Definition And Use Of Additive Genetic Effects For Genetic Improvement Of Populations

P. Bijma[*]

Defining the magnitude of additive genetic effects

The goal of animal breeding is to generate response to selection. Thus a key question is: "what are the additive effects that can be used to generate response". Response to selection follows from regressing the genetic mean of the population on the selection criterion, giving

$$\Delta\overline{P}_G = i\rho\sigma_G\,, \qquad (1)$$

where i is intensity, ρ accuracy and σ_G the standard deviation of G-values among individuals. Beware that G is defined here as the heritable effect determining the population mean, not to be confused with $G = A + D + I$. Equation [1] applies to any selection strategy and inheritance model, since it equals the first term of Price's Theorem (Price 1970).

The issue here is: what is the σ_G in Equation 1? In the classical model, where $P = A + E$, response equals the change in mean breeding value, $\Delta\overline{P}_G = \Delta\overline{A}$, so that $G \equiv A$, and $\sigma_G^2 \equiv \sigma_A^2$. Moreover, since $\sigma_P^2 = \sigma_A^2 + \sigma_E^2$, σ_G^2 also equals the (additive) genetic component of phenotypic variance, $Var_G(P) = \sigma_G^2$. Thus, with $P = A + E$, a variance partitioning perspective and a response to selection perspective yield the same definition of genetic variance.

Things become different, however, when trait are affected by multiple individuals, for example with maternal effects or social interactions. With maternal effects, where $P_i = A_{D,i} + A_{M,dam(i)} + E$, i denoting the individual and dam(i) its mother, response equals $\Delta\overline{P}_G = \Delta\overline{A}_D + \Delta\overline{A}_M$. Therefore, $G_i = A_{D,i} + A_{M,i}$, which is entirely a heritable property of i, because i transmits its own genes, not those of its dam. Thus $\sigma_G^2 = \sigma_{A_D}^2 + 2\sigma_{A_{DM}} + \sigma_{A_M}^2$, which differs from $Var_G(P) = \sigma_{A_D}^2 + \sigma_{A_{DM}} + \sigma_{A_M}^2$. Hence, the genetic variance determining response to selection, σ_G^2, differs from the genetic component of phenotypic variance, $Var_G(P)$, while $\Delta\overline{G} = i\rho\sigma_G$ and $\Delta\overline{G} \neq i\rho\sqrt{Var_G(P)}$ (Eaglen and Bijma, 2009). In general, therefore, the genetic variance determining response cannot be obtained by partitioning phenotypic variance into a genetic and residual component. One consequence is that σ_G^2 may exceed σ_P^2, at least in theory. (With maternal effects, the $\Delta\overline{G} = i\rho\sigma_G$ may not surface immediately in the next generation. Nevertheless, because genes mix in the population over time, Equation 1 represents the ultimate response originating from the change in allele frequency due to a cycle of selection.)

[*] Animal Breeding and Genomics Centre, Wageningen University, Marijkeweg 40, 6709PH, Wageningen, The Netherlands.

For genetic improvement, the relevant definition of genetic variance is "the variance among individuals of the heritable effect determining the mean trait value of a population". This σ_G^2 follows from linearizing the population mean trait value into additive genetic effects, and taking the variance thereof, treating it as a property of a single individual. (This is an analogy of the variance in an aggregate genotype). As a somewhat extreme example, consider a sow line for integrated pork production. Interest is in total amount of meat produced from a sow, which is the product of number of offspring and carcass meat yield of those offspring, $P = nC$. The n is a trait of the sow, whereas C may have both a direct and a maternal genetic component. Linearization yields $P_i \approx \bar{n}C_i + \bar{C}n_i + const.$ Response due to selection in the sow line equals $\Delta\bar{P}_G = \bar{n}(\frac{1}{2}\Delta\bar{A}_{C_D} + \Delta\bar{A}_{C_M}) + \bar{C}\,\Delta\bar{A}_n$, where D and M indicate direct and maternal effects, respectively, and the ½ indicates that the sow line contributes half of the genes in the offspring. Thus the genetic term relevant for response equals $G_i = \bar{n}(\frac{1}{2}A_{C_D,i} + A_{C_M,i}) + \bar{C}\,A_{n,i}$, so that $\Delta\bar{G} = \Delta\bar{P}_G$. Taking its variance yields $\sigma_G^2 = \frac{1}{4}\bar{n}^2 Var(A_{C_D}) + \bar{n}^2 Cov(A_{C_D}, A_{C_M}) + \bar{n}^2 Var(A_{C_M}) + \bar{n}\bar{C}Cov(A_{C_D}, A_n) + 2\bar{n}\bar{C}Cov(A_{C_M}, A_n) + \bar{C}^2 Var(A_n)$, where variances and covariances represent the ordinary additive genetic variances and covariances. Hence, response equals $\Delta\bar{P}_G = i\rho\sigma_G$, where ρ is the correlation between the selection criterion and the G-values in the candidates for selection.

In conclusion, the above approach summarizes all heritable components of a trait, such as direct, maternal and social, into a single σ_G^2 expressing the overall opportunity for genetic improvement of the trait. Moreover, it yields a singe accuracy, expressing the overall quality of the information recorded. The approach rests on the variance of the additive genetic components of the trait mean, rather than the additive genetic component of the trait variance.

Utilizing social genetic effects

Interest in social genetic effects is on the increase, as there is evidence that such effects can contribute substantially to σ_G^2 (Craig & Muir, 1996; Chen *et al*., 2008; reviewed in Muir 2005). When individuals are kept in groups of n members, socially affected traits may be modeled as a sum of direct and social effects, $P_i = A_{D,i} + \sum_{n-1} A_{S,j} + \text{terms in } E$, where A_D and A_S are direct and social breeding values, and the sum is over the n–1 group members j of individual i (Griffing, 1967). Thus response equals $\Delta\bar{P}_G = \Delta\bar{A}_D + (n-1)\Delta\bar{A}_S$, so that $G = A_D + (n-1)A_S$, and genetic variance relevant for response to selection equals (Bijma *et al*., 2007)

$$\sigma_G^2 = \sigma_{A_D}^2 + 2(n-1)\sigma_{A_{DS}} + (n-1)^2\sigma_{A_S}^2 .$$

The term $(n-1)^2$ shows that, even when $\sigma_{A_S}^2$ is very small relative to σ_P^2, social effects may contribute substantially to σ_G^2 when estimates come from data with large groups. A

number of studies have suggested large contributions of social effects to σ_G^2 (*e.g.*, Bergsma *et al*., 2008). Theoretically, one expects genes of social effect to harbor more sequence variation than genes of direct effect, because natural selection targets social effects to a lesser extent (Denison *et al*., 2003). Indeed, Cruickshank and Wade (2008) observed greater sequence variation maternal-effect genes than in direct-effect genes in Drosophila.

Utilization of social genetic effects requires adjustment of selection strategies. Traditional selection on individual trait value or EBV targets direct effects only, yielding suboptimal or even negative response (Griffing, 1967). Between-group selection and the use of groups composed of relatives have been advocated as solutions (Griffing, 1976 & 1977). The following investigates the accuracy of those selection methods.

A selection criterion allowing for a varying degree of between-group selection is (Bijma *et al*. 2007)

$$SC_i = P_i + g\sum_{j=1}^{n-1} P_j ,$$

where n is group size, the sum is over the $n-1$ group members j of individual i, and g is the degree of between-group selection, with $g \in [0...1]$. A g = 0 yields $SC_i = P_i$, indicating individual mass selection. A g = 1 yields $SC_i = \sum_{j=1}^{n} P_j$, the summed phenotypes of all n group members, indicating full between-group selection. Accuracy of selection on *SC* yields (derivation not shown)

$$\rho(g,r) = \frac{\left[g + r + (n-2)gr\right]\sigma_G^2 \; + \; (1-g)(1-r)\left[\sigma_{A_D}^2 + (n-1)\sigma_{A_{DS}}\right]}{\sigma_{SC}\sigma_G},$$

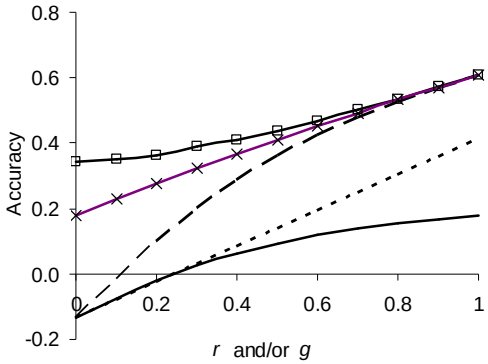

Figure 1: Accuracy of selection for survival in cannibalistic laying hens, as a function of relatedness between group members (*r*) and / or the degree of group selection (*g*). Parameters taken from Peeters *et al*., (this congress). Solid line: r = 0, g varies; dotted line: g = 0, r varies; dashed line: $r = g$ varies; crosses: g = 1, r varies; boxes: optimum index of own performance and group members.

where r is relatedness between group members. The first term in the numerator shows that relatedness and group selection act directly on σ_G^2, and thus contribute to positive accuracy. The second term in the numerator illustrates the risk of negative accuracy of individual selection when direct and social genetic effects are negatively correlated (Griffing, 1967). The numerator is symmetric in g and r, indicating that they have similar effects. Figure 1 illustrates the effects of group selection and relatedness on accuracy, using genetic parameters for survival time in cannibalistic laying hens (Peeters *et al.*, this congress). Accuracy increases less with group selection than with relatedness, because greater g yields greater σ_{SC} which limits the increase (solid vs dotted line).

Often keeping candidates in groups is undesirable, *e.g.*, when recording individual

feed intake or egg number, and breeding relies on phenotypes of relatives. For that case, Ellen *et al*. (2007) showed that positive accuracies are guaranteed when keeping the relatives in family groups, and that limiting accuracies are the same as in classical theory, being 0.5 for HS, 0.71 for FS and 1 for progeny.

When genetic parameters are known, accuracies can be further improved by using BLUP and selecting on $\hat{G} = \hat{A}_{D,i} + (n-1)\hat{A}_{S,i}$ (Muir, 2005). The use of BLUP, however, does not remove the benefit of using related group members (Muir *et al*., this congress). Moreover, selection index calculations indicate that, when using related group members, BLUP and group selection can yield similar accuracy. In Figure 1, for example, accuracy of group selection and of an optimum index are similar when r = 0.5. Hence, when group members are related, benefits of BLUP may primarily come from better accounting for fixed effects and selection, rather than from optimum weighting of direct and social effects (Muir *et al*., this congress). Furthermore, BLUP may be beneficial with common litter covariances among sibs. Using pseudo-BLUP selection index theory (Wray and Hill, 1989), deterministic prediction of the accuracy of BLUP is feasible, but complex with social effects (not shown).

Inherited variability

Breeders have long been interested in increasing uniformity. In the classical model, where $P = A + E$, opportunities for genetic changes in variability are very limited. At best, breeders can approach $\sigma_A^2 \approx 0$, which reduces phenotypic standard deviation by only ~16% when $h^2 \approx 0.3$. There is, however, increasing evidence that σ_E^2 is under direct genetic control (*e.g.* Rowe *et al*., 2006; Ibanez-Escriche *et al.* 2008). In the literature, two classes of models exist. First, models specifying an additive effect on the residual variance, $E = \chi\sqrt{\sigma_{E_{add}}^2 + A_{v_{add}}}$, where χ is a standard normal deviate, and $A_{v_{add}}$ is an additive breeding value for environmental variance (Hill and Zhang, 2004). Second, exponential models, where $E = \chi \exp[\tfrac{1}{2}(\ln(\sigma_{E,\exp}^2) + A_{v,\exp})]$ (SanCristobal-Gaudy *et al*., 1998). The relationship between both models is that $A_{v_{add}} \approx \sigma_E^2 A_{v_{\exp}}$, indicating that the exponential model specifies a multiplicative effect on σ_E^2. Hence, estimates from both models are easily interconverted (Mulder *et al*., 2007). Both models have pros and cons. The exponential model is statistically more correct, since it ensures $\sigma_E^2 > 0$, whereas $\sqrt{\sigma_E^2 + A_v}$ is defined only for $\sigma_E^2 + A_v > 0$. The additive model, however, fits more easily in quantitative genetic models of inheritance and response to selection.

Heritability of σ_E^2, defined as the regression coefficient of σ_E^2 on P^2, appears to be low (~0.03). The genetic coefficient of variation, in contrast, $\sigma_{A_v} / \sigma_E^2$ appears to be substantial (~0.3; Mulder *et al*., 2007). Thus σ_E^2 can in principle be changed considerably relative to its current mean, but it is difficult to obtain high accuracy of selection for σ_E^2.

When σ^2_E is heritable, directional mass selection with p < 50% tends to increase σ^2_E, because individuals with high A_v are more likely to be in the tail of the distribution. Hence, mass selection may unintentionally increase variability. Directional selection on a family mean puts much less pressure on A_v, particularly when families are large. Stabilizing selection tends to reduce σ^2_E, but a lower bound of –1 for selection intensity limits response. Disruptive selection, on the other hand, allows for high positive selection intensities. Hence, when selection relies on own performance information, increasing variability seems feasible, but decreasing it is difficult (Mulder *et al*., 2007).

Genetic improvement of uniformity, therefore, requires the use of family information. The key information source for A_v is the within-family variance. Mulder *et al*. (2007) show that accuracy of selection on within-family variance is similar to classical expressions for accuracy of selection based on relatives. Hence, limiting accuracies for large numbers of relatives, may approach ~0.5 for HS, ~0.7 for FS and ~1 for progeny. Given the low heritability of σ^2_E, however, very large families are needed to approach those limits. Nevertheless, meaningful accuracies can be obtained based on within-family variance, which, combined with the large estimates for σ_{A_v}/σ^2_E, suggest that σ^2_E can be reduced considerably relative to its current value when selecting for lower within-family variance. A selection experiment for lower variability in body weight of broilers would be very useful to test whether realized response agrees approximately with theoretical predictions.

The mechanisms underlying inherited variability are largely unknown at present. Theoretically, there exists a relationship between genotype×environment-interaction and inherited variability. This follows from a simple reaction-norm model, $y|E = \mu + A_L + A_S E + e$, where A_L and A_S are breeding values for level and slope, and E is the environmental variable. Greater A_S indicates greater environmental sensitivity. With $Cov(A_L, A_S) = 0$ and $\mathrm{E}(E) = 0$, phenotypic variance of a genotype equals $Var(y|A_L, A_S) = A^2_S\sigma^2_E + \sigma^2_e$, which increases with A_S. Hence, when such *G*x*E*-interaction is not explicitly modeled, *e.g.* because E is unknown, then genotypes of greater environmental sensitivity appear to have greater residual variance. Thus statistical analysis of inherited variability may pick up hidden *G*x*E*-interaction.

From a *G*x*E*-interaction perspective, sensitive genotypes in good environments are in the upper-tail of the distribution, so that directional mass selection in good environments tends to increase sensitivity. From an inherited variability perspective, variable genotypes are overrepresented in the tails of the distribution, so that directional mass selection tends to increase variability. Hence, both perspectives agree on the consequences of directional selection; with inherited variability originating from hidden *G*x*E*-interaction, directional mass selection favors the sensitive genotypes present in the good environments.

There also seems to be a link between inherited variability and social interactions. In aquaculture, competition for feed inflates size variation among individuals. To limit size variation, regular grading of fish is common. Hence, competition seems to increase variability. In current models of social genetic effects (see above), however, phenotypic variance is independent of the average social breeding value. Hence, in current models, a

reduction in phenotypic variance due to decreased competition seems to require a reduction in social genetic variance. This may largely be an empirical, rather than theoretical, issue.

Optimum selection criteria

In the absence of molecular genetic information, the optimum selection procedures are well known. Breeding value estimation should focus on maximizing accuracy, and selection should focus on maximizing the genetic selection differential while restricting the rate of increase in mean kinship (*e.g.* Meuwissen, 1997). Given a restriction on the rate of kinship, minimum coancestry and factorial mating increase response, particularly in small schemes. When restricting the rate of kinship, there appears to be little trade-off between short-term response, *i.e.*, maximizing today's genetic selection differential, and long-term response.

This is different with molecular information. Maximizing today's genetic selection differential implies maximizing accuracy. Accuracy is maximized when $\hat{g} = E[g \mid data]$, g denoting a marker effect, which requires that apparently smaller effects are regressed stronger (Fernando and Gianola, 1986). Hence, maximizing accuracy of genome wide evaluation requires putting lower weight on smaller effects, such as with BayesB (Meuwissen *et al*., 2001). Moreover, maximizing accuracy requires putting lower weight on rare alleles, which occurs implicitly in methods giving equal *a priori* weight to all loci, such as genomic-BLUP (Meuwissen *et al*., 2001). (This is analogous to estimating sire breeding values from progeny averages; progeny averages are regressed stronger when progeny are fewer). Hence, maximizing accuracy requires lower weights on rare alleles and/or alleles of small effect.

This is precisely opposite to maximizing long-term response. First, expressed relative to their contribution to genetic variance, alleles of smaller effect contribute more to long-term response. Response is proportional to allelic effect, whereas variance is proportional to the square of allelic effect. Hence, the ratio of response over variance due to an additive allele is inversely proportional to the effect of the allele, indicating that most long-term response comes from alleles of small effect. Second, since variance due to an allele is maximal when allele frequency equals 0.5, selection favoring rare alleles increases genetic variance over time, thereby increasing response in later generations. In contrast, selection for alleles currently explaining most variance, *i.e.* those at allele frequency of ~0.5, reduces genetic variance over time.

Hence, with genomic selection, there appears to be a trade-off between short and long-term response (beyond the classical trade-off of response versus rate of inbreeding). Experience in dairy cattle suggests that the theoretically expected short-term superiority of BayesB/C over BLUP does not always happen (VanRaden *et al.* 2009; S. de Roos, pers. comm.), which may be related to the true number of qtl *vs*. the effective number of chromosome segments (Daetwyler *et al*., in press). When both methods are identical in the short term, one expects BLUP to be superior in the long-term response since it puts more weight on alleles of small effect. In addition, to balance short and long-term response, it might be beneficial to put some more weight on rare alleles. This would be opposite to current practice, where rare alleles are sometime omitted because they may reflect typing errors. Stochastic simulations, including simulation of typing errors, may be useful to better understand how to weigh markers depending on their frequency and apparent effect, so as to better balance short *vs.* longer term response.

Solutions offered by genome-wide evaluation

Since the introduction of AI, genomic-wide evaluation (GWE) is the most important development in livestock genetic improvement. Benefits of GWE are greatest in species where phenotypic data becomes available considerably after reproductive age (*e.g.* dairy cattle, not broilers), and for well-defined traits that cannot be recorded on the selection candidates themselves. GWE has become routine practice in dairy cattle, probably because of the value of the individual and because milk yield cannot be recorded on males.

GWE provides solutions for a number of problems. When genotypes and phenotypes can be recorded on breeding goal traits expressed in commercial production environments, then GWE allows direct selection for breeding goal traits. An obvious application is direct breeding for crossbred performance recorded in commercial environments. This allows combining the efficiency of two-tier nucleus schemes with accurate EBVs for breeding goal traits, rather than relying on information of sibs or on purebred performance expressed in good environments.

Benefit of direct selection for breeding goal traits compared to selection based on correlated traits is probably greater than apparent from a comparison of direct and correlated responses. This is because selection for correlated traits optimizes the organism for the wrong goal, leading to inefficiency. For example, increasing lean meat yield by selecting for growth rate will also increase fat in the carcass, which has no use but carries a cost. In general, selection for correlated traits will create costly changes to the organism, that have no use for the breeding goal. Such cost may go unnoticed when they surface only in commercial environments. Hence, when direct and correlated responses for the goal trait are similar, direct selection is to be preferred by far. Moreover, direct selection is much less sensitive to estimation errors in genetic parameters. Hence, GWE seems to be very useful in cases where direct selection is not possible in classical breeding schemes, probably more useful than suggested by selection index calculations. Furthermore, when GWE replaces sib-information, benefits will be greater than suggested by the difference in accuracy, because GWE yields greater response than sib schemes when compared at the same rate of inbreeding. Hence, in pig and poultry breeding, large scale recording of crossbred phenotypes (an maybe genotypes) seems to be a condition for future commercial success of breeding companies.

GWE extends readily to socially affected traits. A comparison of estimated genetic parameters for direct and social effects on survival in cannibalistic laying hens shows large differences between purebred parental lines (Ellen *et al*. 2008) and their crossbred offspring (Peeters *et al*, this congress). The genetic correlation between direct and social effects appears to be considerably negative in crossbreds, while around zero in purebreds. Moreover, estimated parameters differ between reciprocal crosses. Those result indicate *G*x*E*-interaction between purebreds and crossbreds. GWE based on phenotypes recorded on crossbreds can be used to predict breeding values of nucleus individuals for direct and social effects referring to crossbred performance. Hence, combining GWE with social-effect models is promising to reduce mortality due to cannibalism in commercial herds.

When genotyping becomes cheaper, GWE offers increased possibilities for having separate breeding programs for different environments, such as organic versus conventional farming or seasonal *vs*. year-round calving. Compared to traditional schemes relying on nucleus information, benefits of having separate breeding programs is greater with GWE because differences in response between environments will be greater. GWE combined with the shift

of breeding goals towards more emphasis on functional traits may accelerate a trend towards more breeding programs, because functional traits often show greater *G*x*E*-interaction (*e.g.*, compare longevity *vs.* yield in dairy cattle).

This research was financially supported by the Dutch science council (NWO) and was co-ordinated by the Netherlands Technology Foundation (STW).

References

Bergsma, R., Kanis, E., Knol, E. F. *et al*. (2008). *Genetics*, 178:1559-1570.

Bijma, P., Muir, W. M., and Van Arendonk, J. A. M. (2007). *Genetics*, 175: 277-288.

Chen, C. Y., Kachman S. D., Johnson, R. K., *et al*. (2008). *J. Anim. Sci*. 86:2525-2530.

Craig, J.V., and Muir, W.M. (1996). *Poultry Sci.* 75:294-302.

Cruickshank, T., and Wade, M. J. (2008). *Evolution and Development*, 10:583-590.

Denison, R. F., Kiers, E. T., and West, S. A. (2003). *Quart. Rev. Biol.* 78:145–168.

Eaglen, S. A. E., and Bijma, P. (2009). *J. Dairy Sci.*, 92:2229-2237.

Ellen, E. D., Muir, W. M., Teuscher, F. *et al*., (2007). *Genetics* 176:489-499.

Ellen, E. D., Visscher, J., Van Arendonk, J. A. M., *et al.* (2008). *Poultry Sci.*, 87:233-239.

Fernando R.L., and Gianola, D. (1986). *Theor. Appl. Genet.*, 72:822-825.

Griffing, B. (1967). *Aust. J. Biol. Sci.*, 10:127-139.

Griffing, B. (1976). *Genetics* 82:703-722.

Griffing, B. (1977). In *Proc. Int. Conf. Quant. Gen. Ames, Iowa*, pages 413-434.

Hill, W. G., and Zhang, X.S. (2004). *Genet. Res*., 83:121-132.

Ibanez-Escriche, N. Sorensen, D., *et al.*, (2008). *Genetics*, 180:2209-2226.

Meuwissen, T. H. E. (1997). *J. Anim. Sci*., 75:934-940.

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). *Genetics*, 157:1819-1829.

Muir, W. M. (2005). *Genetics*, 170:1247-1259.

Mulder, H. A., Bijma, P., and Hill, W.G. (2007). *Genetics*, 175:1895-1910.

Price, G. R. (1970). *Nature* 227:529-531.

Rowe, S. J., White, I. M. S., Avandano, S., *et al*. (2006). *Genet. Sel. Evol*., 38:617-635.

Sancristobal-gaudy, M., Elsen, J. M., Bodin, L, *et al*. (1998) *Genet. Sel. Evol*. 30:423-451.

VanRaden, P.M., Van Tassel, C. P., Wiggans, G. R. *et al.*, (2009) *J. Diary Sci*. 92:16-24.

Wray, N. R., and Hill, W. G. (1989). *Anim. Prod*. 49:217-227.