APY inverse of genomic relationship matrix: Field data analyses

Ignacy Misztal, Ivan Pocrnic, Daniela Lourenco, Yutaka Masuda

University of Georgia

Using recursion to compute the inverse of the genomic relationship matrix

I. Misztal,*[1] A. Legarra,† and I. Aguilar‡
*Department of Animal and Dairy Science, University of Georgia, Athens 30602-2771
†INRA, UR631-SAGA, BP 52627, 31326 Castanet-Tolosan Cedex, France
‡Instituto Nacional de Investigación Agropecuaria, Las Brujas 90200, Uruguay

Hot topic: Use of genomic recursions in single-step genomic best linear unbiased predictor (BLUP) with a large number of genotypes

B. O. Fragomeni,*[1] D. A. L. Lourenco,* S. Tsuruta,* Y. Masuda,* I. Aguilar,† A. Legarra,‡ T. J. Lawlor,§ and I. Misztal*
*Department of Animal and Dairy Science, University of Georgia, Athens 30602
†Instituto Nacional de Investigacion Agropecuaria, Canelones, 90200, Uruguay
‡INRA, UMR1388 GenePhySE, Castanet Tolosan, 31326, France
§Holstein Association USA Inc., Brattleboro, VT 05302

GENETICS | INVESTIGATION

Inexpensive Computation of the Inverse of the Genomic Relationship Matrix in Populations with Small Effective Population Size

Ignacy Misztal[1]
Animal and Dairy Science, University of Georgia, Athens, Georgia 30602

GENETICS | INVESTIGATION

The Dimensionality of Genomic Information and Its Effect on Genomic Prediction

Ivan Pocrnic,*[1] Daniela A. L. Lourenco,* Yutaka Masuda,* Andres Legarra,† and Ignacy Misztal*
*Department of Animal and Dairy Science, University of Georgia, Athens, Georgia 30602, and †Institut National de la Recherche Agronomique, GenPhySE, F-31326 Castanet-Tolosan, France

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	2.4 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:
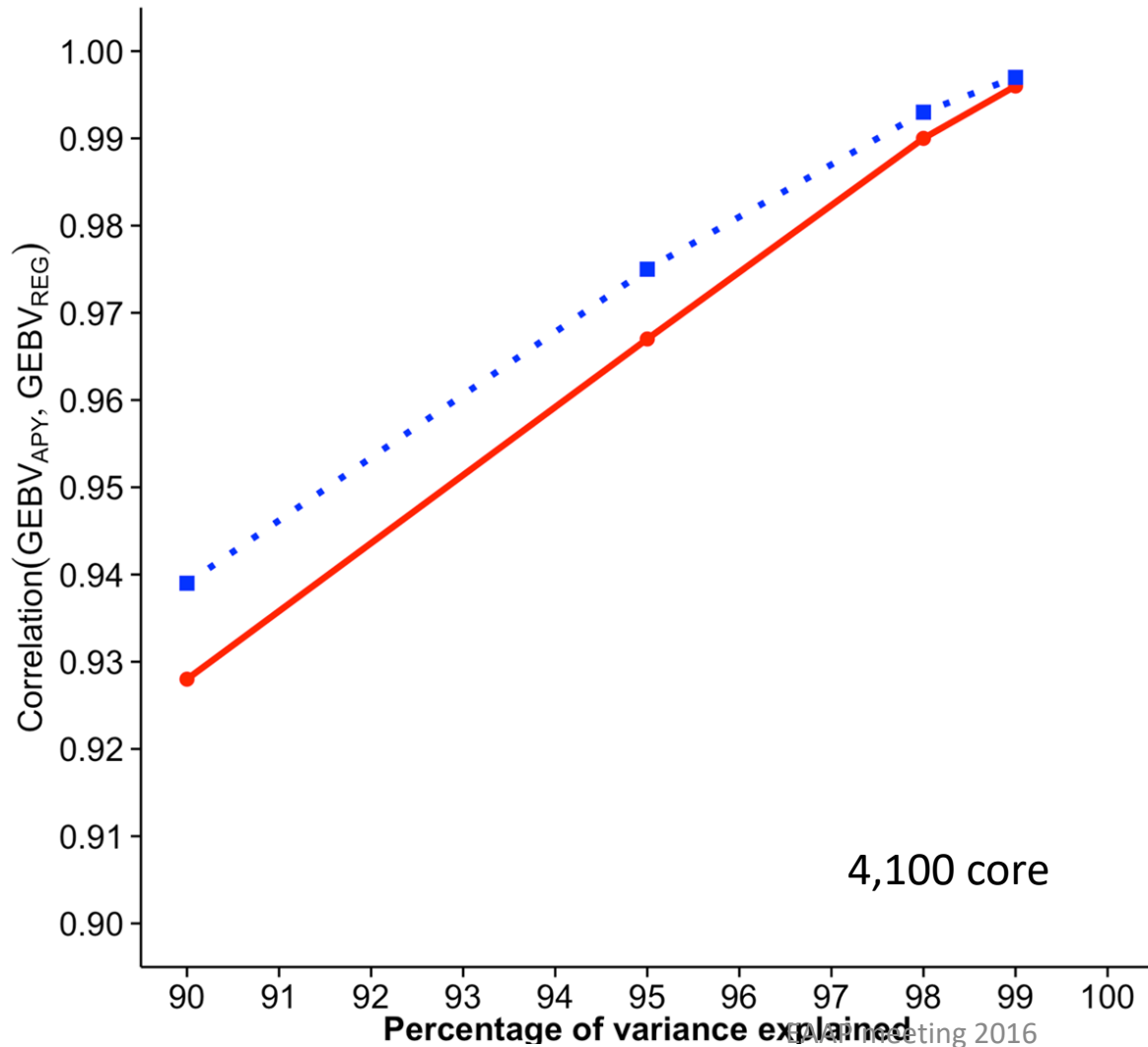
    corr (y-Xb, GEBV)

Number of eigenvalues in G to explain given fraction of variability
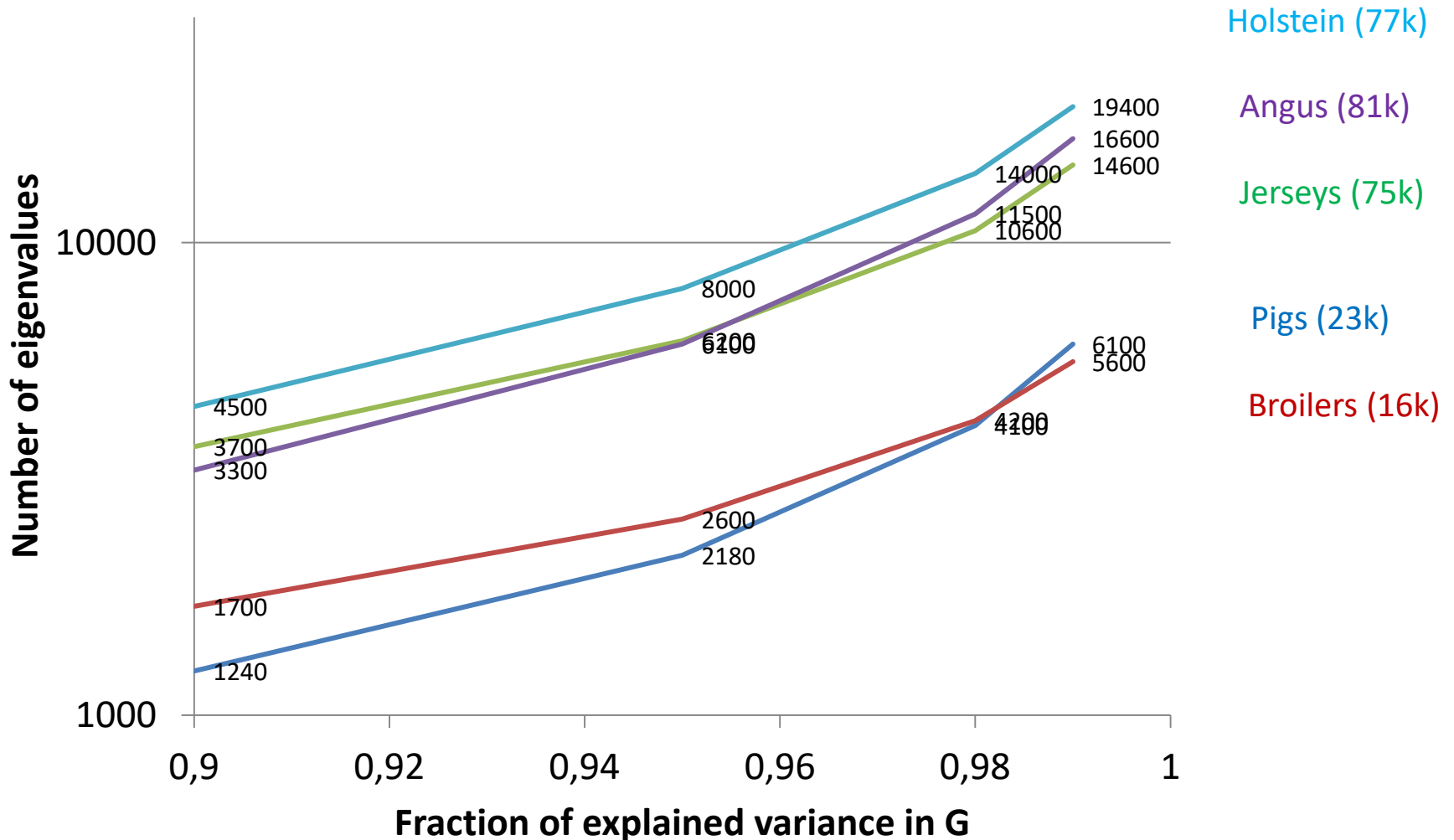
EAAP meeting 2016

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

Postweaning Gain
Weaning Weight
Birth Weight

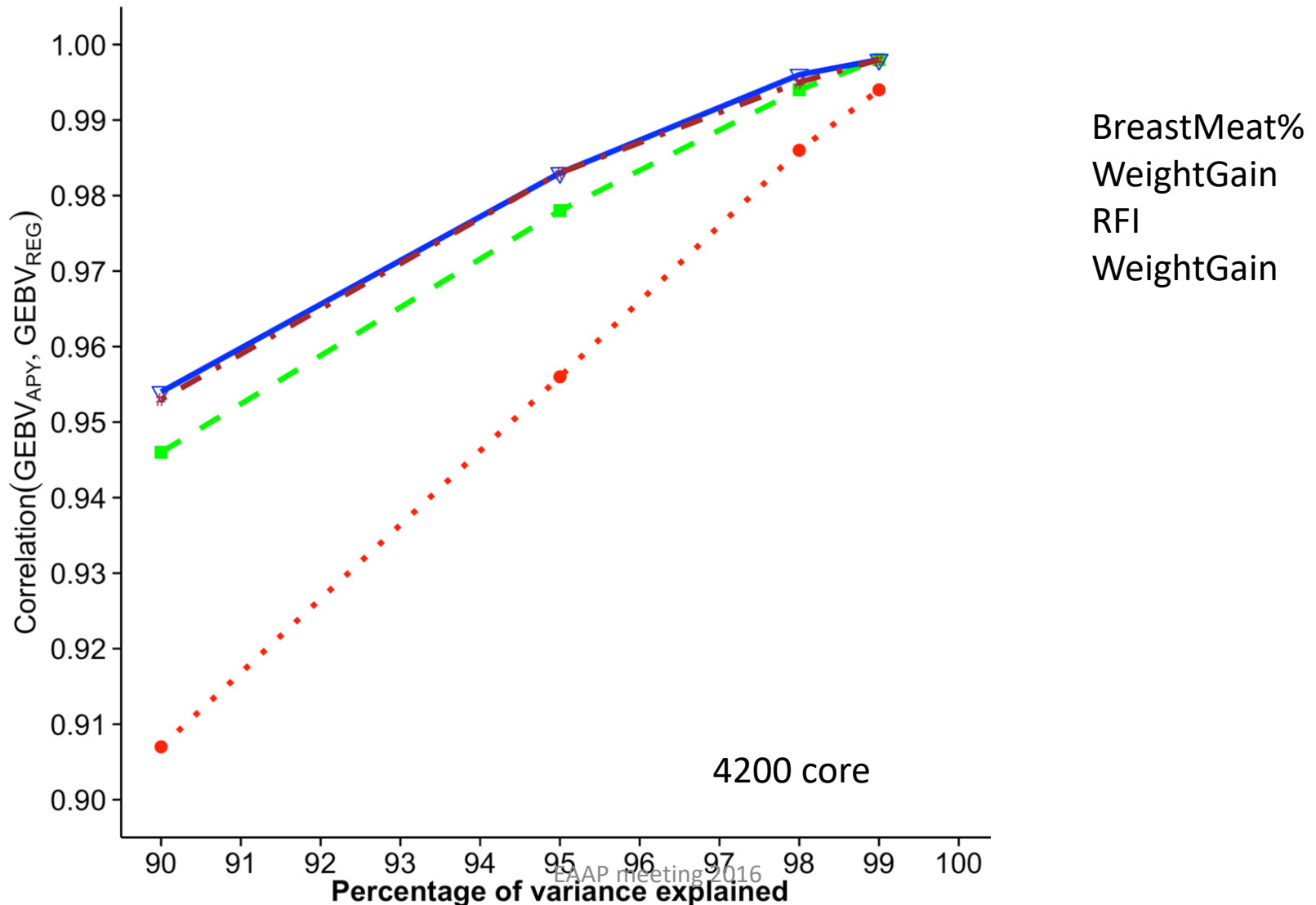10,600 core
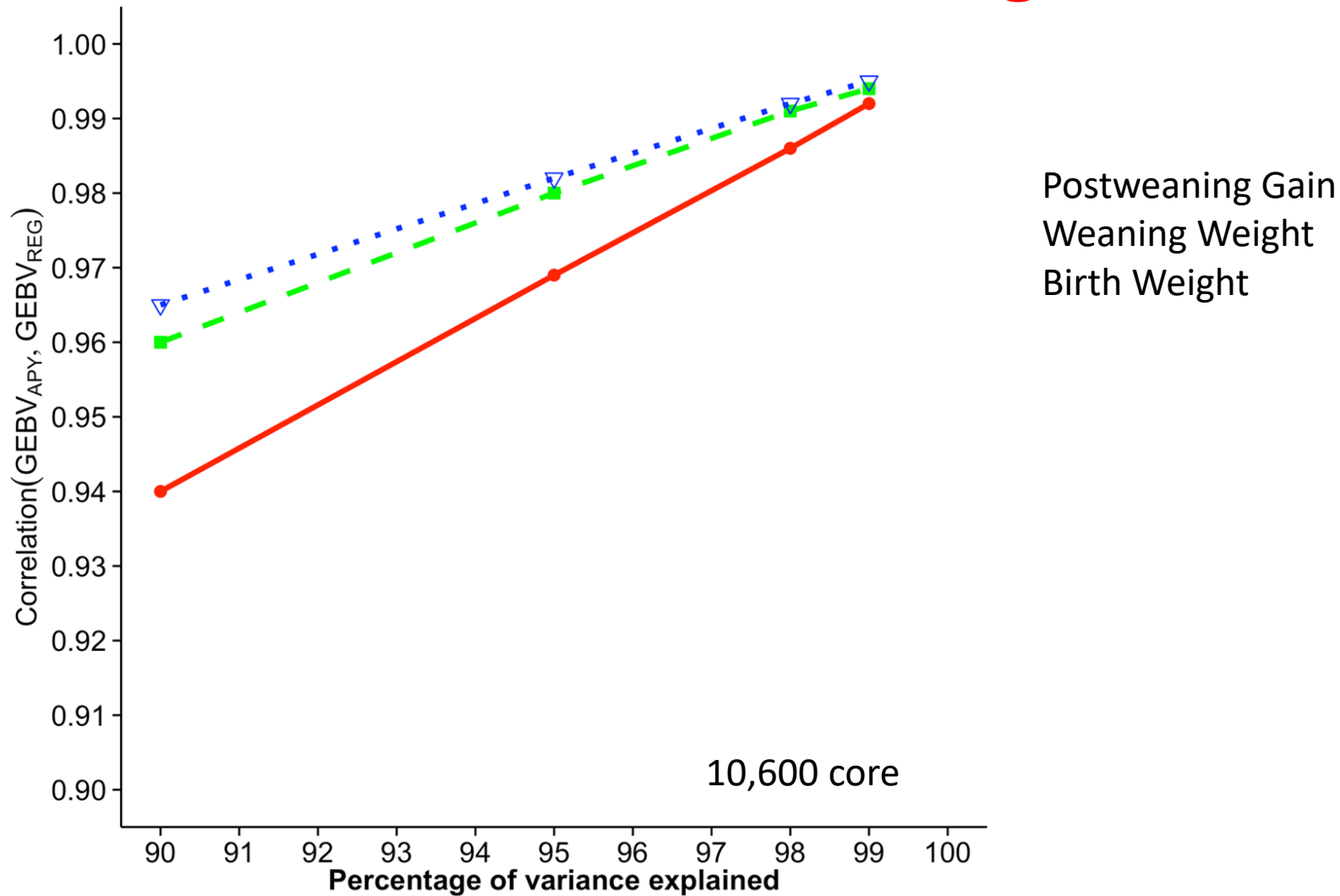
Correlations between "regular" and "APY" GEBV - Pigs

Stillbirth
Litter Size
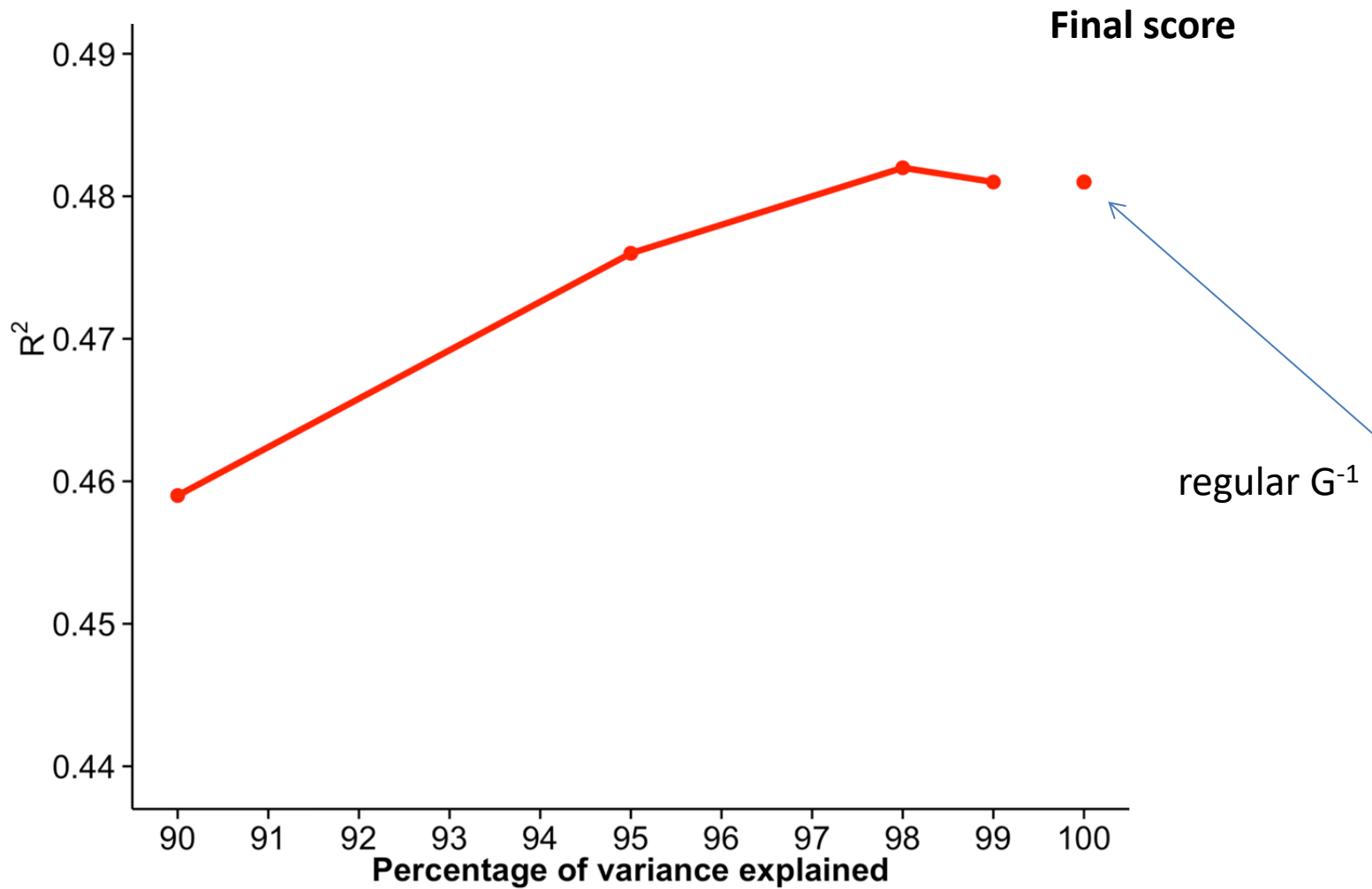
About 2500 well chosen
optimal (Ostersen et al., 2016)

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

BreastMeat%
WeightGain
RFI
WeightGain

4200 core

Reliabilities - Holsteins

Final score

regular G^{-1}

R²

Percentage of variance explained

Core: 4500   8000   14,000 19,400

Reliabilities - Jerseys
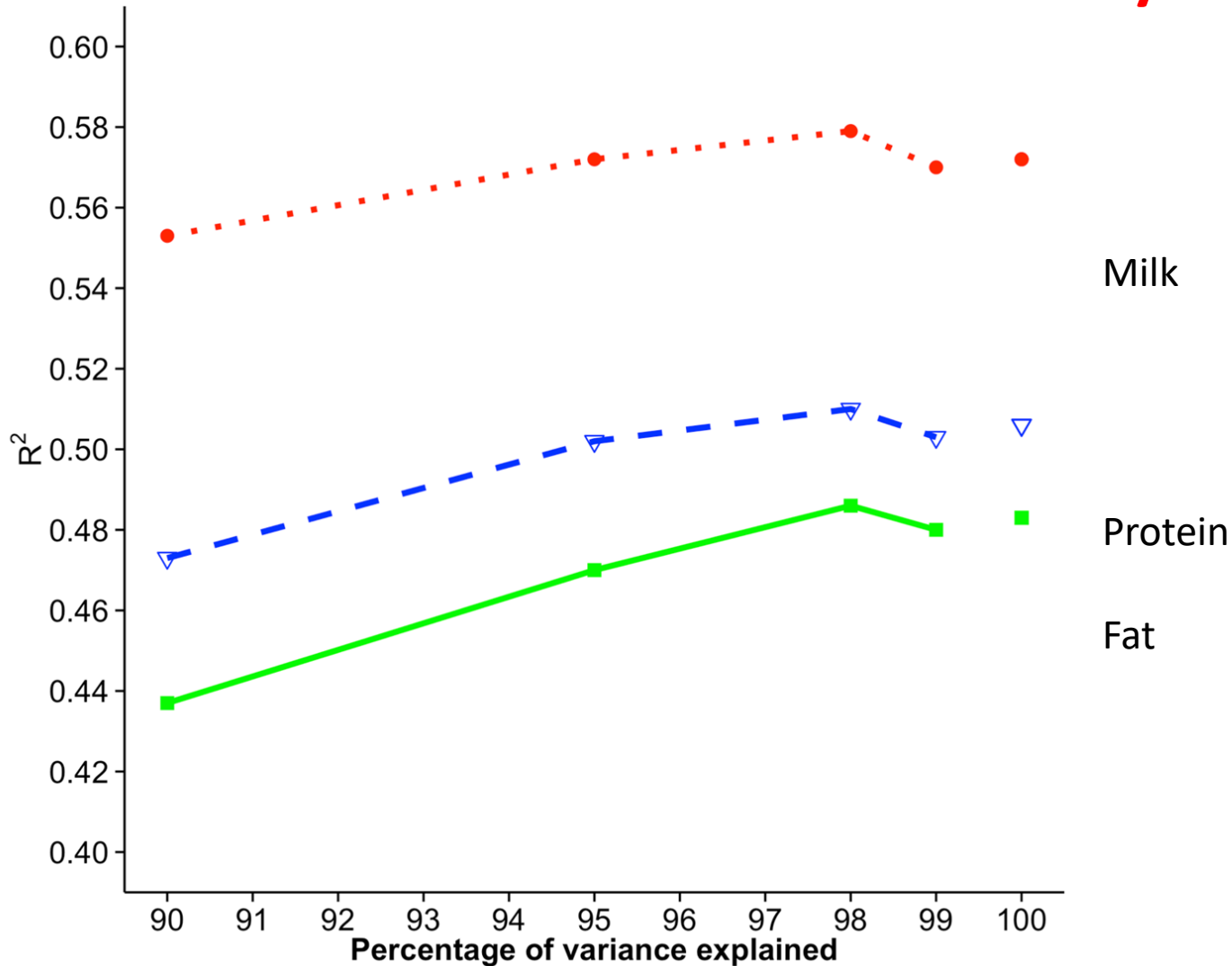
Milk

Protein

Fat

Core: 3300   **6100**   EAAP meeting 2016   **11,500**

Predictability - Angus

Birth Weight
Weaning Weight

Postweaning Gain

Reliability = (predictability/h)^2
Predictability model dependent
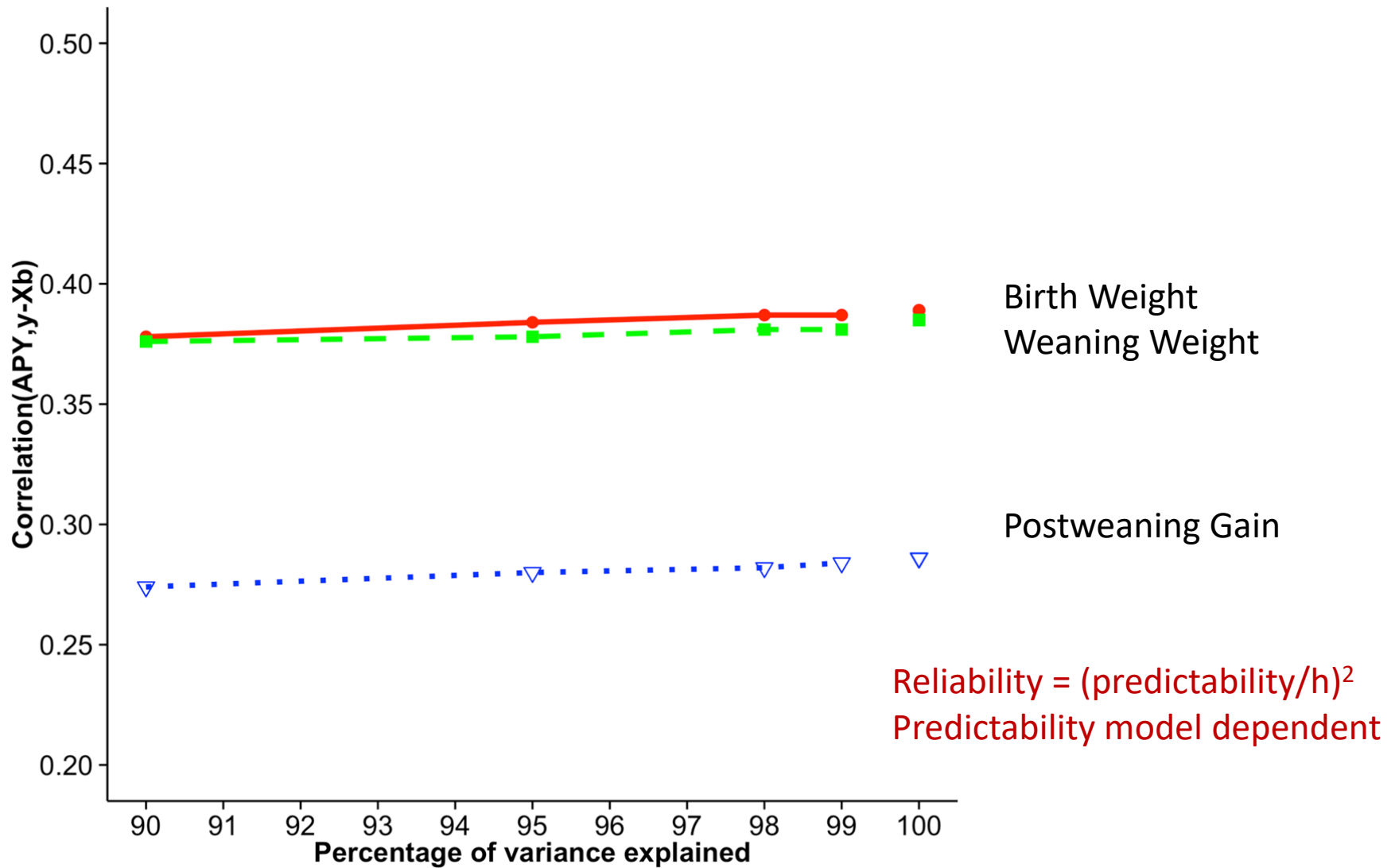
Core: 3600　　　　**6200** EAAP meeting 2010 **10,600**

Predictability - Pigs

Litter size
Stillbirth
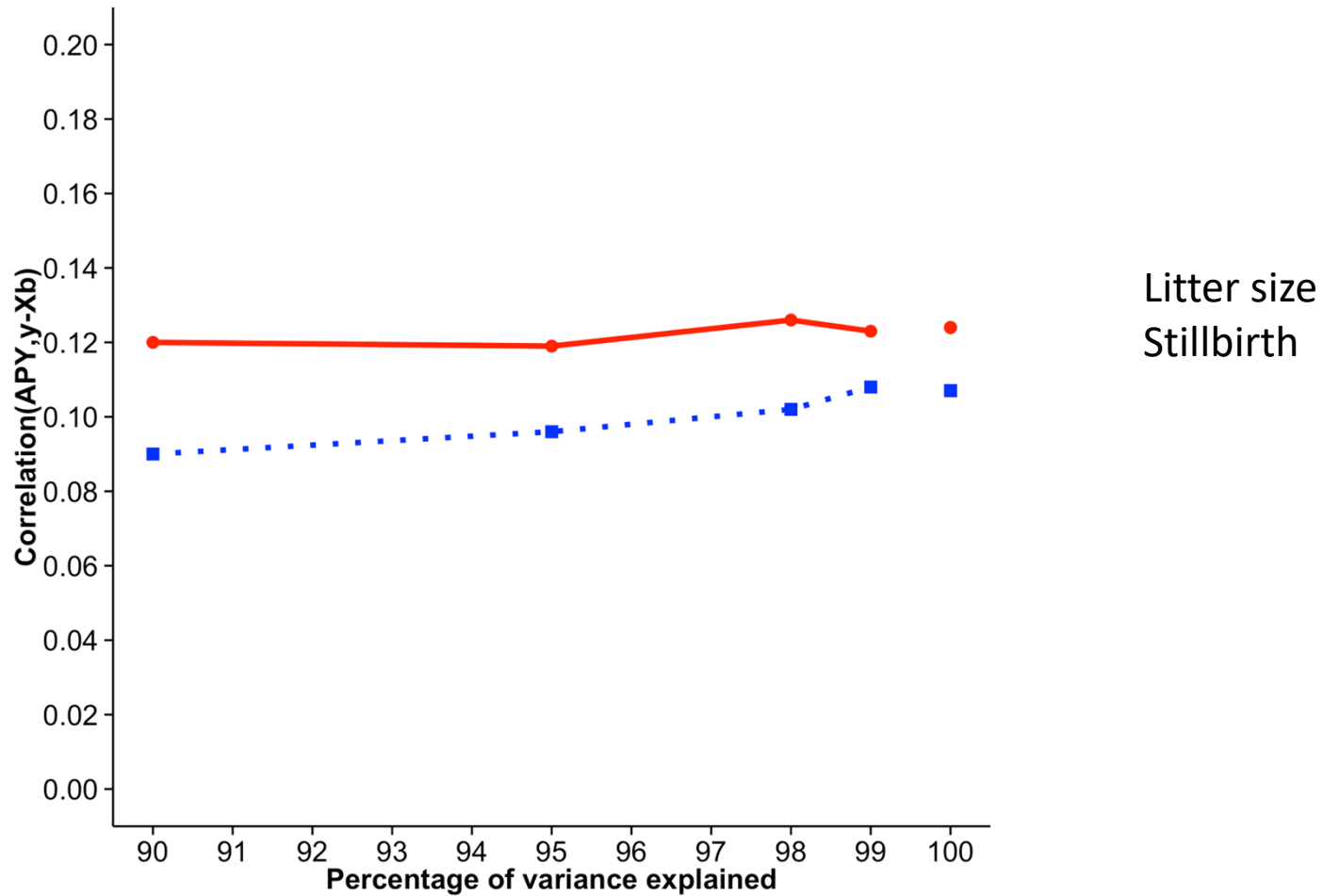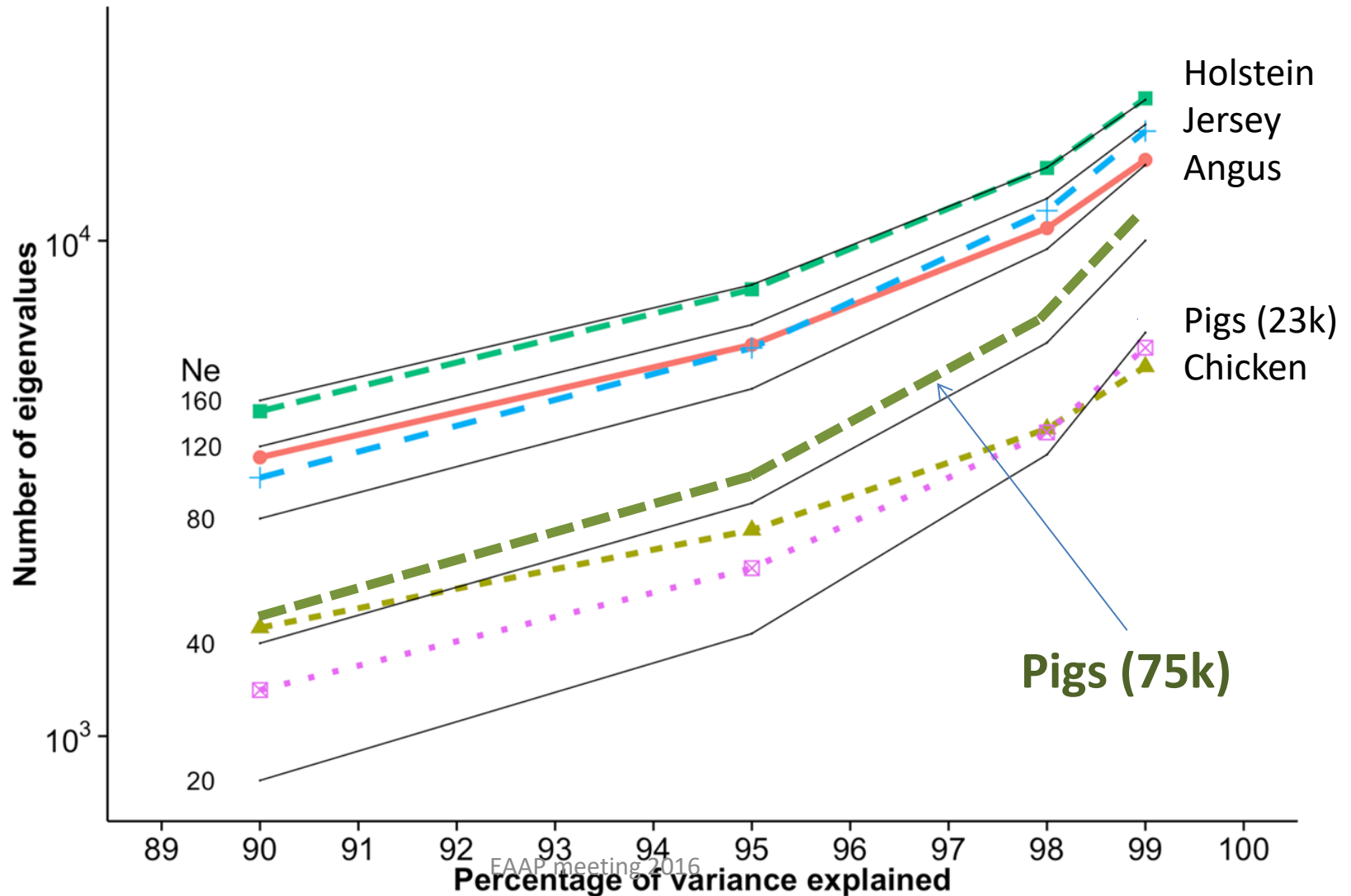
Predictability - Broilers
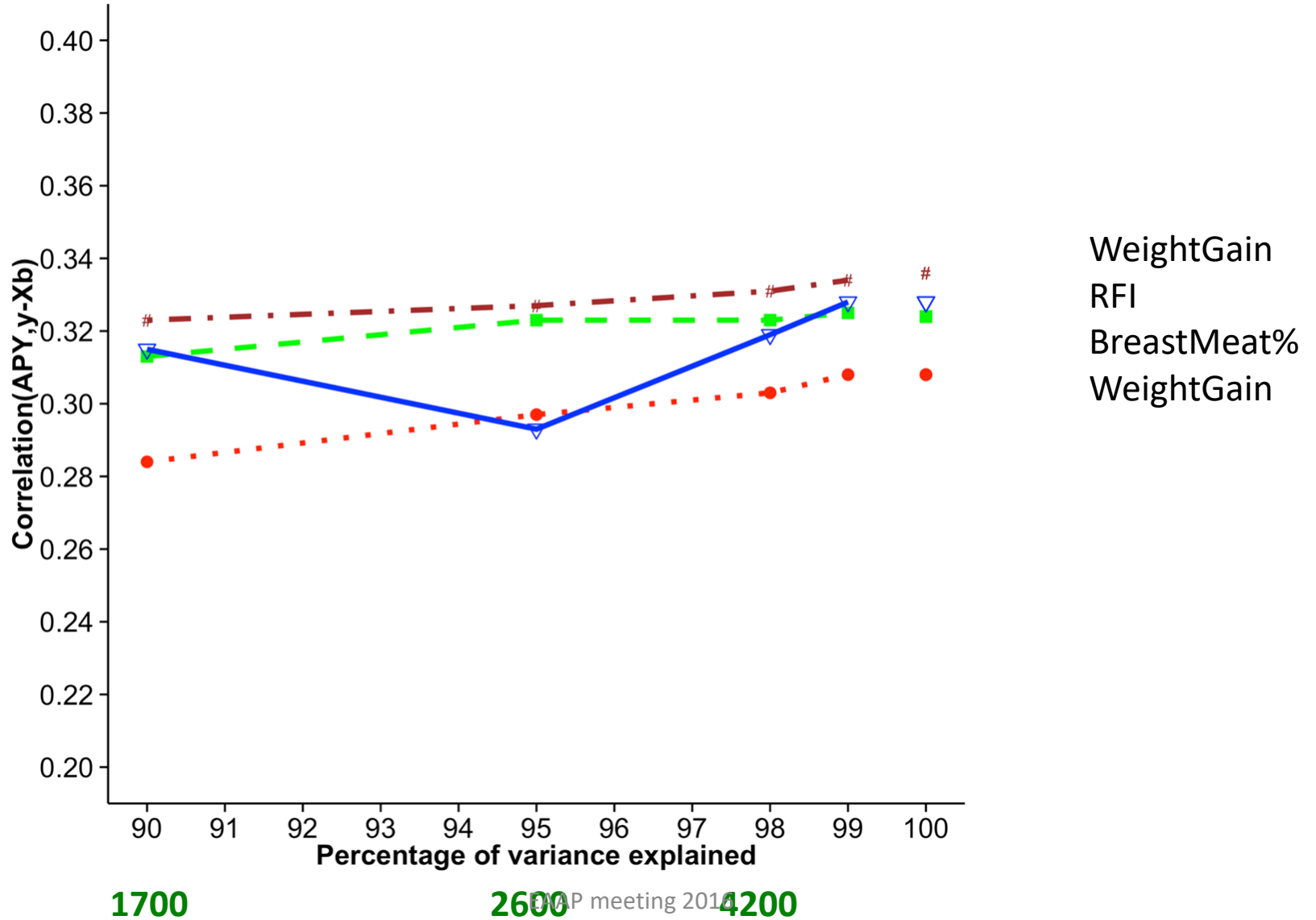
Number of eigenvalues in G to explain given fraction of variability

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
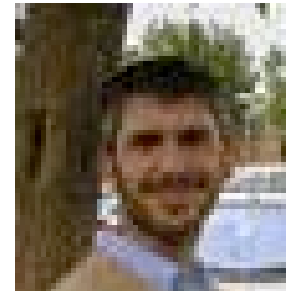
Breno Fragomeni

Ivan Pocrnic

Daniela Laurenco

Yutaka Masuda

Andres Legarra

EAAP meeting 2016