60th Annual EAAP Meeting, August 24-27, 2009, Barcelona, Spain
session S.53 "Free communications in genetics (methodology)"

Gametic Gene Flow Method Accounts for Genomic Imprinting

Vinzent Börner & Norbert Reinsch *

Introduction The economic success of commercial breeding programmes depends heavily on the transfer of genes of selected breeding animals to the production stage where traits of interest are realised. Until the 1980's it was a paradigm that the expression of genes in offspring is independent of parent-of-origin. First indications that this is not the case arose from transgenic mouse experiments (Sapienza *et al.*, 1987). A process called genomic imprinting was found to be responsible for a total or partial silencing of gene expression through DNA methylation during gametogenesis. The DNA methylation during gametogenesis depends on the parental sex and is reversible only during gametogenesis in offspring. First indications about the possible role of genomic imprinting in livestock came from variance component estimations in pigs describing significant parts of the phenotypic variance as induced by imprinted loci. Beside these quantitative approaches, quantitative trait loci (QTL) and genes coding for traits of economic relevance have been found to be influenced by genomic imprinting. According to these findings, it is may be necessary to account for genomic imprinting in defining the breeding goal and in genetic evaluation. A flexible and general approach might be the simultaneous calculation of two breeding values for a single animal by an extended BLUP-model (Reinsch & Guiard, 2005; Neugebauer *et al.*, 2009), whereby one reflects the breeding value of the animal as a dam and the other of the animal as a sire. However, for calculating a total merit index out of these two breeding values, economic weights for both are a necessary prerequisite.

The aim of our work was to derive relative economic weights for both breeding values of a genomic imprinting affected trait. For this purpose, the gene flow method (Hill, 1974) was extended to a gametic level and applied to a hypothetical pig breeding programme.

---

*Norbert Reinsch, Research Institute for the Biology of Farm Animals (FBN), Wilhelm-Stahl-Allee 2, 18196 Dummerstorf, Germany, email: reinsch@fbn-dummerstorf.de, phone: +49 38208 68900, fax: +49 38208 68902

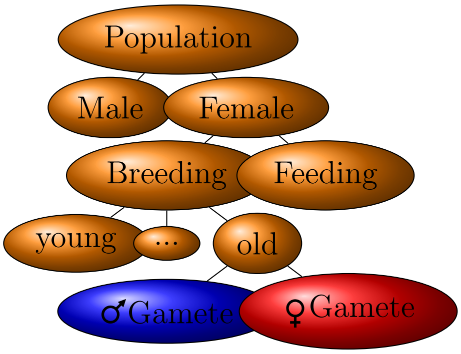


Figure 1: dividing a population

Additionally, it is demonstrated that this approach is also useful for modelling the loss of heterozygosity by a single round of selection in populations with overlapping generations.

Methods A breeding population can be divided into classes (C) of animals. The main characteristics of these classes are sex (s), use (u), age (a) and the parent-of-origin of the genes forming the individuals g (C_{suag}, see Figure 1, $g = 1$ for male and 2 for female origin). As only one round of selection and the following spread of genes of the selected animals within the population over time is considered, the selected individuals are called founders ($\widetilde{F}$). To specify, from which C_{sua} the founders are being selected, the abbreviation is extended to $\widetilde{F}_{sua}$. Every $\widetilde{F}_{sua}$ is the origin of one or many selection paths xy, where x is the selected group ($\widetilde{F}_{sua}$ itself) and y is the target group (e.g. boars to sows, x=boars, y=sows). If a $\widetilde{F}_{sua}$ founds a certain xy, it is denoted as $\widetilde{F}_{sua_{xy}}$. Every $\widetilde{F}_{sua_{xy}}$ can consist of one or many individuals selected from the same C_{sua}.

According to the gene flow method (Hill, 1974), a transmission matrix P^* can be derived on the gametic level. P^* consist of reproduction rows and is a Markov chain transmission matrix. The values within the reproduction rows represent the transmission probabilities of the genome of the C_{sua1} and C_{sua2} of already existing C_{sua}'s to the C_{sua1} and C_{sua2} of a new born C_{sua}. All rows of P^* must sum up to one. The matrices Q^* and E^*_{xy} are derivatives of the matrix P^*, containing only ageing rows or the reproduction row of the selection path xy.

A vector $n^*_{xy_0}$ is set up containing only zeros except a one at each of both C_{sua1} and C_{sua2} positions of the C_{sua} from which the founders were selected. The vector $m^*_{xy_0}$ is a null vector, and $m^*_{xy_t}$ reflects the proportion of the genome of the $\widetilde{F}_{sua_{xy}}$ on the genome of the C_{suag}'s after t time units due to the selection path xy. $m^*_{xy_t}$ vectors are calculated as follows:

$$
\begin{aligned}
n^*_{xy_t} &= Q^* \times n^*_{xy_{t-1}}, && (1)\\
m^*_{xy_t} &= E^*_{xy} \times n^*_{xy_{t-1}} + P^* \times m^*_{xy_{t-1}}. && (2)
\end{aligned}
$$

For calculating the trait realisation at time t, two vectors r_{kg} are used, one for the genome on the paternally inherited gametes and one for the genome one the maternally inherited gametes. The vector r_{k1} for calculating the weighted sum of genome proportions of $\widetilde{F}_{sua_{xy}}$ in the C_{sua1}'s contains zeros at the positions where the $m^*_{xy_t}$ vectors contain the values of the C_{sua2}'s. The same applies to r_{k2} being used to calculate the weighted sum of genome proportions of $\widetilde{F}_{sua_{xy}}$ in the C_{sua2}'s. It contains zeros at positions where the $m^*_{xy_t}$ vectors contain the values of the C_{sua1}'s.

The realisation $R_{xy_{ktg}}$ of the trait k at a certain time point t due to the genome proportions of $\widetilde{F}_{sua_{xy}}$ being inherited by a dam or a sire from $t-1$ to t is calculated as:

$$R_{xy_{ktg}} = m^{*'}_{xy_t} \times r_{kg}. \qquad (3)$$

The genetic net present value $g_{xy_{kg}}$ as with interest rate p capitalised $R_{xy_{ktg}}$ values is derived by the formula

$$g_{xy_{kg}} = \textstyle\sum_{t_0}^{t_{max}} R_{xy_{ktg}} \times \frac{1}{(1+p)^t}, \qquad (4)$$

where t_{max} is the planning horizon and p is the interest rate. These $g_{xy_{kg}}$'s are the weighting coefficients for the breeding values of trait k of the individuals as a sire and a dam. They are necessary to calculate a combined breeding value for trait k. The sum of these $g_{xy_{kg}}$'s is the correction factor for the economic weight of trait k in the selection index.

Inbreeding As the cumulated genome proportions of a group of founder individuals $\widetilde{F}_{sua}$ in the C_{suag}'s due to calculation of m^*_{xt} vectors are known, a vector $\Delta F_{t_{\widetilde{F}_{sua}}}$ representing the average increase of inbreeding coefficients of every C_{sua} at time unit t can be calculated. The values of $\Delta F_{t_{\widetilde{F}_{sua}}}$ reflect the increase in inbreeding at time point t due to the selection of a certain $\widetilde{F}_{sua}$. If all individuals within a certain group of founders $\widetilde{F}_{sua}$ are not related, are not inbreed and have the same chance of reproduction, $\Delta F_{t_{\widetilde{F}_{sua}}}$ is calculated by

$$\Delta F_{t_{\widetilde{F}_{sua}}} = m^*_{xt_1} \cdot m^*_{xt_2} \times \tfrac{1}{2} \times \tfrac{1}{N_{\widetilde{F}_{sua}}}, \qquad (5)$$

where $m^*_{xt_1}$ is a vector containing only the C_{sua1} values and $m^*_{xt_2}$ is a vector containing only the C_{sua2} values of the m^*_{xt} vector in time unit t. Thus, the dimension of $m^*_{xt_1}$ and $m^*_{xt_2}$ is half the dimension of the m^*_{xt} vector. $N_{\widetilde{F}_{sua}}$ is the number of individuals of the

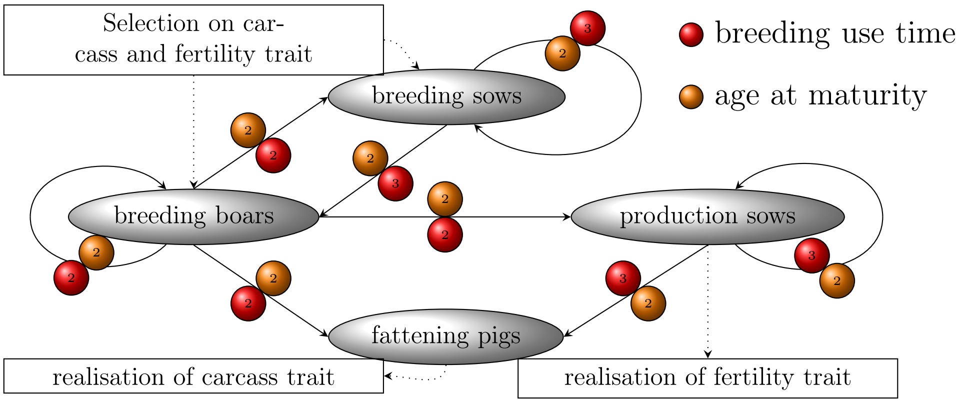


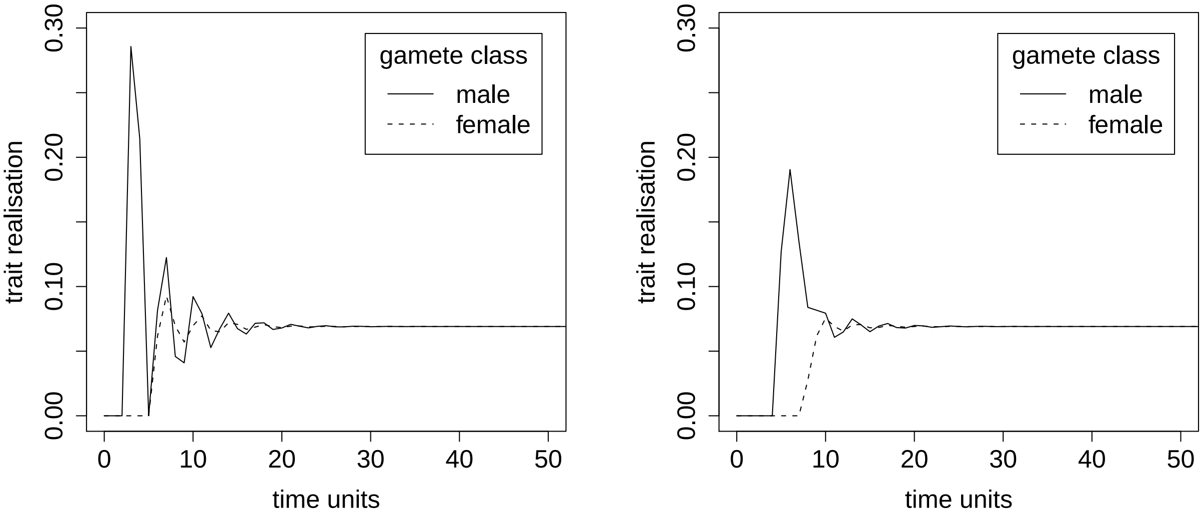
Figure 2: the breeding scheme

$\widetilde{F}_{sua}$. Note that $m^*_{x_{t_1}} \cdot m^*_{x_{t_2}}$ is the Hadamard product of the mentioned vectors. For a more general derivation see Börner & Reinsch (2009).

Application to a Breeding Programme We considered a hypothetical pig breeding programme of the structure given in Figure 2. The survival rate of boars and sows after maturity is 75 %. The traits of interest are litter size and backfat thickness. The realisation of litter size occurs in females at breeding (10 %) and production stage (90 %), whereas backfat thickness is realised only by fattening animals. The length of the planning horizon is either 5, 10, 15 and 30 time units. The interest rate is 5 % per year and, therefore, roughly 2.5 % per time unit.

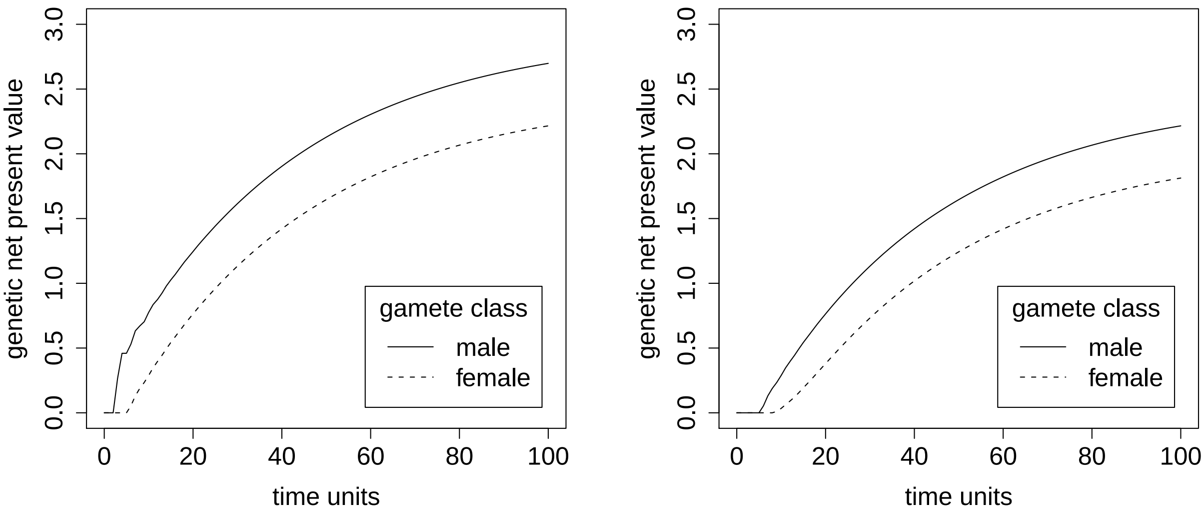
Results Figures 3(a) and 3(b) illustrate the trait realisation values $R_{x_{ktg}}$ due to the cumulated sire derived selection paths subject to trait and planning horizon. As expected from the breeding scheme and, therefore, possible gene flow, the fluctuations of $R_{x_{ktg}}$ values within the first 10 time units are enormous. In later time units the proportions of the founder group genome inherited by a sire or a dam are equal for every C_{sua}.

Figures 4(a) and 4(b) show plots of the $g_{x_{kg}}$'s of backfat thickness for the cumulated sire derived selection paths and cumulated breeding dam derived selection paths. The founder genome proportions in the male gamete class are more important for both because they arrive earlier at trait realising animals. The maternally inherited founder genome arrives at trait realising animals with a certain time lag. This creates a sustained difference in genetic net present values even if the planning horizon is unlimited (results not shown). As the $g_{x_{kg}}$'s increase asymptotically in time, the difference between both can be regarded as fixed if the planning horizon is long enough. Considering backfat thickness, this fixed difference between the $g_{x_{k1}}$ and the $g_{x_{k2}}$ is about 20 % for both cumulated selection paths. Comparing cumulated selection paths, asymptotical genetic net present values of backfat



(a) Backfat thickness (b) Litter size

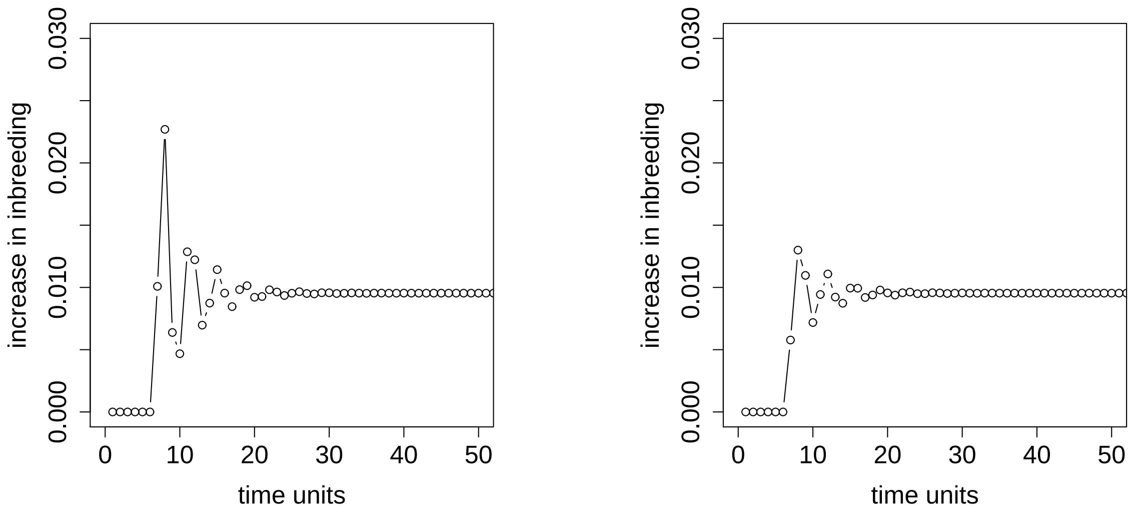
Figure 3: Trait realisation subject to planning horizon and gamete class for the cumulated sire derived selection paths.



(a) Backfat thickness - cumulated sire derived paths (b) Backfat thickness - cumulated breeding dam derived paths

Figure 4: Genetic net present values subject to trait, selection paths, planning horizon and gamete class.

thickness due to cumulated breeding dam derived selection paths are always lower than due to the cumulated sire derived selection paths. This is explained by specificities of the breeding programme (a breeding dam is not producing fattening pigs).



(a) Primary selected boars

(b) Primary selected breeding sows

Figure 5: Inbreeding coefficient subject to planning horizon for new born boars due to different primary selected animals.

Figures 5(a) and 5(b) illustrate the development of the inbreeding coefficient of new born boars due to founding boars and founding breeding sows when only one animal was initially selected. The inbreeding reaches a maximum in time unit eight and then is stabilised at roughly 1 %. The long term increase of the inbreeding coefficient induced by selecting founding boars is equal to the increase induced by selecting founding breeding sows. The short-term increase within the first 10 time units often exceeds the long term increase considerably. The peak caused by founding boars exceeds the long term increase by more than 100 %, whereas the peak caused by founding breeding sows exceeds the long term increase by about 20 %.

Discussion The results from our hypothetical breeding scheme clarify that the breeding values of an individual as a sire or a dam can not always be equally weighted. In complex nucleus breeding schemes the gene flow of primary selected animals to trait realising C_{sua}'s depends heavily on the special construction of the scheme. The complexity of such breeding schemes can lead to counter-intuitive results. This is exemplified by our hypothetical breeding scheme. The breeding values of the breeding dams as sires should be weighted higher than their breeding values as dams. Additionally, sires selected in the nucleus are used for producing trait realising C_{sua}'s and its female precursors. Due to this construction, the males have breeding values as a sire and a dam which are not realised at the same time. As almost every breeding scheme has its own specificities, computations are necessary to derive the exact values of realisation for both breeding values.

As now the probability that a C_{sua} receives the same genes from its sire and dam is calculable, fluctuations of inbreeding coefficients can be modelled for every C_{sua} and the whole planning horizon. The results of our calculations show clearly that increase of the inbreeding coefficients of special combinations of C_{sua}'s and time unit can exceed its average increase by far. Thus, a possible inbreeding depression might decrease the real trait realisation especially in early time units.

In conclusion, relative weighting factors for parent-specific additive genetic effects can be derived appropriately with a gametic version of the discounted gene flow method. Moreover, gametic gene flow may also be useful for predicting short-term fluctuations of inbreeding coefficients.

Acknowledgement This work was financed by by the German Federal Ministry for Education and Research (BMBF), project "FUGATO-Brain".

References

Börner, V. & Reinsch, N. (2009), Gametic gene flow method accounts for genomic imprinting. *J. Anim. Breed. Genet.* Accepted for publication.

Hill, W. G. (1974), Prediction and evaluation of response to selection with overlapping generations. *Anim. Prod.* **18**, 117–139.

Neugebauer, N., Räder, I., Schild, H. J., Zimmer, D. & Reinsch, N. (2009), Evidence for parent-of-origin effects on genetic variability of beef traits. *submitted* .

Reinsch, N. & Guiard, V. (2005), Mixed mendelian and imprinted inheritance, unpublished.

Sapienza, C., Peterson, A. C., Rossant, J. & Balling, R. (1987), Degree of methylation of transgenes is dependent on gamete of origin. *Nature* **328**(6127), 251–254, URL http://dx.doi.org/10.1038/328251a0.