# APY inverse of genomic relationship matrix:  Field data analyses

Ignacy Misztal, Ivan Pocrnic, Daniela Lourenco, Yutaka Masuda

University of Georgia

# Questions

- Do real-data studies agree with theory and simulations?

- Any implications?
  - Are simulation results useful to estimate effective population size in real populations?
  - Optimal chip size for species?
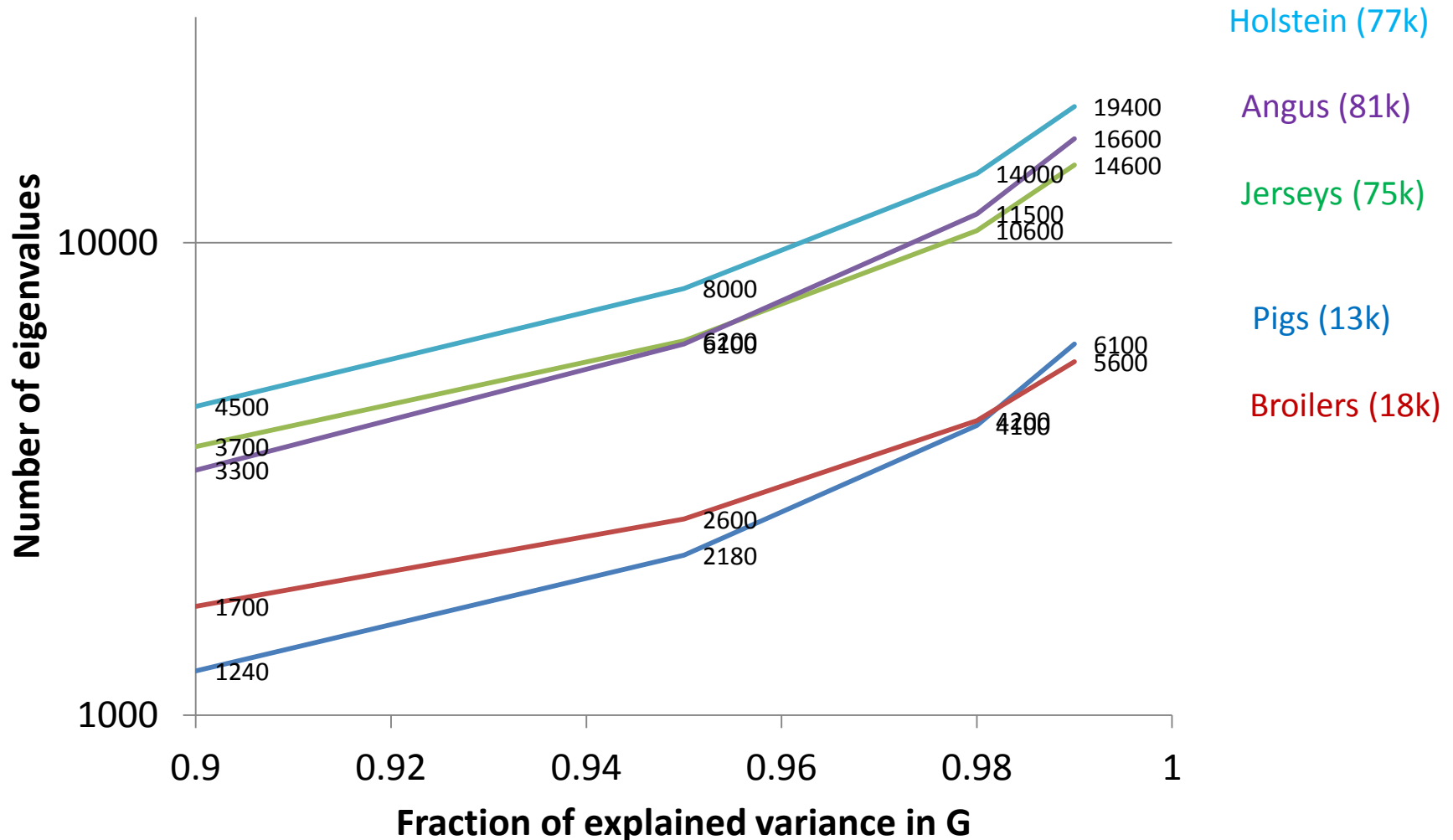  - Impact on GWA

# Data sets

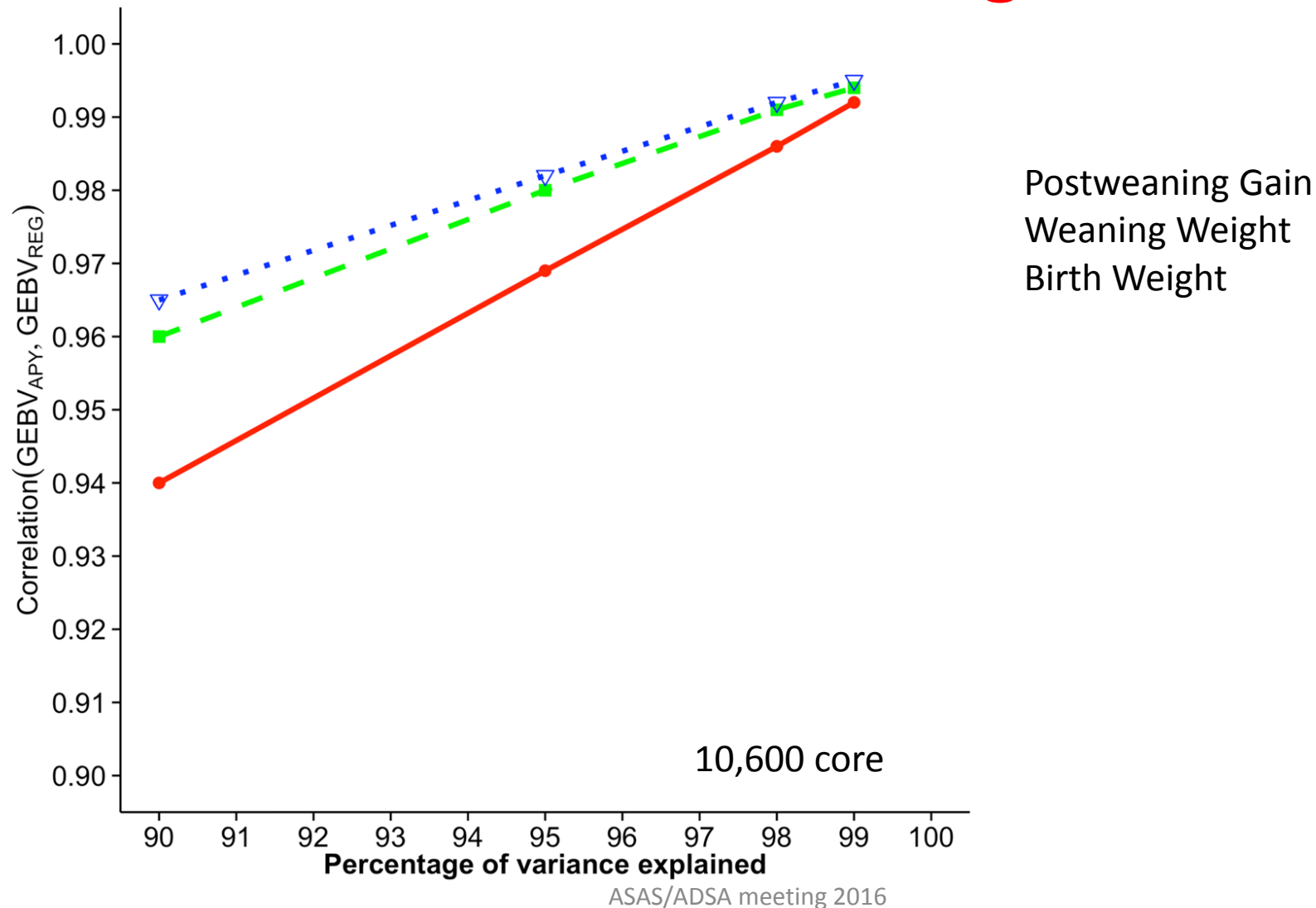| Specie | Trait | Animals | Genotypes | SNP after editing |
|--------|-------|---------|-----------|-------------------|
| Holsteins | Final score | 10.7 M | 77k | 61k |
| Jerseys | Milk, Fat, Protein | 6.7 M | 75k | 61k |
| Angus | Birth weight Weaning weight Yearling weight | 8.2 M | 81k | 38k |
| Pigs | Litter size Number of stillborn | 2.4 M | 23k | 37k |
| Chicken | Body weight | 199 k | 16k | 39k |

# Computations

- Eigenvalue decomposition of raw G

- Single-step GBLUP with APY and regular inverse

- Validation with APY and regular inverse
  - Based on $R^2$ in dairy
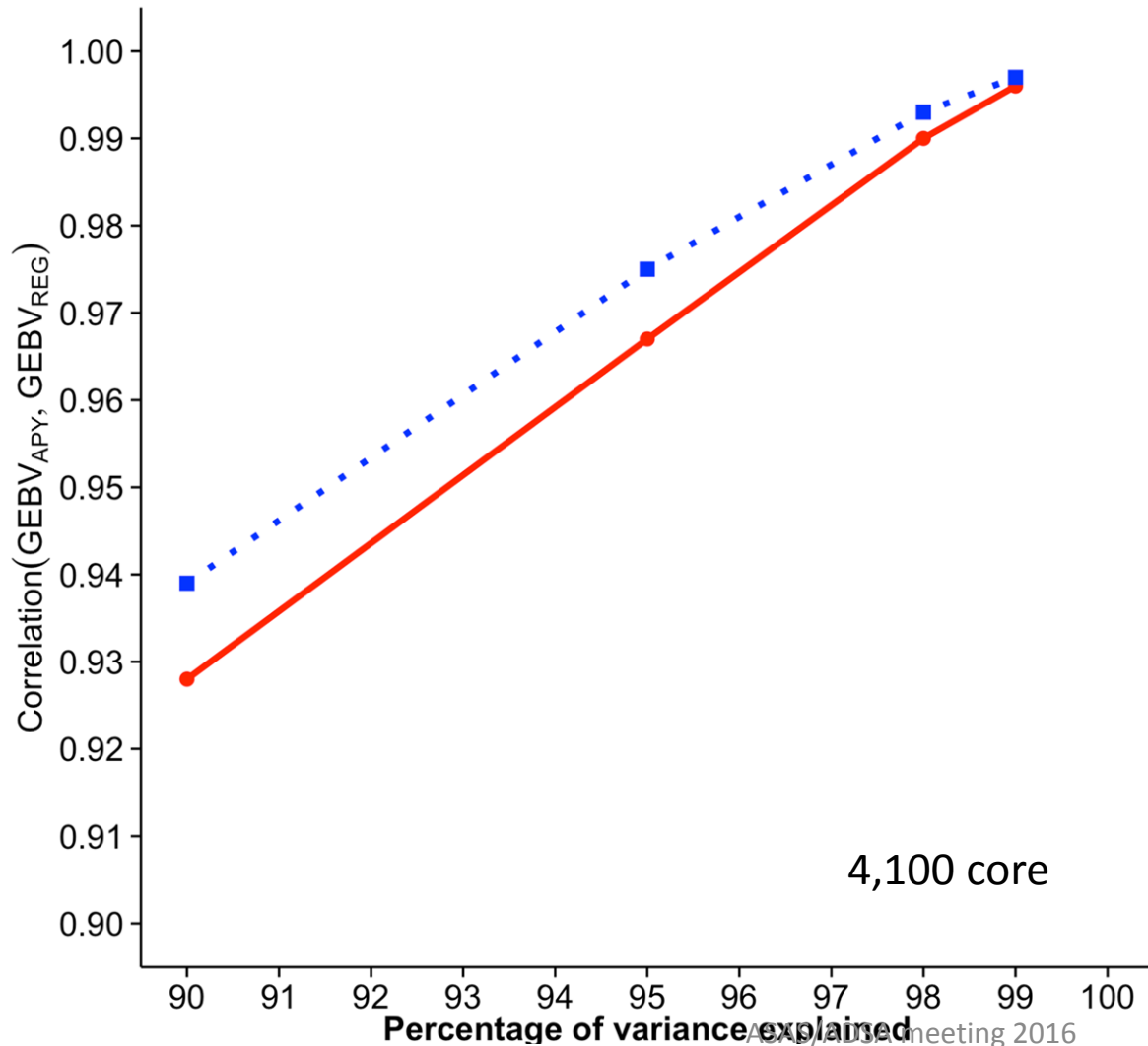  - Based on predictivity in Angus, pigs and broiler chicken:

    **corr (y-Xb, GEBV)**

# Number of eigenvalues in G to explain given fraction of variability



**Number of eigenvalues** vs **Fraction of explained variance in G**

Holstein (77k)
Angus (81k)
Jerseys (75k)
Pigs (13k)
Broilers (18k)

Data labels: 19400, 16600, 14000, 14600, 11500, 10600, 8000, 6200, 6100, 5600, 4500, 3700, 3300, 2600, 2180, 1700, 1240, 4100

# Correlations between "regular" and "APY" GEBV - Angus



Postweaning Gain
Weaning Weight
Birth Weight

10,600 core

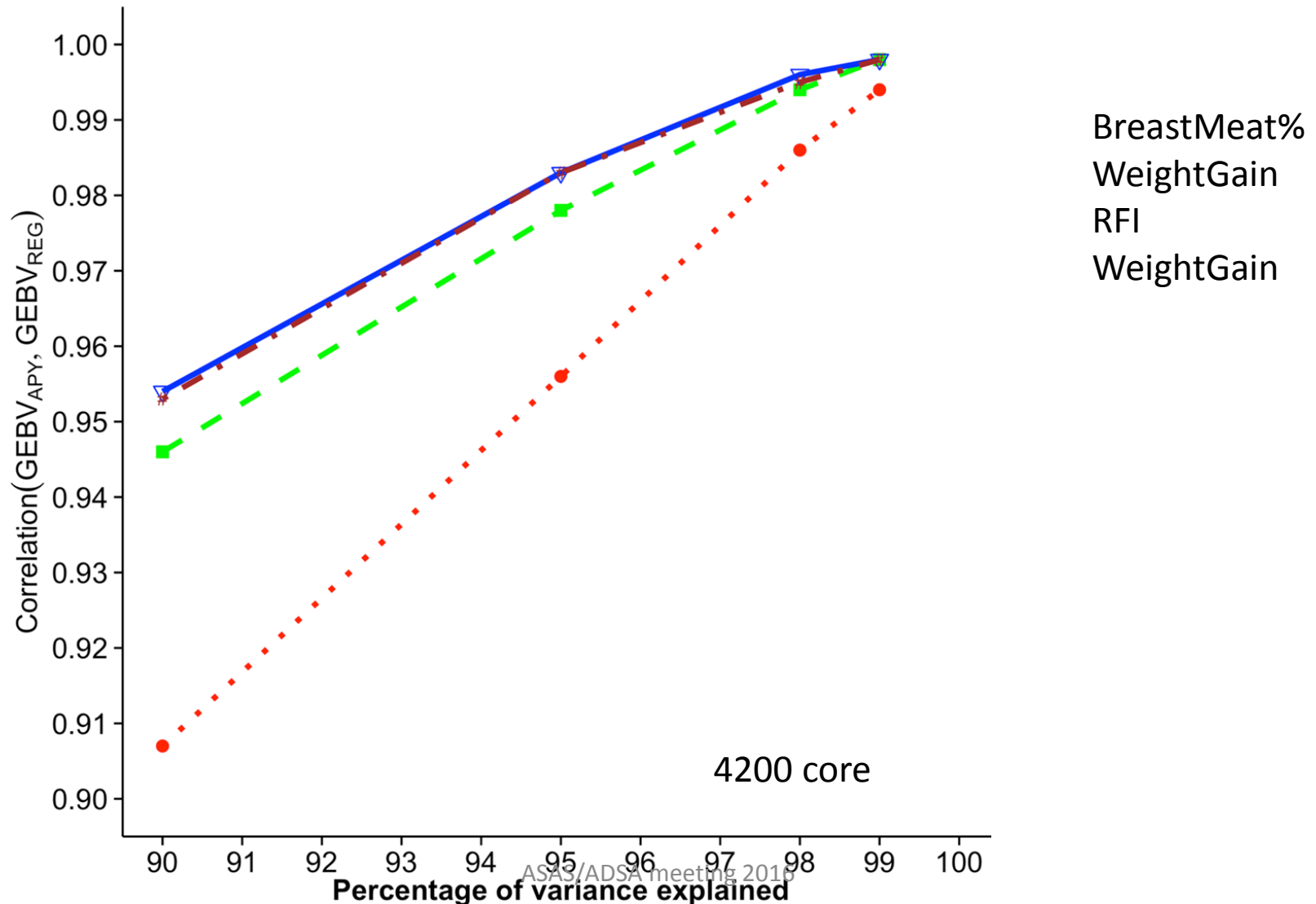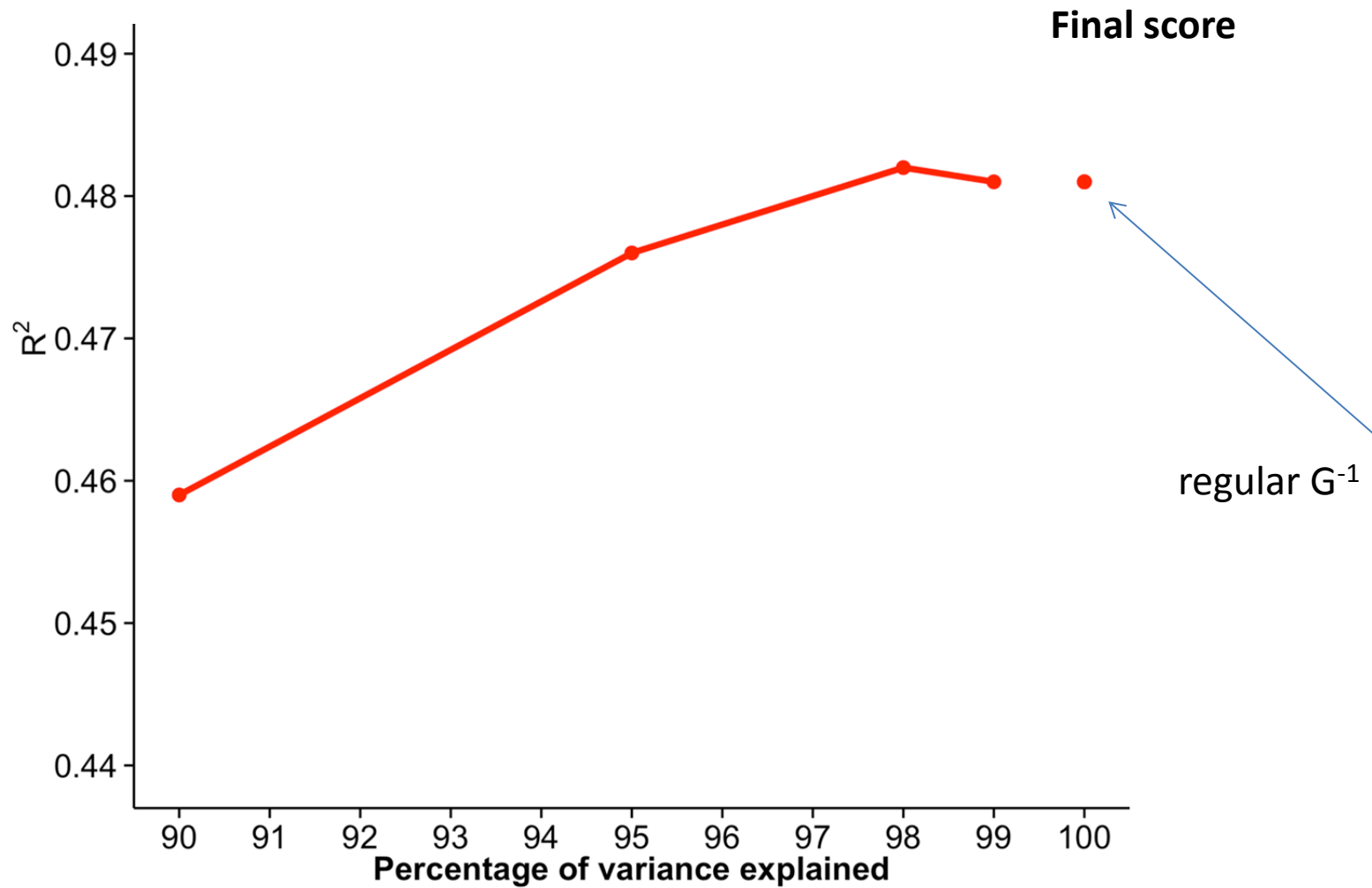# Correlations between "regular" and "APY" GEBV - Pigs



Stillbirth
Litter Size

About 2500 well chosen
Optimal  (Ostersen et al., 2016)

4,100 core

# Correlations between "regular" and "APY" GEBV - Broilers



BreastMeat%
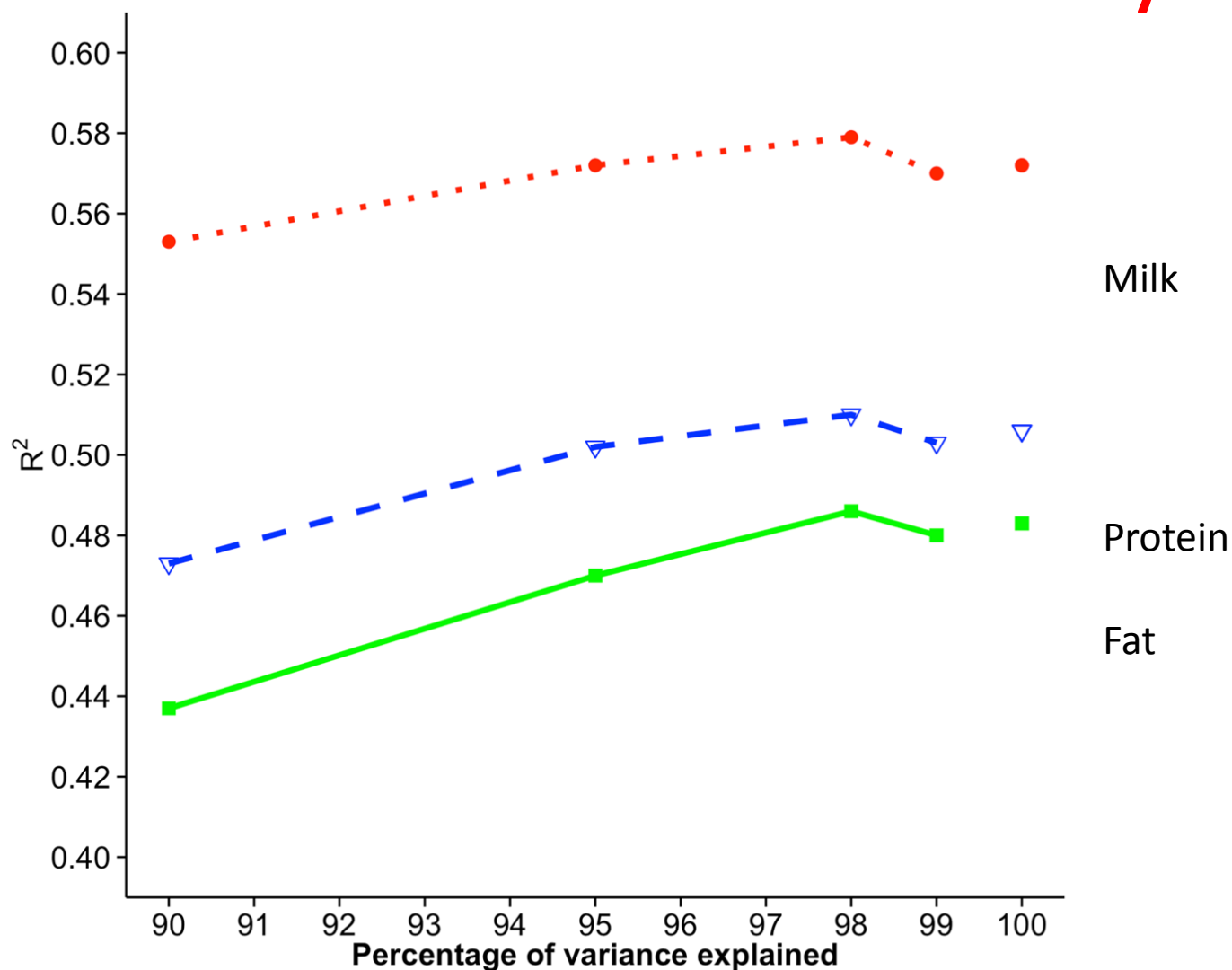WeightGain
RFI
WeightGain

4200 core

# Reliabilities - Holsteins



**Final score**

$R^2$

regular $G^{-1}$

Percentage of variance explained

ASAS/ADSA meeting 2016

Core:  4500                    8000        14,000  19,400

# Reliabilities - Jerseys



Milk

Protein

Fat

**Percentage of variance explained**

**Core: 3300**          **6100**    ASAS/ADSA meeting 2016   **11,500**

# Predictability - Angus

# Predictability - Pigs



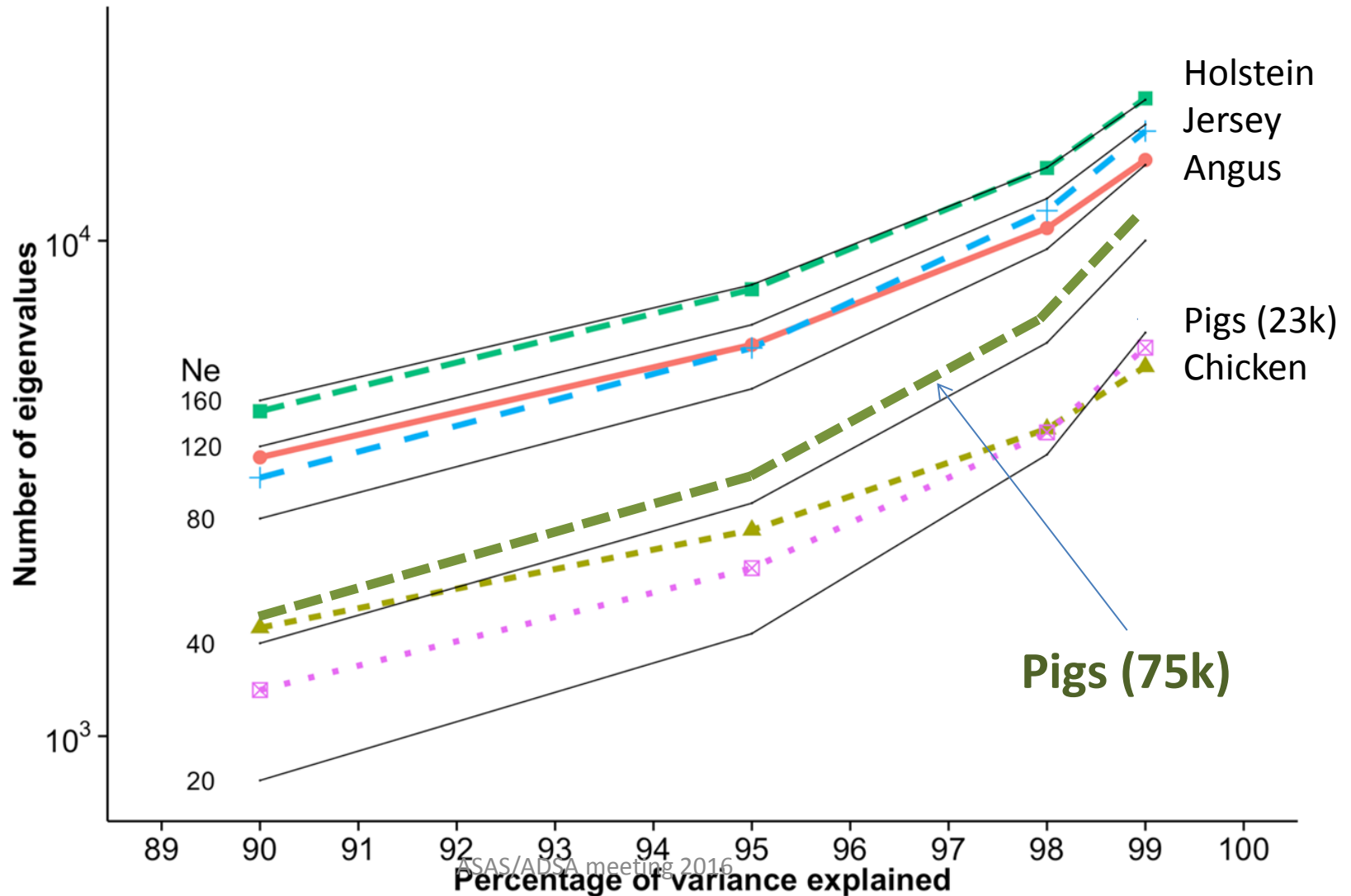Litter size
Stillbirth
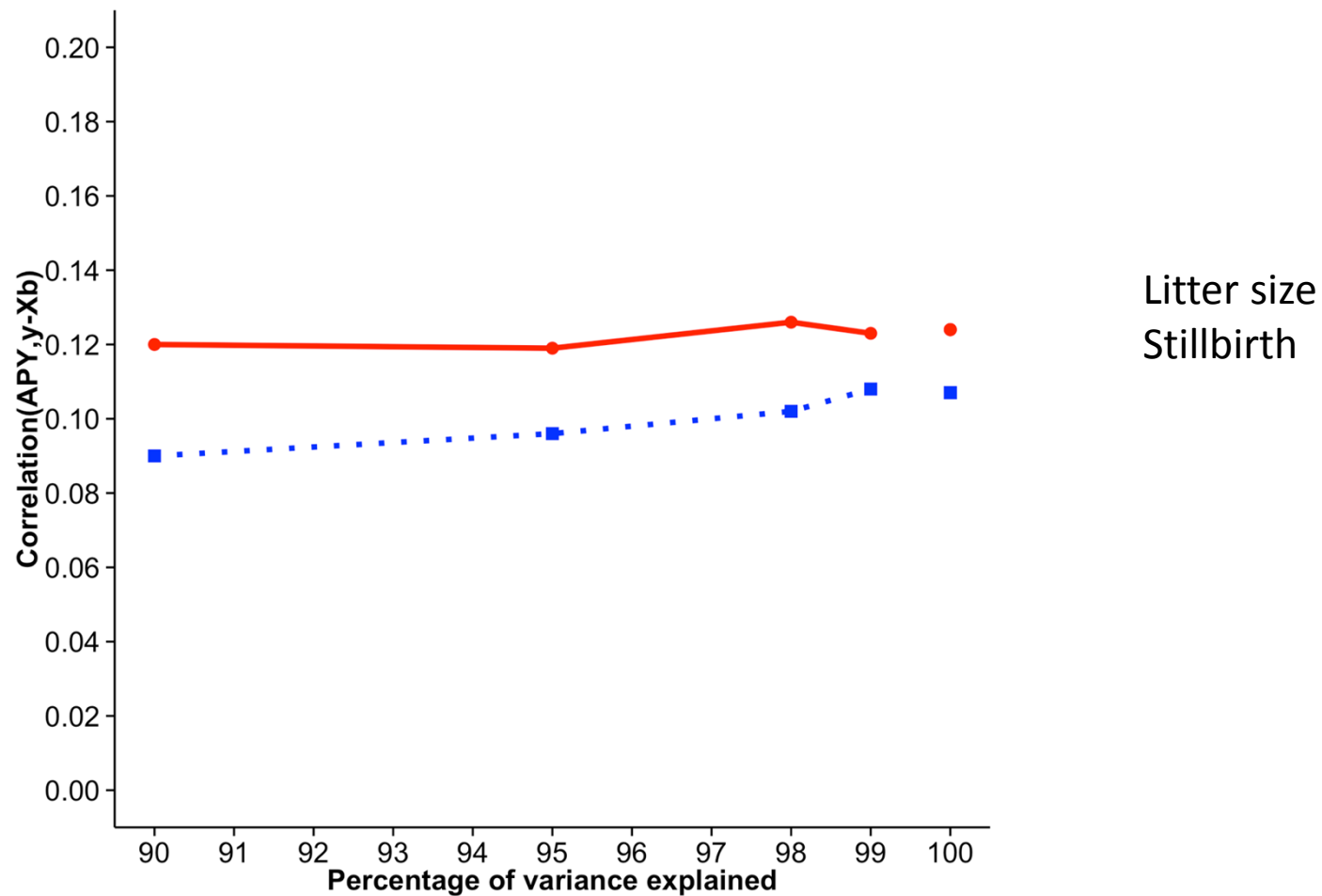
# Predictability - Broilers
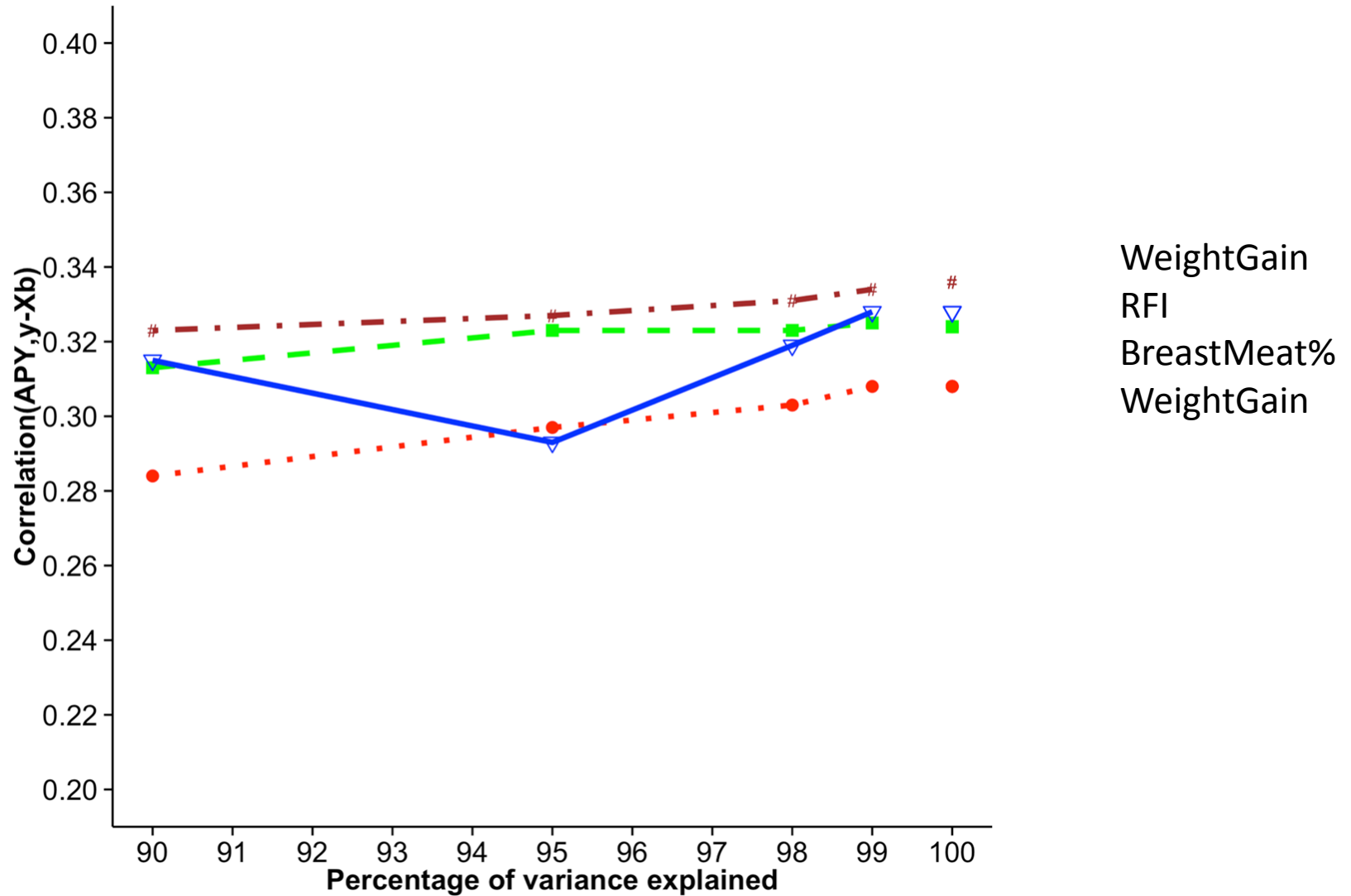


WeightGain
RFI
BreastMeat%
WeightGain

# Number of eigenvalues in G to explain given fraction of variability

# Estimated effective population size

- Me ~ NeL

  Me – chromosome segments, Ne – effective population size, L – Length in Morgans

- Interpolation based on 90%

| Specie | Ne at L=30 Morgans |
|--------|--------------------|
| Holsteins | 149 |
| Jerseys | 101 |
| Angus | 113 |
| Pigs | 32   (43 at L=20 Morgans) |
| Chicken | 44 |

# Estimated number of core animals and optimal number of SNP

- About 12 SNP per one chromosome segment (MacLead et al., 2005) to detect 90% transitions

| Specie | Range of Me (95-99%) | Number of SNP (12 x Me) |
|---|---|---|
| Holsteins | 8k-14k | 100-180k |
| Jerseys | 6k-12k | 70k-150k |
| Angus | 6k-11k | 70k-130k |
| Pigs | 2k-6k | 24k-72k |
| Chicken | 3k-6k | 36K-72k |

# Comments and Questions

- Is choice of core animals critical?
  - Random choice better than oldest or youngest (Ostersen et al., 2016)

- Dimensionality of weighted G (Fragomeni et al., 2016)
  - With weighted G, dimensionality smaller
  - With all 100 QTL known, 0.99 accuracy with 200 core animals (Fragomeni et al., 2016)

- Is resolution of GWAS by average size of chromosome segments (90-98% eigenvalues)?

# Conclusions

- APY works well across species

- Optimal number of core animals from about 4,000 (broilers) to 18,000 (Holsteins)

  - Little reduction in accuracy with half that number

- Estimates of effective population size from 40s (pigs and chicken) to 150 (Holsteins)
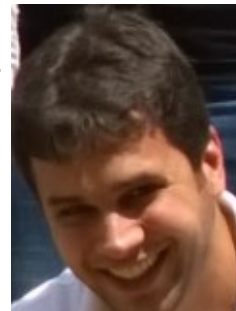
# Acknowledgements

- Grants from Holsteins Assoc., Angus Assoc., Cobb-Vantress, Zoetis, Smithfield, PIC,…
- AFRI grant 2015-67015-22936 from USDA NIFA
- Paul VanRaden, George Wiggans, CDCB
- Collaborators

Shogo Tsuruta

Ignacio Aguilar
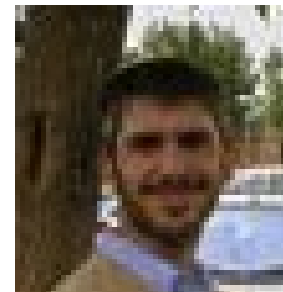
Breno Fragomeni

Ivan Pocrnic

Daniela Laurenco

Yutaka Masuda

Andres Legarra

ASAS/ADSA meeting 2016