Genomic prediction using haplotypes in Brown Swiss

Franz R. Seefried
2016 Interbull Meeting
Puerto Varas, Chile
Prediction based on haplotypes

Opportunities:
- Reduction of variables
- Increased accuracy
- Higher ability to identify mutations

Challenges – Questions:
- Definition haplotype:
  - Phasing – which algorithm
  - Length (absolute based on LD)
  - Pseudo-SNPs or multiallelic
- Haplotype labelling
Haplotype labelling

1. Phasing genotypes

a. Unphased Data (Genotype)
- AA
- AB
- BB
- AA
- AB
- AB
- AA

b. Phased Data (Haplotypes)
- A
- A
- B
- B
- A
- B
- A
- A

Figure modified from: http://www.chromosomechronicles.com/2009/09/30/use-family-snp-data-to-phase-your-own-genome/
Haplotypes labelling

1. Phasing genotypes
2. Define length of haplotype blocks

b. Phased Data (Haplotypes)
Haplotype labelling

1. Phasing genotypes
2. Define length of haplotype blocks
3. Label haplotype alleles

b. Phased Data (Haplotypes)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>B</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Hap_All1_1</th>
<th>Hap_All1_2</th>
<th>Hap_All2_1</th>
<th>Hap_All2_2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alleles</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure modified from: http://www.chromosomechronicles.com/2009/09/30/use-family-snp-data-to-phase-your-own-genome/
# Animals and traits

<table>
<thead>
<tr>
<th></th>
<th>Non-return rate 56 heifer (NRH)</th>
<th>Somatic cell score (SCS)</th>
<th>Stature (STA)</th>
<th>Protein yield (PY)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
<td>2,018</td>
<td>4,786</td>
<td>5,294</td>
<td>5,128</td>
</tr>
<tr>
<td>Validation</td>
<td>240</td>
<td>560</td>
<td>596</td>
<td>596</td>
</tr>
</tbody>
</table>
Phenotypes

- Minimal reliability
- Deregressed BV (Garrick et al., 2009)
- Youngest bulls as validation
- Prediction accuracy: \( \text{cor(deregBV, GEBV)} \)
Methods

• 50K genotype data (40,636 SNPs)
• Phasing: Beagle
• Haplotype labelling: Ghap (R-package)
• Prediction: GenSel:
  – BayesCpi
  – BayesC

\[ y = 1' + X + e \]

\[ = 0 \text{ with probability } \]

\[ ^i \sim N(0, \frac{s_{SNP}^2}{SNP}) \text{ with probability } (1 - p) \]
Results

- Haplotype definitions:
  - 1 Mb
  - 250 kb
  - LD blocks from plink
  - LD blocks including single SNP
  - Haplotype alleles with a frequency of less than 0.1% were excluded
Results

Numbers of variables

- 1M
- 250kb
- plink
- plink + SNP
- 50k SNPs

Frequency of effect allele

- 1M
- 250kb
- plink
- plink + SNP
- 50k SNPs

Estimates for Pi

- PY
- STA
- SCS
- NRH
Accuracy genomic prediction

Accuracy

PY      STA      NRH      SCS

1Mb  250kb  Plink  Plink+single_SNP  GenSel 50K SNPs
Conclusions

• Reduction of variables: No reduction with 50K data
• Little increase in accuracy depending on trait
• Accuracy by trait: Different haplotype labelling method for highest accuracy
Acknowledgement

B. Bapst, B. Gredler, J. Moll, U. Schnyder

C. Flury, H. Signer-Hasler

D. Garrick

C. Stricker

F. Schmitz-Hsu
Acknowledgement

Arbeitsgemeinschaft der Österreichischen Braunviehzüchter

Arbeitsgemeinschaft Deutsches Braunvieh

Associazione nazionale allevatori bovini della razza Bruna, Italy

Brown Swiss Association of the US

Brune Genetique Services, France

Braunvieh Schweiz

Zveza rejcev govedi rjave pasme Slovenije

TUM München (R. Fries)

Tierzuchtforschung e.V. (I. Russ)

BOKU Wien (J. Sölkner)

Uni Milano (A. Bagnato)

FiBL (A. Bieber)
Thank you for your attention