Impact of imputing markers from a low density chip on the reliability of genomic breeding values in Holstein populations

Romain Dassonneville
INRA GABI G²B
Institut de l’élevage

Rasmus Brøndum
Aarhus University
Faculty of Science and Technology

F. Guillaume,
V. Ducrocq,
S. Fritz, UNCEIA

T. Druet,
Univ Liège

B. Guldbrandtsen,
M. Lund,
G. Su
Study of the Illumina 3K chip

EuroGenomics collaboration

Objectives:

- Measure imputation error rate
- Study impact on GEBV reliability
- Study influence of reference population size

Previous studies → 3K custom in silico chip

Our study → commercially available 3K chip

Zhang and Druet, 2010
Weigel et al, 2010
Data

Reference population

training

validation

54K

54K

masking

3K

Imputation

comparison

imputed 54K
### Number of animals and number of markers used

*Including bulls with partially reconstructed genotypes*

<table>
<thead>
<tr>
<th>National</th>
<th>EuroGenomics</th>
<th>No. of Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training</td>
<td>Validation</td>
</tr>
<tr>
<td>Nordic</td>
<td>3,058</td>
<td>1086</td>
</tr>
<tr>
<td>France</td>
<td>3,071/3,505*</td>
<td>966</td>
</tr>
</tbody>
</table>

*Including bulls with partially reconstructed genotypes*
Imputation Method

• Combination of the DAGPHASE 1.1 and Beagle 2.1.3 software

   - Pedigree based family information (segregation rules)
   - Population Linkage Disequilibrium

   *Druet et Georges, 2010*
   *Browning and Browning, 2007*
   *PHASEBOOK package*
Results

Imputation accuracy

<table>
<thead>
<tr>
<th>Reference Population</th>
<th>National</th>
<th>Eurogenomics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sire not in ref.</td>
<td>94,5</td>
<td>96,1</td>
</tr>
<tr>
<td>Sire in ref.</td>
<td>96</td>
<td>97,9</td>
</tr>
</tbody>
</table>

Influence of genotyped ancestors

<table>
<thead>
<tr>
<th>Nordic Population</th>
<th>Sire not in ref.</th>
<th>Sire in ref.</th>
<th>Sire and Maternal grandsire in ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>91,7</td>
<td>95,5</td>
<td>95,7</td>
</tr>
</tbody>
</table>

France: more 3K markers, more genotyped ancestors
Data for genomic evaluation

Traits studied:

- protein yield ($h^2=0.3-0.39$)
- somatic cell count (SCC, $h^2=0.15$)
- fertility (Non Return Rate or Conception rate, $h^2=0.02$)
- udder depth ($h^2=0.36-0.37$)
Genomic evaluation model

Nordic:

- GBLUP

\[ y = 1 \mu + Zg + e \]

VanRaden, 2008

French:

- GMAS, QTL-BLUP combines LDLA and EN

\[ y = 1 \mu + Zu + \sum_{i=1}^{nQTL} (h_{i1} + h_{i2}) + e \]

Boichard et al., WCGALP 2010

u: polygenic effect based on pedigree

LDLA : Linkage Disequilibrium
Linkage Analysis
EN : Elastic Net
Results – Reliability of DGV

Reliability = squared correlation (DGV, deregressed proofs) for validation population
Mean over the 4 traits
Results – Reliability of GEBV

Reliability = squared correlation (GEBV, DYD) for validation population
Mean over the 4 traits
Conclusion - discussion

- Imputation accuracy = 2.5-5% = close to litterature
  - Commercially available 3K chip contains less markers after editing
  - Bigger reference population size
  - Beagle and DAGphase: efficient imputation softwares

- Reliability of GEBV:
  - Reliability of GEBV based on imputed genotypes slightly lower
  - Correlations (GEBV-50K, GEBV-3Kimp) are high (0.91-0.97)

- Low density chip imputed to 50 K:
  - Feasible alternative for pre-selection of young animals
  - Attractive tool for a large screening of the female population
If you want some more details

Article published in Journal of Dairy Science, July 2011

Effect of imputing markers from a low-density chip on the reliability of genomic breeding values in Holstein populations

R. Dassonneville,*†1,2 R. F. Brøndum,‡1,2 T. Druet,§ S. Fritz,# F. Guillaume,*† B. Guldbrandtsen,‡ M. S. Lund,‡ V. Ducrocq,* and G. Su‡
*INRA, UMR1313 Génétique Animale et Biologie Intégrative (GABI), 78350 Jouy-en-Josas, France
†Institut de l’Élevage, 149 rue de Bercy, 75595 Paris, France
‡Aarhus University, Faculty of Science and Technology, Department of Genetics and Biotechnology, DK-8830 Tjøle, Denmark
§Unit of Animal Genomics, GIGA-Research and Department of Animal Production, Faculty of Veterinary Medicine, University of Liège, B-4000 Liège, Belgium
#Union Nationale des Coopératives d’Élevage et d’Insémination Animale (UNCEIA), 149 rue de Bercy, 75595 Paris, France
Impact of imputing markers from a low density chip on the reliability of genomic breeding values in Holstein populations

Romain Dassonneville
INRA GABI G²B
Institut de l'élevage

Rasmus Brøndum
Aarhus University
Faculty of Science and Technology

F. Guillaume,
V. Ducrocq,
S. Fritz, UNCEIA

T. Druet,
Univ Liège

B. Guldbrauntsen,
M. Lund,
G. Su

Imputation accuracy

% alleles correctly imputed

Nordic  French

National  Eurogenomics

reliability

0,41  0,38
0,54  0,48

50 K  3K imp  50 K  3K imp

National ref  Euro ref

questions?