Variation in genome sharing among non-inbred pigs

N.S. Forneris¹, J.P. Steibel³, C.W. Ernst³, R.O. Bates³, J.L. Gualdrón Duarte¹, R.J.C. Cantet¹,²*

Prediction of BVs with a genomic relationship matrix depends on how precise the genome sharing of relatives with the same pedigree relationship is measured.

The realized values of genome sharing deviate due to Mendelian sampling and linkage.

Two approaches: IBD and eIBS methods
Objectives

1. To estimate the empirical variation in genome sharing of relatives from a pig population, using either identity-by-descent (IBD) or identity-by-state (IBS) based estimators.

2. To compare the estimated values of the variances against their theoretical values (Hill and Weir, 2011) for different pedigree relationships.
- 411 pigs
- Genotypes: Porcine SNP60 Beadchip.
- 84254 pairwise relationships

Data

4 F₀ Duroc ♂  15 F₀ Pietrain ♀

4 F₀ Duroc ♂  15 F₀ Pietrain ♀

56 F₁ (50 ♂ x 6 ♀)

60K chip

336 F₂

http://www.nlldb.gov.lk/animals/pig%20-%20Duroc.jpg,
http://www.hesbayebrabanonne.be/IMG/jpg/cochon.jpg,
http://www.illumina.com/products/porcinesnp60_dna_analysis_kit.ilmn
**Unilineal relatives**

*Lineal descendants:*

**PO:** Parent-offspring (AK)

**GG:** Grandparent-grandoffspring (AS)

*Half sibs and their descendants:*

**HS:** Half sibs (OP)

**HUN:** Half uncle-nephew (QZ)

**HFC:** Half cousins (XZ)

*Descendants of full sibs:*

**UN:** Uncle-nephew (KT)

**FC:** First Cousins (SV)

**Bilineal relatives**

**FS:** Full sibs (KL)

**DHFC:** Double half first cousins (XJ)

**3FC:** Three-way cousins (iiE)

**DFC:** Double first cousins (ST)

**EHS:** Half sibs, mothers(fathers)HS

**TQS:** Half sibs, mothers(fathers) FS

(SX)
Genomic relationship estimates

IBD-based

\[
\hat{G}_{xy}^{(IBD)} = \frac{1}{L} \sum_{k=1}^{K} l_k \left[ \frac{1}{2} P(\text{IBD}_k = 1 \mid \mathcal{M}) + P(\text{IBD}_k = 2 \mid \mathcal{M}) \right]
\]

- length of each segment delimited by 2 SNP
- posterior probabilities of IBD sharing

IBS-based

\[
\hat{G}_{xy}^{(IBS)} = \left( \frac{Z Z'}{\sum_{k=1}^{K} 2p_k q_k} \right)_{xy}
\]

- HMM approach
- accounts for pedigree and LD between SNP

- Theoretical variance: formulae from Hill & Weir (2011) for non inbred individuals based on genetic maps.
Results
### Correlations among methods

**Probability > | r | under H<sub>0</sub> : ρ = 0**

<table>
<thead>
<tr>
<th></th>
<th>gIBD</th>
<th>gIBS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>relationship</td>
<td><strong>0.875</strong></td>
<td><strong>0.797</strong></td>
</tr>
<tr>
<td><em>(a&lt;sub&gt;ij&lt;/sub&gt;)</em></td>
<td><strong>P &lt; 0.0001</strong></td>
<td><strong>P &lt; 0.0001</strong></td>
</tr>
<tr>
<td>gIBD</td>
<td><strong>0.721</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>P &lt; 0.0001</strong></td>
<td></td>
</tr>
</tbody>
</table>
## Standard Deviations of actual relationships

<table>
<thead>
<tr>
<th>Relationship</th>
<th>$R$</th>
<th>$N$</th>
<th>Theoretical</th>
<th>IBD</th>
<th>IBS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PO</strong></td>
<td>0.500</td>
<td>784</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0573</td>
</tr>
<tr>
<td><strong>GG</strong></td>
<td>0.250</td>
<td>1344</td>
<td>0.0437</td>
<td>0.0456</td>
<td>0.0993</td>
</tr>
<tr>
<td><strong>HS</strong></td>
<td>0.250</td>
<td>7061</td>
<td>0.0352</td>
<td>0.0353</td>
<td>0.0609</td>
</tr>
<tr>
<td><strong>HFC</strong></td>
<td>0.062</td>
<td>22944</td>
<td>0.0235</td>
<td>0.0351</td>
<td>0.0495</td>
</tr>
<tr>
<td><strong>FS</strong></td>
<td>0.500</td>
<td>639</td>
<td>0.0498</td>
<td>0.0577</td>
<td>0.0826</td>
</tr>
<tr>
<td><strong>TQS</strong></td>
<td>0.375</td>
<td>816</td>
<td>0.0449</td>
<td>0.0480</td>
<td>0.0711</td>
</tr>
<tr>
<td><strong>DFC</strong></td>
<td>0.250</td>
<td>544</td>
<td>0.0395</td>
<td>0.0478</td>
<td>0.0581</td>
</tr>
<tr>
<td><strong>DHFC</strong></td>
<td>0.125</td>
<td>5408</td>
<td>0.0332</td>
<td>0.0450</td>
<td>0.0504</td>
</tr>
</tbody>
</table>
Empirical distribution for estimated actual relationships

As expected from theory, for IBD-based estimates the distributions are + skewed as $R$ becomes smaller.
The overlap in the amount of sharing of quite different pedigree relationship classes was higher for IBS estimates.
Correlations with true relationships
Simulation results

<table>
<thead>
<tr>
<th>True relationship</th>
<th>gIBD</th>
<th>gIBS</th>
<th>gIBS &quot;Tunning&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.79</td>
<td>0.0036</td>
<td>0.65</td>
<td>0.0049</td>
</tr>
<tr>
<td>0.0036</td>
<td>0.67</td>
<td>0.0046</td>
<td></td>
</tr>
</tbody>
</table>
Final comments

- The SD(gIBD) was always smaller than the SD(gIBS), being on average, 18.5% and 70.7% higher respectively, when compared to the theoretical SD.

- Results suggest that the IBD-based method can detect more accurately the degree of genome sharing between relatives and could be used to compute realized relationships for predicting BVs with genomic information.

- gIBD was more strongly correlated to true relationship than gIBS with simulated data.
Thank you!
As expected from theory, as animals become less related, the SD becomes smaller whereas the CV becomes larger.

CV(IBD) is always closer to the theoretical value than CV(IBS).
IBD-based estimated SD and map length

IBD-based estimated SD behaves as its theoretical value: it decreases with increased chromosome map length (Figure A).

For relationships with the same mean, the estimated SD declines less rapidly with map length for lineal descendants than for those involving half sibs, showing the fastest decline for descendants of full sibs (B).