A practical approach for optimised partitioning of genomic relationship across chromosomes

Majid Khansefid[1] and Mohammad H. Ferdosi[2]

[1] AgriBio Centre for AgriBioscience, Agriculture Victoria, Bundoora, Australia

[2] Animal Genetics and Breeding Unit, University of New England, Armidale, Australia

Introduction

In genomic best linear unbiased prediction (GBLUP) models, genotyped markers are used to make a single genomic relationship matrix (GRM) and consequently each marker contributes similarly in explaining the genetic variance of traits.

Alternatively, GRMs made from markers located on each chromosome can be fitted in a GBLUP model. However, fitting multiple GRMs in a model is not always useful because not all chromosomes contain major genes.

Materials and Methods

- **Phenotypes:** Final weight of 2,578 Hereford cattle
- **Genotypes:** 26,073 SNPs
- **Models:** The following four models were fitted in ASReml;

1. $y = g_{\text{All Chrs}} + e$
2. $y = g_{\text{Chromosome } i} + e$
3. $y = g_{\text{All Chromosomes – Chromosome } i} + e$
4. $y = g_{\text{Chromosome 6}} + g_{\text{Chromosome 20}} + g_{\text{Others}} + e$

$g \sim N(0, ZZ'\sigma^2_g)$, where Z and $ZZ'/2\Sigma q_{SNP.a}(1- q_{SNP.a})$ are the matrix of SNP counts and GRM, respectively and i is the chromosome number (1 to 29).

Results

- SNPs on chromosome 6 and 20 explained 20% and 23% of the total genetic variance for final weight, respectively (**Model 2**).
- Excluding markers on chromosome **6** or **20** from the GRM (**Model 3**), significantly **reduced** the Log-Likelihood (p<0.01) compared to the model in which GRM was constructed from all markers (**Model 1**).
- In **Model 4**, the GRMs for chromosome 6, 20 and other than 6 and 20, explained 22%, 26% and 52% of the genetic variance, respectively.
- The Log-Likelihood was significantly **improved** (p<0.01) in **Model 4** compared to **Model 1**.

Conclusions

- Excluding one chromosome from the GRM in turn could be a practical approach to find the chromosomes which contain major genes.
- Fitting multiple GRMs constructed from every chromosome which contribute to explaining genetic variance and also include the remaining SNPs in a separate GRM, could improve the Log-Likelihood of the prediction model compared to fitting all SNPs in a single GRM.

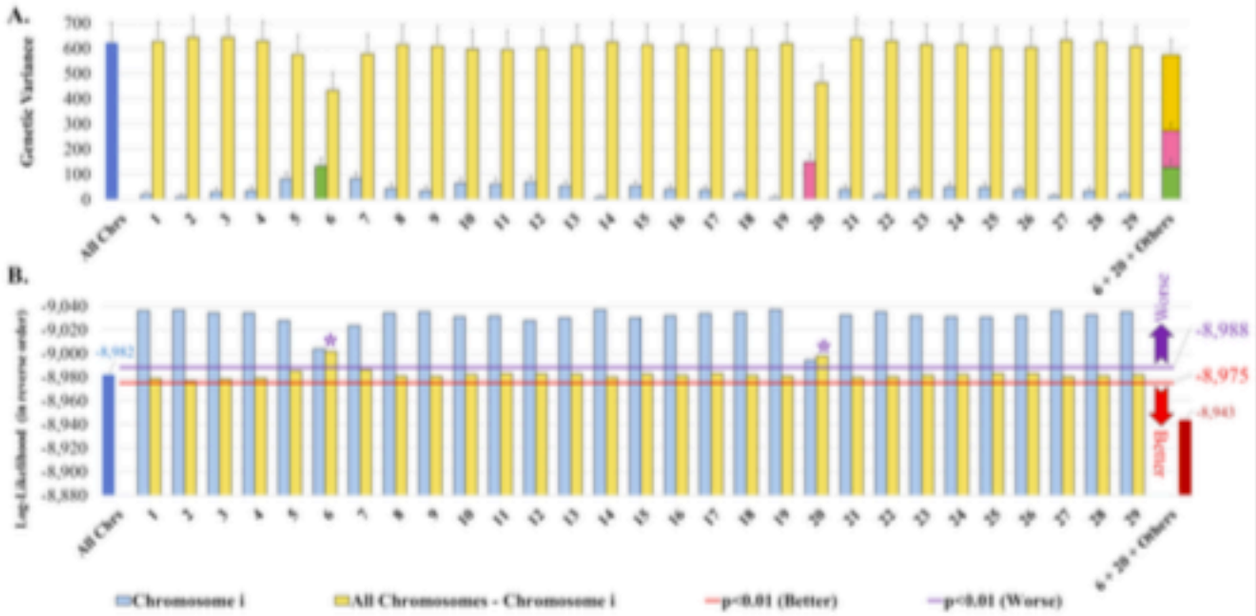

Figure 1. The estimated additive genetic variance (A) and Log-likelihood (B) of different models when a genomic relationship matrix (GRM) was constructed from markers from each chromosome (Chromosome i) separately, or markers of chromosomes other than each chromosome (All Chromosomes – Chromosome i). The lines in red and purple show the significance level (p<0.01) for the models fitted better and worse, respectively, compared to the base model (All Chrs: all markers were used to construct the GRM). The highest Log-likelihood was achieved when markers on chromosome 6, 20 and the remaining chromosomes were fitted in the model as three separate components.

Acknowledgements: The authors wish to thank Australian Hereford society for providing genotype and phenotype for this study.