Proc. Assoc. Advmt. Anim. Breed. Genet. 26: 467-470

IMPLEMENTATION OF GENOMIC MATE SELECTION AND OTHER NON-ADDITIVE ISSUES

B.P. Kinghorn[1,2] and A.J. Kinghorn[2]

[1] University of New England, Armidale, NSW, 2350 Australia
[2] BSK Software

SUMMARY

Increasing attention is being paid to our ability to influence mendelian segregation variance as a function of matched parental genetic markers. This results in 'Look Ahead Mate Selection' to target longer-term genetic gains. This paper addresses this and other aspects of genomic mate selection, in a new implementation method that can also accommodate other non-additive and compatibility issues, including animal behaviour compatibility, non-linear indices, transport costs and, in general, specification of must-have and must-not-have matings.

INTRODUCTION

The standard way of handling trait EBVs in a breeding program assumes that progeny genetic merit is predicted as the average of the parental EBVs. This is true for most traits. However, this is not appropriate for some traits and other factors – wherever we have knowledge that the progeny from a specific mating are predicted to deviate from their parental mean in some known way, eg. due to genetic dominance, or increased segregation variance.

This knowledge cannot be incorporated into any datafile that has one line of data per candidate. For this reason, we need to read these non-additive components from a separate file for each trait, which contains the predicted non-additive deviation for each possible pairing of a male candidate and a female candidate.

As some form of mate selection based on genomic information seems a most likely application, we refer to this approach as genomic mate selection (GMS), even though it can be applied without genomic information, and indeed without any genetic components. There is an increasing interest in GMS, with frequent new academic publications on how to use genomic information for increasing both short-term and longer-term genetic gains:

Short-term gains: Total genetic merit in progeny. Refined predictions of progeny merit can involve dominance components, and possibly some epistatic components, at nominated QTL. It could also involve genome-wide heterozygosity, although that may be less powerful. The Expected Cross Value of Ahadi *et al*. (2024) goes beyond this to consider the value of a mating pair as a function of allelic information and transmission probabilities for each partner, increasing the desirable allele frequency.

Long-term gains: Look Ahead Mate Selection (LAMS). "Genomic Mating uses information in a similar fashion to genomic selection but includes information on complementation of parents to be mated" (Akdemir and Sanchez 2016). This is in fact GMS, as there is selection of the parents as well as mate allocation among them. How these mate selections are chosen depends on several issues, including the number of generations ahead that we use to define as in our utility function. Niehoff *et al*. (2024a) illustrate scenarios in which different selection decisions should be made today according to the different objectives of aiming 1, 2 or 3 generations ahead. This is essentially driven by Mendelian Sampling Variance (MSV), which can be affected by:

- Selection decisions, where some individuals are expected to yield greater genetic variation among their progeny, even under random mating. An **Aa** heterozygote is expected to yield more genetic variation in progeny than offspring of **AA** or **aa** homozygotes.

- Mate selection decisions, where, in addition to selection effects, certain pairs synergise to give more variation among progeny. A cross **Aa** x **Aa** is expected to yield **AA**, **Aa** (or **aA**), and **aa** progeny in 1:2:1 ratio, which maximises genetic variation in the progeny.

Of course, mate selection has the most impact, and indeed when we use genomic information to increase genetic variation in the progeny, we are largely aiming at mating of heterozygotes genome-wide. This extra genetic variation leads to increased genetic gains in later generations, under Look Ahead Mate Selection.

Niehoff *et al*. (2024a) describe a number of objective functions that invoke LAMS, leading ultimately to increased genetic merit in great-grandprogeny. Tobias Niehoff has contributed notes in Kinghorn and Kinghorn (2025) that show how to invoke these objective functions in our software. Additionally, Niehoff *et al*. (2024b) generalise their approach to aim **n** generations ahead.

MATERIALS AND METHODS

As noted above, handling any type of non-additive effect usually requires more than the typical animal dataset that has one line of data per candidate. Where the dataset is to be used to make selection and mating decisions, these data lines typically include information on parent IDs, trait EBVs, genetic markers and usable information such as age and location. Non-additive effects that depend on mate allocation decisions require either storage of such information in files of dimension male candidates x female candidates, to cover all possible mate allocations, or calculation of non-additive effects on-the-fly as a function of information supplied in the main data file.

For the latter, calculation of predicted progeny merit is a simple function of the parents mated, either for each mating pair (eg. a non-linear selection index) or across all mating pairs (e.g. corrective mating about a target mean). These are simply covered by trait management tools (Kinghorn and Kinghorn 2025). In this paper we concentrate on non-additive effects that are not simply derived from additive effects, and require separate calculation and storage.

Classically, genetic merit in progeny is estimated as:

$$ProgenyMerit = \frac{A_{sire} + A_{dam}}{2} + D_{sire,dam}$$

where *A* and *D* are additive and dominance effects, respectively. Terms for epistatic effects can be added to this, but these are left out for simplicity. In general, we use this:

$$ProgenyMerit = \frac{EBV_{sire} + EBV_{dam}}{2} + Deviation_{sire,dam}$$

where the EBVs can be replaced by factors such as animal age, location etc, or more typically by 0's, where there is no additive effect involved. The *Deviation* can be any effect that can be calculated for each *sire* + *dam* combination. This could be a predicted progeny mean dominance deviation, a predicted MSV for the resulting full-sib family, or even the transport cost associated with implementing a mating between this sire and dam.

The three examples in the last sentence can all be included as non-additive 'traits' in a single analysis, with competition between them for emphasis in decision making, just as for normal additive traits.

Implementation in our software is simply achieved by making the name for such a trait start with "GMS" (Eg. GMS_BirthWt) and providing a file containing the non-additive effects for the male and female candidates for selection (in this case, named exactly "candGMS_BirthWt.txt").

For simple illustration, we here use progeny inbreeding from pedigree as the non-additive trait "GMS_PedInb". In the main datafile, the EBVs for this 'trait' are set to zero – as inbreeding is not heritable and is a function of parental pairs, not individual parents. The file candGMS_PedInb.txt looks like this, with males in rows and females in columns:

candGMS_PedInb	Lilianna	Guadalupe	Rayna	Chanel	Lyra	
Samir	0.155	0.050	0.063	0.059	0.054	...
Magnus	0.027	0.079	0.027	0.026	0.021	...
Jaziel	0.035	0.034	0.038	0.026	0.108	...
etc.						

RESULTS AND DISCUSSION

Figure 1 shows the results for the simple illustrative example. The 'trait' GMS_PedInb was chosen as the result can be compared to Progeny Inbreeding as reported separately in the application. GMS_PedInb has been managed to give a bimodal distribution in progeny, with 5 matings in a high mode around F=0.14 and 45 matings in a lower mode around F=0.06. As can been seen, the progeny distribution is the same as for progeny inbreeding, with a visual difference only because of the numbers of histogram bars differing between the graphs.

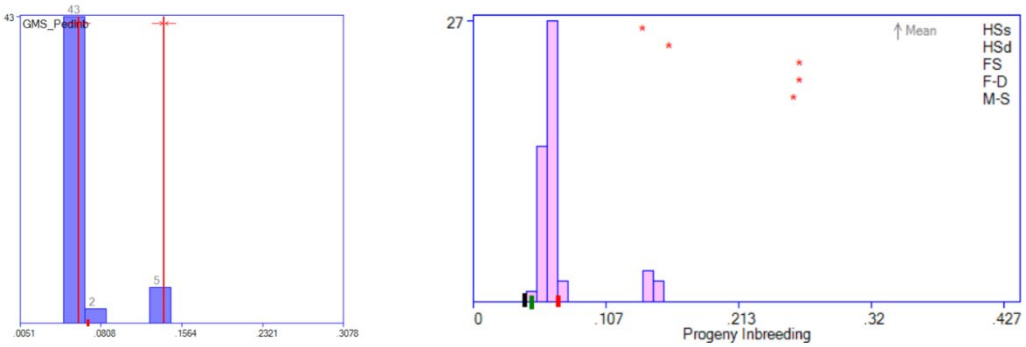

Figure 1. Results from targeting bimodality with 10% of families showing high values for the trait GMS_PedInb (left graph). As expected, the result is identical to Progeny Inbreeding (right graph), with difference in presentation only, due to the numbers of histogram bars used

The GMS approach can be used to accommodate many issues, some of which are demonstrated in this video https://youtu.be/XmFcKYBcJvc. Of course, the candGMS files of non-additive effects have to be generated in advance of the analysis, using the data resources available and probably appropriate scripting/coding to calculate each entry. Example applications include:

- LAMS, as described above.
- Phenotypic effects on mating/pregnancy/birth performance (eg. A small male might not reach a big female, or a big bull on a small cow could lead to dystocia).
- Animal behaviour compatibility scores for each possible male x female mating.
- The predicted transport cost associated with making each possible male x female mating.
- Imprinting effects, whereby some aspect of genetic merit depends on the gender of each parent involved – predicted merit is not simply an average of parental additive effects.
- Must-Have matings: In the file of effects for this 'trait' set all values to zero, except for the must-have matings, set these to 1. Then manage this trait to give the full number of must-have matings represented – or a lesser number as desired.
- Must-Not-Have matings: As above, except set value to 1 for the must-not-have matings, then manipulate to have that trait at value zero for all matings, or a greater number if desired.

Dynamic non-additive effects: Moeinizade *et al*. (2019) aimed a number of generations ahead using an approximate objective function of genetic merit in the last generation, involving simulation. However, for this method, "The contribution of a breeding parent is evaluated based on not only the favorable alleles that it carries but also the favorable alleles that it carries but are missing in other

selected breeding parents." This means that the elements of the GMS matrix are dynamic, and cannot be read-in up-front. For such cases, further development would be required to give 'on-the-fly' calculation of the required elements, or indeed merit across all matings in light of the full mating set. Alternatively, direct calling of external routines that calculate such values can be developed.

CONCLUSION

The GMS approach described in this paper is simple to implement for most non-additive effects, and gives access to a wide range of technical, logistical and cost functions that involve a specific non-additive effect for each possible *sire x dam* pairing. Some of these functions extend the use of genomic information beyond genomic EBVs, giving some prospects to increase the returns from investment in genotyping animals.

REFERENCES

Ahadi P., Balasundaram B., Borrero J.S. and Chen C. (2024) *Heredity* **133**: 113.
Akdemir D. and Sánchez J.I. (2016) *Front. Genet*. **7**: 210.
Kinghorn B.P. and Kinghorn A.J. (2025) Instructions for Matesel. https://matesel.com/content/documentation/MateSelInstructions.pdf
Moeinizade S., Hu G., Wang L. and Schnable P.S. (2019) *G3-Genes Genom. Genet.* **9**: 2123.
Niehoff T.A.M., ten Napel J., Bijma P., Pook T., Wientjes Y., Hegedűs, B. and Calus M.P.L. (2024a) *Genet. Sel. Evol*. **56**: 41
Niehoff T.A.M., ten Napel J. and Calus M.P.L. (2024b) *G3-Genes Genom. Genet.* **14**: jkae205.