




Estimation of heritability for dairy traits, combining pedigree with dense snp information on some animals


R.F. Veerkamp, H.A. Mulder and M.P.L. Calus



Animal Breeding & Genomics Centre





Introduction

- Genetics of feed utilisation complex: milk, fat and protein production, feed intake, BCS and liveweight
- Interesting for agroproductie 21^{ste} century:
 - Energy balance: robustness, health & fertility 
 - Feed inefficiency: greenhouse gas emissions
 - Feed efficiency: economic efficiency
- “Something new to be learned in this area?”
→ better insight using using genomics



Questions

- GWAS
- but also utilise SNP information to improve estimation of genetic parameters?
 - Replace pedigree by genomic relationships?
 - Utilise information on individual SNP effects?
- Difficulty: animals with scarce and expensive historical phenotypes but without DNA available
 - Combine genotyped and none genotyped relationships?

Data

- Dutch dataset: 639 Holstein first lactation heifers, fed *ad libitum* TMR
- Recording of milk yield, milk composition, liveweight and feed intake
→ average first 15 weeks of lactations
- Ca 580 animals DNA
→ genotyped Illumina 50k SNP panel
- Quality control checks → 517 animals left






General statistical model

- ASREML used to estimate h^2 and s.e.

$$y_{ij} = \mu + \text{fixed_effects}_j + a_i + e_{ij}$$

- Fixed effects: year-quart. (26) and age calving (3)
- a_i is the random genetic effect of the i th animal;
 $\text{Var}(a) = A\sigma_a$ or $G\sigma_a$ or $H\sigma_a$ or $WG\sigma_a$






Statistical models

- Relationship matrix:
 - A based on pedigree (639 phenotyped animals plus 3363 in pedigree)
 - G from SNPs (517 genotyped+pheno. animals)

$$G = \frac{ZZ'}{2 \sum p_i(1-p_i)} \quad (\text{VanRaden 2008})$$

- H^{-1} combine A^{-1} and G^{-1} (639 animals)
(Aguilar et al. 2010)

Statistical models



- WG: weighted genomic relationship matrix
- Step 1: estimate SNP effects BAYESC

$$y = 1_n \mu + \text{fixed_effects} + \sum_{j=1}^m (X_j (q_j v_j)) + e$$

- Step 2: weighted G

$$WG = \frac{ZDZ'}{2 \sum p_i(1-p_i)}$$



D diagonal matrix weights per SNP: squared allele substitution effects, rescaled to be 1.0 across all loci.

Results (1)



	h ²				SE				
	RM	A	A	G	H	A	A	G	H
# phenotypes	517	639	517	639	517	639	517	639	639
Milk	0.48				0.13				
Fat %	0.89				0.10				
DMI	0.83				0.11				
LW	0.50				0.13				

Differences A&G paper Verbyla et al WCGALP

Results (2)

RM	G	WG	WG	WG	WG	WG
P. Prob.	none	p>0.10	p>0.05	p>0.01	p>0.001	All
# SNP	-	3	7	116	4237	43011
h ²	0.77					
SE	0.08					

Conclusion

- Using SNP relationships improves estimation of variance components even when 517 phenotypes were used instead of 639 phenotypes
- Combining SNP and pedigree information gave best estimate
- Differences between estimates with G and A
 - Scale A and G to same base population?
 - IBD v. IBS (Janss)
 - Also an issue when creating H matrix

