Genetic connectedness in the U.S. sheep industry

R.M. Lewis, L.A. Kuehn, H. Yu,
G. Morota & M.L. Spangler

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

s_1 s_2
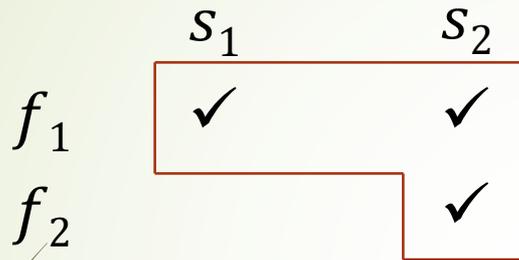
f_1 ✓   ✓

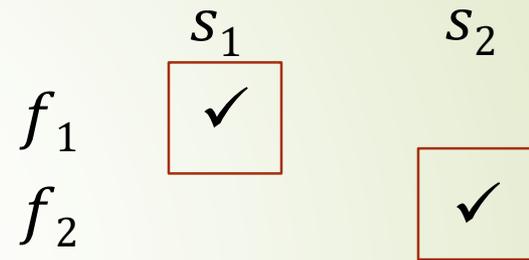f_2 ✓

Connected

Sire model:

$$y = Xf + Zs + e$$

- f and s fixed
  - $f_1 - f_2$ estimable
  - $s_1 - s_2$ estimable
- f fixed; s random
  - $f_1 - f_2$ estimable
  - $s_1 - s_2$ estimable

My perspective

s_1 s_2

f_1 ✓ ✓

f_2 ✓

Connected

s_1 s_2

f_1 ✓

f_2 ✓

Disconnected

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:

$$y = Xf + Zs + e$$

	s_1	s_2
f_1	✓	
f_2		✓

Disconnected

My perspective

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

$$f_1 \quad \boxed{\checkmark}$$
$$f_2 \qquad\qquad \boxed{\checkmark}$$

$$s_1 \qquad\qquad s_2$$

Disconnected

Sire model:

$$y = Xf + Zs + e$$

Animal model:

$$y = Xf + Za + e$$

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

$$d_{ij} = PEV(\hat{a}_i - \hat{a}_j)$$
$$= PEV(\hat{a}_i) + PEV(\hat{a}_j) - 2\text{PEC}(\hat{a}_i, \hat{a}_j)$$
$$= \left(C_{ii}^{22} + C_{jj}^{22} - 2C_{ij}^{22}\right)\sigma_e^2$$

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)

$$CD_{ij} = 1 - \lambda \frac{C_{ii}^{22} + C_{jj}^{22} - 2C_{ij}^{22}}{\boxed{K_{ii} + K_{jj} - 2K_{ij}}} \quad \text{where} \quad K = A \text{ or } G$$

Accounts for reduced variability in true breeding values due to relationships

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)

$$CD_{ij} = 1 - \lambda \frac{C_{ii}^{22} + C_{jj}^{22} - 2C_{ij}^{22}}{K_{ii} + K_{jj} - 2K_{ij}}$$

- Correlation (Lewis et al., 1999)

$$r_{ij} = \frac{C_{ij}^{22}}{\sqrt{C_{ii}^{22} \times C_{jj}^{22}}}$$
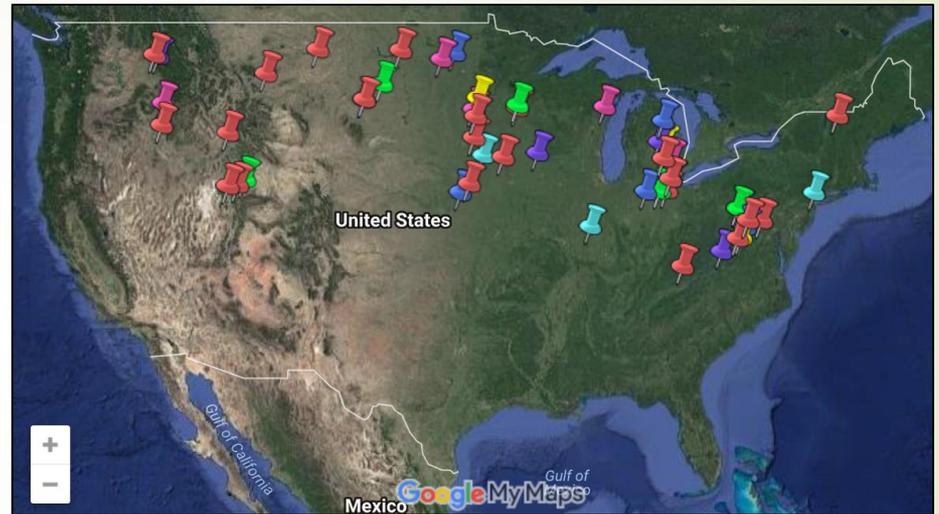
U.S. sheep industry

- 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

Variable	Suffolk	Targhee
No. animals	55,599	75,155
No. sires	2,989	1,294
No. dams	13,534	17,465
No. flocks	105	56
Sire family size		
Average	16.7	54.3
Largest	355	386

Weights

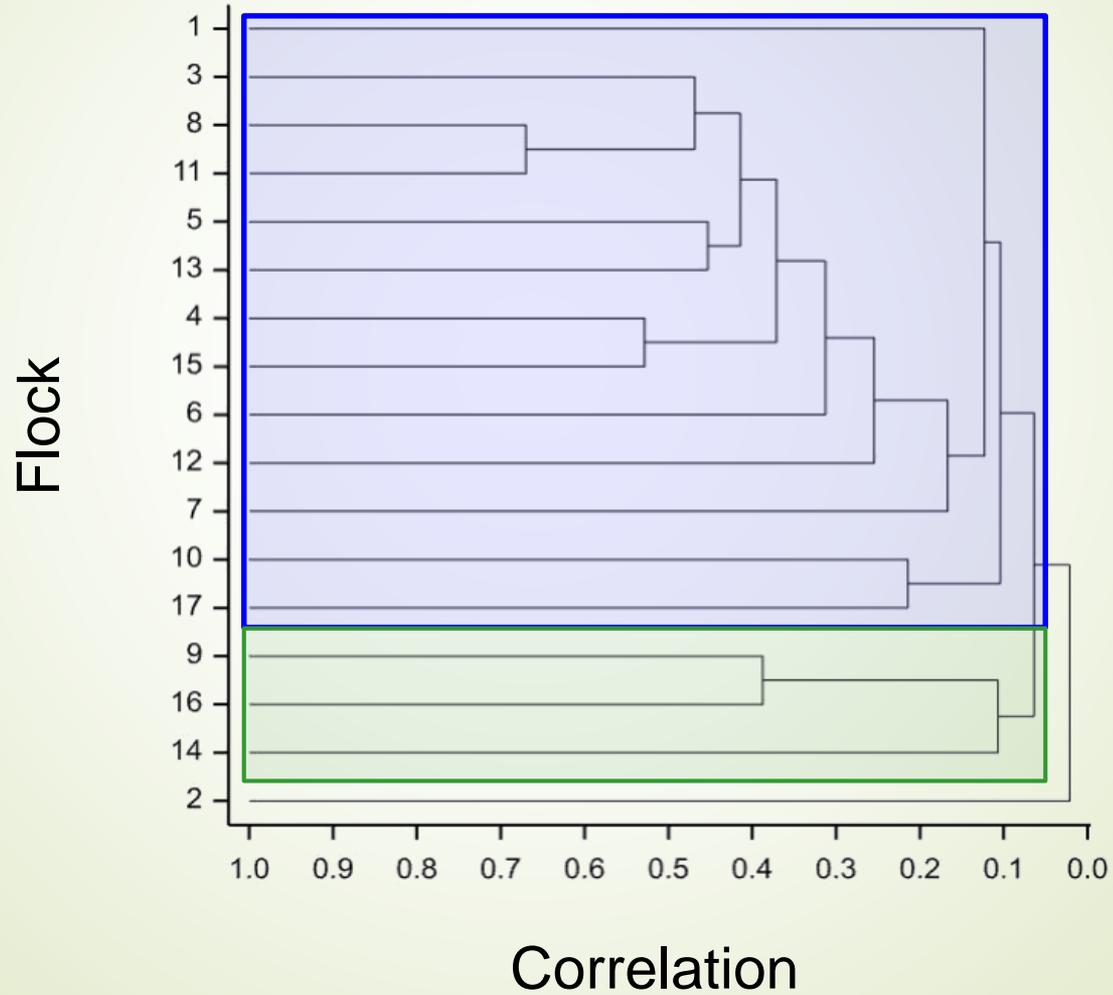| Variable	Suffolk[‡]	Targhee[§]
No. lambs	36,409	48,014
No. flocks		
Total	101	41
Active	18	17
Average		
Age (day)	63.0	130.1
Weight (kg)	30.7	36.6

[‡] At weaning: heritability of 0.15
[§] At post-weaning: heritability of 0.10

Suffolk
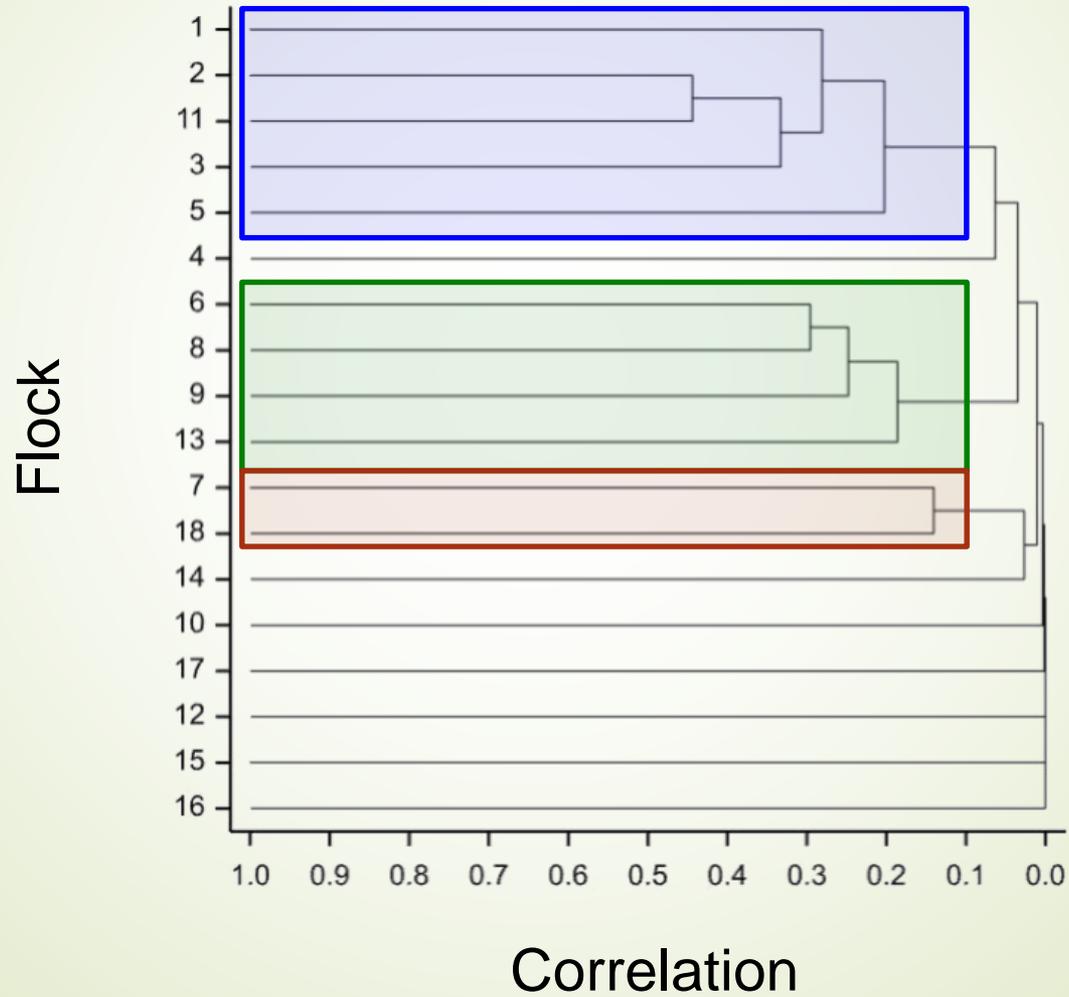
Targhee

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)
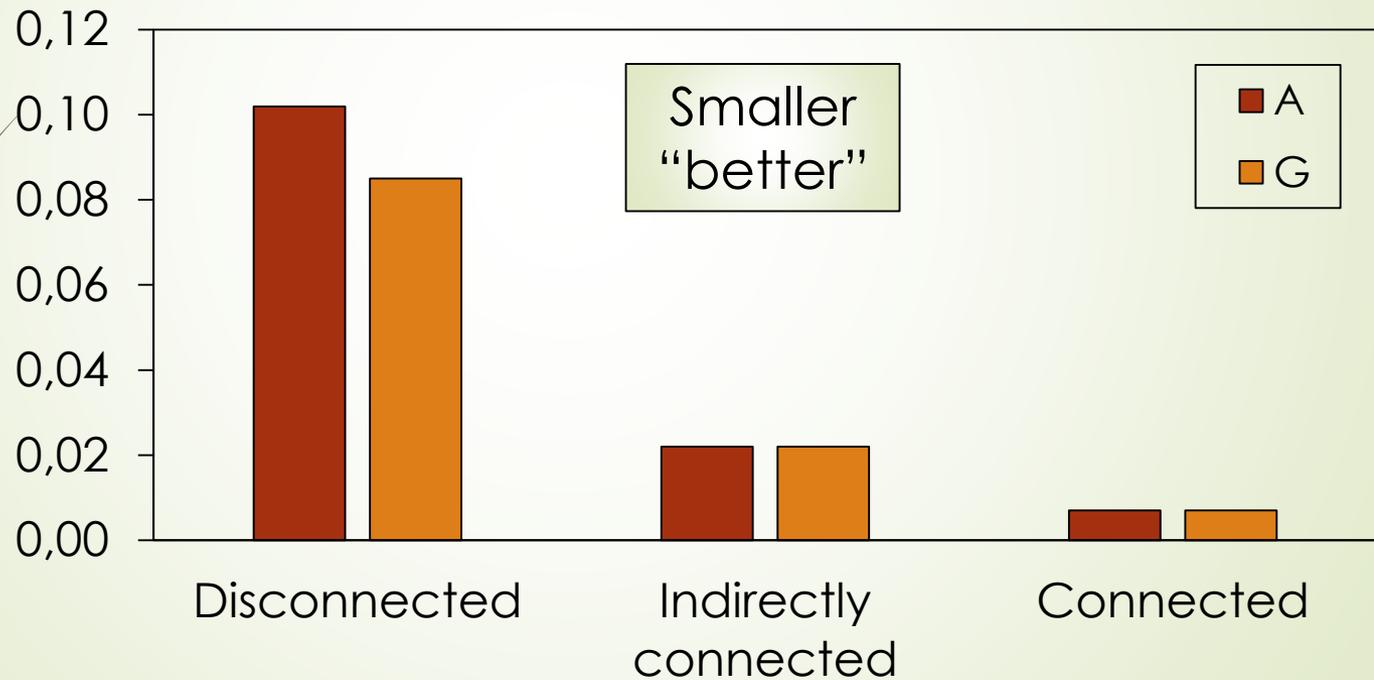
Genomic relatedness

- 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

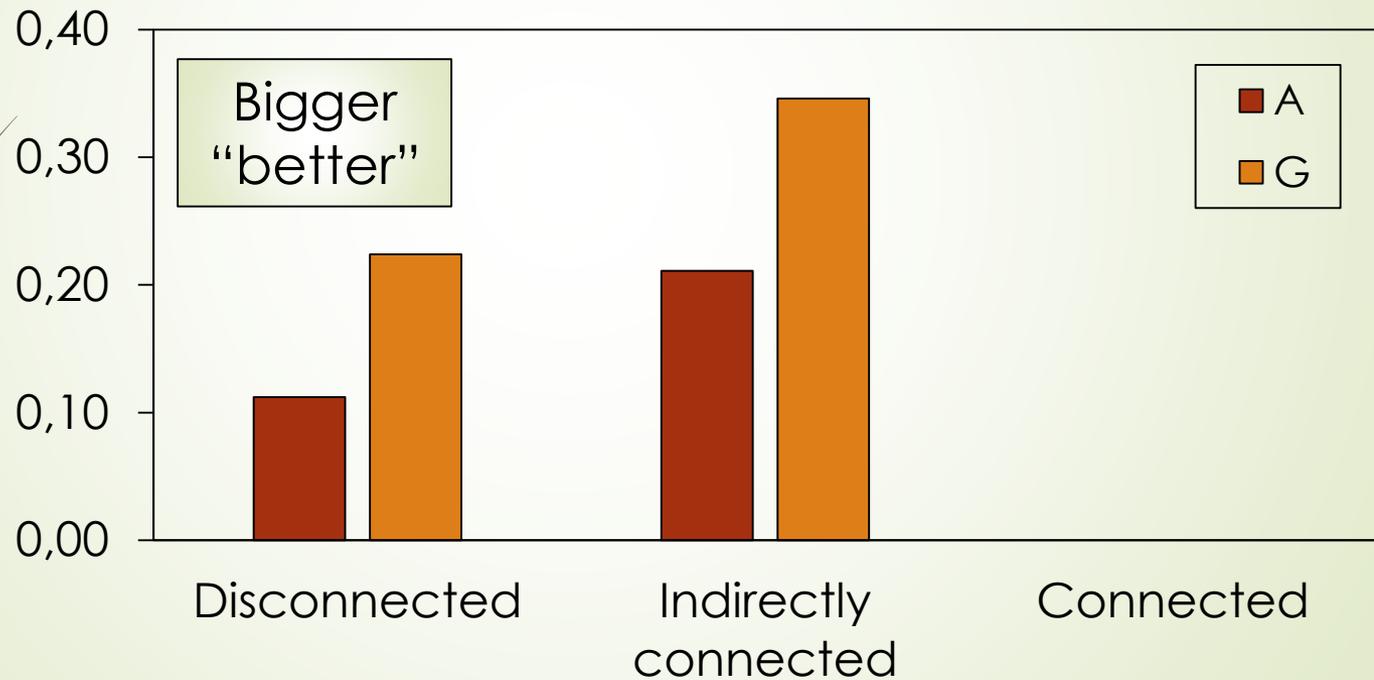# Genomic relatedness

PEV of difference

Smaller "better"

Legend: A, G

$h^2 = 0.2$

Genomic relatedness

Coefficient of determination

$$h^2 = 0.2$$

Genomic relatedness

Coefficient of determination

Bigger "better"

Reflects reduced variability due to relationships

Y-axis: 0,40 — 0,30 — 0,20 — 0,10 — 0,00

X-axis categories: Disconnected | Indirectly connected | Connected
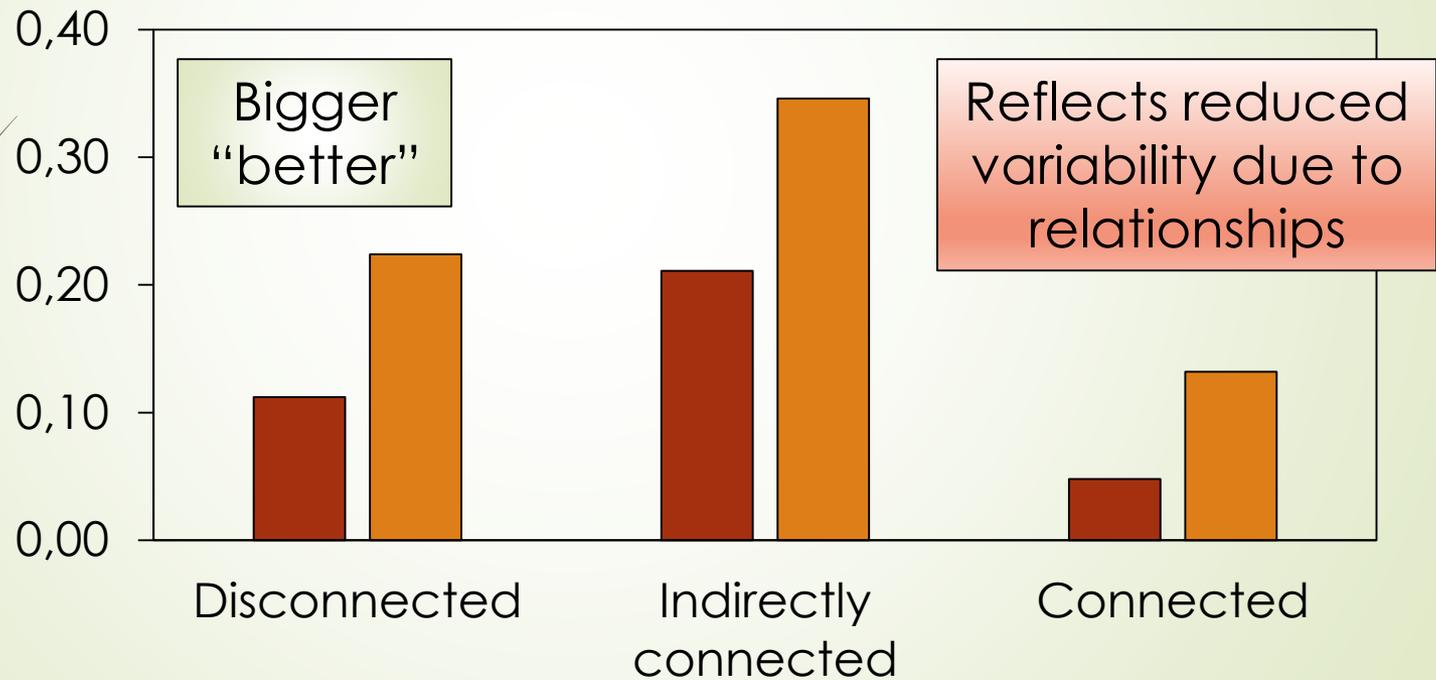
$$h^2 = 0.2$$

Genomic relatedness

- Provides relationships between distant individuals

- Captures Mendelian sampling

Panel	No. tests
Parentage	3,185
LD (15k)	387
50k	968
HD (600K)	1,100
Total	*5,640*

U.S. sheep genotyping

Targeting genotyping

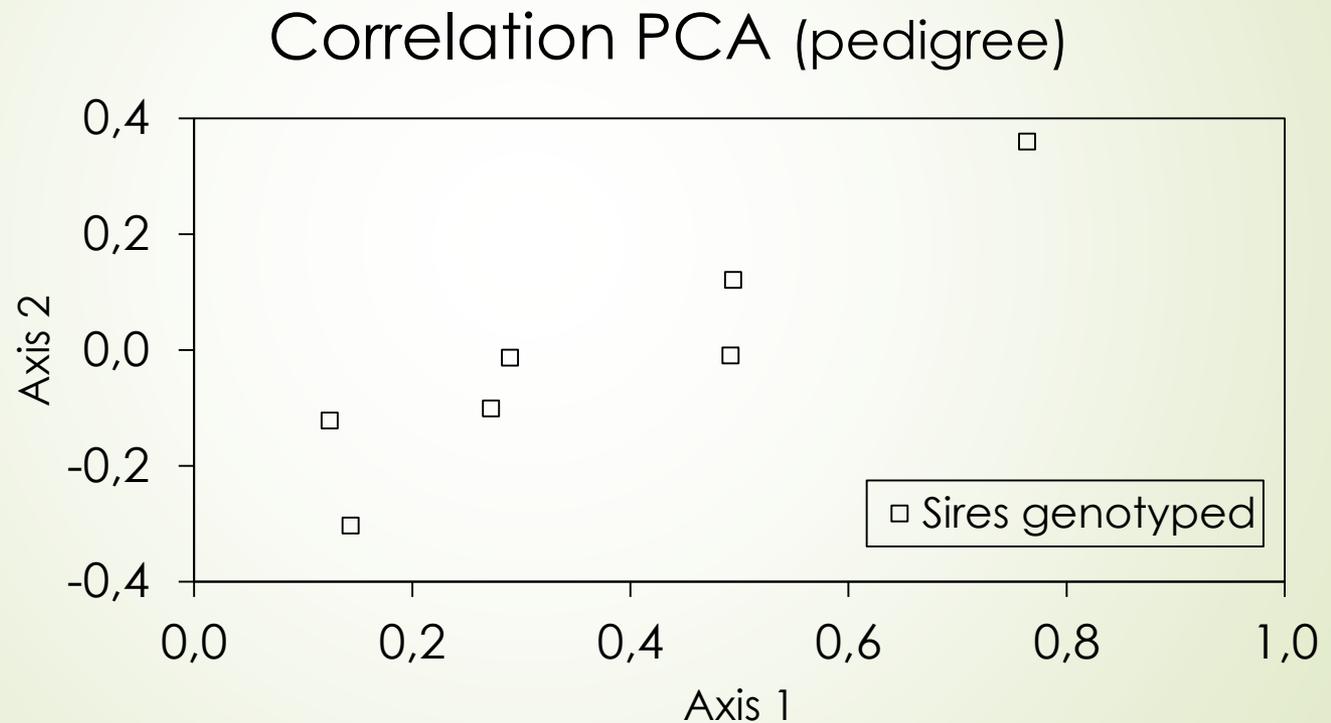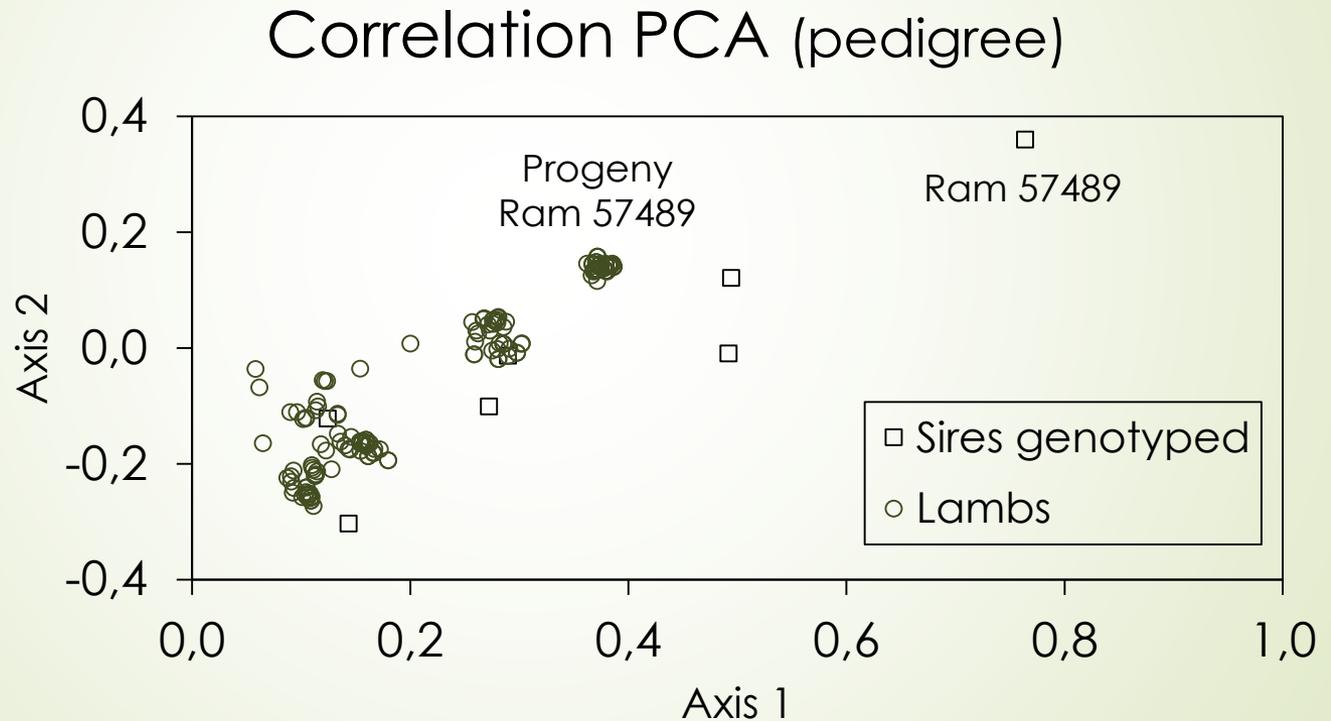- GWAS for fecal egg count at weaning (WFEC) in Katahdin sheep

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
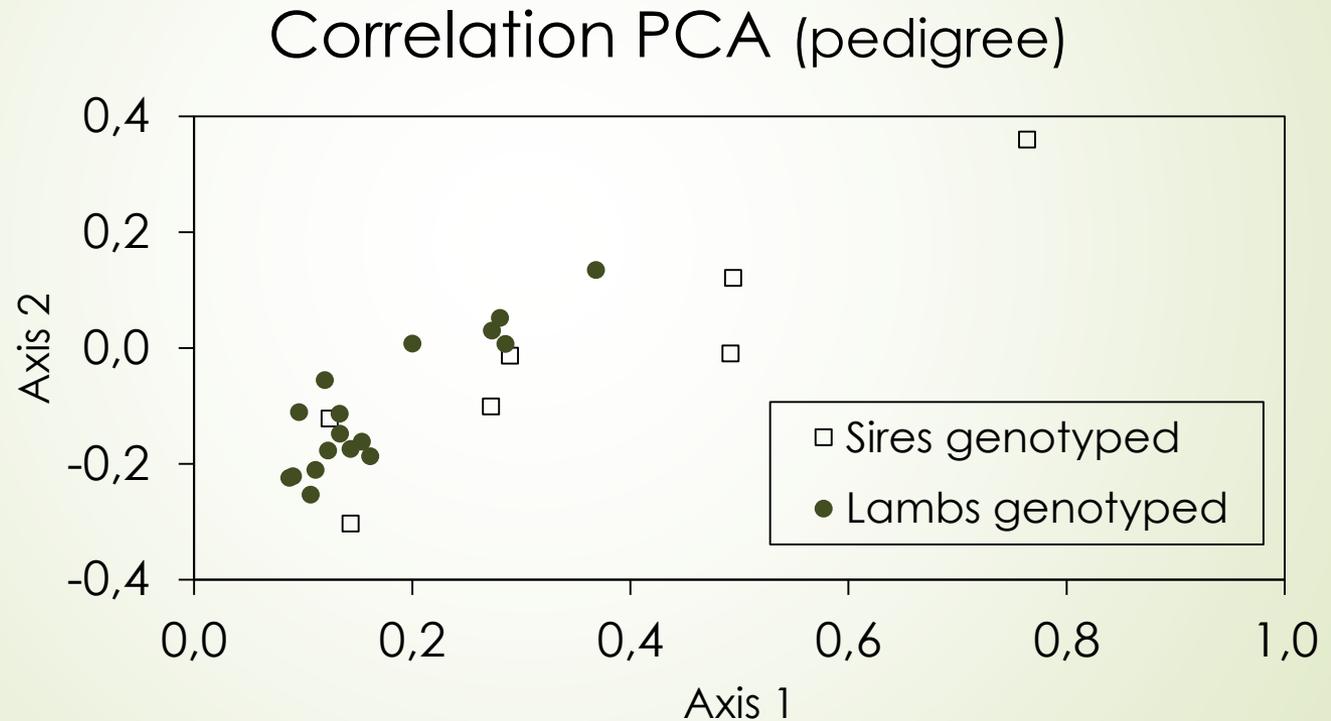 - Within upper or lower 10% on WFEC EBV

Targeting genotyping

Correlation PCA (pedigree)

Low EBV WFEC category

Targeting genotyping

Correlation PCA (pedigree)

Low EBV WFEC category

Targeting genotyping

Correlation PCA (pedigree)
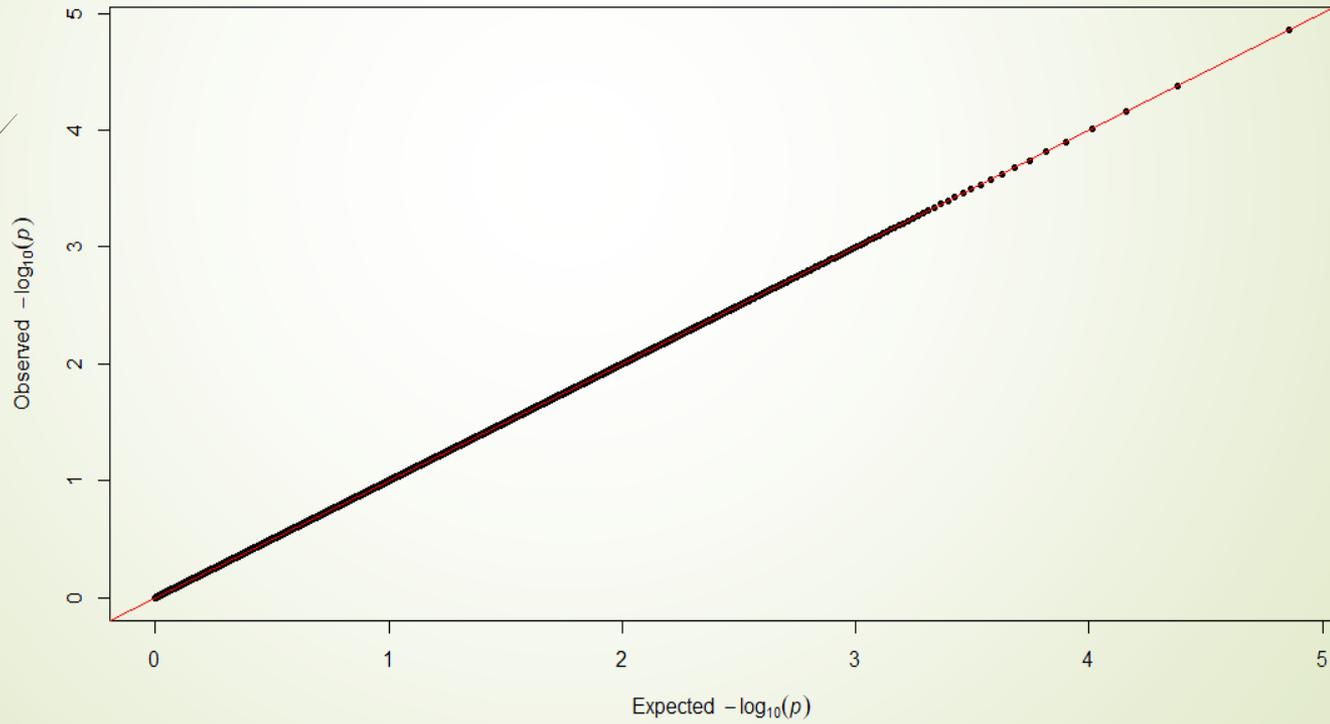

Low EBV WFEC category

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 ...

Considerations on Genetic Connectedness Between Management Units Under an Animal Model[1]

B. W. Kennedy and D. Trus[2]

Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, Ontario, Canada N1G 2W1

- SAC (now SRUC)& Roslin Institute
  - Ron Crump
  - Robin Thompson
  - Geoff Simm