Genetic connectedness in the U.S. sheep industry


69th Annual Meeting of EAAP
Dubrovnik, Croatia
Aug. 28, 2018
Grateful

- For the invitation (Joanne Conington)
- To ASAS and EAAP for financial support
Today’s talk

Will not include discussion of:

- Theoretical considerations of connectedness statistics
- Genomic selection (at least not directly)
- U.S. politics
Today’s talk

Will include discussion of:
- My perspective on connectedness
- Connectedness among flocks within the U.S. sheep industry
- Potential to strengthen connectedness with genomic relatedness
- Use of connectedness to target sheep for genotyping
My perspective

Sire model:

\[ y = Xf + Zs + e \]
My perspective

Sire model:
\[ y = Xf + Zs + e \]
My perspective

- $f$ and $s$ fixed
  - $f_1 - f_2$ not estimable
  - $s_1 - s_2$ not estimable
- $f$ fixed; $s$ random
  - $f_1 - f_2$ estimable
  - $s_1 - s_2$ estimable

Sire model:

$$y = Xf + Zs + e$$
My perspective

- Comparison of $s$ possible with caveats
  - Prediction error variance increases
  - Presumes average merit of $s$ equal in each $f$
    - Biased if otherwise

Sire model:
$$y = Xf + Zs + e$$

Animal model:
$$y = Xf + Za + e$$

Disconnected

Comparison of $s$ possible with caveats

➢ Prediction error variance increases
➢ Presumes average merit of $s$ equal in each $f$
❖ Biased if otherwise
Prediction error variance

“...connectedness itself is not the major issue per se.”

“... how it contributes to prediction error variance (PEV) ... is”

Thus, PEV “... is a logical measure of the quality of connectedness.”

(Kennedy and Trus, 1993)
Prediction error variance

- With selection, wish to compare (estimated) breeding values
  \[ \hat{a}_i - \hat{a}_j \]

- Connectedness based on PEV of that difference

\[
\begin{align*}
  d_{ij} &= \text{PEV}(\hat{a}_i - \hat{a}_j) \\
         &= \text{PEV}(\hat{a}_i) + \text{PEV}(\hat{a}_j) - 2\text{PEC}(\hat{a}_i, \hat{a}_j) \\
         &= (C_{ii}^{22} + C_{jj}^{22} - 2C_{ij}^{22}) \sigma_e^2
\end{align*}
\]
Connectedness statistics

- PEV of difference
  
  \[ d_{ij} = \left( C_{ii}^{22} + C_{jj}^{22} - 2C_{ij}^{22} \right) \sigma_e^2 \]

- Coefficient of determination (Laloë, 1993)
  
  \[ CD_{ij} = 1 - \lambda \frac{C_{ii}^{22} + C_{jj}^{22} - 2C_{ij}^{22}}{K_{ii} + K_{jj} - 2K_{ij}} \]
  
  where \( K = A \) or \( G \)

Accounts for reduced variability in true breeding values due to relationships
Connectedness statistics

- **PEV of difference**
  \[ d_{ij} = (C_{ii}^{22} + C_{jj}^{22} - 2C_{ij}^{22})\sigma_e^2 \]

- **Coefficient of determination** (Laloë, 1993)
  \[ CD_{ij} = 1 - \lambda \frac{C_{ii}^{22} + C_{jj}^{22} - 2C_{ij}^{22}}{K_{ii} + K_{jj} - 2K_{ij}} \]

- **Correlation** (Lewis et al., 1999)
  \[ r_{ij} = \frac{C_{ij}^{22}}{\sqrt{C_{ii}^{22} \times C_{jj}^{22}}} \]
U.S. sheep industry

- National Sheep Improvement Program
  - Since 1987, provided genetic evaluation services in U.S. sheep
  - Since 2010, partnership with MLA for genetic evaluation via LAMBPLAN
U.S. sheep industry

- Dynamic
  - Typically small flocks
  - Spread across wide geography

Suffolk breeders in NSIP
U.S. sheep industry

- **Pedigree**
  - **Suffolk**
    - Since 1960
  - **Targhee**
    - Since 1977
U.S. sheep industry

- Pedigree
  - Suffolk
    - Since 1960
  - Targhee
    - Since 1977

- Weights
  - Suffolk
    - Weaning since 1983
  - Targhee
    - Post-weaning since 1983
# Pedigree

<table>
<thead>
<tr>
<th>Variable</th>
<th>Suffolk</th>
<th>Targhee</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. animals</td>
<td>55,599</td>
<td>75,155</td>
</tr>
<tr>
<td>No. sires</td>
<td>2,989</td>
<td>1,294</td>
</tr>
<tr>
<td>No. dams</td>
<td>13,534</td>
<td>17,465</td>
</tr>
<tr>
<td>No. flocks</td>
<td>105</td>
<td>56</td>
</tr>
<tr>
<td><strong>Sire family size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>16.7</td>
<td>54.3</td>
</tr>
<tr>
<td>Largest</td>
<td>355</td>
<td>386</td>
</tr>
</tbody>
</table>
## Weights

<table>
<thead>
<tr>
<th>Variable</th>
<th>Suffolk ‡</th>
<th>Targhee §</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. lambs</td>
<td>36,409</td>
<td>48,014</td>
</tr>
<tr>
<td>No. flocks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>101</td>
<td>41</td>
</tr>
<tr>
<td>Active</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (day)</td>
<td>63.0</td>
<td>130.1</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>30.7</td>
<td>36.6</td>
</tr>
</tbody>
</table>

‡ At weaning: heritability of 0.15
§ At post-weaning: heritability of 0.10
Targhee

Correlation

Flock
Genomic relatedness

- Provides relationships between distant individuals that appear disconnected according to the available pedigree information
- Captures Mendelian sampling that is not present in pedigree relationships (Hill and Weir 2011)
Genomic relatedness

Cattle data

- 1,929 from base to 5th generation (Wimmer et al., 2015)
- 500 with ≈7k SNP

Clustered into 8 dissimilar groups on A

Simulated management units

-Disconnected: 8 units
-Indirectly connected: 2 units; 1/3 across
-Connected: 2 units; equally allocated

(Yu et al., 2017)
Genomic relatedness

PEV of difference

Disconnected
Indirectly connected
Connected

Smaller “better”

$h^2 = 0.2$
Genomic relatedness

Coefficient of determination

\[ h^2 = 0.2 \]
Genomic relatedness

Coefficient of determination

\[ h^2 = 0.2 \]
## Genomic relatedness

- Provides relationships between distant individuals
- Captures Mendelian sampling

<table>
<thead>
<tr>
<th>Panel</th>
<th>No. tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parentage</td>
<td>3,185</td>
</tr>
<tr>
<td>LD (15k)</td>
<td>387</td>
</tr>
<tr>
<td>50k</td>
<td>968</td>
</tr>
<tr>
<td>HD (600K)</td>
<td>1,100</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5,640</strong></td>
</tr>
</tbody>
</table>

U.S. sheep genotyping
Targeting genotyping

- GWAS for fecal egg count at weaning (WFEC) in Katahdin sheep
Targeting genotyping

- GWAS for fecal egg count at weaning (WFEC) in Katahdin sheep

- Criteria
  - Recorded in NSIP and had DNA sample
  - Accuracy WFEC EBV ≥ 50%

- Selected 24 high and 24 low WFEC EBV animals from 1,111 that met criteria
  - Within upper or lower 10% on WFEC EBV
Targeting genotyping

Correlation PCA (pedigree)

Low EBV WFEC category
Targeting genotyping

Correlation PCA (pedigree)

Low EBV WFEC category

Axis 1

Axis 2

Progeny Ram 57489

Ram 57489

Sires genotyped
Lambs
Targeting genotyping

Correlation PCA (pedigree)

Low EBV WFEC category
Q-Q plot

WFEC GWAS
Summing up

- Connectedness varies in U.S. sheep breeds
  - Impacts extent of attention needed
- Genomic relatedness improves connectedness
  - But currently little genotyping in U.S. sheep industry
- Still, pedigree-based connectedness can assist in targeting sheep for genotyping
Thanks for listening and to ...