Royal Veterinary College
University of London

Modelling aspects of association studies for quantitative traits in livestock

Geoff Pollott

Motivation

- Association analyses – common method for dissecting complex traits in humans
- SNP markers in livestock allow use of association analyses of production traits
- Data and DNA available from an extensively recorded group of dairy herds in South-East England
- Scaling up from few SNPs to genome-wide SNPs
- What are appropriate methods and models to use on these data?

RVC

What can we learn from human association studies?

- Population structure a confounding factor
- Cryptic relatedness
- Correction for multiple testing
- Models used very simple compared to those traditionally used in analysis of livestock data

RVC

Binary traits

- Common for disease traits
- Chi-squared approach e.g. PLINK
- 2 x 2 (allele) or 3 x 2 (genotype) table analysed
- Mixed model used if ‘fixed’ effects thought to affect incidence of condition e.g. ASReml
- Example: Calf mortality found to be affected by candidate SNPs in the leptin gene (Brickell *et al.*; JDS, 2010)

RVC

Quantitative traits

- Models in ‘human’ packages too simple
- Mixed model methodology applicable using packages like ASReml – animal model + SNP effect
- Results for early growth, IGF-1 and fertility in heifers reported by Clempson *et al.* (2010a and b) at BSAS and WCGALP for candidate SNP in Leptin and mitochondrial genes
- Is mixed-model approach (e.g. with ASReml) suitable for genome-wide association studies (GWAS)?

RVC

Issues with GWAS of QT

RVC

Appropriate computing strategy

- 1) ASReml in R or other appropriate programming framework - restart from previous SNP solution

- 2) Three-step approach
- Step 1 – fit full model without the SNP effect and save residuals
- Step 2 – fit SNP effect to residuals one at a time
- GRAMMAR (Aulchenko et al, 2007) (ASReml plus PLINK another option)
- Step 3 – reanalyse significant SNP with full model in ASReml

Use of 7-SNP windows to find true SNP

RVC

Using heifer daily milk production (kg) as an example with the A59V leptin SNP

	SNP effect P value	Genotype effect CC	CT	TT
Mixed model in ASReml	**0.035**	**26.3^{a}**	**26.2^{a}**	**23.0^{b}**
Analysis of residuals	**0.041**	**22.1^{a}**	**21.8^{a}**	**18.2^{b}**
Inclusion of more recorded heifers without SNP data*	**0.021**	**22.2^{a}**	**21.6^{b}**	**18.0^{c}**

* 5 more SNPs (out of 40) became significant with this method

RVC

Substitution effect or genotype effect SNP model using SNP A59V

	SNP effect P value	Genotype effect CC	CT	TT
Milk/d (kg)	**0.035**	**26.3^{a}**	**26.2^{a}**	**23.0^{b}**
Crown rump length - 15mo (cm)	**0.034**	**171.5^{a}**	**174.0^{b}**	**169.6^{c}**

RVC

Multiple testing correction

- Bonferroni correction – too conservative – Type 1 errors
- Genome-wide and chromosome wide P values
- Q value and FDR approach
- Permutation testing

Unresolved issue

RVC

Testing interactions

- Mixed model or analysis of residuals approach allows for testing SNP x SNP interactions
- Fit model with 2 SNPs plus interaction
- Significant interaction indicates possibility of epistasis

RVC

An example interaction – height at 220d (cm)

	Main effect	Genotype SNP A AA	Aa	aa
Main effect		**107.0[a]**	**107.3[a]**	**108.1[b]**
Genotype SNP B - BB	**107.8[a]**	**106.6**	**107.0**	**107.9**
Bb	**107.6[a]**	**107.2**	**107.7**	**105.7**
bb	**108.1[b]**	**107.2**	**106.3**	**110.0**

RVC

Conclusions

- Models and computing platforms for GWAS available
- Correction for multiple testing an issue
- Testing for epistasis tedious but potentially interesting

RVC

Acknowledgements

Prof. Claire Wathes
Nicola Bourne
Jessica Brickell
Andrew Clempson

RVC

Thank you for your attention

RVC