Partitioning of variance between multiple relationship matrices in BLUP analyses

P.M. Gurman[1], L. Li[1], A.A. Swan[1], N. Moghaddar[2], J.H.J. van der Werf[2]

[1] *Animal Genetics and Breeding Unit, University of New England, Armidale, 2351, Australia*
[2] *School of Animal Science, University of New England, Armidale, 2351, Australia*

Introduction

- GWAS analyses have identified SNP markers more predictive for specific traits
- Adding these to the current genomic relationship matrix (GRM) would weight these SNPs equally with all other SNPs, which may dilute their predictive ability
- Can these selected SNPs be separated into a second genomic relationship matrix (GRM)?

Aim

- **Explore the impact of using two GRMs in univariate REML analyses**

Methods

- 31.6k genotyped terminal sire breed sheep recorded for:
  - Intramuscular fat (imf), 10.6k obs
  - Shear force (sf5), 10.6k obs
  - Carcase eye muscle depth (cemd), 10.8k obs
  - Carcase fat (ccfat), 10.7k obs
  - Carcase weight (cwt), 11.0k obs
  - Post weaning weight (pwt), 10.6k obs
- SNP Markers split into three groups:
  - M_r: regular unselected SNPs, 55382
  - M_t: selected (top) SNPs, 4514
  - M_{rt}: regular and selected SNPs, 59896
- Univariate REML analyses performed using GCTA
- Fixed Effects: contemporary group
- Random Effects:
  - A: Numerator relationship matrix for genotyped animals constructed from full pedigree.
  - G_r, G_t, G_{rt}: GRM constructed using M_r, M_t, M_{rt}, respectively
  - Genetic groups fitted to account for breed structure
- All GRMs calculated according to Gurman *et al.* (2019) accounted for breed structure.
- Models compared:
  - A = pedigree only
  - $A + G_r$ = pedigree + regular SNP GRM
  - $A + G_r + G_t$ = pedigree + regular and top SNP GRMs
  - $A + G_{rt}$ = pedigree + combined GRMs

Results

- Detailed variance estimates presented for imf in Table 1 and for pwt in Table 2.
- Models with higher log-likelihoods do not translate to higher heritability (h^2)

Table 1: Variance component estimates for imf. Results include: proportions of genetic variance explained by random effects (λ_), heritability (h^2) and log-likelihoods relative to first model (LogL).

imf	λ_A	λ_{G_r}	λ_{G_r}	$\lambda_{G_{rt}}$	h^2	LogL
A	1				0.6	0
$A + G_r$	0.41	0.59			0.62	194
$A + G_r + G_t$	0.36	0.35	0.29		0.59	363
$A + G_{rt}$	0.33			0.67	0.62	267

Table 2: Variance component estimates for pwt. Results presented in same format as Table 1.

pwt	λ_A	λ_{G_r}	λ_{G_r}	$\lambda_{G_{rt}}$	h^2	Logl
A	1				0.14	0
$A + G_r$	0.09	0.91			0.24	90
$A + G_r + G_t$	0.16	0.45	0.39		0.24	156
$A + G_{rt}$	0.06			0.94	0.25	118

- Proportions of variance explained by each effect in $A + G_r + G_t$ two GRM model presented in Table 3
- Highest proportions explained by A for imf, G_r^* for cwt and pwt or G_t^* for sf5, cemd and ccfat

Table 3: Proportions of genetic variance explained for each genetic effect in $A + G_r + G_t$

	imf	sf5	cemd	ccfat	cwt	pwt
A	0.36	0.32	0.34	0.16	0.38	0.16
G_r^*	0.35	0.13	0.16	0.37	0.48	0.45
G_t^*	0.29	0.55	0.55	0.46	0.14	0.39

Discussion

- Model with two GRMs produced higher log likelihood
- Model with G_{rt} produced higher h^2 for all traits

Conclusions

- **The proportion of genetic variance explained by the relationship matrices changes by trait**
- **Models with high heritability do not equate to high log likelihood.**
- **Variance component estimates can be used for BLUP estimation**