Estimation of genetic parameters for body weight of the Goettingen minipig with random regression models

F. Köhn,* A. R. Sharifi,* Š. Malovrh,† and H. Simianer*

*Institute of Animal Breeding and Genetics, University of Göttingen, 37075 Göttingen, Germany; †Animal Science Department, Biotechnical Faculty, University of Ljubljana, 1230 Domžale, Slovenia

ABSTRACT: The Goettingen minipig is a laboratory animal especially developed for medical research. For easy and comfortable handling during experiments, and to minimize costs, a low BW is essential. To breed for an even smaller minipig, genetic parameters for BW were estimated using a random regression model (RRM). The RRM was calculated using random animal, common litter environment, and permanent environment effects, respectively. Regressions for the random effects in the RRM were modeled using Legendre polynomials from second to fourth order of fit in different combinations. The model was applied to a data set that focused on the time period from 30 to 400 d of age. Eight age classes were built to consider heterogeneous residual variances. The heritabilities were moderate and ranged from 0.211 (375 d of age) to 0.254 (275 d of age). The variances initially decreased and then increased toward the end of the examined time period for permanent environment and litter effects. Genetic and phenotypic correlations between BW in different age classes decreased with increasing distance between age classes. The major eigenfunction showed positive values throughout the whole trajectory (i.e., a selection for low BW had positive effects on this trait throughout the whole range of time). On the basis of the estimated genetic parameters, a breeding scheme can be created to develop genetically smaller Goettingen minipigs in the future.

Key words: body weight, genetic parameter, Goettingen minipig, random regression

INTRODUCTION

The Goettingen minipig is a laboratory animal that has a widespread popularity in medical and pharmaceutical research because of its shared anatomic and physiologic characteristics with humans. In general, swine are becoming more and more important as nonrodent models in medical experiments. Major advantages are that pigs are not considered companion animals as are dogs, and they do not have physical similarities or phylogenetic proximity to humans as do primates. With respect to ethics and animal welfare, pigs therefore are a much less disputed model species in medical research (Gad, 2007). The most important advantage of the minipig in comparison with normal pigs is the smaller body size and the resulting low BW of 35 to 45 kg for adult minipigs (Bollen et al., 1998). The low BW is a necessary trait because handling in medical experiments is more easily facilitated, and the costs for housing, feeding and dosing are lower. At present there is no line of minipig that has been selected for a low BW, and the inheritance of low BW of Goettingen minipigs at different ages is still unknown. It is necessary to have a genetically small minipig, not simply a pig of reduced weight from restricted feeding (Brandt et al., 1997).

Body weight is a trait where the phenotype of an animal can be represented by a continuous function of time. Thus, this trait is characterized by a trajectory with a theoretically infinite number of measurements. Therefore, an appropriate model is one that considers a complex covariance structure. In the infinite-dimensional approach, the covariance structure is modeled as a covariance function (CF; Kirkpatrick et al., 1990). The CF quantifies the variation of the individual growth curves about the population mean. A useful tool for the estimation of CF is the random regression model (RRM; Van der Werf, 2001). The aims of this study were the estimation of genetic parameters for the trait BW with a RRM and conclusions about prospects of selection response for low BW.

MATERIALS AND METHODS

Animal Care and Use Committee approval was not obtained for this study because the data were obtained...
from the database Naviision from Ellegaard Goettingen Minipigs ApS, Dalmose, Denmark.

Body weight data were provided from Ellegaard Goettingen Minipigs ApS, where 2 subpopulations of the Goettingen minipig base population are housed in 2 units that are entirely separated from each other, both physically and with respect to gene flow between them. The data were acquired from 1995 to 2005. The minipigs were weighed routinely at various intervals, without provision for fasting before weighing. All pigs were weighed at birth, then at weaning (21 to 28 d of age), and again at 8 wk of age, when they left the rearing unit. Later, all minipigs were weighed once each month, and each minipig was weighed before it was sold. Unlike slaughter pigs, minipigs are not kept until a certain age or weight, but are sold for different purposes at different ages. Therefore, the number of available weights per pig varied considerably.

In total, the original data set contained 199,764 BW records of 33,749 animals. Because of scarcity, BW measured after 400 d of age were excluded from the analyses. Outliers of the data set were also excluded using the Studentized residual. This resulted in a total of 180,092 BW records of 32,510 animals.

A data set, with a focus on 30 to 400 d of age, was prepared out of this data for the random regression analysis. The BW from d 0 to 29 were excluded because they are influenced by the parity and the lactation length of the sow and cannot be classified as independent BW traits like those after weaning.

The residuals, or the so-called temporary environmental effect, should not be assumed to be homogeneous over the observed range of time (Schaeffer, 2004). After a graphical analysis of the residuals, they were considered independently distributed, with heterogeneous variances. Changes in the residual variances with age were modeled as a step function with different age classes. Therefore, the number of available BW measurements per pig varied considerably.

In total, the original data set contained 199,764 BW records of 33,749 animals. Because of scarcity, BW measured after 400 d of age were excluded from the analyses. Outliers of the data set were also excluded using the Studentized residual. This resulted in a total of 180,092 BW records of 32,510 animals.

A data set, with a focus on 30 to 400 d of age, was prepared out of this data for the random regression analysis. The BW from d 0 to 29 were excluded because they are influenced by the parity and the lactation length of the sow and cannot be classified as independent BW traits like those after weaning.

The residuals, or the so-called temporary environmental effect, should not be assumed to be homogeneous over the observed range of time (Schaeffer, 2004). After a graphical analysis of the residuals, they were considered independently distributed, with heterogeneous variances. Changes in the residual variances with age were modeled as a step function with different age classes. Therefore, the number of available BW measurements per pig varied considerably.

Table 1. Age classes, age range (d), number of records, mean BW (kg), and variance of BW (kg²) per age class

<table>
<thead>
<tr>
<th>Age class</th>
<th>Age range, d</th>
<th>No. of records</th>
<th>Mean BW, kg</th>
<th>Variance of BW, kg²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30 to 60</td>
<td>5,792</td>
<td>3.33</td>
<td>1.04</td>
</tr>
<tr>
<td>2</td>
<td>61 to 100</td>
<td>5,761</td>
<td>5.28</td>
<td>1.54</td>
</tr>
<tr>
<td>3</td>
<td>101 to 150</td>
<td>5,855</td>
<td>8.36</td>
<td>3.24</td>
</tr>
<tr>
<td>4</td>
<td>151 to 200</td>
<td>5,821</td>
<td>11.84</td>
<td>5.61</td>
</tr>
<tr>
<td>5</td>
<td>201 to 250</td>
<td>5,639</td>
<td>15.25</td>
<td>7.82</td>
</tr>
<tr>
<td>6</td>
<td>251 to 300</td>
<td>4,105</td>
<td>18.69</td>
<td>9.06</td>
</tr>
<tr>
<td>7</td>
<td>301 to 350</td>
<td>3,012</td>
<td>21.82</td>
<td>10.14</td>
</tr>
<tr>
<td>8</td>
<td>351 to 400</td>
<td>2,038</td>
<td>24.86</td>
<td>10.48</td>
</tr>
</tbody>
</table>

where \( t_{\text{min}} \) is the youngest age, here 30 d of age, and \( t_{\text{max}} \) is the oldest age, here 400 d of age, respectively. The random regression animal model was

\[
y_{ijkl}(t) = \mu + S_i + U_j + Y_k + \sum_{r=1}^{3} \beta_r t^r + \sum_{s=1}^{r} \alpha_{sijkl} t^s + \sum_{s=1}^{r} \rho_{sijkl} \phi_s(t^s) + \epsilon_{ijkl},
\]

where \( y_{ijkl}(t) \) = weight of animal \( l \) at age \( t \) within sex \( i \), unit \( j \), and birth year \( k \); \( S_i \) = fixed effect of sex \( i \); \( U_j \) = fixed effect of unit \( j \); \( Y_k \) = fixed effect of birth year \( k \); \( \beta_r \) = fixed cubic regression coefficient; \( t = \text{age in d} \); \( t^s \) = age standardized to the range –1 to 1; \( \phi_s \) = value of the \( s \)-th LP at standardized age \( t^s \); \( \alpha_{sijkl} \) = random regression coefficient for additive genetic effects; \( \gamma_{sijkl} \) = random regression coefficient for permanent environmental effects; \( \rho_{sijkl} \) = random regression coefficient for common environmental effects for litter; and \( \epsilon_{ijkl} \) = random measurement error.

The fixed regression with the fixed effects sex, unit, and birth year was calculated with a polynomial of third order of fit. This polynomial model was chosen as that with the best fit after applying different linear and nonlinear functions to growth data of Goettingen minipigs (Köhn et al., 2007). The random regression coefficients for additive genetic effect, permanent environmental effect, and common litter environmental effect, were fitted with LP from quadratic (LP 2) with 3 terms to quartic (LP 4) LP with 5 terms. Every possible combination of LP for the random effects was calculated.

The final pedigree contained 35,066 animals. The 6,713 animals that were represented in this study were offspring of 332 sires and 1,030 dams.

Covariance components for the random regression coefficients estimated via REML were obtained by using VCE-5 (Kovac et al., 2002). With the procedure IML from the software package SAS (SAS Inst. Inc., Cary, NC), the eigenfunctions with corresponding eigenvalues were calculated for the covariance components of the random regression coefficients. Eigenfunctions and eigenvalues can provide insight into the potential for genetic change in the average growth curve of the whole
Genetic parameters for body weight

Figure 1. Variances (kg²) for random and residual effects estimated with random regression model (LP 2). pe = permanent environment.

Population. They are calculated from the covariance matrix of the random animal effects. To each eigenfunction, a specific eigenvalue is associated. This eigenvalue represents the amount of variation explained by the corresponding eigenfunction. Accordingly, a large eigenvalue represents considerable genetic variation for pattern of growth and changes in the growth curve that can be modified by selection (Bermejo et al., 2003).

Jamrozik and Schaeffer (2002) outlined different test criteria like Akaike’s information criterion or the information criterion based on Bayes statistics rank models in a different way. Due to the fact that it is unclear which criterion is the best for choosing the right model in this study, the model comparison was carried out by using the estimated mean squared error, as was done by Meyer (2005b).

RESULTS

After examining the estimated variances and variance proportions for all models, the best fitting RRM was the model with a polynomial of third order of fit for the fixed effects and LP of second order of fit (LP 2) for all random effects. In the following, only the results of this model are presented.

The variances of random animal effect increased up to an age of 125 d of age (Figure 1). At 175 d of age a very low value was estimated, and afterwards the variances increased again. The variances for litter, permanent environment, and residual effect increased steadily over the examined time period.

The heritabilities increased up to 275 d of age with a lower value at 80 d of age and decreased afterwards (Figure 2). The proportions of variance for the litter effect decreased continuously up to 275 d of age and increased again afterwards. For the permanent environmental effect the proportions of variance increased steadily. The proportions of the residual variance increased and decreased over the examined time period but on a more or less constant level between 0.14 (45 d) and 0.24 (80 d).

As expected, the genetic and phenotypic correlations between BW in different age classes decreased continuously with increasing distance between the age classes (Figure 3). The estimates of the first eigenfunction were positive throughout (Figure 4). As a consequence, selection on the trait low BW at any age will change BW in the period of 30 to 400 d of age in the same direction. The first eigenvalue explained 89.93% of the genetic variation. As outlined by Kingsolver et al. (2001), eigenvalues indicate the amount of variance explained by its associated eigenfunction. The second eigenfunction explained 9.45% of the genetic variation. In this case, a selection before 305 d of age on the trait low BW leads to a negative effect on this trait from 305 to 400 d of age. The third eigenfunction explained only 0.62% of

Figure 2. Variance proportions for the random and residual effects estimated with random regression model (LP2). pe = permanent environment.

Figure 3. Genetic (above diagonal) and phenotypic (below diagonal) correlations between BW in different age classes.
the genetic variation and can be neglected. It can be concluded that the genetic variation associated with the second and third eigenfunction is small compared with the genetic variation associated with the first eigenfunction.

DISCUSSION

The importance of Goettingen minipigs in medical research is increasing every year. To ensure a genetically small pig and to decrease BW of Goettingen minipigs in the future, a breeding strategy for the trait low BW has to be developed. This is the first genetic study made for BW of Goettingen minipigs in such detail. Additionally, this data set is unique in pig research due to the high number of animals that had at least 5 weight recordings over a wide time range. If the analyzed trait is measured along a trajectory, which is the case with BW measurements, the covariance structure should take the ordering of the measurements in time into account. The RRM is able to give the covariances of a trait that is measured at different ages as a function of these ages. There is no need for an adjustment of ages with this model, and as a consequence, errors associated with an adjustment are avoided (Albuquerque and Meyer, 2001).

The main advantages of RRM in comparison to the conventionally used multiple trait model are a smaller number of parameters to describe longitudinal measurements, smoother covariance estimates, no regular measurement schedules are needed, as well as the possibility to estimate covariance components and predict breeding values at any point along the trajectory. The RRM enable modeling and investigation of growth curves as a function of age for individual animals.

The RRM have been used for the analysis of test day records in dairy cattle (Schaeffer and Dekkers, 1994; Jamrozik et al., 1997; Van der Werf et al., 1998; Veerkamp and Thompson, 1999). Test day records are a very typical example for a repeated measured trait in animal breeding because milk yield per test day changes during the lactation. The same can be observed for BW of animals that are also changing gradually until a plateau at maturity is reached (Arango and Van Vleck, 2002). Thus, growth is a longitudinal process with many measurements that are highly correlated. Many studies used RRM for the estimation of genetic parameters of growth in beef cattle (Meyer, 1999, 2000, 2001; Albuquerque and Meyer, 2001; Nobre et al., 2003b; Arango et al., 2004; Bohmanova et al., 2005; Krejcova et al., 2005) and sheep (Lewis and Brotherstone, 2002; Fischer et al., 2004). In the last few years RRM were also applied for the investigation of genetics in pig growth and feed intake (Huisman et al., 2002; Malovrh, 2003; Huisman et al., 2005).

When including random effects in the model, the total variation can be separated in within- and between-animal variation. Normally, nonlinear functions like the Gompertz, von Bertalanffy, Brody, and Richards function are used to describe growth curves in pigs (Köhn et al., 2007). They are preferred because of their well-founded biological meaning. When using nonlinear functions, the growth curve has to be modeled and the growth parameters of the particular function have to be estimated for each animal. In a second step, the estimation of environmental effects and variance components has to be conducted. The problem is that the information on relatives is not considered while estimating the growth curve. Further, errors that are typically correlated in repeated measures analysis are not taken into account with these traditional approaches (Albuquerque and Meyer, 2001). These disadvantages are the reasons for choosing linear polynomial models for the estimation of variance components with RRM. Kirkpatrick et al. (1990) proposed to use orthogonal polynomials of standardized units of time as covariates for the random regressions. The most common type of orthogonal polynomials used is the LP. In this study, a polynomial of third order was used for modeling the fixed part of the regression and LP were used to model the regressions of random effects.

The heritabilities for BW were moderate in this study (Figure 2). The decrease for the estimated heritability at 80 d of age is due to high residual variance ratios at this age. This can be explained by a lower number of analyzed weight recordings in the second age class and a higher variance in weights compared with the first age class.

Huisman et al. (2002) calculated lower heritabilities of 0.17 (70 d of age) to 0.18 (190 d of age). They estimated the heritabilities with the weight data of 1,315 boars. The RRM was fit as a sire model and not as an animal model as in this study. Therefore, the results are not completely comparable with our results. In the study of Malovrh (2003) 3,819 boars of 4 German pure-bred lines for fattening pigs were weighed 6 times in an age range from 30 to 167 d of age. A RRM with animal, litter, and permanent environment as random effects was applied. In the study of Malovrh (2003), the heritabilities estimated were higher than in our study. They
ranged from 0.30 at 31 d of age to 0.40 at 145 d of age. The proportion for the litter effect decreased over time, as was found in this study.

Additionally, Malovrh (2003) tested the influence of maternal genetic effect on pig growth. She found out that the maternal genetic effect accounted for only a small proportion of variance compared with the other random effects. Thus, the maternal genetic effect can be neglected in growth studies for pigs if the pigs are weaned at an age of 21 to 28 d. However, the maternal genetic effect on growth traits has to be included in the model for growth analysis in sheep and beef cattle. Their offspring are weaned at later ages and the influence of the dam is more important for their growth development.

Another important fact is the exclusion of birth weight in the analysis. Implausible values of variance components often occur at extreme ages like at birth. This is mainly the problem with high order polynomials (Meyer, 2005a). Thus, including birth weights increased the order of fit for the best fitting polynomial model in a significant way in the study of Meyer (2001).

Apiolaza et al. (2000) also mentioned that RRM have problems of fitting data with small variances as it is typical for birth weights. So it is assumed that RRM growth underlies a different growth process than postnatal growth. The study of Lewis and Brotherstone (2002) supports this hypothesis. Their investigation of genetic parameters for growth in sheep from 2 to 159 d of age led to the conclusion that live weight in early ages seems to be under different genetic control than live weight in later ages.

The estimation of genetic parameters with RRM also provides an insight into the effects of selection across the growth trajectory by examining the eigenfunctions and eigenvalues. In this study the first eigenfunction had the highest associated eigenvalue. It explained 89.93% of the genetic variation. Because of low values for the second and third eigenfunction, it is sufficient to focus on the first eigenfunction. It indicates that selection between d 30 to 400 changes the pattern of growth in the same direction, i.e., selection on low BW at d 30 will also lead to a low BW at d 400. Thus, a successful selection on low BW is possible.

As outlined by Nobre et al. (2003a), random regression analysis requires structured data. Otherwise a poor fit for some curves modeled with RRM may be the consequence. Thus, for the estimation of genetic parameters with RRM on the basis of weight recordings with an uneven distribution as it is given in original, not adjusted data sets with minipig BW, an adjustment of the results based on estimates from multiple trait models and literature information could be necessary. Another problem is a possible poor fit of growth curves estimated with RRM using LP. The LP require large data sets with an almost homogeneous distribution of weight recordings for a successful estimation of variance components (Miształ, 2006). Points on the growth curve estimated with LP on the basis of only a few records very often contain artifacts. This is especially the case at the extremes of the growth curve. A possible alternative to LP is the use of splines. Splines are piecewise polynomials [i.e., curves modeled from pieces of polynomials of low degrees and joined together at knots (Meyer, 2005a)]. They are able to estimate genetic parameters in a more robust way even if the distribution of weight recordings is very heterogeneous. However, with this method other problems do occur. A lack of smoothness for the predicted breeding values can be observed, and the number and position of the knots have to be chosen very carefully to avoid artifacts (Miształ, 2006). Nevertheless, Iwaisaki et al. (2005) report a simpler implementation of RRM with splines when compared with LP.

On the basis of the results of this study it is possible to construct a breeding scheme with a focus on the reduction of BW. As outlined by Brandt et al. (1997), there is a good potential for the selection on low BW due to moderate heritabilities and a high genetic variation in BW especially in later ages. This conclusion is also supported by the results of this study. Therefore, breeding progress for the trait low BW is expected.

**LITERATURE CITED**


