SABRE

CUTTING EDGE GENOMICS FOR SUSTAINABLE ANIMAL BREEDING

Prediction of haplotypes with missing genotypes and its effect on marker-assisted breeding value estimation

Han Mulder, Mario Calus and Roel Veerkamp

Animal Breeding and Genomics Centre

herman.mulder@wur.nl

Session 33, presentation 6, abstract 3584

ANIMAL SCIENCES GROUP
WAGENINGEN UR

4th SABRE Conference / 60th Annual Meeting EAAP, Barcelona, 24-27 August 2009

Food Quality and Safety

Acknowledgements

- CRV, Hendrix Genetics and IPG for financial support.
- Egbert Knol, Dieuwke Roelofs-Prins, Marc Rutten, Chris Schrooten, Addie Vereijken, Martin Lidauer, Ismo Stranden and Robin Thompson are thanked for helpful discussions about this study.
- MiXBLUP software commercially available at the end of 2009 (John.Voskamp@wur.nl)
- These results are part of the SABRE research project that has been co-financed by the European Commission, within the 6th Framework Programme, contract No. FOOD-CT-2006-016250.

Introduction

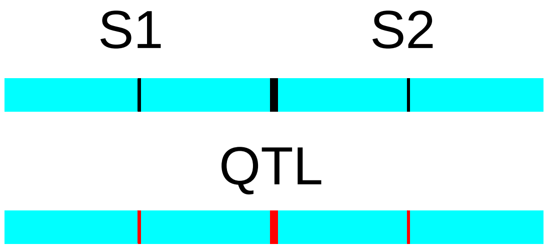
- Some genes or some QTL regions are finemapped
  - Examples genes: DGAT1 in dairy cattle, IGF2 in pigs
- Breeding value estimation
  - Classically without gene/QTL effects
  - Marker/gene-assisted breeding value estimation on limited scale
- Problem: only part of a population is genotyped, e.g. males/sires

- Solution: prediction of genotypes or haplotypes, e.g. with efficient BLUP-method (Gengler et al., 2007, 2008)

Objective

- To develop a method to predict haplotypes of animals that are not genotyped using mixed model equations by extending Gengler’s method
- To investigate the use of these predicted haplotypes in marker-assisted breeding value estimation

Definition of haplotypes



Two SNP markers that flank a QTL

			Number of haplotype copies (*nhc*)			
Animal	Haplotype 1	Haplotype 2	HAP1 (11)	HAP2 (12)	HAP3 (21)	HAP4 (22)
1	11	11	2	0	0	0
2	11	12	1	1	0	0
3	11	21	1	0	1	0
4	11	22	1	0	0	1

Prediction of haplotypes

- Model for one haplotype

$$nhc_i = \mu_{nhc_i} + d_i + e_{nhc_i}$$

$$\begin{bmatrix} \mathbf{1'1} & \mathbf{1'M} \\ \mathbf{M'1} & \mathbf{M'M} + \lambda \mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \mu_{nhc} \\ \mathbf{d}_y \\ \mathbf{d}_x \end{bmatrix} = \begin{bmatrix} \mathbf{1'nhc_y} \\ \mathbf{M'nhc_y} \end{bmatrix}$$

- d_y = EBV for nhc for genotyped animal
- d_x = EBV for nhc for not genotyped animal
- $\lambda = \sigma^2_{e_{nhc}} / \sigma^2_{a_{nhc}} = 0.01 / 0.99$
  - Allows for genotyping errors, haplotyping errors and recombination

Marker-assisted BLUP

- Random regression on predicted number of haplotype copies:

$$y = \mu + u_{pol} + \sum_{i}^{n} \left((\mu_{nhc_i} + d_i) \times b_i \right) + e$$

$$\sigma^2_{b_i} = 0.5\sigma^2_{A_{qtl}} \qquad \sigma^2_{u_{pol}} = \sigma^2_{A_{pol}}$$

- Haplotypes based on:
  - Neighboring marker (nm)
  - 2 flanking markers (hap2)
  - 4 markers, 2 on each side of QTL (hap4)

Simulation

- Simulation
  - Bi-allelic QTL with only additive effects
  - Polygenic additive genetic effect
  - 1, 2 or 4 polymorphic flanking SNP markers
  - 100 generations to get LD between markers and QTL
  - 5 generations of data: 50 sires and 250 dams producing 2000 offspring (10,000 total).
  - Sires in generation 1-4 and males in generation 5 are genotyped
  - Distance between markers and heritability are varied

Results: accuracy of predicted haplotypes

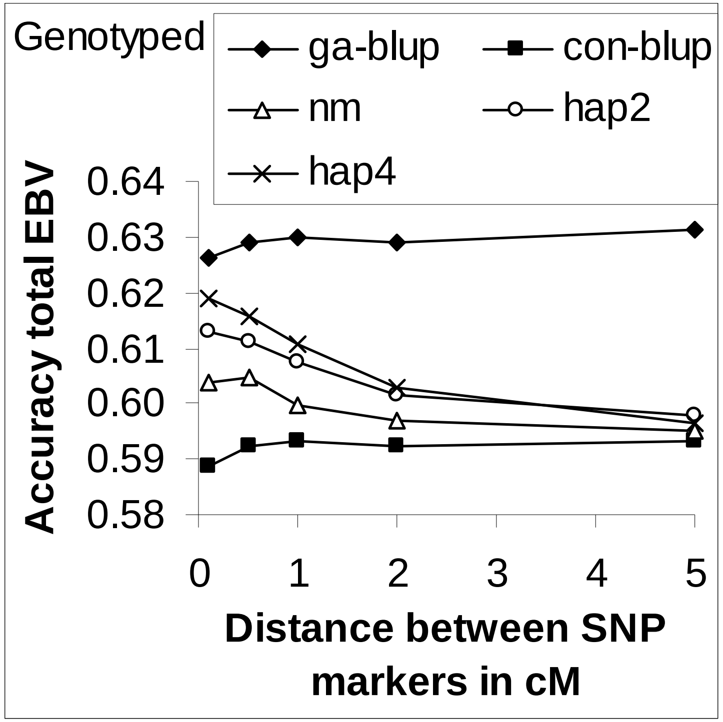
Marker-distance is 0.1 cM

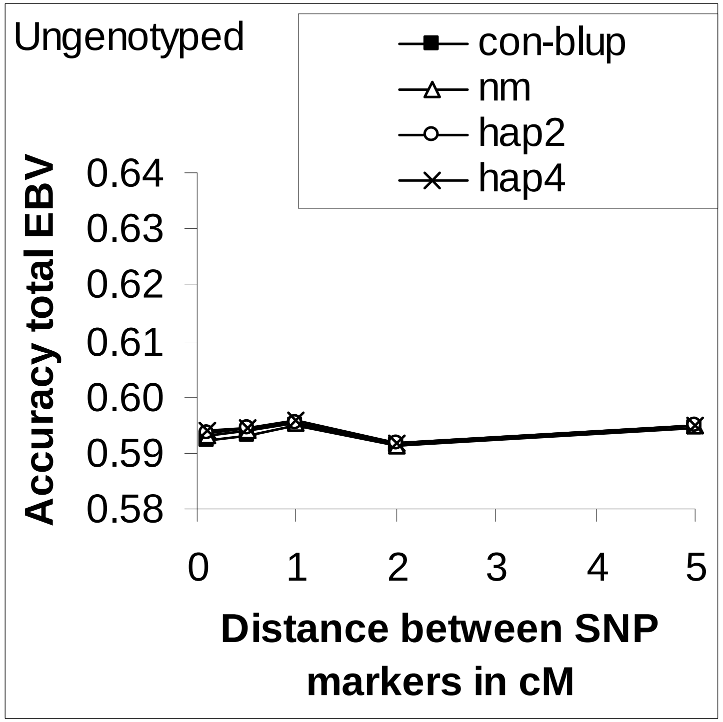
Haplotype	Correlation true and estimated nhc
nm	0.64
hap2	0.63
hap4	0.59

Conclusion: Accuracy nhc prediction reduced for hap4

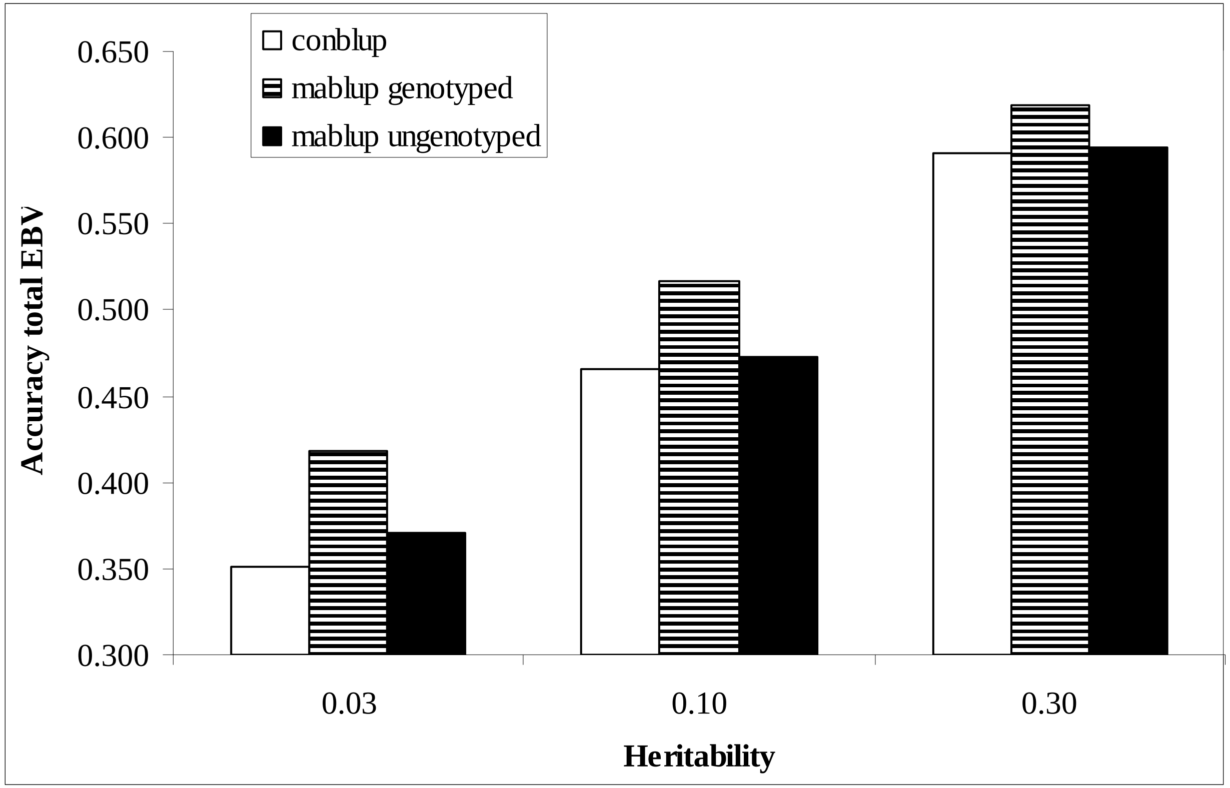
Results: accuracy of total-EBV

Males are genotyped; females are not genotyped;
Heritability = 0.3; QTL-variance =0.15





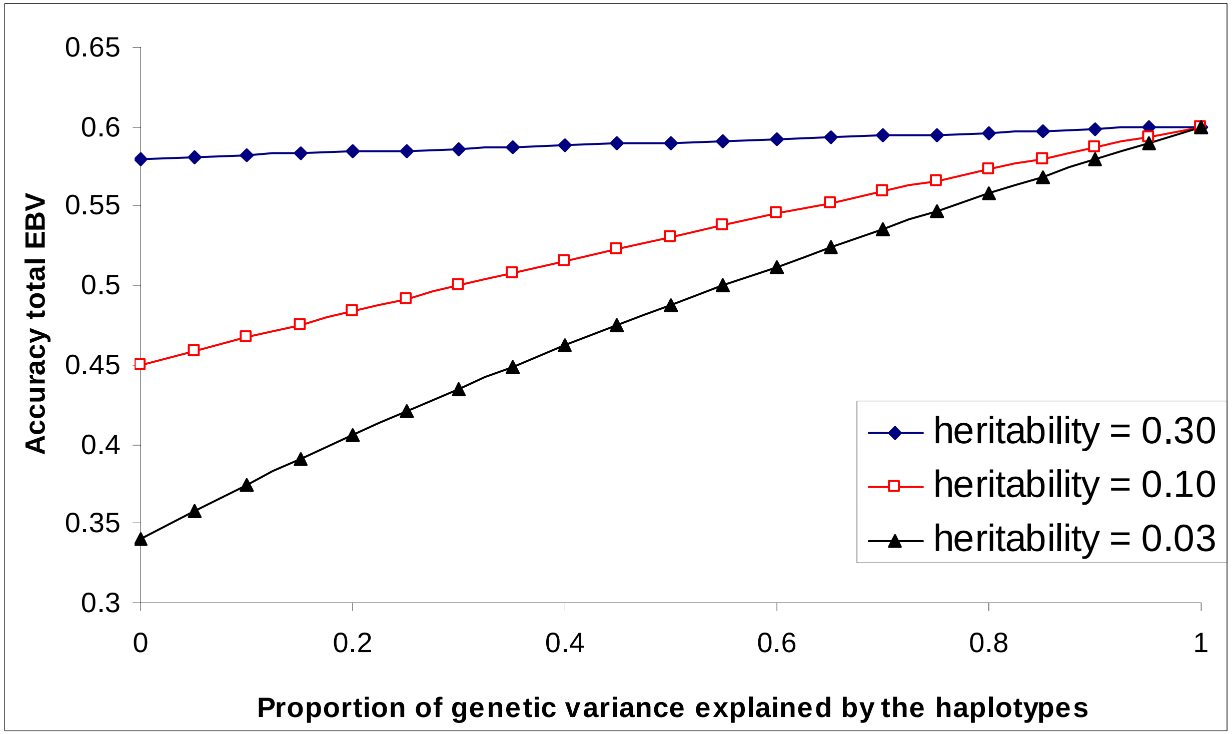
Results: effect of heritability



Conclusion: Ungenotyped animals benefit from marker-assisted BLUP for low heritable traits

Translation to genomic selection

$$r_{totalEBV} = \sqrt{(1-q^2)r^2_{A_{pol}} + q^2 r^2_h}$$



Conclusion

- Prediction of haplotypes possible within mixed model framework
- Haplotypes consisting of 4 markers are best
- Method works best with dense markers, i.e. <1 cM between markers
- Increase in accuracy mainly for genotyped animals

- For traits with low heritability also ungenotyped animals benefit from marker-assisted or genomic breeding value estimation using predicted haplotypes.

Thank you for your attention!

Conclusion

- Prediction of haplotypes possible within mixed model framework
- Haplotypes consisting of 4 markers are best
- Method works best with dense markers, i.e. <1 cM between markers
- Increase in accuracy mainly for genotyped animals

- For traits with low heritability also ungenotyped animals benefit from marker-assisted or genomic breeding value estimation predicted haplotypes.