} b_{im} X_{ml}^i + e_{ml} \qquad [1]$$

where y_{ml} is the phenotypic value of progeny *l* of genotype (sire) *m*, μ is the overall mean, β_i are fixed regression coefficients, X_{ml}^i is the environmental value for progeny *l* of genotype (sire) *m*, *i* is the order of polynomial regression model, a_m is the breeding value for genotype (sire) *m*, and b_{im} are random regression coefficients for genotype (or animal) m, e_{ml} is the residual and *n* is the max order of the regression model, when n is 1, it is a linear RN model; when n >1, it is a polynomial regression model with n-1 *th* order.

The random effects a and b_i are assumed to follow N(0,**G**) with **G** matrix

$$\mathbf{G} = Var\begin{bmatrix} a \\ b_1 \\ \vdots \\ b_{n-1} \end{bmatrix} = \begin{bmatrix} \sigma_a^2 & \sigma_{a,b_1}^2 & \cdots & \sigma_{a,b_{n-1}}^2 \\ \sigma_{a,b_1}^2 & \sigma_{b_1}^2 & \cdots & \sigma_{b_1,b_{n-1}}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{a,b_{n-1}}^2 & \sigma_{b_1,b_{n-1}}^2 & \cdots & \sigma_{b_{n-1}}^2 \end{bmatrix} \otimes \mathbf{A} \text{ or } \mathbf{I}$$

where σ_a^2 is the additive genetic variance of level (random intercept of RN), $\sigma_{b_i}^2$ is the additive genetic variance of random slope, **A** is the numerator relationship matrix if pedigree information is used, otherwise it is replaced by identity matrix **I**.

The random residual is assumed to follow N(0,$\mathbf{I}\sigma_e^2$) when homogeneous variance is fitted in the model. Otherwise heterogeneous variance is assumed with NID(0,$\sigma_{e_j}^2$), *j*=1,2,…,*n*. *n* is the number of the environment groups. Some studies calculated $\sigma_{e_j}^2$ by a function of the associated environmental covariate (Mattar et al., 2011).

The heritability in environment *k* can be estimated as

$$\mathrm{h}_k^2 = \frac{\mathbf{x}_k' \mathbf{G} \mathbf{x}_k}{\mathbf{x}_k' \mathbf{G} \mathbf{x}_k + \sigma_{\mathrm{e_k}}^2} \qquad [2]$$

where **G** is the (co)variance matrix for the random coefficients with order n x n, $\boldsymbol{x}_k$is a column vector with elements $\{x_k^i\}$ for $i = 0, \cdots, n-1$.

Breeding value of genotype (sire) *m* in environment *k* can be estimated as

$$BV_m | x_k = a_m + \sum_{i=1}^{n} b_{im} X_{mk}^i \qquad [3]$$

Phenotypic value of genotype (sire) *m* in environment *k* can be estimated as

$$PV_m | x_k = \mu + \sum_{i=1}^{n} \beta_i X_{mk}^i + a_m + \sum_{i=1}^{n} b_{im} X_{mk}^i \qquad [4]$$

The genetic correlation between environment *k1* and *k2* is

$$\mathrm{r}_{g_{k1,k2}} = \frac{\mathbf{x}_{k1}' \mathbf{G} \mathbf{x}_{k2}}{\sqrt{\mathbf{x}_{k1}' \mathbf{G} \mathbf{x}_{k1} \mathbf{x}_{k2}' \mathbf{G} \mathbf{x}_{k2}}} \qquad [5]$$

When more than one environmental descriptor is used in the model, Eq [1] can be extended to:

$$y_{ml} = \mu + \sum_{i=1}^{n_i} \beta_i X_{ml}^i + a_m + \sum_{i=1}^{n_i} b_{im} X_{ml}^i + \sum_{j=1}^{n_j} \gamma_j Z_{ml}^j + \sum_{j=1}^{n_j} r_{jm} Z_{ml}^j + e_{ml} \quad [6]$$

where Z_{ml} is additional environmental value, γ_j and r_{jm} are fixed and random regression coefficients for environment *Z*, n_i is the order of regression model for environment X, n_j is the order of regression model for environment *Z*.

Eq [6] can be further extended to include multiple traits in one analysis:

$$\mathbf{y}_{\mathrm{ml}} = \boldsymbol{\mu} + \sum_{i=1}^{n_i-1} \boldsymbol{\beta}_i x^i + \boldsymbol{a}_m + \sum_{i=1}^{n_i-1} \boldsymbol{b}_{im} x^i + \sum_{j=1}^{n_j-1} \boldsymbol{\gamma}_j z^j + \sum_{j=1}^{n_j-1} \boldsymbol{r}_{jm} z^j + \boldsymbol{e}_{ml} \quad [7]$$

where$\boldsymbol{y}$, $\boldsymbol{\mu}$, $\boldsymbol{\beta}$, $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{r}$, $\boldsymbol{e}$ are column vectors with dimension size of number of traits instead of scalar in single trait analysis. The (co)variance matrix **G** for the random coefficients with two traits will be $\begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix}$, then the covariance between environments *j* and *k* becomes:

$$\mathbf{x}_{\mathrm{j}}' \mathbf{G} \mathbf{x}_{\mathrm{k}} = [\, 1\ x_j\ 1\ x_j\,] \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} \boldsymbol{x_k} \quad [8]$$

Model comparison was conducted either by the deviance information criterion (DIC) for Gibbs Sampler method (Mattar et al., 2011) or likelihood ratio test (LRT) for REML methods (Kolmodin et al., 2002).

Some researchers have argued that RN models have some limitations in some cases (Crossa, 1990). For example, when only extreme environments are included in the data, the fitting of RN model is not convincing and not suitable. Further, if the environmental descriptor is calculated as the average of all genotypes in that environment, it is not independent of the genotype performance in the model.

Models with principal components analysis in plant breeding

Plant breeders have been interested in studying the genotype by environment interaction for a long time (Burdon, 1978; Hill, 1975; Kang, 1998; Ramburan et al., 2011). One of the advantages for plant breeders is that they can have large family sizes and field trials are normally well designed leading to data structure that is much more balanced than animal data. Besides the RN model, other models based on principal components analysis, such as AMMI (additive main effect and multiplicative Interaction) (Zobel et al., 1988) and GGE (genotype main effect plus genotype-environment interaction) (Crossa and Cornelius, 1997) have been widely used in plant breeding to investigate GxE. Both models start with the basic model for G x E of:

$$y_{ij} = \mu + G_i + E_j + GE_{ij} + e_{ij} \quad [9]$$

where y_{ij} is the phenotypic value of genotype *i* in environment *j*, μ is the overall mean, G_i is genotype *i* effect, E_j is environment *j* effect, GE_{ij} is the genotype *i* in environment *j* interaction effect. AMMI and GGE decompose the GE_{ij} or $G_i + GE_{ij}$ terms by using the first few principal components,

In AMMI model,

$$GE_{ij} = \sum_{k=1}^{n} \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij} \qquad [10]$$

In GGE model ,

$$G_i + GE_{ij} = \sum_{k=1}^{n} \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij} \qquad [11]$$

where λ_k is the singular value of the axis in the principal component analysis, α_{ik} is the eigenvector of the genotype *i* environment for the *kth* axis, which can be interpreted as the genotypic sensitivity of genotype *i* to environmental variable γ_{jk}. Alternatively, γ_{jk} is the eigenvector of the environment *j* for the *kth* axis, which can be interpreted as the environmental potential of environment *j* to genotype α_{ik}, *n* is the number of principal components in the model, ε_{ij} is the residual term. The results of these analyses are often represented in graphs showing the position (or loading) of each of the input variables with respect to the principal components, which are drawn as the axes (Yan et al., 2007).

As both models use all the correlations among genotypes and among environments, it has been demonstrated to be better able to explain GxE than other methods like regression method in the area of plant breeding (Wu and Ying, 2001). So far, little information is available about applying AMMI or GGE model to animal breeding data, since data structures from animal breeding are quite different from plant breeding in terms of size of family, data balance and definition and number of environments, it is worthwhile comparing this kind of method with RN models more commonly used in animal breeding studies investigating GxE.

2.2. ***Variation in environments and breed by environment interactions***

Description of data

Data comprising Duroc (DU), Large White (LW) and Landrace (LR) pigs recorded from 2000 to 2010 were used. These herds are part of the across-herd genetic evaluations of the National Pig Improvement Program (NPIP) in Australia. Only herds with at least 1000 animals for each breed and at least two breeds were used for these analyses. On average, pigs were recorded for lifetime average daily gain (ADG) and back fat (BF) at a live weight of 92.8 kg. For both traits, only records that were within the range of four standard deviations from the mean of each trait within breed were considered in these analyses. After editing, the data contained 265,165 pigs belonging to nine herds, of which 130,766 were female and 134,399 were male pigs. The pedigree file was generated based on all animals with observations leading to a total of 268,989 animals from 2,394 sires and 12,363 dams.

Models for analyses of environmental variable

The two definitions of CG explored in this study were herd by birth month (HBM) and herd by birth week (HBW). Four quantities were used to compare the properties of the two alternative environmental descriptors. These were the phenotypic mean of a CG (Mean) and least squares means (LSM1, LSM2, LSM3) obtained from three linear models fitted to ADG and BF using PROC GLM in SAS (1999). The three models were:

LSM1: $y_{ijkl} = \mu + Common\ Effects + hbm(hbw)_k + e_{ijkl}$

LSM2: $y_{ijkl} = \mu + Common\ Effects + hbm(hbw)_k + sire_j + e_{ijkl}$

LSM3: $y_{ijkl} = \mu + Common\ Effects + hbm(hbw)_k + animal_l + e_{ijkl}$

where y_{ijkl} is the phenotype (ADG or BF) of animal *l* of sire *j* of breed *i* in herd-by-birth month (or herd-by-birth week group) *k*, μ is the overall mean, $Common\ Effects$ are the fixed effects of sex, birth parity and breed fitted for both traits and weight at test fitted as a linear covariate for BF only, as well as the random common litter effect, $hbm(hbw)_k$ is the fixed effect of herd-by-birth month (or herd-by-birth week) group *k*, $sire_j$ is the random effect of sire *j*, $animal_l$ is the random effect of animal *l*, e_{ijkl} is the random residual associated with the observation.

Statistical models used for reaction norm analyses

A fixed regression on the environmental variable X_{ijk} within breed was added to all models to obtain reaction norms for breed. The statistical models used in this analysis included two animal models (Model A and B) and a sire model (Model C) to evaluate the use of alternative genetic models. An overall fixed regression on the environmental variable X_{ijk} used in Model B and Model C replaced HBM group in Model A to evaluate alternative methods to account for the overall changes along the environmental trajectory.

Model A: $y_{ijkl} = \mu + Common\ Effects + hbm_k + b_{0i} + b_i X_{ijk} + animal_l + e_{ijkl}$

Model B: $y_{ijkl} = \mu + Common\ Effects + bX_{ijk} + b_{0i} + b_i X_{ijk} + animal_l + e_{ijkl}$

Model C: $y_{ijkl} = \mu + Common\ Effects + bX_{ijk} + b_{0i} + b_i X_{ijk} + sire_j + e_{ijkl}$

where b is overall fixed linear regression coefficients of y_{ijkl} on X_{ijk}, X_{ijk} is the environmental descriptor for progeny *k* of sire *j* of breed *i* , which is defined as the phenotypic mean (Mean) or least squares means (LSM1, LSM2, and LSM3) of HBM group, b_{0i} is the intercept of breed *i*, b_i is the fixed linear regression coefficient from the regression of y on X_{ijk} for breed *i*. The intercept for each breed replaced the fixed effect of breed in the model. All the other notations are the same outlined as above. All reaction norm analyses were performed with the ASReml statistical software package (Gilmour et al., 2009). Pedigree information was used to calculate the numerator relationship matrix (NRM) when an animal model (Model A and B) was fitted.

Variation in environments

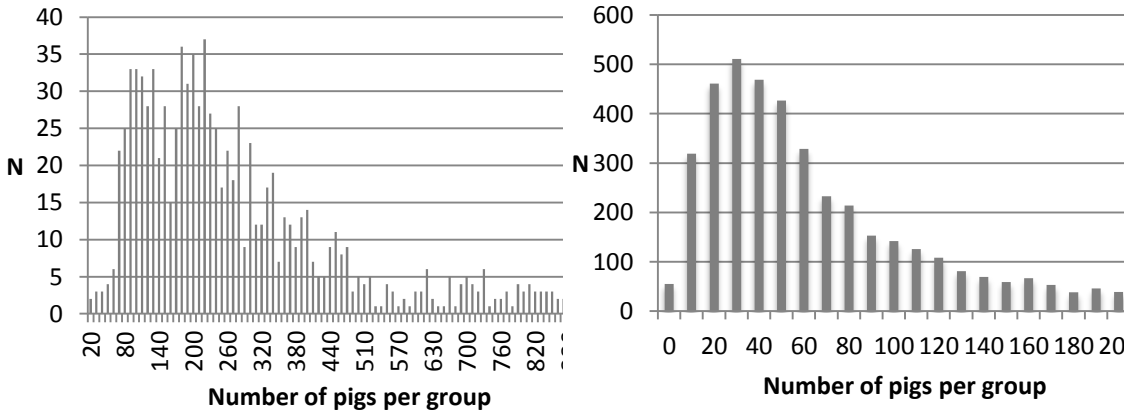
Most records were available for Large White pigs while Duroc was the smallest breed with approximately 33 000 records available (Table 1). Duroc had the lowest weight and age at recording which were highest for Large White. Growth rate was similar across breeds and Landrace had a slightly lower backfat than the other two breeds.

Table 1. Descriptive statistics for weight at test (WT), age at test (AGE), average daily gain (ADG), and back fat (BF) for the three breeds Duroc (DU), Landrace (LR) and Large White as well as the total data set (All) (means; standard deviations).

Breed	N*	WT (kg)	AGE (days)	ADG (g/day)	BF (mm)
DU	33 677	90.1; 12.7	139.0; 14.3	649.1; 72.7	10.5; 2.0
LR	87 987	92.9; 13.9	142.7; 16.9	652.0; 72.8	10.3; 2.0
LW	143 501	93.3; 13.5	144.5; 17.8	648.0; 73.3	10.7; 2.1
All	265 165	92.8; 13.6	143.2; 17.2	649.5; 73.1	10.5; 2.1

*N: Number of pigs

For the CG definition of HBM, there were 950 groups with an average number of animals of 279 ranging from 16 to 1071 pigs per group. When HBW was defined as the CG, 4084 groups were formed with an average number of 65 animals per group ranging from 1 to 320 pigs per group. The distribution of the number of animals per group is shown in Figure 1(a) for HBM and in Figure 1(b) for HBW. Both distributions were characterized by a tail representing larger CG sizes outside the normal distribution.



(a) (b)

Figure 1. Number of pigs (N) per herd-by-birth month group (a) and herd-by-birth week group (b)

Analysis of variance

The fixed effect model explained 21.4% of the variation for ADG and 21.8% of the variation for BF using the HBM environmental descriptor. The coefficients of determination increased to 25.0% for ADG and 25.1% for BF when HBW was used as the CG definition.

Distribution of mean performance of contemporary groups

The mean performance per CG was normally distributed for both definitions of CG (HBM and HBW) for both traits (Figure 2). Although the number of CGs was about a quarter for the HBM definition in comparison to the HBW CG, the spread in mean phenotypic performance (Mean) and least squares means from linear models not

fitting genetic effects (LSM1) or using a sire model (LSM2) were similar for both definitions of CG ranging from about 540 g/day to 720 or 740 g/day for ADG. Least squares means from the animal model (LSM3) were more normally distributed with slightly less variation in mean performance of CGs ranging from 540 g/day to 720 g/day for ADG. There were less CGs with higher least squares means for the animal model indicating that the higher mean for these groups obtained from alternative models may have included some additive genetic effects that were better accounted for in the animal model in comparison to the other models.

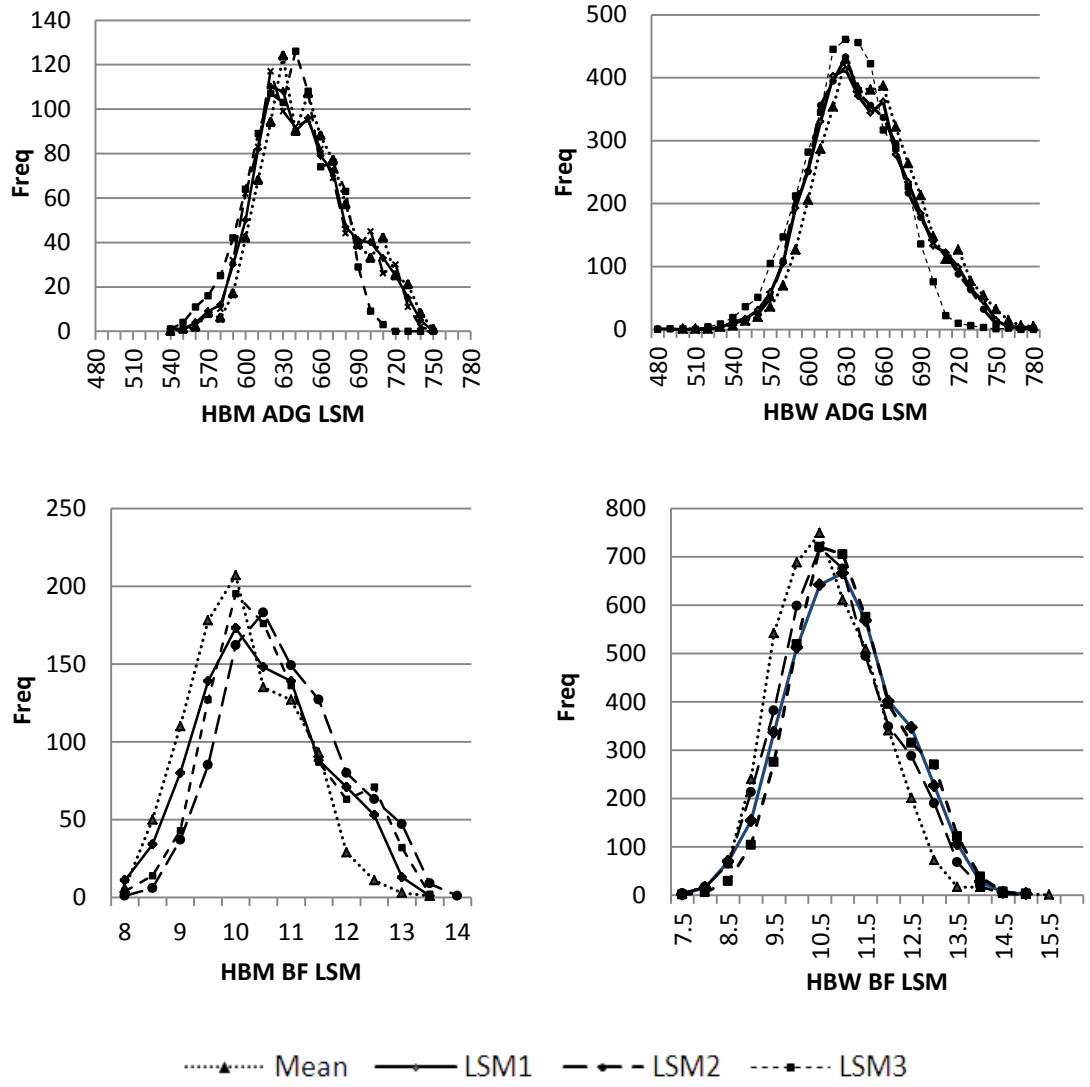


Figure 2. Distribution of means or least squares means for average daily gain (ADG) and back fat (BF) for herd-by-month (HBM) and herd-by-week (HBW) CGs derived from different models (Mean: mean phenotypic performance, LSM1: least squares means of linear mixed model fitting fixed effects only; LSM2: least squares means from sire model, LSM3: least squares means from animal model).

As to BF, the distribution of CG predictors based on the three least squares means (LSM1, LSM2, LSM3) were similar for both definitions of CG, with means and least squares means for HBM and HBW ranging from about 8.0 cm to 14.0 cm. The mean

of the mean phenotypic performance (Mean) was around 0.56 cm less compared to the mean of three least square means due to the adjustment for other effects fitted in models.

The number of pigs available for each CG determines the standard errors of least squares means. The association between the number of pigs per CG and standard error (s.e.) of least squares means is shown in Figures 3a and 3b for ADG and in Figures 3c and 3d for BF for both definitions of CG using the animal model (LSM3). Contemporary groups with small number of pigs were not excluded. These graphs illustrate how standard errors of CGs were reduced for larger groups. This information may be used to choosing the optimal cut-off point in regard to small CG sizes for each definition of CGs. Further, considerably variation in standard errors was apparent for CGs with similar sizes. These differences were due to different data structures. Exclusion of sires with few progeny and poor cross-classification of other effects with CGs is expected to reduce this variation in standard errors for CGs with similar number of pigs per group.

Variation in least squares means of contemporary groups within herds and within years

In order to detect sire by environment interactions a good spread of progeny records across environmental conditions is required. In pig breeding programs in general, sires may only be used within a single herd over a relatively short time period of a few months. Therefore, variation in least squares means of CGs obtained from the animal model (LSM3) was further explored by evaluating within-herd and within-year variation in least squares means. The spread of least squares means showed a range of 109 g/day to 162 g/day for ADG (Figure 4a) and 3.0 cm to 4.9 cm for BF (Figure 4b) within years indicating that most variation observed in least squares means across year was also apparent within years for both traits. Also, there was a considerable spread of least squares means for HBM CGs within herds for both ADG and BF. A range of 87 to 145 g/day for ADG (Figure 4c) and a range of about 1.6 cm to 3.5 cm for BF (Figure 4d) were found within herds. The variation of BF between herds was greater than between years, which was partly due to variation in weight at test between herds. The distribution of least squares means is shown for ADG for four example herds in Figure 5 as a further illustration.

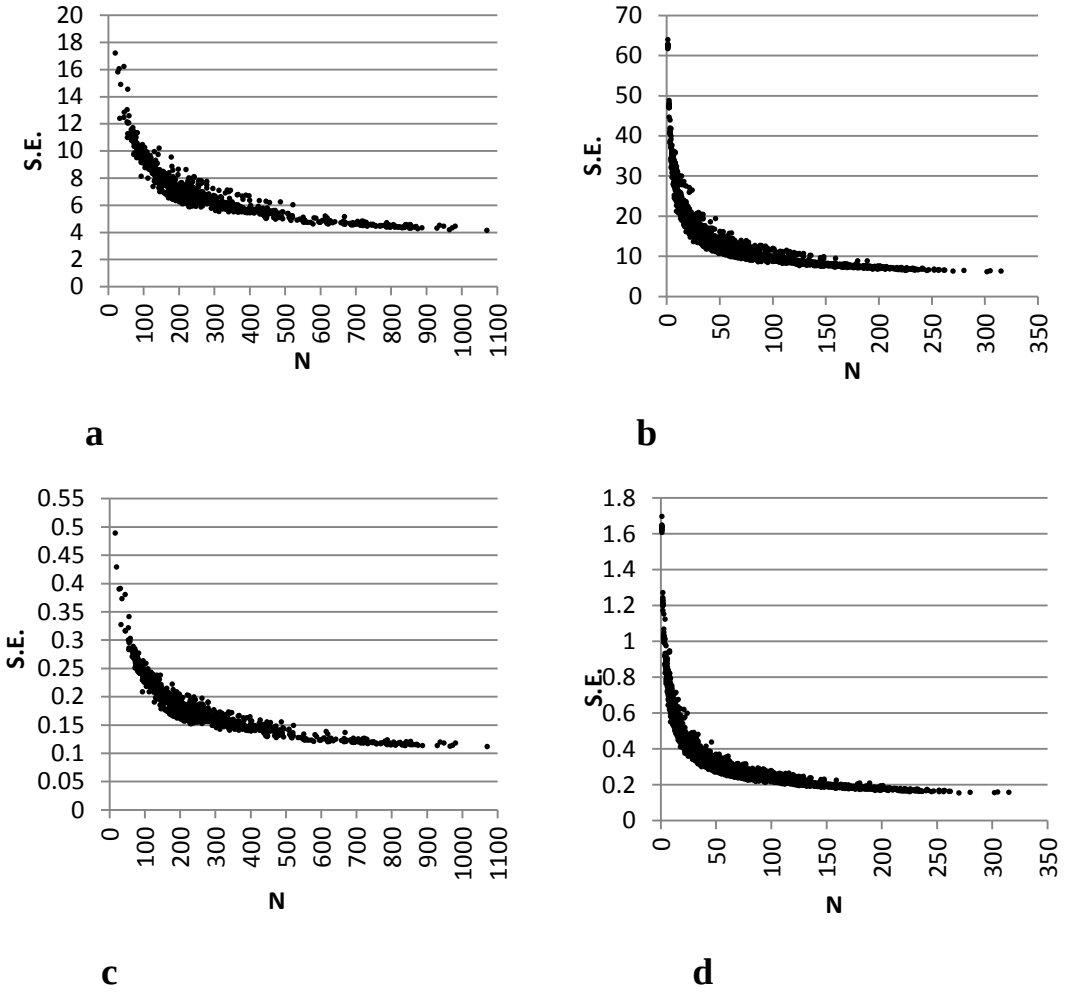


Figure 3. The effect of number of animals per group (N) on standard errors (S.E.) of least squares means for average daily gain (ADG, Figures 3a, 3b) and back fat (BF, Figures 3c, 3d) using either herd-by-month (HBM) or herd-by-week (HBW) as a definition of CGs applying an animal model (LSM3).

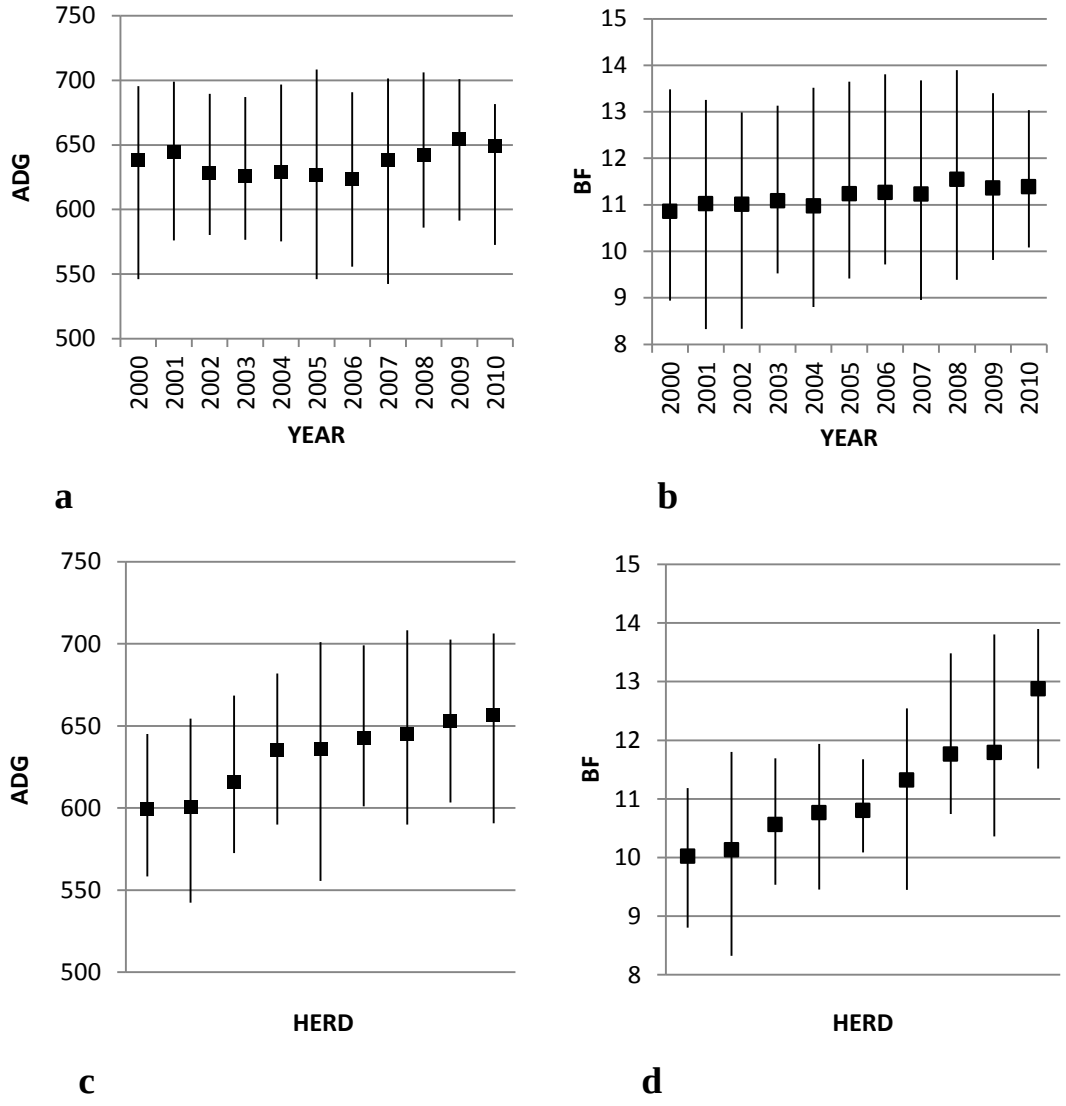


Figure 4. Mean and range of least squares means (LSM3) for average daily gain (ADG) and back fat (BF) of herd-by-month CGs within years (Figures 4a, 4b) and within herds (Figures 4c, 4d).

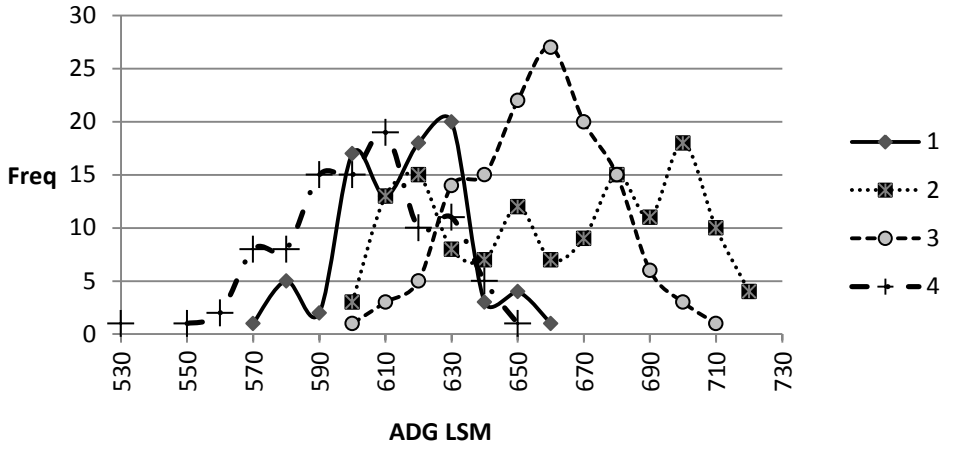


Figure 5. Frequency of least squares means (LSM) of average daily gain (ADG) for herd-by-month contemporary groups for four example herds.

Breed by environment interactions

Three alternative models (Models A, B, C) were used that fitted fixed intercept and slope within breeds to evaluate breed effects and the environmental sensitivity of each breed. Based on the previous results, only the CG definition of HBM was used as the environmental descriptor. The environmental variable was based on phenotypic means (Mean) or least squares means obtained from the three linear models (LSM1, LSM2 and LSM3).

Breed solutions for the intercept from the fixed regression of the observations on environmental variable showed that Landrace was the fastest growing breed, while Duroc had the lowest growth rate (Table 2). Large White was the leanest breed although the difference to the other two breeds was small with solutions ranging from 0.04 to 0.33 mm for Landrace and Duroc using alternative models and environmental variables. Using an environmental variable that also included genetic components (Mean, LSM1) affected breed solutions for the intercept for both traits although changes in these solutions were generally not significant.

The fixed regression coefficients from the regression of observations on environmental variables were close to 1 (0.95-1.04) for both traits using either an animal or sire model fitting alternative descriptors of the environmental variable based on least squares means (LSM1,LSM2 and LSM3) (Table 2). The fixed regression coefficients from the regression of observations on phenotypic means (Mean) was 0.83 for BF using an animal model indicating that phenotypic means was not a good environmental variable as the other least squares means for BF. The regression coefficient was reduced because variation in means included variation in weight at recording BF, which was adjusted for in the reaction norm model used to analyse BF. The solutions for the fixed regression coefficients of breed on the environmental variable were 0.0 for Large White and varied from -0.07 to -0.09 for Landrace and from -0.12 to -0.17 for Duroc for ADG. These results indicate that Large White was more sensitive to changes in the environment than Landrace or Duroc, which were least sensitive to changes in environmental conditions for ADG.

For BF, Landrace was the least sensitive breed based on an animal model. Regression coefficients obtained for Landrace using an animal model differed significantly from Large White although estimates of regression coefficients did not differ significantly from Duroc.

Finally, regression coefficients differed using either an animal model versus a sire model for both traits. The direction of the change in regression coefficients differed between Landrace versus Duroc pigs. In Duroc pigs, regression coefficients had a larger magnitude from sire models which was not observed for Landrace.

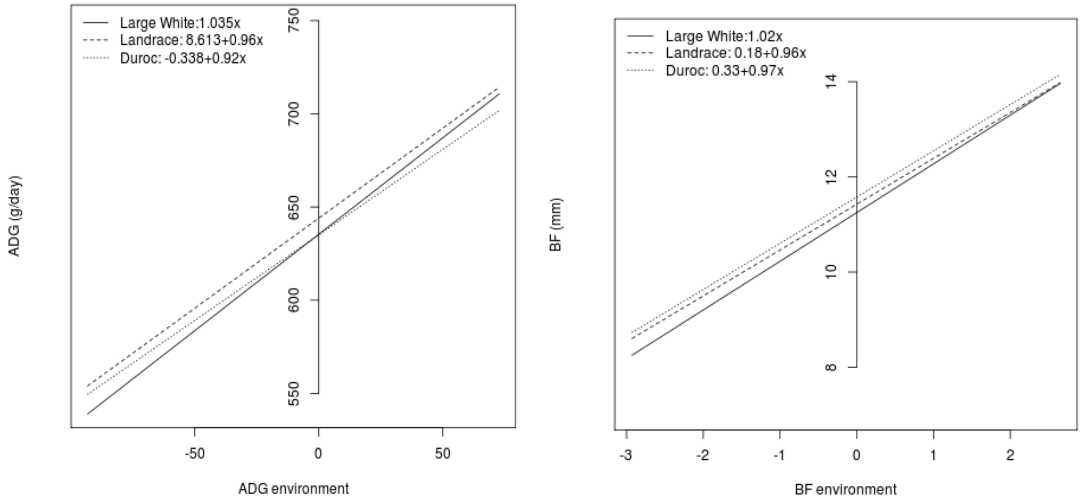
Reaction norms for ADG and BF predicted for three breeds for ADG and BF from Model B are shown in Figure 6 illustrating a breed by environment interaction. Large White grew faster than Duroc in more favourable environments, while Duroc grew faster than Large White in unfavourable environments (Figure 6a). The magnitude of breed by environment interactions was lower for BF and there was no re-ranking of breeds across the environmental trajectory for BF (Figure 6b).

Genetic parameters

The heritability estimate for ADG was 0.22 using traditional models (LSM3, Model A). Fitting a breed by environment interaction did not affect this heritability estimate as long as an environmental variable was used that did not include any genetic effects (LSM2, LSM3). Using environmental variables that included genetic effects (Mean, LSM1) reduced heritability estimates to 0.16 for the animal model and 0.20 for the sire model. Heritability estimates were less variable for BF across models ranging from 0.29 to 0.34 for alternative animal and sire models. It should be noted that breed by environment interaction was not significant and therefore alternative methods to account for any breed by environment interaction should lead to similar results.

Estimates of the common litter effect varied from 0.11 to 0.13 using an animal model and from 0.16 to 0.17 applying alternative sire models for ADG. Common litter effects were also lower using an animal model (range: 0.06 to 0.08) than a sire model (0.12) for BF.

Similar phenotypic correlations and correlations for the common litter effect were found between ADG and BF using an animal and or sire model with estimates of 0.17±0.00 and 0.17±0.01, respectively. Genetic correlations between these two traits were much more different between models with estimates of 0.24±0.02 for the animal model and 0.02±0.03 for the sire model.



a **b**

Figure 6. Reaction norms for Duroc, Landrace and Large White for average daily gain (ADG) and back fat (BF) along the environmental trajectory quantified by least squares means from an animal model (LSM3) of herd-by-month contemporary groups.

Influence of alternative data restrictions on parameter estimates

There were no significant differences between regression coefficients for breed using different subsets of data based on imposing limits on either the minimum number of animals per HBM to at least 50 or 100 pigs or alternatively, imposing limits on the minimum standard error of least squares means for the HBM CG of to at least 12, 10 or 8 (Table 3).

Table 2. Parameter estimates and their standard error (in subscription) using alternative models: fixed slope of the reaction norm (b0) and solutions for the intercept (in) and slope (b) for Landrace (LR) and Duroc (DU), heritability (h^2) and common litter effect (c^2)

Trait	Model	Model	EV	b0	in_{LR}	in_{Du}	b_{LR}	b_{DU}	h^2	c^2
		LSM3	-		$8.601_{1.867}$	$-0.651_{2.517}$			$0.22_{0.006}$	$0.12_{0.002}$
		Model A	LSM3		$8.428_{1.866}$	$-0.485_{2.514}$	$-0.087_{0.191}$	$-0.138_{0.192}$	$0.22_{0.006}$	$0.12_{0.002}$
	Animal	Model B	LSM3	$1.035_{0.011}$	$8.613_{1.795}$	$-0.338_{2.408}$	$-0.069_{0.020}$	$-0.117_{0.028}$	$0.22_{0.006}$	$0.11_{0.002}$
ADG	**model**	Model B	LSM1	$0.977_{0.011}$	$6.331_{1.624}$	$-4.991_{2.160}$	$-0.084_{0.019}$	$-0.133_{0.026}$	$0.16_{0.005}$	$0.13_{0.002}$
		Model B	Mean	$0.980_{0.011}$	$5.479_{1.625}$	$-5.075_{2.162}$	$-0.089_{0.019}$	$-0.147_{0.026}$	$0.16_{0.006}$	$0.13_{0.002}$
	Sire	Model C	LSM2	$1.047_{0.011}$	$9.034_{0.892}$	$-6.413_{1.186}$	$-0.078_{0.018}$	$-0.174_{0.025}$	$0.22_{0.009}$	$0.16_{0.002}$
	model	Model C	LSM1	$1.024_{0.011}$	$8.826_{0.863}$	$-6.211_{1.150}$	$-0.075_{0.018}$	$-0.171_{0.025}$	$0.20_{0.009}$	$0.16_{0.002}$
		LSM3	-						$0.33_{0.007}$	$0.06_{0.002}$
		Model A	LSM3		$0.166_{0.057}$	$0.322_{0.077}$	$-0.074_{0.021}$	$-0.063_{0.026}$	$0.33_{0.007}$	$0.06_{0.002}$
	Animal	Model B	LSM3	$1.023_{0.011}$	$0.176_{0.054}$	$0.326_{0.074}$	$-0.060_{0.019}$	$-0.052_{0.025}$	$0.33_{0.006}$	$0.06_{0.002}$
BF	**model**	Model B	LSM1	$0.951_{0.010}$	$0.152_{0.052}$	$0.333_{0.070}$	$-0.068_{0.019}$	$-0.012_{0.025}$	$0.29_{0.006}$	$0.07_{0.002}$
		Model B	Mean	$0.829_{0.012}$	$0.092_{0.056}$	$0.326_{0.076}$	$-0.122_{0.019}$	$-0.046_{0.028}$	$0.34_{0.006}$	$0.08_{0.002}$
	Sire	Model C	LSM2	$1.021_{0.009}$	$0.070_{0.029}$	$0.320_{0.037}$	$-0.039_{0.019}$	$-0.102_{0.022}$	$0.37_{0.013}$	$0.12_{0.001}$
	model	Model C	LSM1	$0.992_{0.009}$	$0.036_{0.028}$	$0.283_{0.036}$	$-0.069_{0.018}$	$-0.093_{0.022}$	$0.33_{0.012}$	$0.12_{0.002}$

Note: ED, environmental variable; breed solutions are expressed relative to Large White with values of $in_{LW}=0_0$ and $b_{LW}=0_0$

Table 3. Parameter estimates and their standard error (in subscription) using an animal model applying different criteria for data elimination: fixed slope of the regression (b0) and solutions for the intercept (in) and slope (b) for Landrace (LR) and Duroc (DU), heritability (h^2) and common litter effect (c^2)

Criterion*	**N**	**b0**	**$b0_{LR}$**	**$b0_{Du}$**	**b_{LR}**	**b_{Du}**	**h^2**	**c^2**
N < 50	264809	$1.035_{0.011}$	$8.671_{1.796}$	$-0.330_{2.413}$	$-0.070_{0.020}$	$-0.117_{0.028}$	$0.22_{0.006}$	$0.11_{0.002}$
N < 100	256329	$1.040_{0.012}$	$8.326_{1.820}$	$0.571_{2.453}$	$-0.078_{0.020}$	$-0.115_{0.028}$	$0.22_{0.006}$	$0.11_{0.002}$
LSM SE >12	264476	$1.035_{0.011}$	$8.614_{1.798}$	$-0.204_{2.413}$	$-0.070_{0.020}$	$-0.117_{0.028}$	$0.22_{0.006}$	$0.11_{0.002}$
LSM SE>10	259598	$1.038_{0.012}$	$8.310_{1.811}$	$-0.269_{2.433}$	$-0.075_{0.020}$	$-0.117_{0.028}$	$0.22_{0.006}$	$0.11_{0.002}$
LSM SE >8	235659	$1.043_{0.013}$	$8.836_{1.937}$	$1.337_{2.578}$	$-0.085_{0.021}$	$-0.116_{0.030}$	$0.23_{0.006}$	$0.11_{0.002}$

* N: number of pigs per contemporary group is below 50 or 100; LSM SE: standard error of least squares means

2.3. *Sire by environment interactions*

Introduction

The phenotype of a genotype might be described as function of its environment variable. The phenomenon that different genotypes have different responses to the same environment or alternatively, the same genotype has different responses to different environments is called phenotypic plasticity, which is often described as reaction norms. Reaction norm model have been used in pig breeding studies (i.e. Hermesch et al., 2006, Knap and Guo, 2008, Lewis et al., 2010, Bergsma et al., 2011) and have been found to offer a powerful and flexible approach to detect and measure genotype by environment interaction (GxE).

It was the objective of this to apply reaction norm models to investigate sire by environment interactions for growth rate and back fat in Australian pigs.

Description of data and environmental variables

The same data as described in Section 2.1. were used for analyses of sire by environment interactions. Further, two subsets were generated by removing sires with less than 50 progeny (subset 1) or sires with less than 100 progeny (subset 2). These limits were imposed in order to investigate the influence of number of animals per sire on the accuracy of genetic parameter estimations.

The environment contemporary group (CG) was defined as herd-by-birth month (HBM) in these analyses. There were 950 HBM CGs with an average number of 279 animals ranging from 16 to 1071 pigs per group. A linear sire model was fitted to obtain least squares means (LSM) for each CG as environmental descriptors which were used in reaction norm models. The models for least squares means were:

$$y_{ijk} = \mu + Common\ Effects + (age) + hbm_k + s_{ai} + e_{ijk}$$

where y_{ijk} is the phenotype (ADG or BF) of progeny j of sire i in herd-by-birth month group (hbm) k, μ is the overall mean, $Common\ Effects$ are the fixed effects of sex, birth parity and breed as well as random common litter effect all fitted for both traits while weight at test (wt) was fitted as a linear covariate for BF only, hbm_k is the fixed effect of herd-by-birth month group k, s_{ai} is the sire i effect. Two alternative least squares means for ADG were obtained by either fitting age as linear covariate or not fitting in the model, e_{ijk} is the random residual associated with the observation.

Statistical models for reaction norm analyses

A series of sire models were used in this analysis. Model A was a standard model without fitting a reaction norm on sires. A random regression on an environmental variable X_{ij} within sire was added to the model in Model B. An overall fixed regression on an environmental variable X_{ij} replaced herd-by-birth month group (hbm) in Model C, D and E. An extra environmental variable Z_{ij} was added to the Model E. When the trait ADG was analysed, the primary environmental variable was LSM of ADG for HBM groups and the second environmental variable was LSM of BF for HBM groups. When the trait BF was analyzed, the primary and secondary environmental variables were vice versa. Model F included a quadric term of primary

environmental variable based on Model D. Models D and E were used to evaluate alternative fixed effect models for ADG that either fitted age at test as a linear covariate or omitted this adjustment.

Model A: $y_{ijk} = \mu + Common\ Effects + hbm_k + s_{ai} + e_{ijk}$

Model B: $y_{ijk} = \mu + Common\ Effects + hbm_k + s_{ai} + s_{bi}X_{ij} + e_{ijk}$

Model C: $y_{ijk} = \mu + Common\ Effects + b_{Fx}X_{ij} + s_{ai} + e_{ijk}$

Model D: $y_{ijk} = \mu + Common\ Effects + (age) + b_{Fx}X_{ij} + s_{ai} + s_{bi}X_{ij} + e_{ijk}$

Model E: $y_{ijk} = \mu + Common\ Effects + (age) + b_{Fx}X_{ij} + b_{Fz}Z_{ij} + s_{ai} + s_{bi}X_{ij} + s_{ci}Z_{ij} + e_{ijk}$

Model F: $y_{ijk} = \mu + Common\ Effects + b_{Fx}X_{ij} + b_{Fz}X_{ij}^2 + s_{ai} + s_{bi}X_{ij} + s_{ci}X_{ij}^2 + e_{ijk}$

where s_{bi} is the random regression coefficient of y_{ijk} on X_{ij} for sire i, , s_{ci} is the random regression coefficient of y_{ijk} on Z_{ij} for sire i, b_{Fx} is fixed linear regression coefficients of y_{ijk} on X_{ij}, b_{Fz} is fixed linear regression coefficients of y_{ijk} on Z_{ij}, X_{ij} and Z_{ij} are the primary and secondary environmental descriptor respectively for progeny j of sire i, which were calculated as least squares means of HBM group from the sire model outlined above and mean centered. All the other notations are the same outlined as above.

The random effects s_{ai}, s_{bi} and s_{ci} were assumed to follow N(0,**G**):

$$\mathbf{G} = \begin{bmatrix} \sigma_a^2 & \sigma_{a,b} & \sigma_{a,c} \\ \sigma_{a,b} & \sigma_b^2 & \sigma_{b,c} \\ \sigma_{a,c} & \sigma_{b,c} & \sigma_c^2 \end{bmatrix}$$

where σ_a^2, σ_b^2 and σ_c^2 are the additive sire variance and slope variance for primary and second environmental variables, respectively. Models A to E have different **G** structures depending on the number of random environmental variables in each model. For example, Model A and C have only first row and first column, Model B and D have the first two rows and columns remaining, while Model E has the full **G** structure. The random birth litter effects was assumed to follow N(0, $\mathbf{I}\sigma_l^2$), where **I** is the identify matrix. The residual e_{ijk} was assumed to follow N(0, $\mathbf{I}\sigma_e^2$) for all the models. Further, heterogeneous residual variances were fitted using Model D and E for both traits dividing observations into eight different groups based on the primary environmental variable for each trait.

For Model D and E, the heritability in environment *k* can be estimated as

$$\mathrm{h}_k^2 = \frac{\mathbf{x}_k'\mathbf{G}\mathbf{x}_k \times 4}{\mathbf{x}_k'\mathbf{G}\mathbf{x}_k + \sigma_\mathrm{l}^2 + \sigma_\mathrm{e}^2}$$

where $\boldsymbol{x}_k$is a column vector with elements $\{1, x_k\}$ for Model D and with elements $\{1, x_k, z_k\}$ for Model E. When heterogeneous residual variances were fitted, the residual variance predicted for each CG was used. The predicted residual variance for each CG was obtained by estimating residual variances for each stratified group (eight levels) and then using first (ADG not adjusted

for age) or second order (ADG, BF) polynomial equations to predict residual variances for all CGs across the environmental trajectory.

Model D was selected to run bi-variates analysis for ADG and BF. The (co)variance matrix **G** for the random effects s_{ai} and s_{bi} with two traits was

$$G = \begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} = \begin{bmatrix} \sigma^2_{a1} & \sigma_{a1,b1} & \sigma_{a1,a2} & \sigma_{b1,a2} \\ \sigma_{a1,b1} & \sigma^2_{b1} & \sigma_{a1,b2} & \sigma_{b1,b2} \\ \sigma_{a1,a2} & \sigma_{b1,a2} & \sigma^2_{a2} & \sigma_{a2,b2} \\ \sigma_{a1,b2} & \sigma_{b1,b2} & \sigma_{a2,b2} & \sigma^2_{b2} \end{bmatrix},$$

and then the genetic correlation between two traits for Model D becomes:

$$\mathbf{x}'_{\mathrm{k1}}\mathbf{G_{12}}\mathbf{x}_{\mathrm{k2}} = [\,1, x_{k1}]\mathbf{G_{12}}\begin{bmatrix} 1 \\ x_{k2} \end{bmatrix}$$

$$\mathrm{r_g} = \frac{\mathbf{x}'_{\mathrm{k1}}\mathbf{G_{12}}\mathbf{x}_{\mathrm{k2}}}{\sqrt{\mathbf{x}'_{\mathrm{k1}}\mathbf{G_{11}}\mathbf{x}_{\mathrm{k1}}}\sqrt{\mathbf{x}'_{\mathrm{k2}}\mathbf{G_{22}}\mathbf{x}_{\mathrm{k2}}}}$$

where $\boldsymbol{x}_{k1}$ and $\boldsymbol{x}_{k2}$ are column vectors $\{1, x_{k1}\}$ and $\{1, x_{k2}\}$ with environmental variables as their elements for the first and second trait, respectively.

All reaction norm analyses were applied with ASReml statistical software package (Gilmour et al., 2009).

Data summary

There were 2394 sires in the data in total (Table 4). Constraining the number of progeny per sire to 50 (subset 1) and 100 (subset2), reduced the number of sires to 937 and 1470 in subset 1 and subset 2, respectively. The average numbers of progeny per sire varied from 110.7 to 220.1 for the three data sets.

Table 4. Summary of three data sets used in the analyses

Data	No. of records	No. of sires	Average no. of pigs per sire	Minimum no. of pigs per sire	Maximum no. of pigs per sire
Full	265,103	2394	110.7	1	1789
Subset 1	245,947	1470	167.3	50	1789
Subset 2	206,260	937	220.1	100	1789

Although there were fewer records available in subset 1 and subset 2 (Table 5), the least squares means (LSM) of HBM groups were similar for all three data sets. The mean LSM was 644 g/d for ADG and 11 mm BF and also the ranges of LSM were similar between data sets varying from 539.6 to 737.7 g/day and 8.19 to 13.82 mm for ADG and BF, respectively. Similar standard deviations of LSM for HBM of about 32.2 g/day and 0.91 mm for ADG and BF were found in the full data set and subset 1, which were slightly greater than the value for subset 2 of about 31.0 g/day for ADG and 0.87 mm for BF.

Table 5. Descriptive statistics for environmental variable for ADG and BF for three data sets

Trait	Data	Mean	S.D.	Minimum	Maximum
ADG	Full	644.4	32.4	539.6	737.7
	Subset 1	644.2	32.1	539.6	737.7
	Subset 2	644.0	31.0	539.6	737.7
BF	Full	11.0	0.91	8.19	13.8
	Subset 1	11.0	0.90	8.19	13.8
	Subset 2	11.0	0.87	8.19	13.8

The association between environmental descriptors for ADG and BF was affected by the adjustment of ADG for age and BF for weight. The correlation between LSMs of HBM for ADG and for BF was 0.30 when both traits were not adjusted and 0.12 when only BF was adjusted for weight (Figure 7a, 7c). However, a negative correlation of -0.24 was found when both traits were adjusted (Figure 7b). Distributions of ADG unadjusted for age and BF adjusted for weight at test were more normally distributed than ADG adjusted for age or BF not adjusted for weight. There were more animals with higher LSMs of HBM groups for ADG when ADG was adjusted for age and more animals with higher LSM of HBM groups for BF was not adjusted by weight.

Homogeneous versus heterogeneous residual variances

The criteria to combine observations into groups are shown in Table 6 for ADG and Table 7 for BF. The number of animal in each group differed between groups with more animals in the middle groups because a fixed range was chosen for the environmental variable of each trait.

The observed variance for each environmental group reduced for environmental groups with higher growth rate for both trait definitions of ADG (Figure 8a, 8b). There was a considerable linear decrease in residual variances across the environmental trajectory for ADG adjusted for age (Figure 8a). When ADG was not adjusted for age, the change in residual variances was less variable across environmental groups (Figure 8b). For BF, the observed variance increased linearly for the first five environmental groups followed by a decrease in observed variance for each environmental group with higher BF LSM (Figure 8c). This trend in variances across the environmental trajectory was less pronounced for residual variance. The linear or second order of polynomial equations fitted to predict heterogeneous variances across the environmental trajectory for ADG and BF shown in Figure 8 were used to derive heritability estimates for different environments.

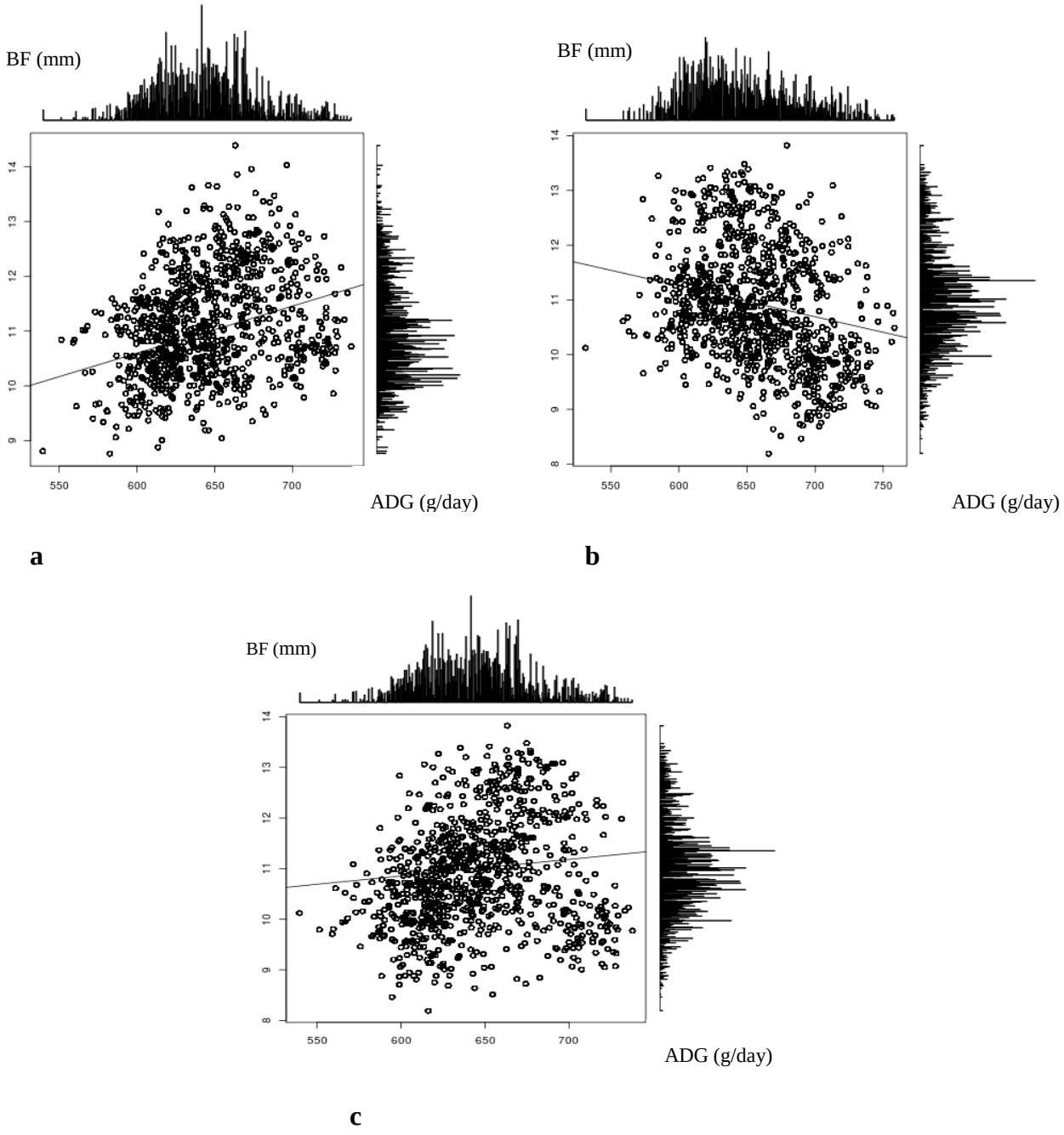


Figure 7. Distribution of least squares means (LSM) of average daily gain (ADG) and back fat (BF) for herd-by-month groups (HBM); (a) ADG unadjusted for age and BF unadjusted for weight at test (r=0.30); (b) ADG adjusted for age and BF adjusted for weight at test (r=-0.24); (c) ADG unadjusted for age and BF adjusted for weight at test (r=0.12).

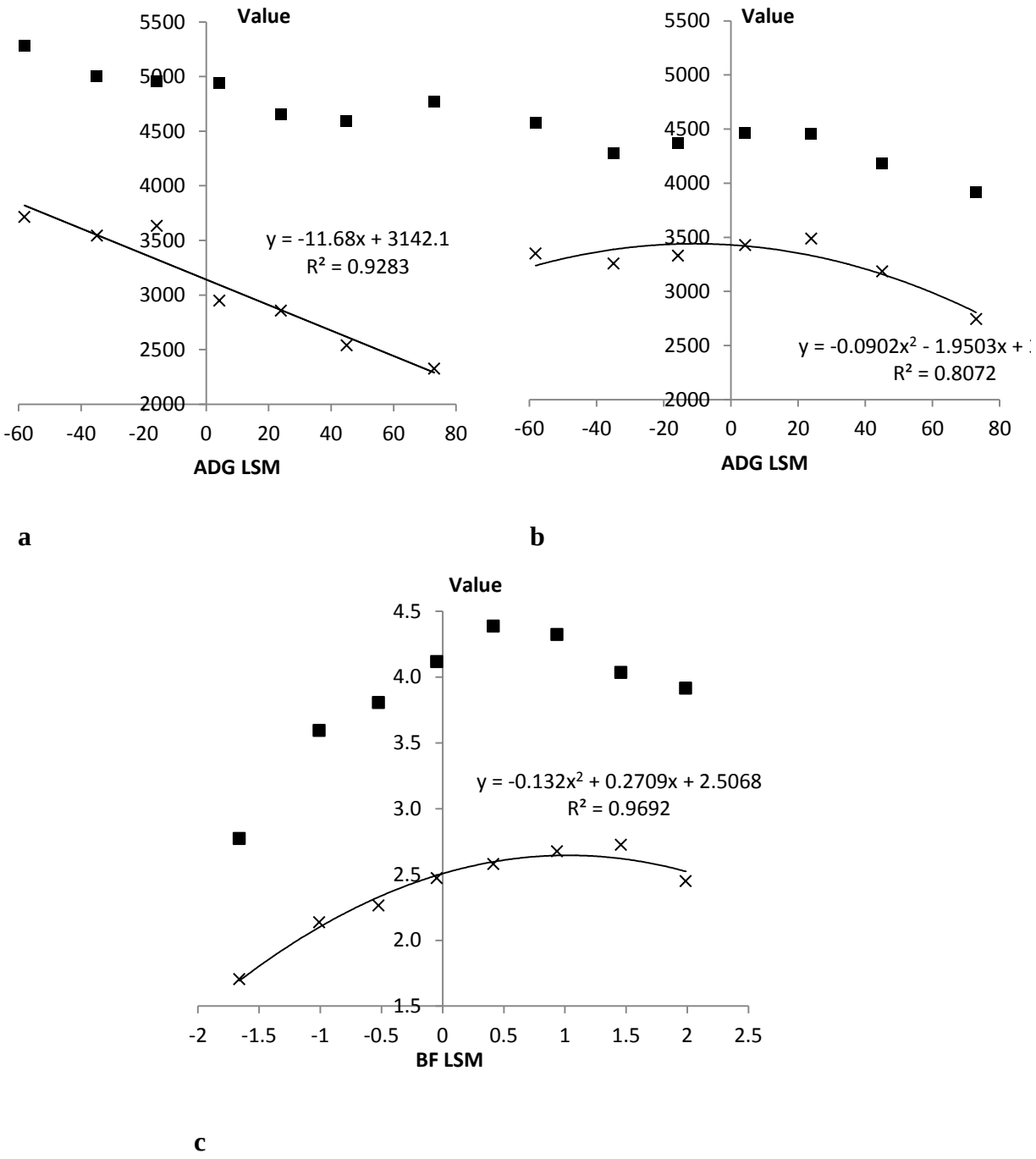


Figure 8. Observed phenotypic variance (▪) and heterogeneous residual variance (×) of each environmental group (using Model D); (a) ADG adjusted for age; (b) ADG unadjusted for age; (3) BF. The x-axis was mean centered (0= 644.2 g/day for ADG and 0=11.0 mm for BF).

Table 6. Number of animals, range of environmental variables in seven groups for ADG and ADG adjusted for age

Group	Environmental variable (g/day)	Number of animals (ADG adjusted for age)	Number of animals (ADG)
1	<600	21419	17490
2	600 - 620	51889	41424
3	620 - 640	50448	52841
4	640 - 660	40041	61755
5	660 - 680	33341	40790
6	680 - 700	24255	17136
7	>700	24554	14511

Table 7. Number of animals, range of environmental variables in eight groups for BF

Group	Environmental variable (mm)	Number of animals
1	<9.7	15343
2	9.7 -10.2	30641
3	10.2 - 10.7	50040
4	10.7 -- 11.2	54469
5	11.2 - 11.7	48902
6	11.7 - 12.2	18607
7	10.2 - 12.7	15659
8	>12.7	12286

Variance components for sire by environment interactions

Similar estimates of heritabilities and common litter effects were found for different models or for different data sets analysed with the same model (Table 8). Estimates only varied from 0.20 t0 0.23 for heritablities and from 0.16 to 0.17 for common litter effects for ADG. For BF, ratios of variances were very consistent for both estimates ranging from 0.36 to 0.37 for heritability estimates and being 0.12 for common litter effects.

For ADG, the variances of sire slope on primary environmental variable obtained from Model B and D for the three data sets were significantly different from zero indicating that sires showed different environmental sensitivity (Table 8). The lowly negative genetic correlations between sire intercepts and slopes of -0.07 to -0.18 were not significant at the 0.05 level from Model B and D.

Sire slope variances on the primary environmental variable were not significant for BF based on Model D using the three alternative data sets (Table 8) and Model E fitting either homogeneous

or heterogeneous residual variances (Table 9). All estimates of genetic correlations between sire intercept and sire slope were not significant for BF.

Fitting heterogeneous residual variances in the model significantly increased the log likelihood for both traits (Table 9). However, fitting heterogeneous residual variances did not affect estimates of genetic parameters or fixed regression coefficients significantly.

The fixed regression coefficients from the regression of observations on primary environmental variables were close to 1.0 ranging from 0.99 to 1.02 for both traits (Table 9). In contrast, the fixed regression coefficients from the regression of observations on secondary environmental variables were not significant from zero for ADG and BF.

The variances of sire slopes based on the second environmental variable were significant for both traits indicating that inclusion of a second different environmental variable improved the model.

The genetic correlation between sire intercept and slope for ADG was affected by the linear adjustment for age. This genetic correlation was negative when ADG was adjusted for age with estimates ranging from -0.48 to -0.35 for alternative models. In contrast, this genetic correlation was not significant from zero when ADG was not adjusted for age at test.

The quadric term of the primary environment as an additional environmental variable in Model F was not significant for both traits and results are not shown here.

Table 8. Genetic parameter estimates with standard errors (SE) using models A, B, C and D for average daily gain (ADG, not adjusted for age) and backfat (BF, adjusted for weight): additive genetic variance ($\boldsymbol{\sigma_a^2}$), variance of slope on primary environmental descriptor ($\boldsymbol{\sigma_b^2}$), genetic correlation between sire intercept and slope on primary environmental descriptor ($r_{g_{a,b}}$), variance due to common litter effect ($\boldsymbol{\sigma_c^2}$), residual variance ($\boldsymbol{\sigma_e^2}$) heritability (h^2) and common litter effect (c^2).

Trait	Model	Data	$\boldsymbol{\sigma_a^2}$	SE	$\boldsymbol{\sigma_b^2}$	SE	$\boldsymbol{r_{g_{a,b}}}$	SE	$\boldsymbol{\sigma_c^2}$	SE	$\boldsymbol{\sigma_e^2}$	SE	$\boldsymbol{h^{2^*}}$	SE	$\boldsymbol{c^2}$	SE
ADG	A	full	242.0	11.0					707.3	9.4	3316.7	9.8	0.23	0.01	0.17	0.002
	B	subset1	230.9	11.9	0.013	0.004	-0.16	0.09	703.2	9.8	3325.3	10.2	0.22	0.01	0.17	0.002
	C	subset1	231.1	11.0					676.7	9.4	3325.4	10.2	0.22	0.01	0.16	0.002
	D	full	230.6	10.7	0.009	0.004	-0.07	0.10	672.9	9.1	3316.7	9.8	0.22	0.01	0.16	0.002
	D	subset1	226.3	11.3	0.009	0.004	-0.14	0.10	673.5	9.4	3325.3	10.2	0.21	0.01	0.16	0.002
	D	subset2	216.5	12.7	0.012	0.004	-0.18	0.10	678.0	10.3	3349.0	11.2	0.20	0.01	0.16	0.002
BF	A	full	0.28	0.01					0.371	0.006	2.394	0.007	0.37	0.014	0.12	0.002
	B	subset1	0.28	0.01	0.007	0.003	0.05	0.11	0.368	0.006	2.389	0.007	0.36	0.015	0.12	0.002
	C	subset1	0.27	0.01					0.352	0.006	2.389	0.007	0.36	0.014	0.12	0.002
	D	full	0.28	0.01	0.004	0.002	0.12	0.13	0.352	0.005	2.394	0.007	0.37	0.013	0.12	0.002
	D	subset1	0.27	0.01	0.004	0.002	0.08	0.13	0.351	0.006	2.389	0.007	0.36	0.014	0.12	0.002
	D	subset2	0.27	0.01	0.005	0.002	0.22	0.15	0.348	0.006	2.390	0.008	0.36	0.017	0.12	0.002

Note: h^{2^*} was calculated assuming an average environmental variable when the random reaction norm term within sires was fitted in Models B and D.

Table 9. Fixed regression coefficients on primary (Ave. slope1 (b)) and secondary (Ave. slope2 (c)) environmental variables, sire variances for intercept (V_a) and slope(V_b, V_c) and their genetic correlations ($r_{a,b}$ and $r_{a,c}$) and log likelihood values (LogL) for ADG and BF from Model D and E.

Model*	Ave. slope1(b)	Ave. slope2(c)	V_a	V_b	V_c	$r_{a,b}$	$r_{a,c}$	LogL
				Average daily gain (ADG)				
D, age, hom.	1.00 (0.008)		189.3 (9.88)	0.0075 (0.0026)		-0.48 (0.09)		119.2
D, age, het.	1.02 (0.008)		188.0 (9.80)	0.0079 (0.0026)		-0.43 (0.09)		1435.4
E, age, hom.	1.003 (0.008)	0.1383 (0.348)	181.8 (10.06)	0.0088 (0.0030)	17.74 (4.59)	-0.38 (0.09)	0.05 (0.09)	133.4
E, age, het.	1.014 (0.008)	-0.311 (0.343)	180.7 (9.98)	0.0095 (0.0030)	16.67 (4.32)	-0.35 (0.09)	0.04 (0.08)	1448.4
D hom.	1.00 (0.009)		226.3 (11.25)	0.0086 (0.0036)		-0.14 (0.10)		791.4
D het.	1.02 (0.008)		226.0 (11.26)	0.0093 (0.0036)		-0.12 (0.10)		945.9
E hom.	1.003 (0.009)	-0.043 (0.334)	221.0 (11.39)	0.0069 (0.0036)	13.55 (4.13)	-0.19 (0.12)	0.07 (0.10)	801.1
E het.	1.003 (0.009)	-0.038 (0.334)	220.8 (11.39)	0.0076 (0.0036)	13.38 (4.12)	-0.17 (0.11)	0.07 (0.10)	955.5
				Backfat (BF)				
D wt, hom.	0.994 (0.008)		0.27 (0.01)	0.0041 (0.0023)		0.08 (0.13)		358.4
D wt, het.	0.994 (0.008)		0.27 (0.01)	0.0043 (0.0023)		0.06 (0.13)		1019.6
E wt, hom.	0.999 (0.009)	0.00018 (0.00024)	0.27 (0.01)	0.0025 (0.0024)	5.2E-05 (2.1E-05)	0.41 (0.26)	0.48 (0.12)	366.9
E wt, het.	0.997 (0.008)	0.00031 (0.00024)	0.27 (0.01)	0.0028 (0.0024)	5.1E-05 (2.1E-05)	0.35 (0.21)	0.49 (0.12)	1028.9

Note: Vb: slope variance on primary environmental variable; Vc, slope variance on second environmental variable, for ADG, LSM for BF was second environmental variable, for BF, LSM for ADG was second environmental variable; age: adjusted for age; wt: adjusted for wt; het: heterogeneous residual variances; hom: homogeneous residual varianace.

Estimates of sire intercept and slope as well as range of environmental variable for sires

Fitting one environmental variable

The range of estimates of sire intercepts and slopes were 98.2 g/day and 0.253 g/d per g/d in the LSM for HBM group for ADG not adjusted for age (Table 10). When ADG was adjusted for age, additive genetic variance for sire intercept was reducing the range of estimates of sire intercepts to 89.8. Despite this lower variability in sire intercept, the range of estimates for sire slopes was slightly larger varying from -0.146 to 0.134. Based on the standard errors of sire slopes, only a small proportion of sire had slopes that were significantly different from zero, although differences between sires were found in regard to the response of their progeny to variation in environmental conditions.

Environmental sensitivity of sires can only be estimated for sires with progeny recorded across a wide range of environments. The average range per sire in environmental variables was 56.5 g/day for ADG adjusted for age and 53.9 for ADG varying from 0 and to a maximum of about 160 g/d for both trait definitions for ADG. Estimates of sire intercept and slope for BF varied from -1.43 to 2.00 mm. No genotype by environment interaction was found for BF which is reflected in small estimates of sire slopes ranging from -0.078 to 0.063 mm per mm in LSM for HBM, respectively. Therefore, progeny of sires had similar response to changes in environmental conditions for BF. The average range in the environmental variable was 1.44 mm per sire.

Table 10. Mean, standard deviation (SD), minimum and maximum for sire intercept, slope, standard error of slope and sire environmental range for ADG, adjusted ADG and BF from Model D by fitting heterogeneous residual error using subset 1 (number of sire: 1470)

Trait	Parameter	Mean	SD	Minimum	Maximum
	Sire intercept	0.00	11.0	-41.9	47.9
ADG_adj	Sire slope	0.00	0.037	-0.146	0.134
	s.e. of sire slope	0.08	0.006	0.057	0.089
	Sire environmental range	56.5	29.1	0	168.6
	Sire intercept	0.00	13.0	-46.6	51.9
ADG	Sire slope	0.00	0.028	-0.116	0.137
	s.e. of sire slope	0.09	0.003	0.070	0.096
	Sire environmental range	53.9	23.7	0	155.1
	Sire intercept	0.00	0.46	-1.43	2.00
BF	Sire slope	0.00	0.013	-0.078	0.063
	s.e. of sire slope	0.064	0.002	0.045	0.066
	Sire environmental range	1.44	0.90	0	5.15

Fitting two environmental variables

For ADG, the variance of sire slope on the primary environmental variable (least squares means for ADG of the contemporary group) was significantly different from zero indicating that sires showed different environmental sensitivity. In contrast, sire slope variances on the primary environmental variable were not significant for BF.

The variances of sire slopes based on the second environmental variable were significant for both traits indicating that inclusion of a second different environmental variable improved the model.

The range of estimates of sire intercepts and slopes was 93.3 g/day for sire intercepts and 0.205 g/d per g/d change in the primary environmental variable (LSM3 of contemporary groups) for sire slopes for ADG (Table 11). Based on the standard errors of sire slopes, only a small proportion of sires had slopes that were significantly different from zero, although differences between extreme sires were found in regard to the response of their progeny to variation in environmental conditions.

Environmental sensitivity of sires can only be estimated for sires with progeny recorded across a wide range of environments. The average range per sire in environmental variables was about 54 g/day for ADG varying from 0 and to a maximum of 155 g/d.

Estimates of sire intercept and slope for BF varied from -1.56 to 2.08 mm. No genotype by environment interaction was found for BF which is reflected in small estimates of sire slopes ranging from -0.076 to 0.080 mm per mm in LSM3 for contemporary groups, respectively. Therefore, progeny of sires had similar responses to changes in environmental conditions for BF. The average range in the environmental variable was 1.44 mm per sire.

Table 11. Mean, standard deviation (SD), minimum and maximum for sire intercept, sire slopes and sire environmental range for ADG and BF (number of sire: 1470)

Trait	Parameter	Mean	SD	Minimum	Maximum
ADG	Sire intercept	0.00	12.7	-41.7	51.6
	Sire slope 1	0.00	0.025	-0.102	0.103
	Sire slope 2	0.00	1.079	-5.037	6.778
	Sire environmental range	53.9	23.7	0	155.1
BF	Sire intercept	0.00	0.49	-1.56	2.08
	Sire slope 1	0.00	0.019	-0.076	0.080
	Sire slope 2	0.00	0.001	-0.004	0.004
	Sire environmental range	1.44	0.90	0	5.15

* Sire slope 1: based on primary environmental variable which was LSM3 of contemporary groups of the trait analysed; Sire slope 2: based on second environmental variable which was LSM3 of contemporary group of other trait;

Influence of number of progeny and environmental range per sire on estimates of sire slopes

The estimates of sire slope were predominantly related to the range of environments covered by sires (Figure 9). The proportion of more diverse estimates of sire slopes (positive or negative) increased when a wider environmental range was covered by the sire. Few sires were available with progeny recorded across an environmental range of over 80 g/d and the number of estimates of sire slopes was not sufficient in these data to draw any conclusions for larger environmental ranges.

A minimum of 50 progeny per sire was imposed and larger number of progeny had little effect on diversity of estimates of sire slopes (Figure 10) indicating that 50 progeny per sire is sufficient as long as progeny are recorded over a wide range of environments.

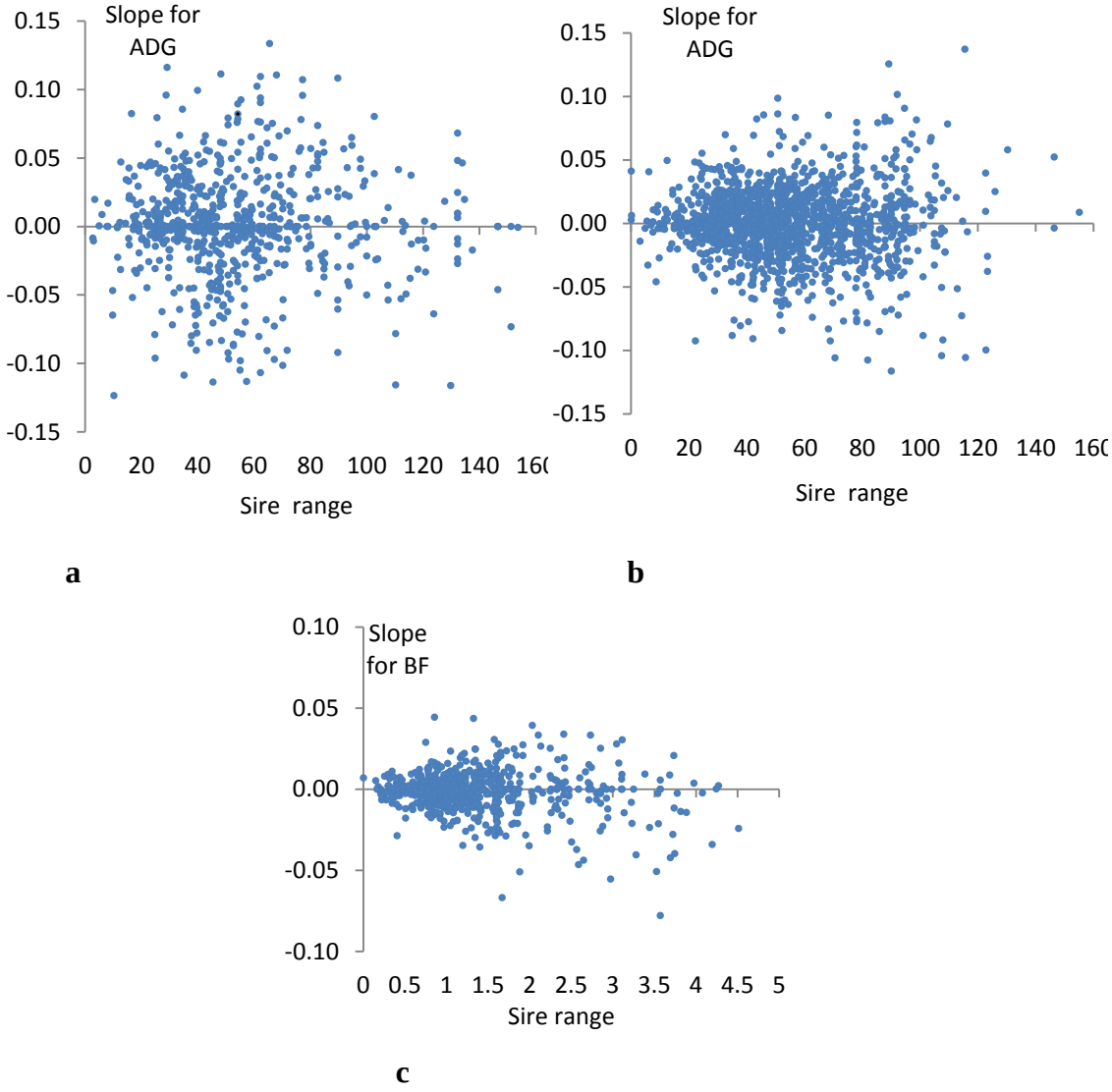


Figure 9. Distribution of estimates of sire slopes versus the environmental range represented by a sire (results from Model D fitting heterogeneous residual and using subset 1). (a) ADG adjusted for age; (b) ADG unadjusted for age; (c) BF.

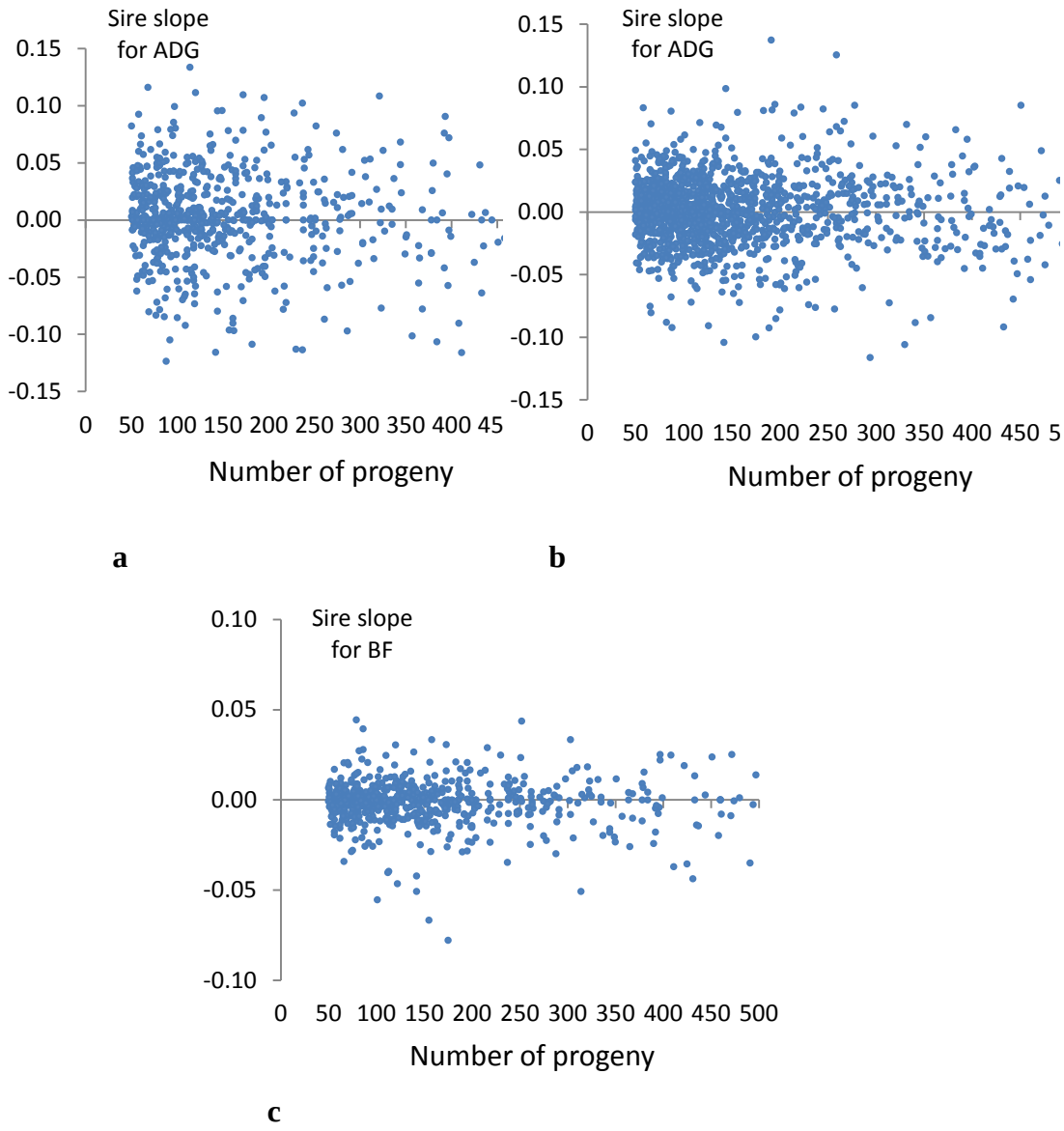


Figure 10. Distribution of sire slopes versus the number of progeny per sire (results from Model D fitting heterogeneous residual and using subset 1). (a) ADG adjusted for age. (b) ADG unadjusted for age. (c) BF.

Illustration of sire intercepts and sire slopes

The level and slope of sires can be used to predict the estimated genetic merit of progeny of sires across environments, which illustrate potential re-ranking of sires across environments. Eight sires were selected to represent a good range of estimates for the intercept and slope for this illustration. Re-ranking of sires across environments is illustrated for ADG defining the environment via the mean growth performance Figure 11) or the average performance in backfat (Figure 12).

Current selection is based on an average environment ranking sires from A to G. Sire C is least robust (most sensitive) to differences in the ADG environment. This sire ranks first in a very good ADG environment (+ 70 g/d) and is only ranked fifth in the inferior ADG environments. In comparison, sire A is the superior sire across all BF environments and shows no genetic sensitivity to changes in the BF environment. It

is sire D that is least robust to changes in the BF environment. This sire is a high-growth sire in a high-BF environment with a low-growth performance in low-BF environment in comparison to other sires. Please note that the ranking of sires in the average environment was the same for the ADG or BF environment and sires were named the same in both Figures.

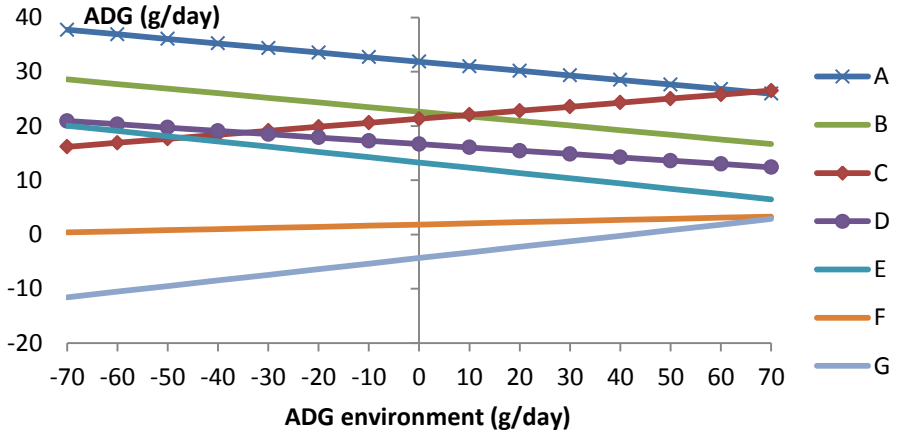


Figure 11. Predicted breeding value for ADG for sires across ADG environments based on estimates of the intercept and slope for each sire from Model D with heterogeneous residuals using subset 1. The x-axis is mean centered (0= 644.2 g/day for ADG)

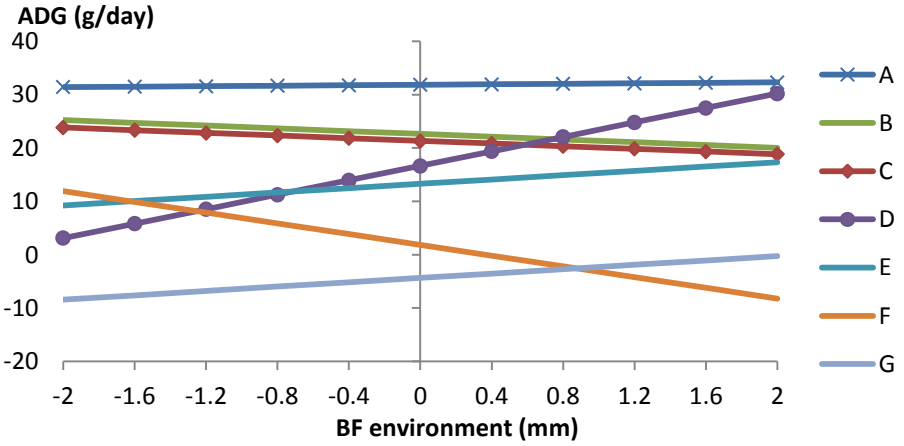


Figure 12. Predicted breeding value ADG for sires across BF environments based on estimates of the intercept and slope for each sire from Model D with heterogeneous residuals using subset 1. The x-axis is mean centered (0=11.0 mm for BF).

Genetic parameters across environmental trajectory

The heritability of ADG adjusted for age decreased from 0.35 to 0.15 across the environmental trajectory when a homogeneous residual variance was used (Figure 13a). This decrease was less pronounced when heterogeneous residual variances were fitted. In comparison, heritability estimates for ADG not adjusted for age were lower in the middle of the environmental range (Figure 13b). Heritability estimates for BF increased slightly when a homogeneous residual variance was used (Figure 13c), while a different pattern was observed for heritability estimates obtained fitting

heterogeneous residual variances. Heritability estimates were higher for low to medium environmental values due to the lower estimates of residual variances at these environments.

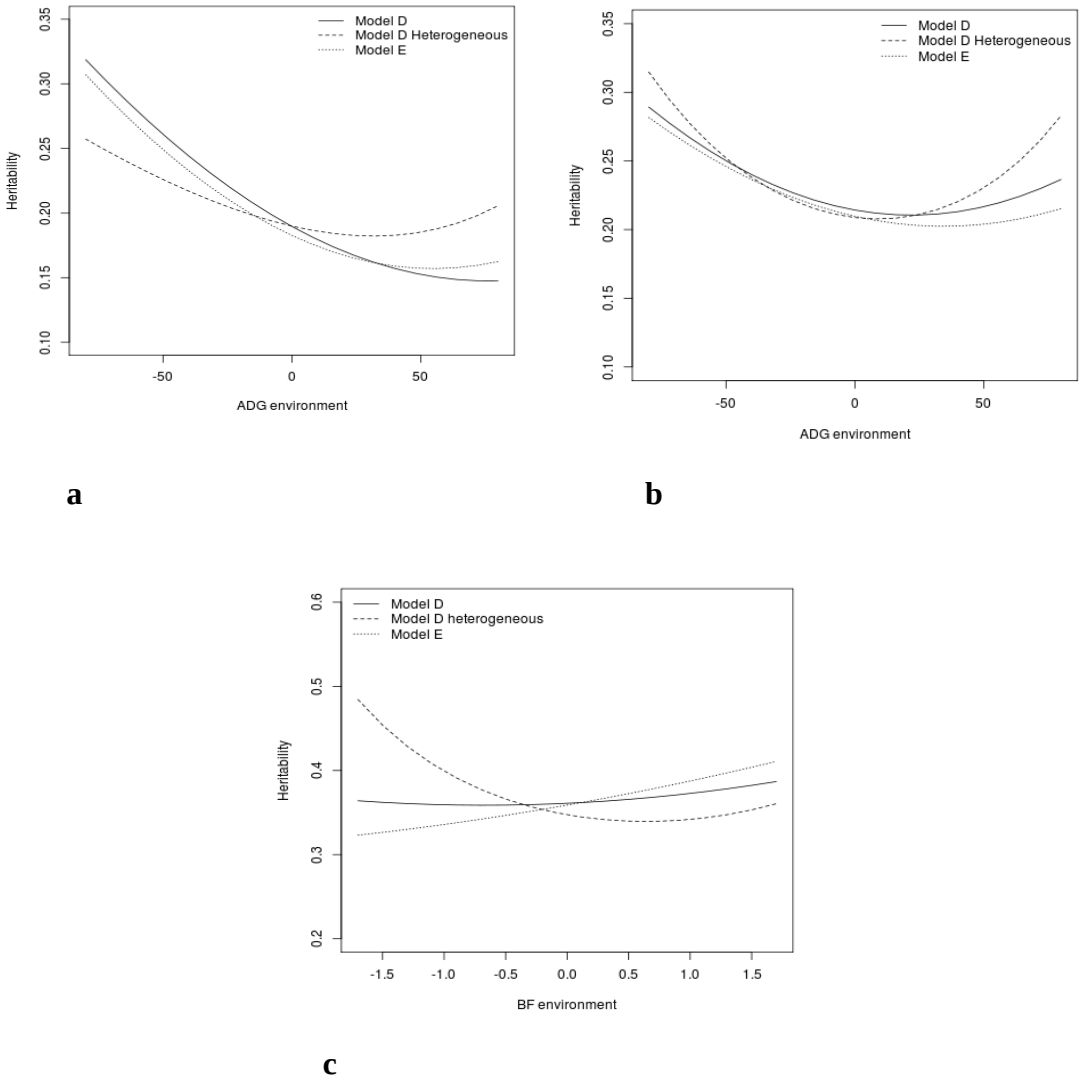
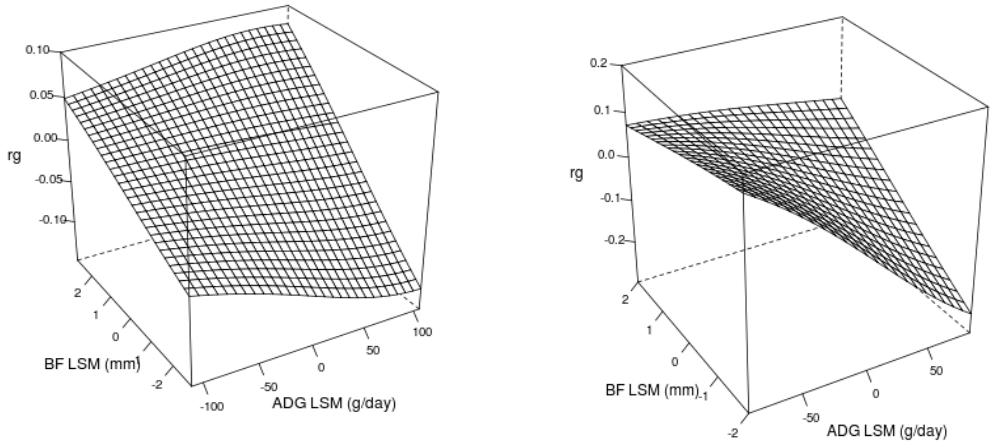


Figure 13. Heritability estimates for average daily gain (ADG) and backfat (BF) across environments from Models D and E fitting either homogeneous or heterogeneous residual variances using subset 1. (a) ADG adjusted for age; (b) ADG unadjusted for age; (c) BF. The x-axis is mean centered (0= 644.2 g/day for ADG and 0=11.0 mm for BF).

Genetic correlations between the two definitions of ADG and BF varied from -0.23 to 0.15 with lower, more favourable estimates for the superior environments for both traits (Figure 14). The pattern of genetic correlations across the environmental trajectories differed for environments with low growth rate and low backfat. When ADG was adjusted for age, it had a higher genetic correlation with BF for these environments. This shift in estimates of genetic correlations was not observed for ADG not adjusted for age.



a **b**

Figure 14. Genetic correlation between average daily gain (ADG) and backfat (BF) across two environments fitting Model D with heterogeneous residual variances s using subset 1. (a) ADG adjusted for age; (b) ADG unadjusted for age. The x-axes were mean centered (0= 644.2 g/day for ADG and 0=11.0 mm for BF)

3. Outputs

The main results from this project have been descripted above. So far, three outputs have been prepared and the main findings have been reported in the Annual Report of the Pork CRC. The main findings were presented at the annual meeting in November 2012. Other refereed international conference and Journal papers are planned.

Outputs

Li, L. and S. Hermesch (2012). Genotypes differ in their response to variation in environments experienced by pigs on farm. 2012 AGBU Pig Genetics Workshop Notes. Editors: S. Hermesch and K. Dobos. Armidale, AGBU, pp**:** 53-60.

Li, L. and S. Hermesch (2012). Variation in environments and GxE for growth and back fat in pigs - Part 1 -. Presentation at AGBU, Armidale, NSW, Australia, 27th August, 2012.

Li, L. and S. Hermesch (2012). Variation in environments and GxE for growth and back fat in pigs - Part 2 -. AGBU, Presentation at Armidale, NSW, Australia, 3rd September, 2012.

4. Application of Research

These analyses demonstrated a considerable spread in environments in high-health farms with good husbandry practices. It is noteworthy that considerable environmental variation was observed within herds and within years.

The unadjusted mean performance of a group of pigs showed a similar distribution than the more complex environmental descriptors based on general linear models. This simple environmental descriptor can easily be derived from standard performance records collected on farms to quantify variation in environmental conditions within herds over time. Therefore, a simple on-farm measure has been developed to quantify environmental conditions.

This simple on-farm measure based on the mean performance of groups of pigs may be extended by using the variation between pigs within a group, as it was found that phenotypic and residual variance was lower in superior environments.

Most breeding companies in Australia have sufficient data to quantify the response of different lines or progeny of sires to variation in the environment. However, breeders will require guidance for the adoption of results as outlined further below.

The immediate impact of this study is on other research studies. The methodology used in this study will be invaluable for further research into disease resilience, disease tolerance and adaptation of animals to heat stress.

The methodology may also be used to obtain EBVs for sires to select for less environmental sensitivity leading to more robust pig genotypes.

5. Conclusions

Variation in environments

The variation in environmental conditions was quantified through predictors of the mean performance of a group of pigs that were born in the same herd during a one-month period of time (contemporary group, CG). Alternative models were used to predict the mean performance of each CG. The distribution of mean performance of CGs showed a spread of about 150 g/day for growth rate and a spread of about 5 mm for backfat across herds and years independent of the model used. This considerable spread in mean performance of CG was observed in high-health herds with good overall husbandry and management conditions showing that even in these overall good conditions on farm, considerable environmental variation existed for individual groups of pigs.

A further important aspect is whether this variation in environments observed across herds is also apparent within herds and within years because pigs may be selected within-herds only at a specific point in time. A similar spread of mean performance within herds and within years was found for growth rate and backfat.

Breed by environment interactions

Breeds differed in their responsiveness to variation in environmental conditions with Large White being the most environmentally sensitive or least robust breed for growth rate and backfat. This breed was the leanest breed in comparison to Landrace and Duroc. The most robust breed was Duroc, which had a similar growth rate as Large White but was characterised by a higher backfat in comparison to the other breeds. Overall, these results support the hypothesis that leaner genotypes tend to be less robust and less able to perform consistently across a range of environmental conditions.

Sire by breed interactions

Extensive analyses were performed using random regression models to evaluate sire by environment interactions. For growth rate, sires differed in the response of their progeny to variation in the environment, which was less apparent for backfat. The implications of data structure in regard to the number of progeny per sire and the range in environmental variation per sire for the ability to estimate genetic differences in environmental sensitivity between sires were outlined. This information is useful for pig breeding companies to setup strategic performance recording procedures for genetic improvement of both, productivity and robustness.

Genetic parameters across environmental trajectory

Heritability estimates differed across the environmental trajectory for growth rate and estimates were higher for inferior environments. The additive genetic variance across the trajectory is based on a function involving the environmental variable and estimates of the intercept and slope of sires. Given these estimates from the reaction norm model additive genetic variance was higher at the lower range of the environmental trajectory. This led to slightly higher heritabilities for inferior environments despite their higher residual and observed variances.

6. Limitations/Risks

Mean performances of groups of pigs is available on commercial farms that do not record individual animals. Therefore, this simple on-farm measure of environmental conditions can be used on commercial farms.

Variation in environments was observed within herds and within years. This implies that selection of genotypes with reduced environmental sensitivity can be based on within-herd genetic evaluations. However, it is beneficial to have a wider range of environments represented by each sire, which requires a sufficient number of progeny recorded across environments. This may limit effective use of this methodology to larger populations.

The models used in these analyses are quite technical and breeders on farm with limited statistical training may find the application of these results on farm challenging. Guidance by experienced geneticists is required for the adoption of results from this study by pig breeding companies as results were quite sensitive to the models used. Breeding companies will require guidance from experienced geneticists for the implementation of results in pig breeding programs.

7. Recommendations

Further research should focus on developing more precise methodology to measure variation in environmental conditions on farm making use of information about seasonal effects, air quality, disease incidence and overall performance of pigs based on multiple traits. Multiple environmental descriptors were significant for growth rate, highlighting the need to develop an overall environmental index that combines multiple environmental factors.

Sires need to be represented across a wide range of environments. Avenues to increase the spread of progeny of sires across environmental conditions should be

explored with breeders in order to improve data structures for future genetic analyses of genotype by environment interactions.

Further, it was shown that residual variance was lower in superior environments in particular for growth rate. Therefore, the environmental descriptor may be extended by taking variation between animals within the same environmental into account. It is a further criterion that may capture environmental conditions better and will become available as more and more automatic weighing devices of individual pigs are installed on farms. Precision pork production will be developed further and measures of variation will be used to optimize pork production beyond current on-farm practices.

Breed and sire by environment interactions were found. The methodology developed in this study can be used by AGBU staff to estimate breeding values (EBVs) for the intercept and slope of sires. Firstly, these EBVs can be used to select sires with superior performance consistently across environments. This is a long-term breeding goal. Secondly, the variation in EBVs for the intercept and slope can be used to select sires whose progeny are best suited for a specific environment. This information should be used by breeders and producers to better match genotypes for specific environments.

8. References

Bergsma, R., and S. Hermesch. "Exploring Breeding Opportunities for Reduced Thermal Sensitivity of Feed Intake in the Lactating Sow." *Journal of Animal Science* 90, no. 1 (2012): 85-98

Black, J L, L R Giles, P C Wynn, A G Knowles, C A Kerr, M R Jones, A D Strom, N L Gallagher, and G J Eamens, eds. *A Review - Factors Limiting the Performance of Growing Pigs in Commercial Environments*. Edited by P D Cranwell. Vol. Eigth Biennial Conference, Manipulating Pig Production Viii. Adelaide, Australia: Australian Pig Science Association, 2001

Bradshaw, A. D. 1965. Evolutionary significance of phenotypic plasticity in plants. Advances in Genetics 13: 115-155.

Bryant, J. R. et al. 2007. Environmental sensitivity in New Zealand dairy cattle. Journal of Dairy Science 90: 1538-1547.

Burdon, R. D. 1978. Genetic correlation as a concept for studying genotype-environment interaction in forest tree breeding. Silvae Genetica 26: 168-175.

Calus, M. P. L., and R. F. Veerkamp. "Estimation of Environmental Sensitivity of Genetic Merit for Milk Production Traits Using a Random Regression Model." *Journal of Dairy Science* 86, no. 11 (2003): 3756-3764.

Calus, M. P. L., P. Bijma, and R. F. Veerkamp. 2004. Effects of data structure on the estimation of covariance functions to describe genotype by environment interactions in a reaction. Genet. Sel. Evol. 36: 489-507.

Calus, M. P. L., J. J. Windig, and R. F. Veerkamp. 2005. Associations among descriptors of herd management and phenotypic and genetic levels of health and fertility. Journal of Dairy Science 88: 2178-2189.

Cardoso, F. F., and R. J. Tempelman. 2012. Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. Journal of Animal Science.

Crossa, J. 1990. Statistical analyses of multilocation trials. Advances in Agronomy 44: 55-85.

Crossa, J., and P. L. Cornelius. 1997. Sites regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of error variances. Crop Science 37: 406-415.

Falconer, D. S. 1990. Selection in different environments – effects on environmental sensitivity (reaction norm) and on mean performance Genetical Research 56: 57-70.

Fikse, W. F., R. Rekaya, and K. A. Weigel. 2003. Genotype x environment interaction for milk production in Guernsey cattle. Journal of Dairy Science 86: 1821-1827.

Gilmour, A. R., Gogel, B. J., Cullis, B. R. and Thompson, R. 2009. ASReml user guide. Release 3.0. VSN International Ltd., Hemel Hempstead, UK. 398 pp.

Hammami, H., B. Rekik, C. Bastin, H. Soyeurt, J. Bormann, J. Stoll, and N. Gengler. "Environmental Sensitivity for Milk Yield in Luxembourg and Tunisian Holsteins by Herd Management Level." *Journal of Dairy Science* 92, no. 9 (2009): 4604-4612.

Haskell, M. J., S. Brotherstone, A. B. Lawrence, and I. M. S. White. 2007. Characterization of the dairy farm environment in great Britain and the effect of the farm environment on cow life span. Journal of Dairy Science 90: 5316-5323.

Hermesch, S. "Genetic Improvement of Lean Meat Growth and Feed Efficiency in Pigs." *Australian Journal of Experimental Agriculture* 44, no. 4-5 (2004): 383-391.

Hermesch, S., A. E. Huisman, B. G. Luxford, and H. U. Graser. 2006. Analysis of genotype by feeding level interaction in pigs applying reaction norm models. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006, 2006.

Hill, J. 1975. Genotype-environment interaction – a challenge for plant breeding. The Journal of Agricultural Science 85: 477-493.

Gilmour, A. R., Gogel, B. J., Cullis, B. R. and Thompson, R. 2009. ASReml user guide. Release 3.0. VSN International Ltd., Hemel Hempstead, UK. 398 pp.

Kang, M. S. 1998. Using genotype-by-environment interaction for crop cultivar development. In: D. L. Sparks (ed.) Advances in Agronomy, Vol 62. Advances in Agronomy No. 62. p 199-252.

Knap, P. W., and G. Su. 2008. Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. Animal 2: 1742-1747.

Kolmodin, R., E. Strandberg, P. Madsen, J. Jensen, and H. Jorjani. 2002. Genotype by Environment Interaction in Nordic Dairy Cattle Studied Using Reaction Norms. Acta Agriculturae Scandinavica, Section A - Animal Science 52: 11-24.

Lewis, C.R.G., W. O. Herring, and K.L. Bunter. "A Random Regression Analysis of Sow Lactation Feed Intake and the Effect of Temperature on Intake." Proceedings

of World Congress for Genetics Applied to Livestock Production, Leipzig, Germany, 2010.

Mattar, M., L. O. C. Silva, M. M. Alencar, and F. F. Cardoso. 2011. Genotype x environment interaction for long-yearling weight in Canchim cattle quantified by reaction norm analysis. Journal of Animal Science 89: 2349-2355.

Pegolo, N. T., L. G. Albuquerque, R. B. Lobo, and H. N. de Oliveira. 2011. Effects of sex and age on genotype x environment interaction for beef cattle body weight studied using reaction norm models. Journal of Animal Science 89: 3410-3425.

Pollott, G. E., and J. C. Greeff. 2004. Genotype x environment interactions and genetic parameters for fecal egg count and production traits of Merino sheep. Journal of Animal Science 82: 2840-2851.

Prunier, A., M. Heinonen, and H. Quesnel. "High Physiological Demands in Intensively Raised Pigs: Impact on Health and Welfare." *Animal* 4, no. 6 (2010): 886-898.

Ramburan, S., M. Zhou, and M. Labuschagne. 2011. Interpretation of genotype by environment interactions of sugarcane: Identifying significant environmental factors. Field Crops Research 124: 392-399.

Strandberg, E. 2006. Analysis of genotype by environment interaction using random regression models.

Strandberg, E., S. Brotherstone, E. Wall, and M. P. Coffey. 2009. Genotype by environment interaction for first-lactation female fertility traits in UK dairy cattle. Journal of Dairy Science 92: 3437-3446.

Su, G. et al. 2006. Bayesian analysis of the linear reaction norm model with unknown covariates. Journal of Animal Science 84: 1651-1657.

Wu, H. X., and C. C. Ying. 2001. Variation on reaction norm in lodgepole pine natural populations. Theoretical and Applied Genetics 103: 331-345.

Yan, W., M. S. Kang, B. Ma, S. Woods, and P. L. Cornelius. 2007. GGE Biplot vs. AMMI Analysis of Genotype-by-Environment Data. Crop Sci. 47: 643-653.

Zobel, R. W., M. J. Wright, and H. G. Gauch. 1988. Statistical analysis of a yield trial. Agronomy Journal 80: 388-393.

Zwald, N. R., K. A. Weigel, W. F. Fikse, and R. Rekaya. 2003. Identification of factors that cause genotype by environment interaction between herds of Holstein cattle in seventeen countries. Journal of Dairy Science 86: 1009-1018.