

An overview on quantitative and genomic tools for utilising dominance genetic variation in improving animal production

Asko Mäki-Tanila

*MTT Agrifood Research Finland, Biotechnology and Food Research, FI-31600 Jokioinen, Finland,
e-mail: asko.maki-tanila@mtt.fi*

In addition to genetic progress made by selection on additive genetic values, short-term gains can be produced by recovering possible inbreeding depression or utilising putative overdominance. These are both caused mainly by dominance genetic variation which can be quantified using mixed model methodology. Inbreeding brings along a requirement for extra parameterisation in expressing and estimating dominance variance. The extra parameters specify how dominance is affecting the mean and (co)variances among inbred animals. The full description for breed crosses contains a very large set of parameters. The benefits from crossbreeding are highest with widely deviating allele frequencies between the breeds. Maximisation of heterosis can be done only on a temporary basis as a continued exploitation leads to stagnation in the overall genetic progress. Therefore efficient methods with immediate returns are needed to find the most promising breeds jointly with the most potential mating pairs. One possibility is the use of genomic tools in assessing populations for crossbreeding and in searching for major genes mediating dominance variation. The analyses are providing markers that can be used in choosing mating pairs that produce desirable dominance deviations in analysed marker brackets. Genome-wide marker sets can be used for discovering genome segments with maximum heterosis effect. The phenotypic records are available for such analyses, soon are also the large marker sets and their genotypes: the analytical tools need developers.

Key-words: animal breeding, genetic improvement, inbreeding, heterosis, crossbreeding, genomic selection

Introduction

Selection is an efficient tool to improve animal production, as the changes achieved are cumulative and permanent. The research and development work consist of profound understanding of genetic variation in production and fitness traits and development of tools to maximise the selection intensity and minimise loss of genetic variation. It is possible to gain extra short-term advantages by choosing mating pairs within a breed or between breeds to maximise the performance in the production animal generation.

Selection is working on the additive genetic differences between animals caused by direct allelic effects. There is much evidence on the existence of variation due to interaction of allelic effects within a locus (dominance) or between loci (epistasis). An indirect evidence on dominance is the commonly detected lowered performance in inbred individuals and its counterpart heterosis. There have been observations on both complete and partial dominance. Overdominance where the heterozygote exceeds the homozygote at a specific locus, is controversial, nor it is clear what the physiological cause of such an effect would be (e.g. Wilkie 1994).

The purpose of this study is to review the methods used to quantify the amount of dominance genetic variation and the tools for its utilisation in the livestock improvement schemes. The first half of the paper is on the implications of dominance in quantitative genetics, starting from observations and recovery of inbreeding depression, then reviewing methods quantifying dominance variation within populations (with allowance for inbreeding) and finally discussing the extensions to crossing populations. The second half presents the state of the art in genomics and its utilisation in characterising quantitative genetic variation and in selection, with a special emphasis on the use of genome-wide marker sets.

Inbreeding depression

Inbreeding increases homozygosity, and thus exposes the harmful effects on fitness of the partially recessive and deleterious mutations. This results in inbreeding depression which under dominance is linearly related to degree of inbreeding (Falconer and Mackay 1996).

The tempting attempts to produce superior uniform production animals by continued inbreeding and selection are not feasible. There are too many recessive genes to be eradicated. Perhaps heterozygosity in itself is a requirement for vigour. Lack of vigour among homozygotes may also be exhibited as increased sensitivity to environmental changes seen as variable performance. (Lerner 1954).

Inbreeding depression is attached to traits physiologically close to reproductive efficiency and longevity of the animal. Natural selection has been acting to increase the reproductive efficiency for countless generations and there is little variation left in additive value for artificial selection to act on. The remaining genetic variation is of non-additive nature seen as substantial depression after populations have experienced some inbreeding. There are several observations on inbreeding depression in farm animals. Inbreeding depression has been shown to decrease milk production by approximately 9 to 26 kg of milk per lactation for each 1% of inbreeding (Thompson et al. 2000 a, b). Wall et al. (2005) found that the difference between a noninbred animal and an animal with an inbreeding coefficient of 10% had a 2.8 d increase in calving interval, a 1.7 d increase in days to first insemination, a 1% increase in probability to return to estrus at first service, 0.03 more inseminations, a 0.27-unit decrease in body condition, and a 0.54 kg decrease in daily milk yield measured around day 110 after calving. Similar findings have been made in pigs. Inbreeding depression, expressed as the performance relative to the mean, per 10% increase of the inbreeding coefficient, was -3.18, -2.31 and -5.37% for weight at 120 d, and -3.16, -2.08 and -6.49% for daily gain in three different populations of the Iberian pig (Fernández et al. 2002). Leymaster and Swiger (1981) found a

0.06–0.39 piglet decrease in the litter size for 3–4 %-unit increase in inbreeding.

Purging or recovering inbreeding depression

Since inbreeding increases the frequency of deleterious homozygous genotypes and exposes the partially recessive and deleterious mutations to selection, it provides an opportunity for a population to be ‘purged’ of its mutational load and the resultant depression in fitness. Templeton and Read (1984) suggested that inbreeding depression can be rapidly and effectively purged by selection with deliberate inbreeding. In principle, such purging might remove most of the harmful recessives and bring fitness traits back to the level where they are in a large, randomly mating population.

Whilst the selection of top breeding animals is based on additive values, the production animals may be produced by maximising heterozygosity and thereby recovering most of inbreeding depression (Fig. 1). The highest heterozygosity in offspring is achieved in subdivided population in two or higher way crosses among the subpopulations (Robertson 1964). Sires are selected from within a subpopulation and then production animals are obtained by rotational/criss-cross breeding design or by sires unrelated to the most popular ones.

Selection of mating pairs seems to give some short-term benefits and only negligible long-term gains. The overall heterozygosity in the interbreeding population is inevitably decreasing if little attention is paid to balanced representation of families and ancestral lines. The reduced heterozygosity means also compromised potential for further selection response. Selection is increasing the rate of inbreeding because offspring of best individuals are more likely to be selected and therefore the contributions of top animals are quickly accumulating (Robertson 1961). Inbreeding and drift increase the variation in the selection response and thereby cause uncertainties in

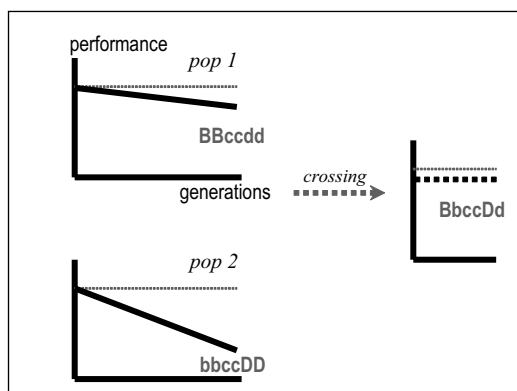


Fig. 1. Over generations, inbreeding increases homozygosity and exposes the harmful effects of recessive deleterious mutations resulting in inbreeding depression. Production animals may be obtained by maximising heterozygosity in two or higher way crosses and thereby recovering most of the inbreeding depression. In the example, after inbreeding one locus is homozygous for a recessive allele in both the populations and therefore the cross fails in recovering all the inbreeding depression.

planning and investing in genetic improvement schemes (Hill 1977). There has been substantial recent research on finding the best possible population structure (mating system and family size) to maximise selection gain and minimise long-term rates of increase in inbreeding and loss of heterozygosity (Wray and Thompson 1990, Meuwissen 1997, Woolliams et al. 1999).

Quantifying the deviations due to dominance

Individuals are resembling each other due to dominance when they share the same allele pair at the loci exhibiting dominance. The type of relatives where such allele pairs are expected to be similar due to common family background, are full-sibs, double first cousins and so on. It is not usually possible to design experiments to estimate the genetic interaction component with any precision. For estimating

covariances due to dominance type of interaction, the most important source of information is the hierarchical structure of full within half sibs, while the estimates from distant relatives involve multiplication of observed covariance by large factors (Falconer and Mackay 1996). There are sufficient numbers of full-sibs in fish, poultry and pigs, while in cattle active use of embryo transfer is needed to achieve a satisfactory number of such groups. If no arrangements are made, the environmental covariance and dominance effects are almost completely confounded among full-sibs.

Dominance variance can be considered as noise and when the effect is included in the statistical model, the accuracy of additive genetic evaluation is improved. In terms of genetic improvement, dominance is, however, relevant in generating/recovering the inbreeding depression and therefore the effect of inbreeding should be integrated in the estimation process. With inbreeding, genetic covariance remains a sum of products of relationship coefficients and (co)variance components. Inbreeding changes the mean and genetic covariance structure of a population. There are interesting historical coincidences in the development of treating dominance and inbreeding in quantitative genetics. To my knowledge the first comprehensive analysis was done by Lagervall (1961) in his thesis work initiated and supervised at Iowa by Lush and Kempthorne and later worked out and better publicised by Harris (1964) of the same research group, and independently in France by Gillois (1964). In addition to additive and dominance variance in a non-inbred population, the extra parameters required are dominance variance and covariance between additive and dominance effects in the completely inbred population, and sum over loci of effects of inbreeding depression and squared effects of inbreeding depression.

Let us assume a locus with several alleles to the measurement, allele frequency x_i and additive effect α_i for the allele i and the corresponding dominance deviation d_{ij} for the allele pair of i and j . Following the presentation by Smith and Mäki-Tanila (1990), after summing over alleles and their pairs, we have additive variance

$\sigma_a^2 = \sum_i x_i \alpha_i^2$, dominance variance $\sigma_d^2 = \sum_{ij} x_i x_j d_{ij}^2$ and the parameters in the completely inbred individuals: the inbreeding depression $u_\delta = \sum_i x_i d_{ii}$, the dominance variance amongst homozygotes $\sigma_\delta^2 = \sum_i x_i d_{ii}^2 - u_\delta^2$ and the covariance between additive and dominance deviations among homozygotes $\sigma_{a\delta} = \sum_i x_i \alpha_i d_{ii}$. In the population with an inbreeding coefficient F the total genetic variance would be approximately (see De Boer and Hoeschele 1993) $(1 + F)\sigma_a^2 + (1 - F)\sigma_d^2 + F\sigma_\delta^2 + 2F\sigma_{a\delta}$. Using the moment expressions, the parameters could be summed over loci. There are some theoretical challenges, as with a very large number of loci the inbreeding depression goes to infinity (Robertson and Hill 1983) – beyond any estimation bracket, like the expected response would do. Some counter examples can be constructed (Smith and Mäki-Tanila 1990) while models with a finite number of loci with graduated effects are more realistic and ease the problem.

Using the same mixed model methodology as in estimating additive genetic values, it is easy to estimate the dominance deviations for individuals (Henderson 1985) and use the information in choosing appropriate mating pairs on top of the selection on additive values. The conventional breeding value estimation model would contain the observations \mathbf{y} explained by the vector \mathbf{a} of random additive genetic effects and \mathbf{e} of random residual effects simultaneously correcting for the fixed environmental effects $\boldsymbol{\beta}$. The fixed effects and random genetic effects are connected to observations with the incidence matrices \mathbf{X} and \mathbf{Z} , respectively, and we have $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e}$. The extension to a model with dominance deviations (vector \mathbf{d}) under inbreeding (vector \mathbf{f} contains the inbreeding coefficients) is straightforward by having the term $\mathbf{Z}(\mathbf{f} u_\delta + \mathbf{d})$ in the model. The expectation for \mathbf{y} is $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{f} u_\delta$, the residual variance is σ_e^2 (no covariances between the residuals). The covariance matrix of additive and dominance effects can be obtained as functions of the five genetic parameters and their corresponding identity coefficients (Gillois 1964). Without computing identity coefficients, Smith and Mäki-Tanila (1990) developed a recursive procedure to compute genotypic covariance directly.

Estimation of σ_d^2 in practice

In predicting dominance genetic deviations or in quantifying the components of genetic variation, the inversion of the dominance relationship matrix is needed. There are no simple rules for inverting the covariance matrix when allowance for dominance and inbreeding are made in the analysis (Smith and Mäki-Tanila 1990). This is restricting the implementation of the full treatment of dominance in populations of the size typical in animal production.

Simulation results on dominance genetic variation have shown that using inbreeding coefficient as a covariate and computing the dominance covariance matrix ignoring inbreeding, give satisfactory results (Uimari and Kennedy 1990, De Boer and Van Arendonk 1992), even when the variation was generated by rare recessives expected to cause the largest departures from conventional predictions (Mäki-Tanila 1982). Hoeschele and VanRaden (1991) have produced a method to directly write the inverse of dominance relationship matrix in non-inbred populations. Because of the progress in computing capacity, estimation of dominance genetic variances with large data sets are now feasible – and also fashionable. Misztal (1997) has developed algorithms to produce solutions for dominance deviations at a cost of less than twice that of computing the breeding value prediction. Subsequently it is possible to predict the dominance deviation for progeny from any mating (Varona and Misztal 1999) allowing the development of algorithms for mate allocation applicable in obtaining the best production animals from the given multiplier candidates.

The computing requirements for the proper consideration of dominance and inbreeding are enormous even for present-day capacities. Recently, Palucci et al. (2007) circumvented the size problem by dividing the population data into subsets and presented the variance components as weighted averages of the subset estimates. They found that for computing the additive and dominance relationships among some 600 000 animals in the pedigree file with the algorithm of Smith

and Mäki-Tanila (1990) would require computing 3.56×10^{11} elements. To store the non-zero coefficients needs many terabytes and several years of computing time. Therefore, the data file was split into four groups of about 50 000 animals each plus their complete pedigree information. The final computing utilized several subsets of 1000 animals (heifers or cows) in which all animals had a non-zero dominance relationship with at least one other animal in the group. Estimates of variances were obtained using Bayesian methodology via Gibbs sampling. The non-additive variance components for the studied fertility traits were as large as or higher than the additive estimates (heritability varied between 0.005 and 0.11).

Extensions to a pair of populations

Heterosis or hybrid vigour in crosses has been long observed. Predicting which lines will give good crosses is a hit and miss affair. Simple dominance is sufficient to give crossbreds better performance than either parent provided that both parents are fixed for the dominant allele at some locus at which the other parent is fixed for the recessive one – that is with one parent being Bc/Bc and the other bC/bC, where B and C are the dominant alleles.

Several breeding procedures have been suggested and used in order to make best use of the heterosis, among them the reciprocal recurrent selection (RRS), proposed by Comstock et al. (1949) makes the most use of quantitative genetic principles. Two non-inbred parental pure-bred populations are maintained. Crosses are made each generation between the populations, cross-bred performance is recorded, and individuals to breed the pure parental populations are chosen on the basis of their performance in crosses.

The process can be presented with the selection index machinery, by replacing genetic variances of purebreds by variances and covariances of pure and cross performance and considering purebred

and crossbred performance as two genetically correlated traits (Wei and van der Werf 1994).

We are interested in crossing two populations (1 and 2) and base the selection within a line on the performance of crossbred individuals. Let us assume a locus with two alleles B and b contributing a ($a > 0$) and 0 to the measurement. The frequency of B is x . Under dominance the heterozygote Bb deviates from the mean of the two homozygotes by ha , where the degree of dominance h could vary from no dominance ($h = 0$) to complete ($h = \pm 1$) and further to overdominance ($|h| > 1$). Considering the locus B/b, after truncation selection (intensity i) the allele frequency in population 1 (phenotypic s.d. = σ) changes by

$$\Delta x_1 = \frac{1}{2}(ia/\sigma)x_1(1-x_1)(1+h(1-2x_2)).$$

Following Comstock et al. (1949) and Hill (1971), the mean of the crossbred progeny is maximised with additivity and partial dominance if both populations are fixed for the favourable allele B. With complete dominance the maximum performance in crossbreds is obtained if either population is fixed for B. When there is overdominance, the heterosis is maximised when $x_1=1$ and $x_2=0$ or *vice versa*. We can compare the efficiency between within line selection and RRS. The mean of the crossbreds will change in pure line selection by

$$[x_1(1-x_1) + x_2(1-x_2)](1+h(1-2x_1))(1+h(1-2x_2))$$

and in RRS

$$x_1(1-x_1)(1+h(1-2x_2))^2 + x_2(1-x_2)(1+h(1-2x_1))^2$$

where the factor $\frac{1}{2}(ia/\sigma)$ has been omitted from both the expressions. Hill (1971) has considered the options with finite populations and concluded: RRS is more effective than pure line selection if there is complete dominance and especially if there is over-dominance. With partial dominance pure line selection is equally efficient. The approach will be useful only if there is overdominance. RRS would require deviating allele frequencies in the populations and will be enhanced by some initial inbreeding perturbing the frequencies. The frequency differences will be increased by the selection when the desired progeny is a heterozygote.

The benefits from crossbreeding are highest with widely deviating allele frequencies between the breeds. Maximisation of heterosis can be done only on a temporary basis as a continued exploitation leads to stagnation in the overall genetic progress. Also elaborated schemes like RRS increase the generation interval. Therefore we need efficient methods with immediate returns to evaluate the most promising breed crosses jointly with the most potential mating pairs.

Starting from the results by Smith and Mäki-Tanila (1990), theory and methods to compute genotypic means and covariances in a two-breed population under dominance inheritance have been presented by Lo et al. (1995). They showed that the genotypic mean is a linear function of five location parameters and that the genotypic covariance between relatives is a linear function of 25 dispersion parameters. These would include the additive (and the corresponding heritability) and dominance variance in both the purebred populations and the variation of contributions from the breeds to the F1 individuals. Inbreeding adds another set of parameters like in the within population case. Recursive procedures are given to compute the necessary identity coefficients. In the absence of inbreeding, the number of parameters for the mean is reduced from five to three and the number for the covariance is reduced from 25 to 12.

The sow longevity traits are typically exhibiting substantial inbreeding and heterosis. Serenius et al. (2006) had a very large data set to estimate the dominance variance in pure and crossbred sow populations. Using Hoeschele and VanRaden (1991) inversion, they found heritabilities ranging from 0.03–0.09 in the purebred to 0.06–0.12 in the crossbred population. The proportion of dominance variance was on average 0.04 in the purebred and 0.10 in the crossbred population.

In conclusion, there is much evidence on dominance type of genetic variation in economically important traits. It could be utilised in crossing lines within and between populations and thereby recovering some of the inbreeding depression caused by harmful recessives. The prediction tools for exploiting heterosis have been developed. The requirements for explicitly determining the variance

components in inbred populations are still beyond the available computing power. Approximations with structuring the genetic variation assuming random mating and using inbreeding as a covariate would yield satisfactory predictions.

From MAS to genomic genetic values

The tools provided by genomics offer new opportunities in understanding and exploiting quantitative genetic variation. Genomics has been used in finding and characterising the major genes behind the variation. The genome screens for QTL (QTL, quantitative trait locus) carried out at MTT Agri-food Research Finland have found several QTL's for production, quality and health traits in dairy cattle (Viitala et al. 2003, Schulman et al. 2004) and in egg-layers (Tuiskula-Haavisto et al. 2002). In several subsequent fine mapping studies the investigations have been able to characterise the actual gene involved (Viitala et al. 2003 and 2006, Farnir et al. 2002, Blott et al. 2003, Honkatukia et al. 2005). In dairy cattle the most efficient designs rely on progeny testing, where only additive genetic effects can be estimated, while in an F2 design typically used in poultry and pigs, also dominant gene action is revealed.

Genes or closely linked markers are useful tools in selection (MAS, marker assisted selection), as they have very little environmental noise, they appear in both sexes and are detectable at an early stage (even in embryos, e.g. Virta et al. 2002) and they can be used in tracing the segregation variation present in a family. Genetic markers are mainly affecting selection accuracy, especially in traits where the accuracy of conventional selection is low. This would typically mean traits with low heritability. Markers can be valuable predictions for traits which can be measured only after slaughter or for traits where records are difficult to find (such as disease traits, feed intake, product quality).

The marker information can be included in the

breeding value estimation by writing (Fig.2.) the model $y = X\beta + Za + Qq + e$, where q refers to the estimated vector of QTL effects and Q is the respective incidence matrix (Fernando and Grossman 1989). $\text{Var}(q) = G \sigma_q^2$ where σ_q^2 is the variance due to the QTL and G is the matrix of probabilities for QTL alleles being identical by descent. These probabilities can be computed from the pedigree, marker and linkage map information.

The marker assisted selection has several drawbacks. The larger the effect of the QTL is, the faster it is fixed by selection and the more it will also reduce the variation in the surrounding genome area (e.g. Pong-Wong and Woolliams 1998). QTL's are usually found interval by interval over the genome with multiple testing with a risk for false positives (Goring et al. 2001). For multiple QTL detection, the simpler methods (e.g. de Koning et al. 2001) are replaced by sophisticated Bayesian methods (Sillanpää and Arjas 1998).

The most widely used genetic markers in farm animals are microsatellites, while SNPs (single nucleotide polymorphisms) are becoming the marker of choice with the advantage of a very high number and density of markers and a high throughput and

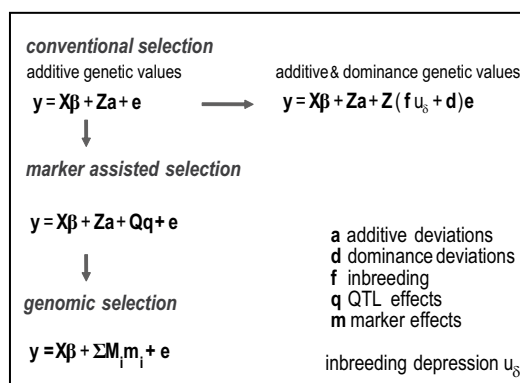


Fig.2. The mixed model methodology is a flexible tool to accommodate the models ranging from simple breeding value prediction to evaluation of dominance deviations and QTL (quantitative trait locus) effects linked to known markers. X, Z, Q and M are the incidence matrices linking the observations to respective effects. The effects of genomewide marker sets are computed through a simple summation over the genome.

low cost. For comparison, in a human population two randomly chosen individuals differ at ~1 in 1000 nucleotides. The genetic diversity of humans is low compared to other (older) species. In cattle and sheep the mean nucleotide diversity is 2–2.5 SNPs per kilobase (cf. Meadows et al. 2004), whilst the chicken estimate is 4–5.5 (International Chicken Genome Sequencing Consortium 2004). With modern DNA chip technology we may analyse up to thousands of loci across the genome. For example, for dairy cattle there is a chip available with 24,000 markers (Affymetrix).

For the future practical application in estimating the genetic values, the most promising approach is utilising simultaneously a vast number of markers over the entire genome (Meuwissen et al. 2001). With a genome-wide set of markers, the estimate of individual's genetic value or genomic estimated breeding value is obtained by summing the marker effects over the genome (Fig.2.). The genomic estimated breeding value for individual i would be summing the marker effects m_j over the genome $\sum m_{ij}$. There is no relationship matrix or identity-by-descent matrix required.

The genotyping work is still expensive and time consuming, so that we currently cannot have more than 1000–2000 animals typed for both pheno- and genotype. We then are facing a serious shortage of degrees of freedom in estimating the marker effects and least squares has to be abandoned, because only the largest effects can be estimated. In the best linear unbiased prediction (BLUP), the haplotype effects can be treated as random effects assuming for simplicity that they all have the same variance. The most promising has been found to be a Bayesian approach allowing for variation in the variance over the genome.

With best linear unbiased prediction and Bayesian method, the accuracy of predicting genetic values for offspring of the recorded animals has been shown to be of the order 0.7–0.8 (Meuwissen et al. 2001), comparable to that at the progeny test. Meuwissen et al. (2001) simulated a population with an effective size 100 over 1000 generations and had genes mediating the quantitative variation spread evenly at 1 cM distance. The effects over

loci had a gamma distribution and were additive over loci. Markers were distributed over the genome like the QTL's and were typed on 200 and 2000 individuals with phenotypic record at two consecutive generations, respectively, before the breeding value prediction was computed for the marker genotyped offspring. When it takes 5–7 years to have the progeny test result for a dairy bull, the genomic genetic value is available at birth. Schaeffer (2006) has shown that in dairy cattle, the genetic change can be enhanced nearly three times at a fraction of the costs of conventional progeny test programmes.

Can we predict genomic heterosis?

Expected reallocation of genetic variation between populations can be expressed with Wright's F coefficients (Falconer and Mackay 1996). Genomic research has proven to be a powerful approach in quantifying the genetic distances between populations, in revealing history of animal populations, number and sites for domestication, population expansions and contractions, selection, origin and mixing of maternal and paternal lineages (see Toro and Mäki-Tanila 2007). Diversity studies at MTT have been carried out in chicken (Vanhala et al. 1998, Rosenberg et al. 2001), in cattle (Kantanen et al. 2000) and in sheep (Tapio et al. 2005).

The effects of inbreeding on mean performance due to dominant genes and on variability between and within lines due to additive genes can readily be deduced from Wright's inbreeding theories, and both are linear functions of inbreeding coefficient. Because heterosis is proportional to the differences in gene frequencies in the parental lines, it is possible to make marker-based prediction of hybrid performance based on genetic distances, despite having only indirect estimates of allele frequencies for the interesting traits *via* the anonymous markers. The prediction is not efficient if the lines are unrelated or originating

from different populations because the associations or linkage disequilibrium (Charcosset and Essioux 1994) between marker and trait loci differ between the populations. Gavora et al. (1996) found high correlation (0.68–0.87) between the marker heterozygosity and performance in egg-layers.

Combining the marker and QTL information together with the dominance interactions, it is possible to make appropriate crosses such that the most desirable genotype is produced. Perez-Enciso et al. (2001) developed methodology for QTL analysis with an allowance for dominance and inbreeding. They used the framework of Smith and Mäki-Tanila (1990) and Lo et al. (1995) and showed how the genetic covariance between any two individuals is expressed as a non-linear function of the probability of up to 15 possible identity modes and the additive and dominance effects, together with allelic frequencies in each of the parental populations. The probabilities of each identity mode are obtained at the desired genomic positions using a Monte Carlo Markov chain method. Clearly simpler and more parsimonious methods are needed.

Xu (2003) stimulated by Meuwissen et al. (2001) has extended a Bayesian analysis to the F₂ population of inbred lines and has successfully also estimated the dominance deviations. The basic model for the observation can be written in terms of the genome wide sums of the deviations from the overall mean as $\sum (2v\beta_i + w\delta_i)$ where the dummy variable $v = +$ or $-\sqrt{2}$ for the homozygotes and 0 for the heterozygote and $w = -1$ for the homozygotes and 1 for the heterozygote. Then the regression coefficient β is half of the difference between the homozygotes and δ is the deviation of the heterozygote from the mean of the homozygotes. However, the cases starting from random breeding populations – in subdivided population or in crossbred animals, typical in animal breeding – need research. The tools are obvious, the phenotypic records are available and soon are also the large marker sets and their genotypes – we have to act.

References

- Blott, S., Kim, J.-J., Moiso, S., Schmidt-Kuntzel, A., Cornet, A., Berzi, P., Cambisano, N., Ford, C., Grisart, B., Johnson, D., Karim, L., Simon, P., Snell, R., Spelman, R., Wong, J., Vilkki, J., Georges, M., Farnir, F. & Coppieters, W. 2003. Molecular dissection of a quantitative trait locus: a phenylalanine to tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. *Genetics* 163: 253–266.
- Charcosset, A. & Essioux, L. 1994. The effect of population structure on the relationship between heterosis and heterozygosity at marker loci. *Theoretical and Applied Genetics* 89: 336–343.
- Comstock, R.E., Robinson, H.F. & Harvey, P.H. 1949. A breeding procedure designed to make maximum use of both general and specific combining ability. *Journal of American Society of Agronomy* 41: 360–367.
- De Boer, I.J.M. & Hoeschele, I. 1993. Genetic evaluation methods for populations with dominance and inbreeding. *Theoretical and Applied Genetics* 86: 245–258.
- De Boer, I.J.M. & Van Arendonk, J.A.M. 1992. Prediction of additive and dominance effects in selected or unselected populations with inbreeding. *Theoretical and Applied Genetics* 84: 451–459.
- De Koning, D.-J., Schulman, N., Elo, K., Moiso, S., Kinoshita, R., Vilkki, J. & Mäki-Tanila, A. 2001. Mapping of multiple quantitative trait loci by simple regression in half-sib designs. *Journal of Animal Science* 79: 616–622.
- Falconer, D.S. & Mackay, T.F.C. 1996. *Introduction to Quantitative Genetics*. 4th edition. Longman, Harlow, Essex, UK. 464 p.
- Farnir, F., Grisart, B., Coppieters, W., Riquet, J., Berzi, P., Cambisano, N., Karim, L., Mni, M., Moiso, S., Simon, P., Wagenaar, D., Vilkki, J. & Georges, M. 2002. Simultaneous mining of linkage and linkage disequilibrium to fine-map QTL in outbred half-sib pedigrees: revisiting the location of a QTL with major effect on milk production on bovine chromosome 14. *Genetics* 161: 275–287.
- Fernández, A., Rodríguez, J., Toro, M.A., Rodríguez, M.C. & Silió, L. 2002. Inbreeding effects on the parameters of the growth function in three strains of Iberian pigs. *Journal of Animal Science* 80: 2267–2275.
- Fernando, R.L. & Grossman, M. 1989. Marker assisted selection using best linear unbiased prediction. *Genetics, Selection, Evolution* 21: 467–477.
- Gavora, J.S., Fairfull, R.W., Benkel, B.F., Cantwell, W.J. & Chambers, J.R. 1996. Prediction of heterosis from DNA fingerprints in chickens. *Genetics* 144: 777–784.
- Gillois, M. 1964. Calcul de coefficients d'identité d'après les liens de parenté. In: *Communication aux Journées d'études de la Commission de Genétique de la Fédération européenne de Zootechnie*, Lisbonn, 11 p.
- Goring, H.H.H., Terwilliger, J.D. & Blangero, J. 2001. Large upward bias in estimation of locus-specific effects from genomewide scans. *American Journal of Human Genetics* 69: 1357–1369.
- Harris, D.L. 1964. Genotypic covariances between inbred relatives. *Genetics* 50: 1319–1348.
- Henderson, C.R. 1985. Best linear unbiased prediction

- of nonadditive genetic merits in noninbred populations. *Journal of Animal Science* 60: 111–117.
- Hill, W.G. 1971. Theoretical aspects of crossbreeding. *Annales Génétique et de Sélection Animale* 3: 23–34.
- Hill, W.G. 1977. Variation in response to selection. p. 343–365. In: Pollak, E., Kempthorne, O., Bailey, T.B. (eds.) *Proceedings of the International Conference on Quantitative Genetics*. 872 p.
- Hoeschele I. & VanRaden P.M. 1991. Rapid inversion of dominance relationship matrices for noninbred populations by including sire-dam subclass effects. *Journal of Dairy Science* 74: 557–569.
- Honkatukia, M., Reese, K., Preisinger, R., Tuiskula-Haavisto, M., Weigend, S., Roito, J., Mäki-Tanila, A. & Vilkki, J. 2005. Fishy taint in chicken eggs is associated with a substitution within a conserved motif of the FMO3 gene. *Genomics* 86: 225–232.
- International Chicken Genome Sequencing Consortium 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432: 695–716.
- Kantanen, J., Olsaker, I., Holm, L.-E., Lien, S., Vilkki, J., Brusgaard, K., Eythorsdottir, E., Danell, B. & Adalsteinsson, S. 2000. Genetic diversity and population structure of 20 North European cattle breeds. *The Journal of Heredity* 91: 446–457.
- Lagervall, P.M. 1961. Quantitative inheritance and dominance. II. The genic and the dominance variance under inbreeding. *Hereditas* 47: 111–130.
- Lerner, I.M. 1954. *Genetic homeostasis*. Oliver and Boyd. Edinburgh. 134 p.
- Leymaster, K.A. & Swiger, L.A. 1981. Selection for increased leanness of Yorkshire swine. III. Inbreeding effects on secondary traits. *Journal of Animal Science* 53: 620–628.
- Lo, L.L., Fernando, R. L., Cantet, R. J. C. & Grossman, M. 1995. Theory for modelling means and covariances in a two-breed population with dominance inheritance. *Theoretical and Applied Genetics* 90: 49–62.
- Mäki-Tanila, A. 1982. *The validity of the heritability concept*. PhD Thesis, University of Edinburgh. 120 p.
- Meadows, J.R.S., Hawken, R.J. & Kijas, J.W. 2004. Nucleotide diversity on the ovine Y chromosome. *Animal Genetics* 35: 379–385.
- Meuwissen, T.H.E. 1997. Maximising the response of selection with predefined rate of inbreeding. *Journal of Animal Science* 75: 934–940.
- Meuwissen, T.H.E., Hayes, B.J. & Goddard, M. E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819–1829.
- Misztal, I. 1997. Estimation of variance components with large-scale dominance models. *Journal of Dairy Science* 80: 965–974.
- Palucci, V., Schaeffer, L.R., Miglior, F. & Osborne, V. 2007. Non-additive genetic effects for fertility traits in Canadian Holstein cattle. *Genetics, Selection, Evolution* 39: 181–93.
- Perez-Enciso, M., Fernando, R.L., Bidanel, J-P. & Le Roy, P. 2001. Quantitative trait locus analysis in crosses between outbred lines with dominance and inbreeding. *Genetics* 159: 413–422.
- Pong-Wong, R. & Woolliams, J.A. 1998. Response to mass selection when an identified major gene is segregating, *Genetics, Selection, Evolution* 30: 313–337.
- Robertson, A. 1961. Inbreeding in artificial selection programmes. *Genetical Research* 2: 189–194.
- Robertson, A. 1964. The effect of non-random mating within lines on rate of inbreeding. *Genetical Research* 5: 164–167.
- Robertson, A & Hill, W.G. 1983. Population and quantitative genetics of many linked loci in finite populations. *Proceedings of the Royal Society, London, Series B* 219: 253–264.
- Rosenberg, N., Burke, T., Elo, K., Feldman, M., Freidlin, P., Groenen, M., Hillel, J., Mäki-Tanila, A., Tixier-Boichard, M., Vignal, A., Wimmers, K. & Weigend, S. 2001. Empirical evaluation of genetic clustering methods using multilocus genotypes from twenty chicken breeds. *Genetics* 159: 699–713.
- Schaeffer, L.R. 2006. Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics* 123: 218–223.
- Schulman, N., Viitala, S., De Koning, D.-J., Virta, J., Mäki-Tanila, A. & Vilkki, J. 2004. Quantitative trait loci for health traits in Finnish Ayrshire cattle. *Journal of Dairy Science* 87: 443–449.
- Serenius, T., Stalder K.J. & Puonti, M. 2006. Impact of dominance effects on sow longevity. *Journal of Animal Breeding and Genetics* 123: 355–61.
- Sillanpää, M.J. & Arjas, E. 1998. Bayesian mapping of multiple quantitative trait loci from incomplete inbred line cross data. *Genetics* 148: 1373–1388.
- Smith S.P. & Mäki-Tanila, A. 1990. Genotypic covariance matrices and their inverses for models allowing dominance and inbreeding. *Genetics, Selection, Evolution* 122: 65–91.
- Tapio, M., Tapio, I., Grisliis, Z., Holm, L.-E., Jeppsson, S., Kantanen, J., Miceikiene, I., Olsaker, I., Viinalass, H. & Eythorsdottir, E. 2005. Native breeds demonstrate high contributions to the molecular variation in Northern European sheep. *Molecular Ecology* 14: 3951–3963.
- Templeton, A.R. & Read, B. 1984. Factors eliminating inbreeding depression in a captive herd of Speke's gazelle. *Zoo Biology* 3: 177–199.
- Thompson, J.R., Everett, R.W. & Hammerschmidt, N.L. 2000a. Effects of inbreeding on reproduction and survival in Holsteins. *Journal of Dairy Science* 83: 1856–1864.
- Thompson, J.R., Everett, R.W. & Wolfe, C.W. 2000b. Effects of inbreeding on reproduction and survival in Jerseys. *Journal of Dairy Science* 82: 2131–2138.
- Toro, M.A. & Mäki-Tanila, A. 2007. Genomics reveals domestication history and facilitates breed development. p. 75–102. In: Oldenbroek, K. (ed.) *Utilisation and conservation of farm animal genetic resources*. Wageningen Academic Publishers, The Netherlands. 232 p.
- Tuiskula-Haavisto, M., Dekoning, D.-J., Honkatukia, M., Schulman, N., Mäki-Tanila, A. & Vilkki, J. 2004. Quantitative trait loci with parent-of-origin effects in chicken. *Genetical Research* 84: 57–66.
- Tuiskula-Haavisto, M., Honkatukia, M., Vilkki, J., De Koning, D.-J., Schulman, N. & Mäki-Tanila, A. 2002. Mapping of quantitative trait loci affecting quality and production traits in egg layers. *Poultry Science* 81: 919–927.
- Uimari, P. & Kennedy, B.W. 1990. Mixed model methodology to estimate additive and dominance genetic values

Mäki-Tanila, A. *An overview on quantitative and genomic tools for utilising dominance genetic variation*

- under complete dominance and inbreeding. *Proceedings of the 4th World Congress on Genetics Applied to Livestock Production XIII*: 297–300.
- Vanhala, T., Tuiskula-Haavisto, M., Elo, K., Vilkki, J., Mäki-Tanila, A. 1998. Evaluation of genetic variability and genetic distances between eight chicken lines using microsatellite markers. *Poultry Science* 77: 783–790.
- Varona, L. and Misztal, I. 1999. Prediction of parental dominance combinations for planned matings, methodology and simulation results. *Journal of Dairy Science* 82: 2186–2191.
- Viitala, S., Schulman, N., De Koning, D.-J., Elo, K., Kinos, R., Virta, A., Virta, J., Mäki-Tanila, A. & Vilkki, J. 2003. Quantitative trait loci affecting milk production traits in Finnish Ayrshire dairy cattle. *Journal of Dairy Science* 86: 1828–1836.
- Viitala, S., Szyda, J., Blott, S., Schulman, N., Lidauer, M., Mäki-Tanila, A., Georges, M. & Vilkki, J. 2006. The role of the bovine growth hormone receptor and prolactin receptor genes in milk, fat and protein production in Finnish Ayrshire dairy cattle. *Genetics* 173: 2151–2164.
- Virta, J., Markola, J., Peippo, J., Markkula, M. & Vilkki, J. 2002. Sex determination of bovine embryo blastomeres by fluorogenic probes. *Theriogenology* 57: 2229–2236.
- Wall, E., Brotherstone, S., Kearney, J.F., Woolliams, J.A. & Coffey, M.P. 2005. Impact of non-additive genetic effects in the estimation of breeding values for fertility and correlated traits. *Journal of Dairy Science* 88: 376–385.
- Wei, M., & van der Werf, J.H.J. 1994. Maximizing genetic response in crossbreds using both purebred and crossbred information. *Animal Production* 59: 401–413.
- Wilkie, A.O.M. 1994. The molecular basis of genetic dominance. *Journal of Medical Genetics* 31: 89–98.
- Woolliams, J.A., Bijma, P. & Villanueva, B. 1999. Expected genetic contributions and their impact on gene flow and genetic gain. *Genetics* 153: 1009–1020.
- Wray N.R. & Thompson R. 1990. Prediction of rates of inbreeding in selected populations. *Genetical Research* 72: 41–54.
- Xu, S. 2003. Estimating polygenic effects using markers of the entire genome. *Genetics* 163: 789–801.

SELOSTUS

Katsaus kvantitatiivisen genetiikan ja genomiikan menetelmistä dominanssista johtuvan geneettisen vaihtelun hyödyntämisessä eläintuotannossa

Asko Mäki-Tanila

MTT Biotekniikka- ja elintarviketutkimus

Jalostusarvojen valinnan avulla tuotetun perinnöllisen edistymisen lisäksi eläinjalostuksessa voidaan saada välittömiä etuja palauttamalla sukusiitostaantuma tai hyödyntämällä oletettua ylidominanssia. Sukusiitostaantumaan ja ylidominanssiin on molempiin syynä pääasiassa dominanssin aiheuttama vaihtelu, jota voidaan arvioida sekamallitekniikalla. Sukusiitoksen takia dominanssivarianssin kuvaamiseen ja arviointiin tarvitaan kuitenkin lisäparametreja, jotka määrittävät, miten dominanssi vaikuttaa sukusiitettujen eläinten keskiarvoon ja (ko)varianssiin. Roturisteytystilanteen kuvaaminen vaatii hyvin suurta parametrijoukkoa. Risteytys on suurinta, kun alleelifrekvenssit rotujen

välillä ovat kaukana toisistaan. Heteroosin maksimointi voi olla vain väliaikaista, koska jatkuva hyödyntäminen hidastaa jalostuksen kokonaiskehitystä. Sen takia lupaavimpien rotujen ja paritusten etsimisessä tarvitaan tehokkaita, välittömiä etuja tuovia menetelmiä. Yksi mahdollisuus on genomiikan käyttö risteytykseen soveltuvien rotujen ja dominanssivarianssiin merkittävästi vaikuttavien geenien seulomiseksi. Analyyseistä saadaan geenimerkkejä, joiden avulla voidaan valita toivotunlaisia dominanssipiikkeitä tuottavia parituksia. Genominlaajuiset merkkijoukot soveltunevat maksimiheteroosin tuottavien perimän alueiden etsimiseen.