Selected projects at University of Georgia

Ignacy Misztal and Daniela Lourenco

Recent projects

- Blupf90 software
  - Convergence improvements
    - Case of #phenotypes << # animals
  - Multiple categorical traits with large data
  - P-values in GWAS with national data sets

- Applications
  - Implementation in dairy
  - Canalization for disease resistance

- Potential negative effects of genomic selection
  - Parameter estimation with large data
- Improvement of accuracies with sequence data
- Explaining peculiarities of GWAS

WHY GWAS IN UGA / BLUPF90 PROGRAMS

Large research interest in GWAS

Limitations for current methods

- Simple models
- Single trait
- Complicated if not all animals genotyped

Can ssGBLUP be used for GWAS?

Genet. Res., Camb. (2012), **94**, *pp.* 73–83. © Cambridge University Press 2012
doi:10.1017/S0016672312000274

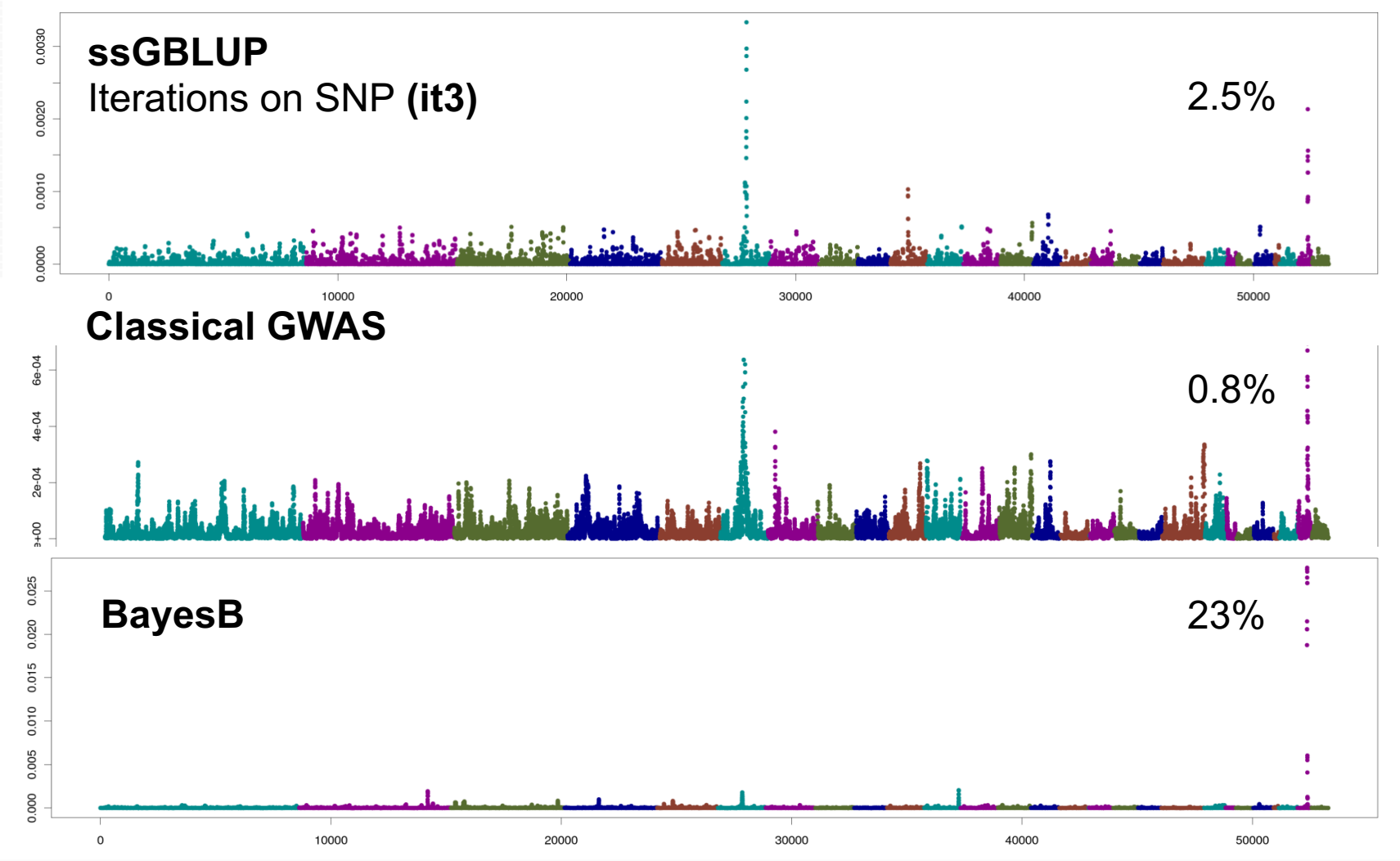
Genome-wide association mapping including phenotypes from relatives without genotypes

H. WANG[1]*, I. MISZTAL[1], I. AGUILAR[2], A. LEGARRA[3] AND W. M. MUIR[4]
[1] *Department of Animal and Dairy Science, University of Georgia, Athens, GA 30602-2771, USA*
[2] *Instituto Nacional de Investigación Agropecuaria, INIA Las Brujas, 90200 Canelones, Uruguay*
[3] *INRA, UR631 Station d'Amélioration Génétique des Animaux (SAGA), BP 52627, 32326 Castanet-Tolosan, France*
[4] *Department of Animal Science, Purdue University, West Lafayette, IN 47907-1151, USA*

(*Received 19 September 2011; revised 8 December 2011, and 9 March 2012; accepted 13 March 2012*)

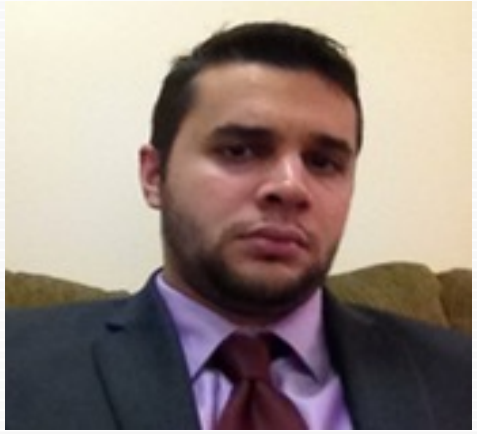
Discrepancies in GWAS methods

Chicken weight



Manhattan plots by % variance explained by SNP windows

INCLUDING SEQUENCE DATA IN US HOLSTEINS



4M records for Stature

3M Cows

4.6M Animals in pedigree

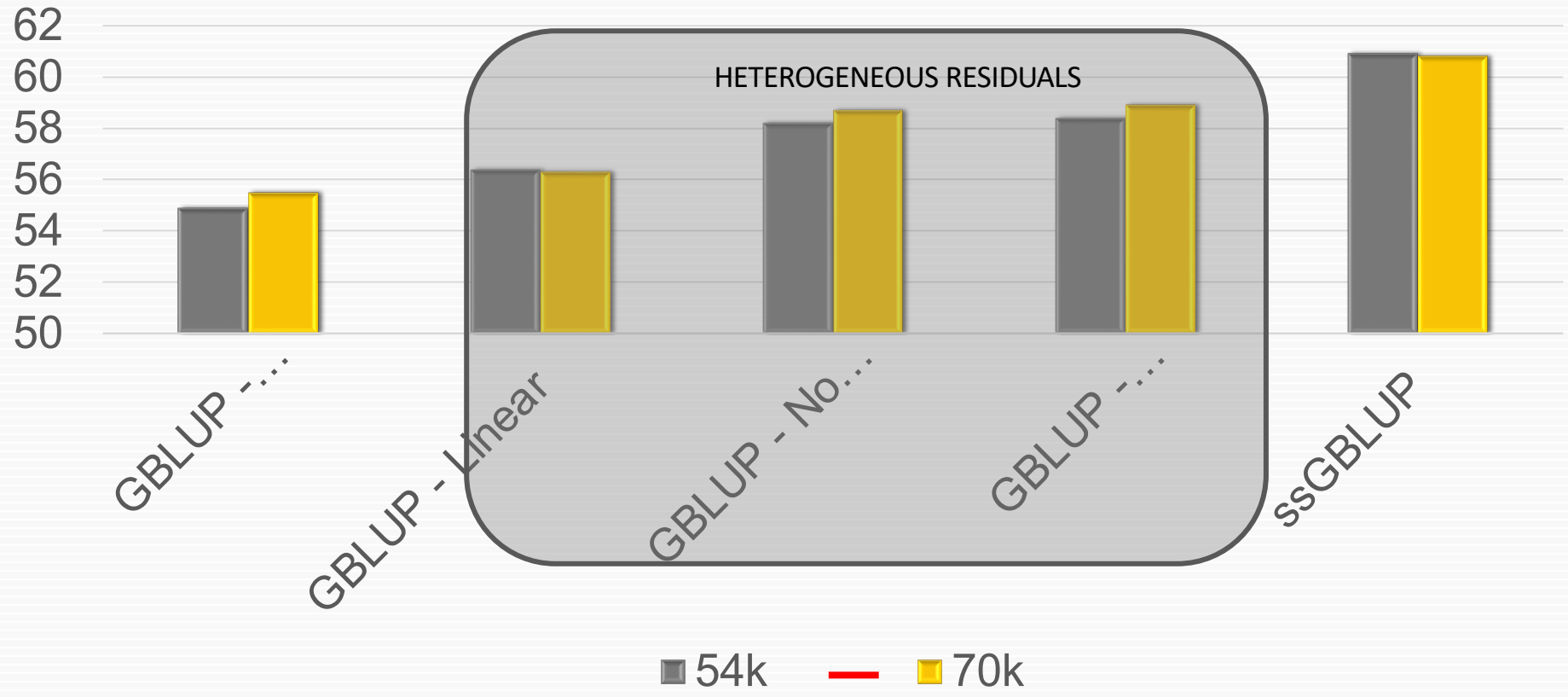
27k Genotyped Sires

54k SNP

54k SNP + 17k Causative Variants (VanRaden et al., 2017)

Fragomeni et al. (2019)

RELIABILITIES WITH DIFFERENT METHODS AND SNP SETS



P-values for GWAS in (ss)GBLUP

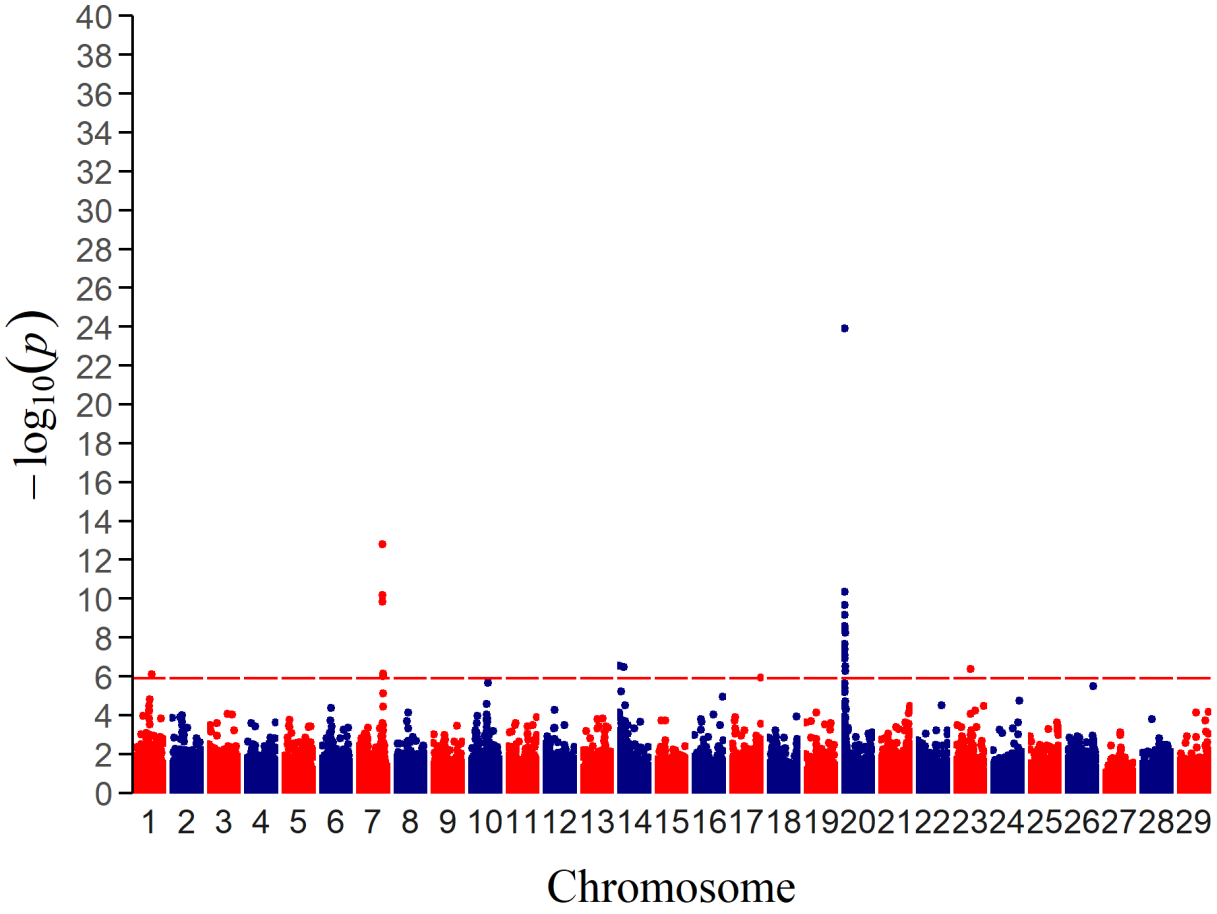
$$pval_i = 2\left(1 - \Phi\left(\left|\frac{\widehat{snp}_i}{sd(\widehat{snp}_i)}\right|\right)\right)$$ (Chen et al., 2017)

If $sd(\widehat{snp}_i)$ approximately constant, Manhattan plots based on $|\widehat{snp}_i|$ and $pval_i$ similar

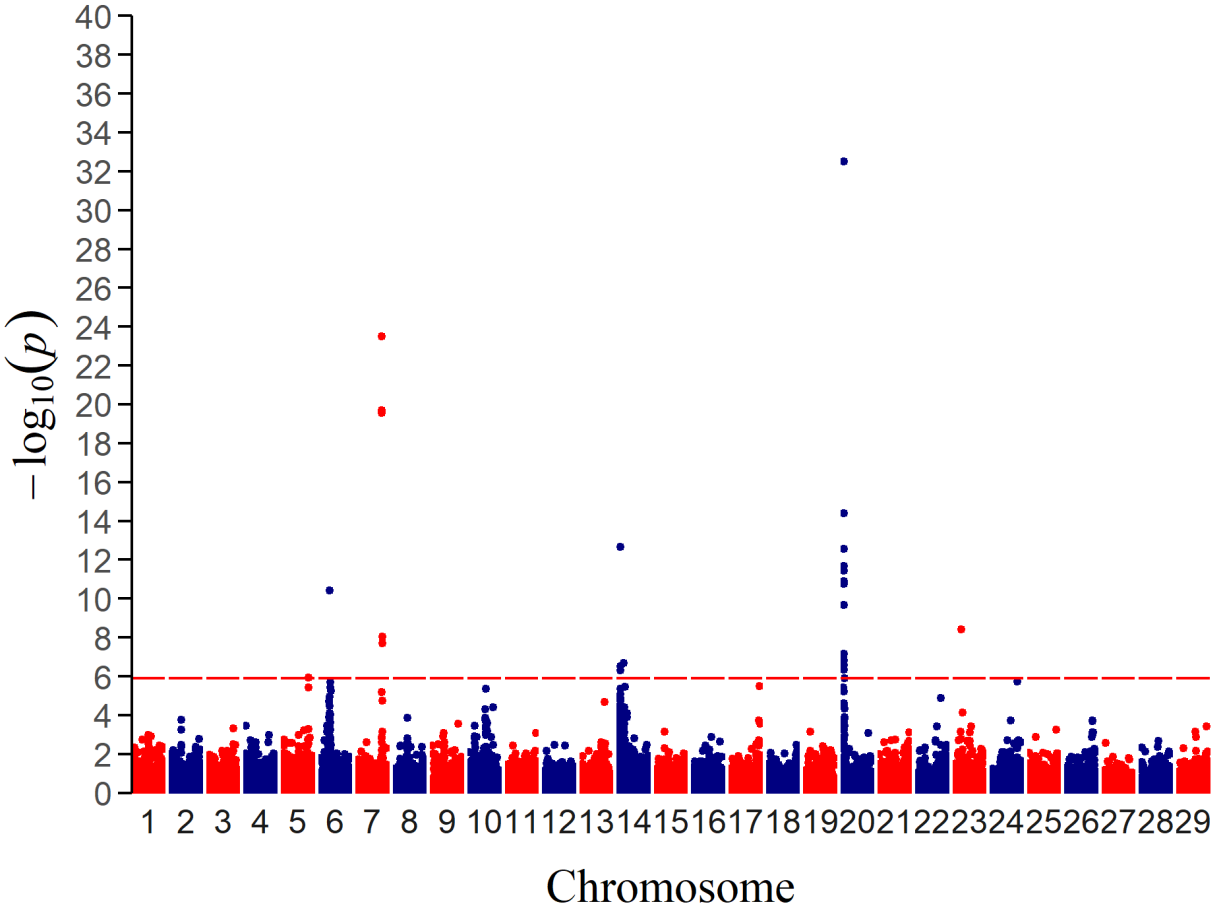
Large data – PEV from accuracy approximations based on APY algorithm(Bermann et al., 2021)

Post-weaning gain in American Angus

50k genotyped animals



500k genotyped animals



Sequence project at Roslin Institute

- Contracts with major companies (including PIC and COBB)
- Partly gov’t supported
- Headed by John Hickey
- 20 students and postdocs

- Steps
  - Imputation to sequence
  - Analyzes

Largest pig sequence data

Total = 379k

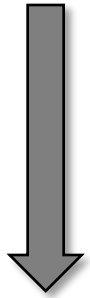
Terminal lines

Jang et al. (2023)
Jang et al. (accepted)

Lines	ADG	BF	ADGX	BFX	Animals in pedigree	Sequenced/ Imputed

Sequence Variants

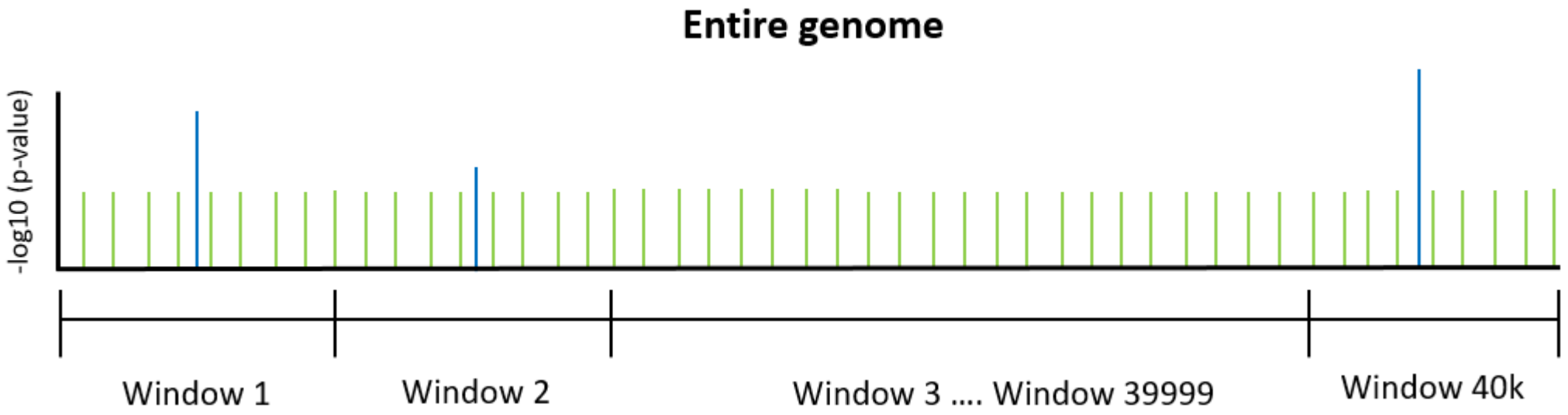
15M to 20M variants ~ 10M segregated across lines



Should we use all 10M?

SNP preselection based on GWAS - I

- **Top 40k**

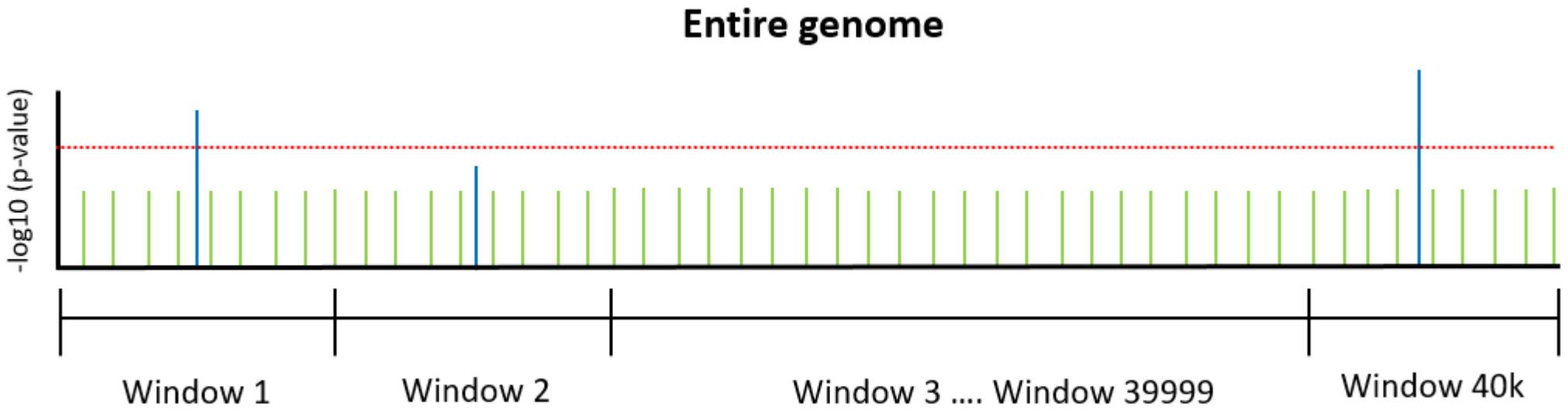


Extracting only 40k SNP: Similar number as the regular SNP chip (~40k)

Ros-Freixedes et al., 2022

SNP preselection based on GWAS - II

- **Chip+Sign**



Extracting only significant ones + 40k SNP chip

Ros-Freixedes et al., 2022

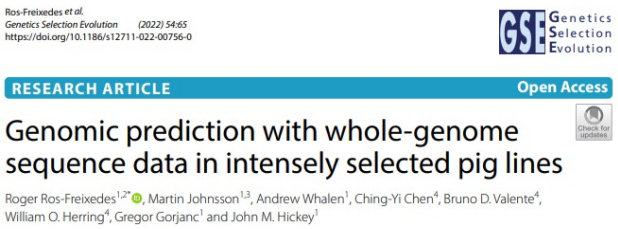
Steps

1) Accuracy of GEBV with SNP preselected from sequence data

- Many animals with sequence

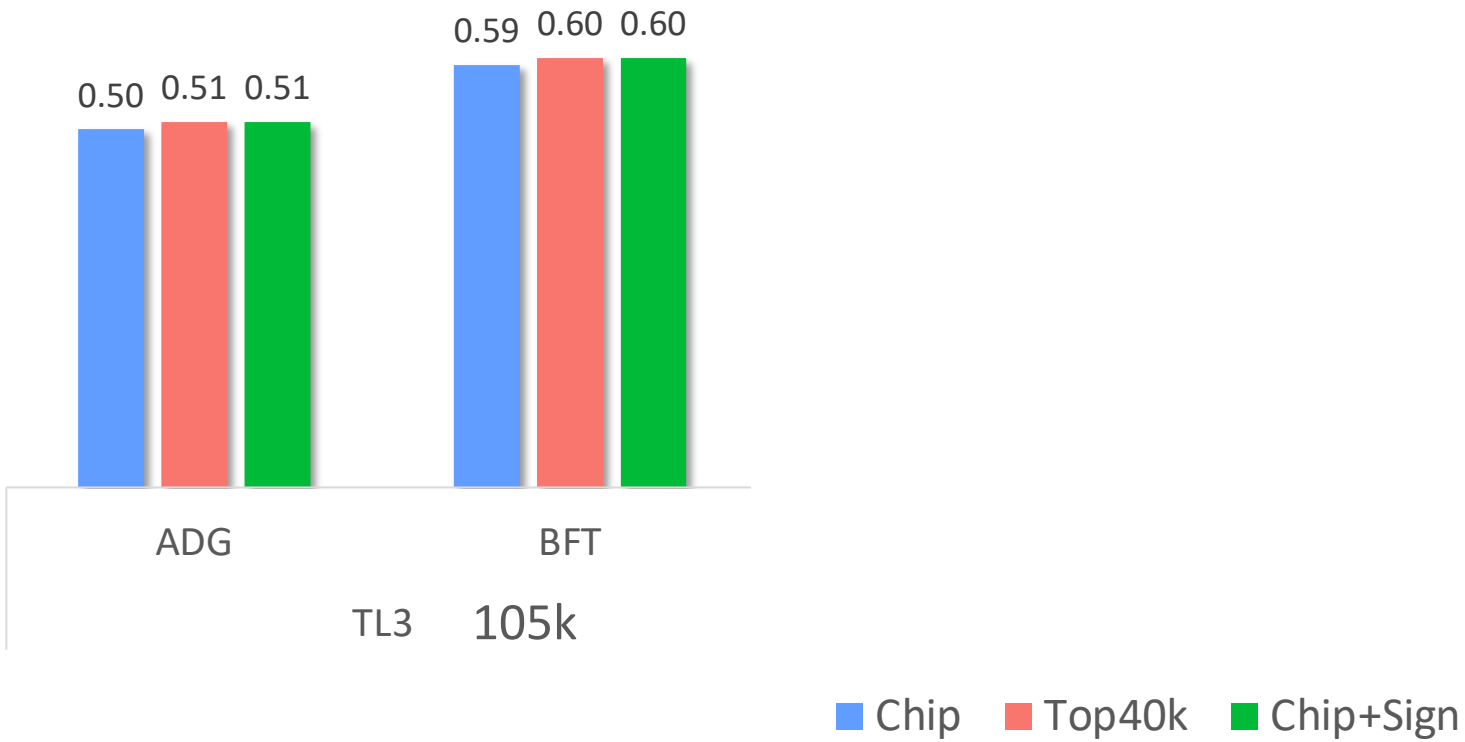
2) Single-line and multi-line ssGBLUP evaluations

3) Compare ssGBLUP with BayesR from Roslin



Step 1 – Accuracy with preselected variants

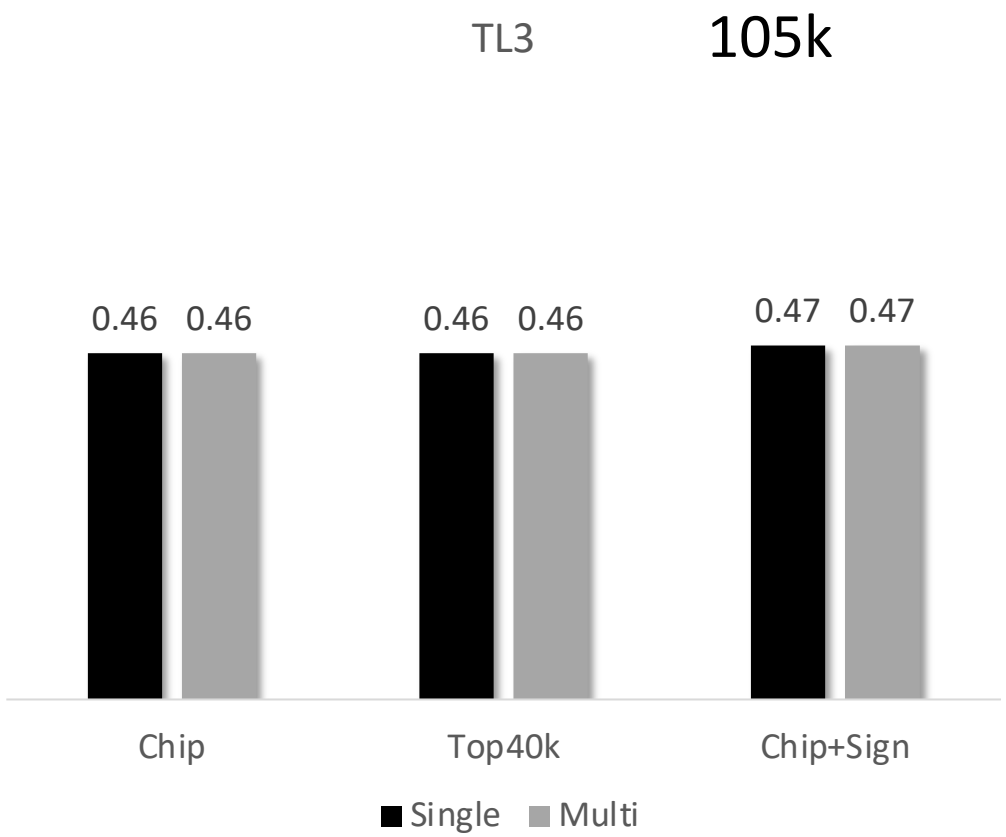
- Prediction accuracy = cor(DEBV, GEBV)

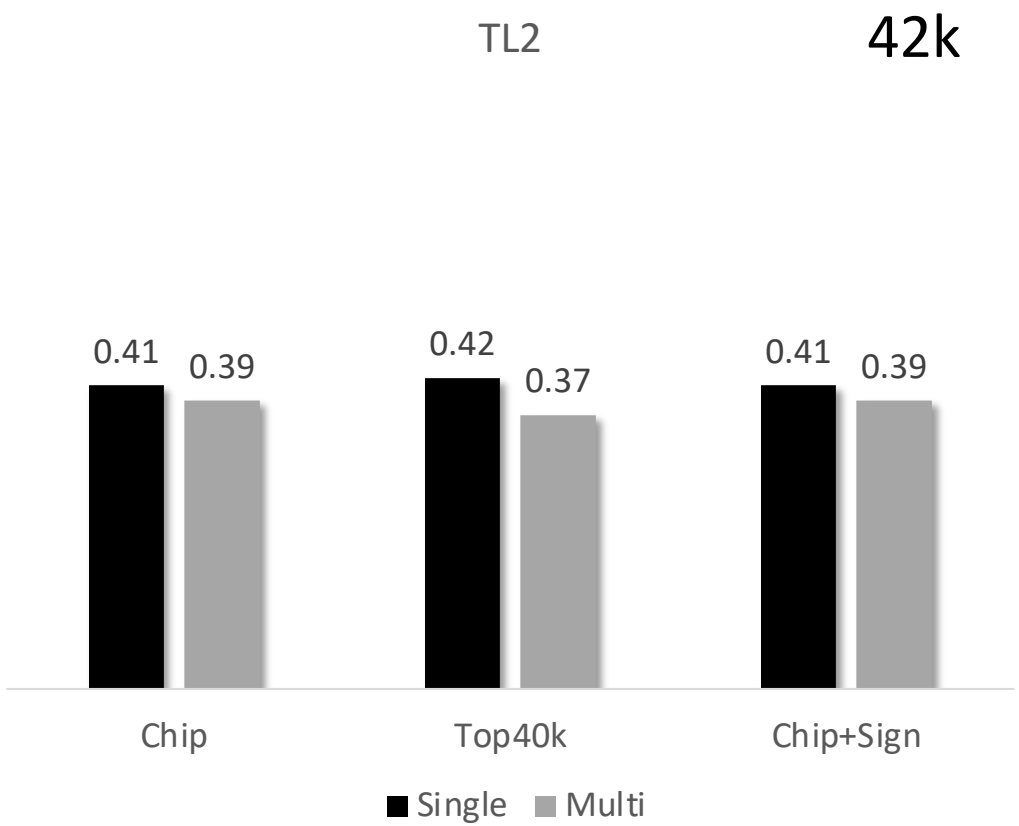


Step 2 – Single vs. Multi-line all traits

- Prediction accuracy = cor(DEBV, GEBV)

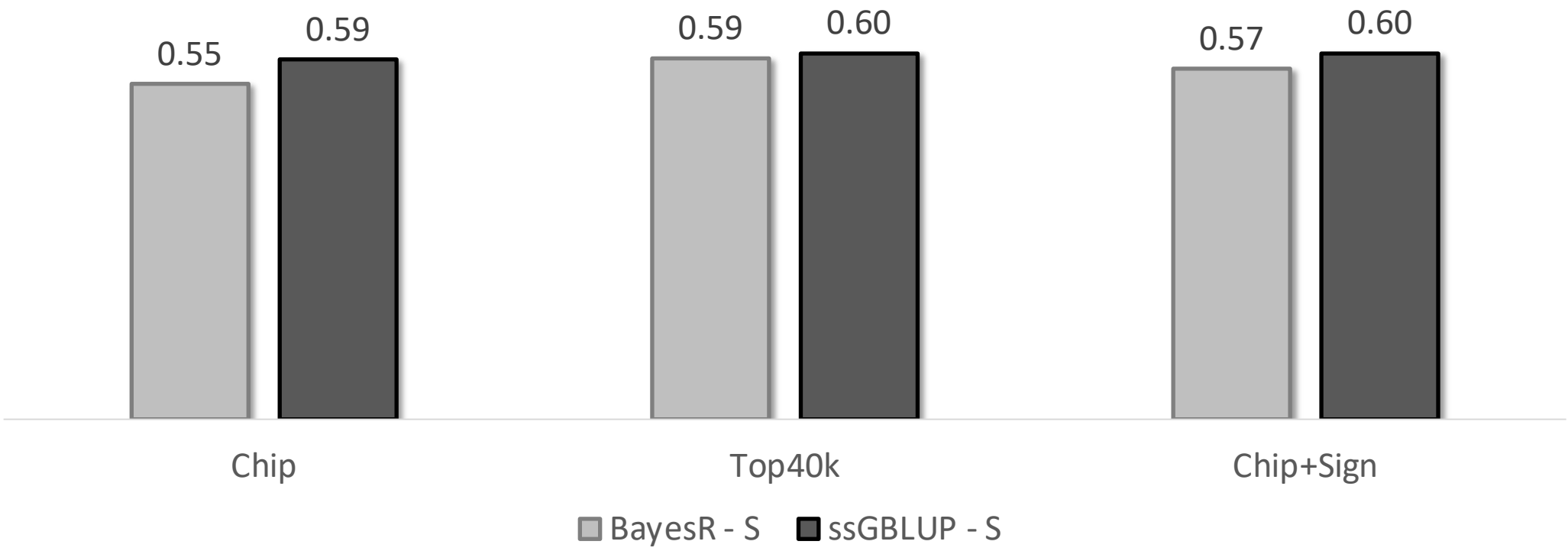
Multi-line GWAS and predictions dominated by TL3





Predictions must be not QTL oriented!

Step 3 - ssGBLUP vs. BayesR

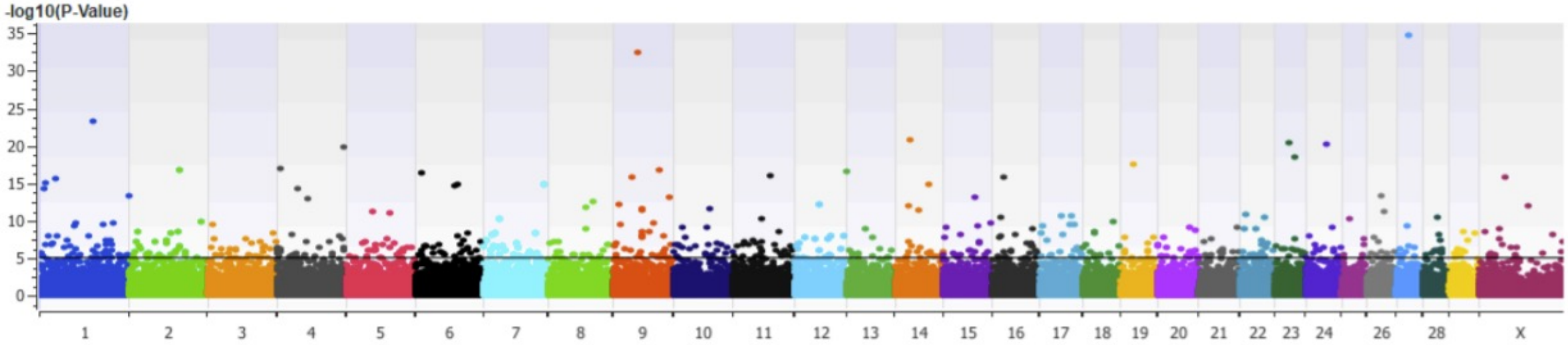


BayesR from Ros-Freixedes et al., 2022

Questions with GWAS and predictions

- Little or no gain with sequence data for ssGBLUP with commercial data
- GWAS by
  - % of variance explained usually per 1Mb
  - p-values
- Few regions explain > 1% additive variance
- Lots of QTLs detected with small data sets
- Fewer QTLs detected with large data

First conception rate on 2k Holstein heifers

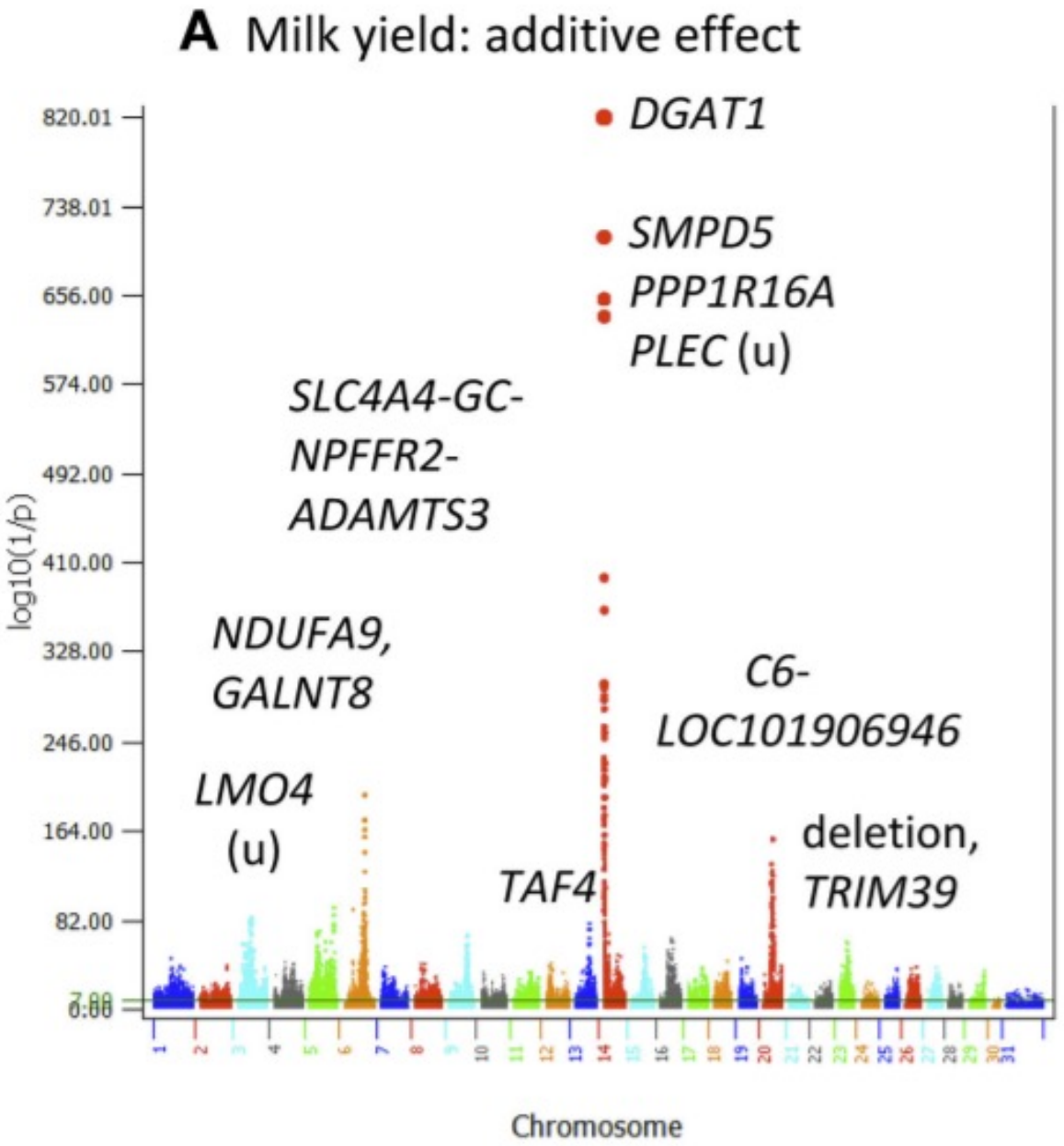


Estimated heritability 36% (normally 1%)

Identified 146 unique loci at $p < 5 \times 10^{-8}$ level

Galliou et al., 2020, https://doi.org/10.3390/genes11070767

GWAS on 294k Holstein cows

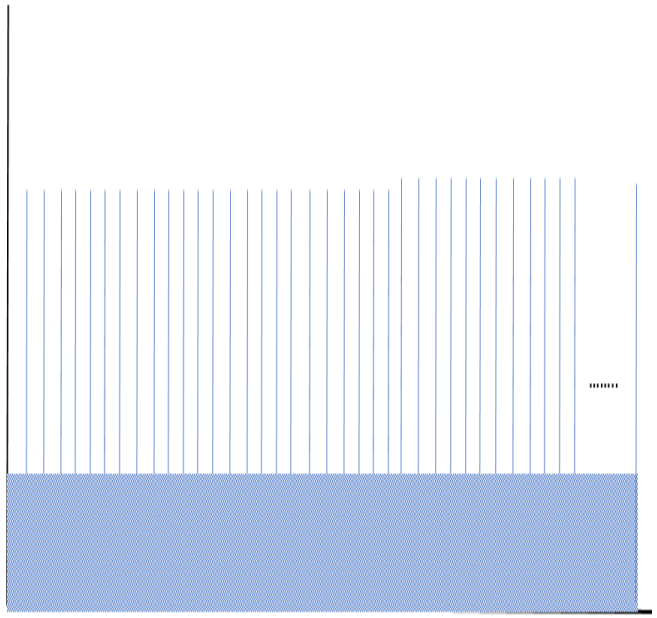


Jiang et al., 2019

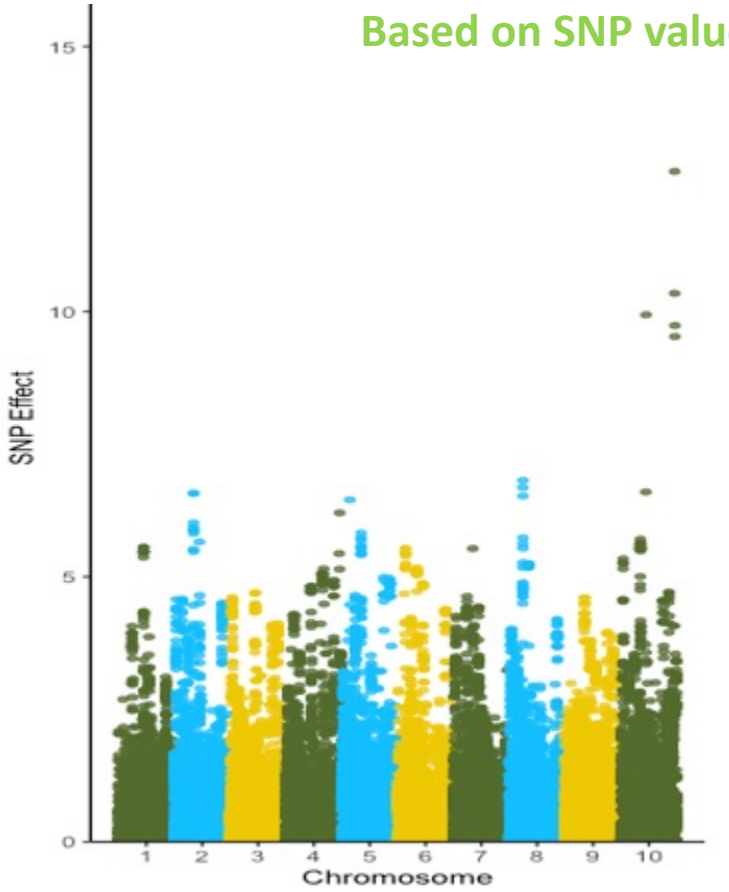
doi: 10.3389/fgene.2019.00412

Manhattan plots for simulated population with 100 identical equidistant QTNs

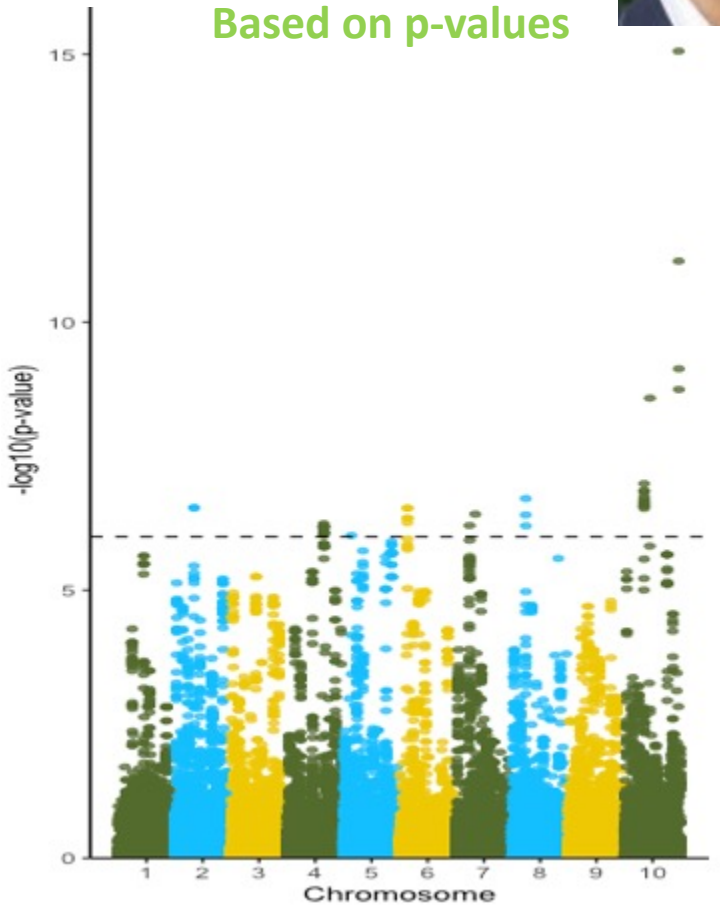
Expectation



Based on SNP values

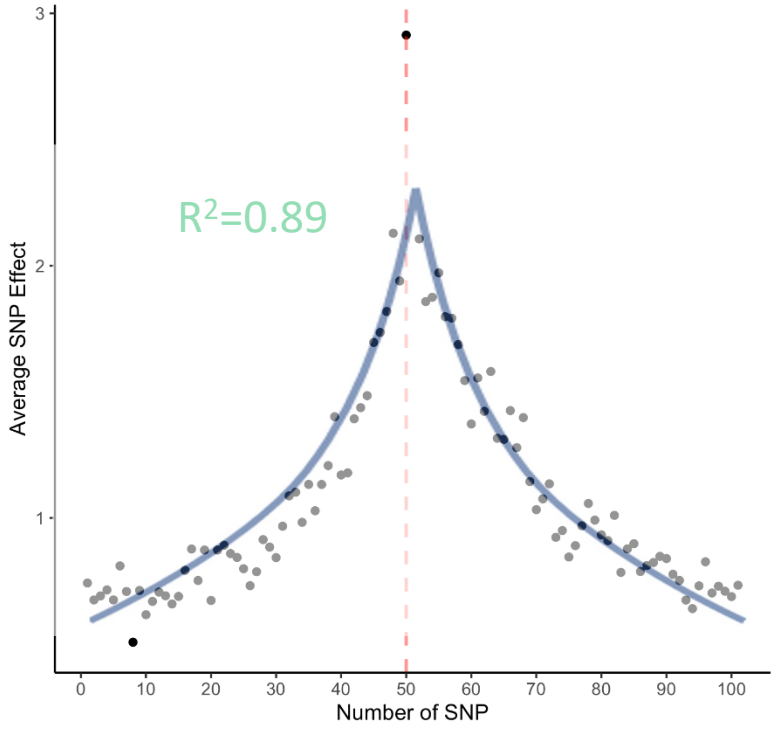


Based on p-values



Work started by Pocrnic et al. (2018)

Plots averaged for 100 QTN



Pairwise linkage disequilibrium curve

~ 2 Mb for cattle

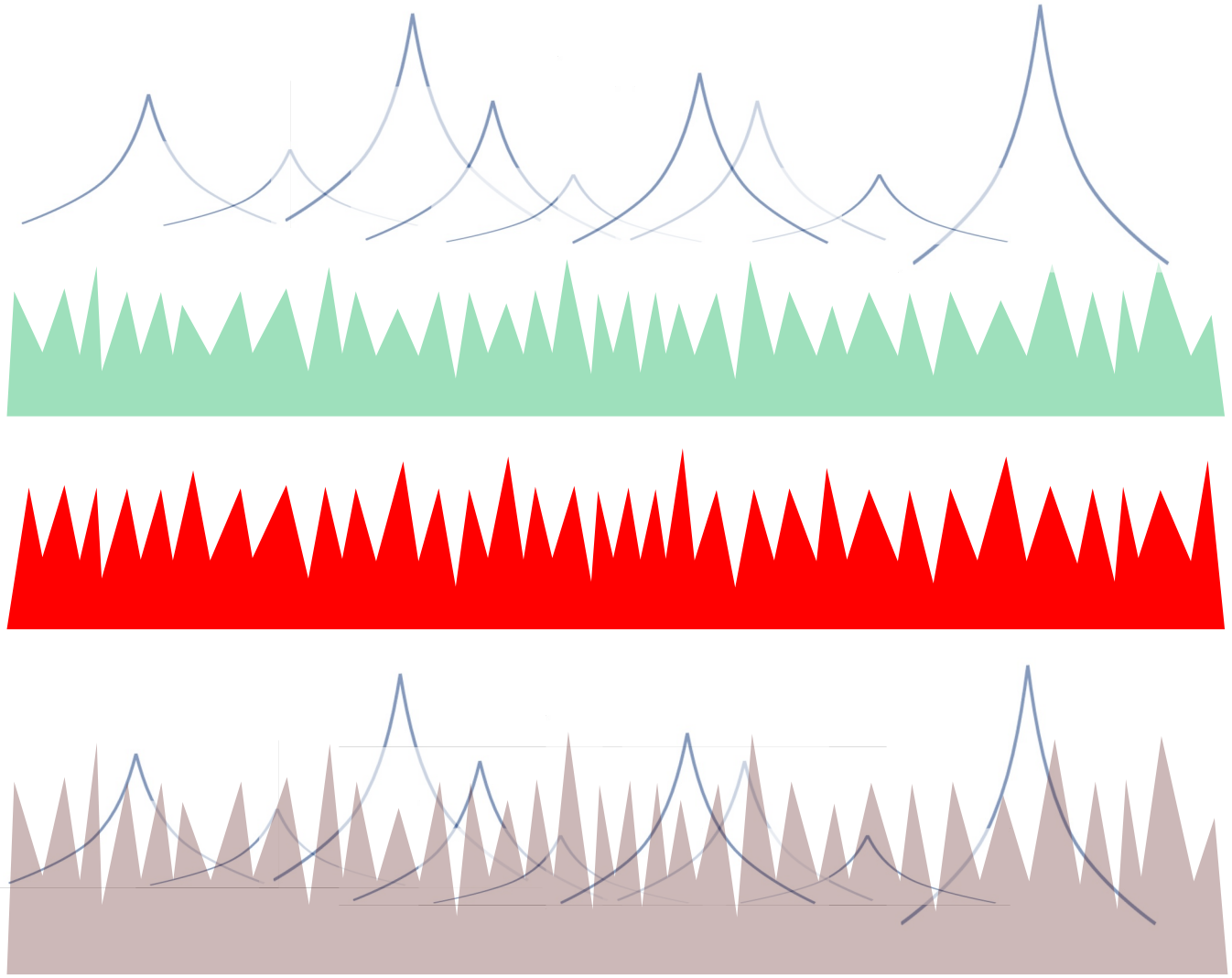
~ 5 Mb for pigs/chickens

~ 15 kb for humans

1/Ne Morgans for 80% QTN variance

Ne - effective population size

What is Manhattan plot composed of?



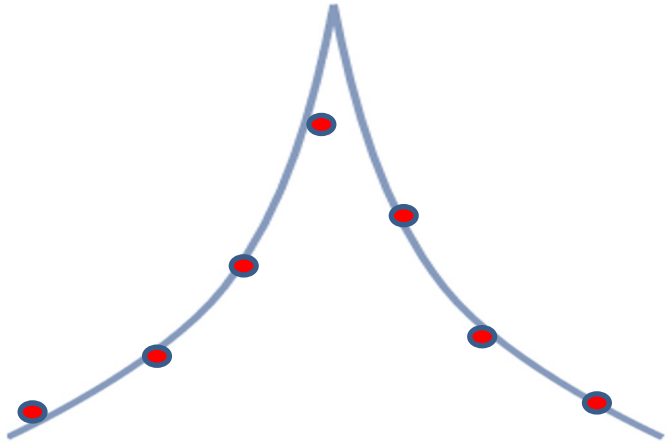
QTNs **Bigger with larger QTN and larger data**

Relationships

Noise **Smaller with more data**

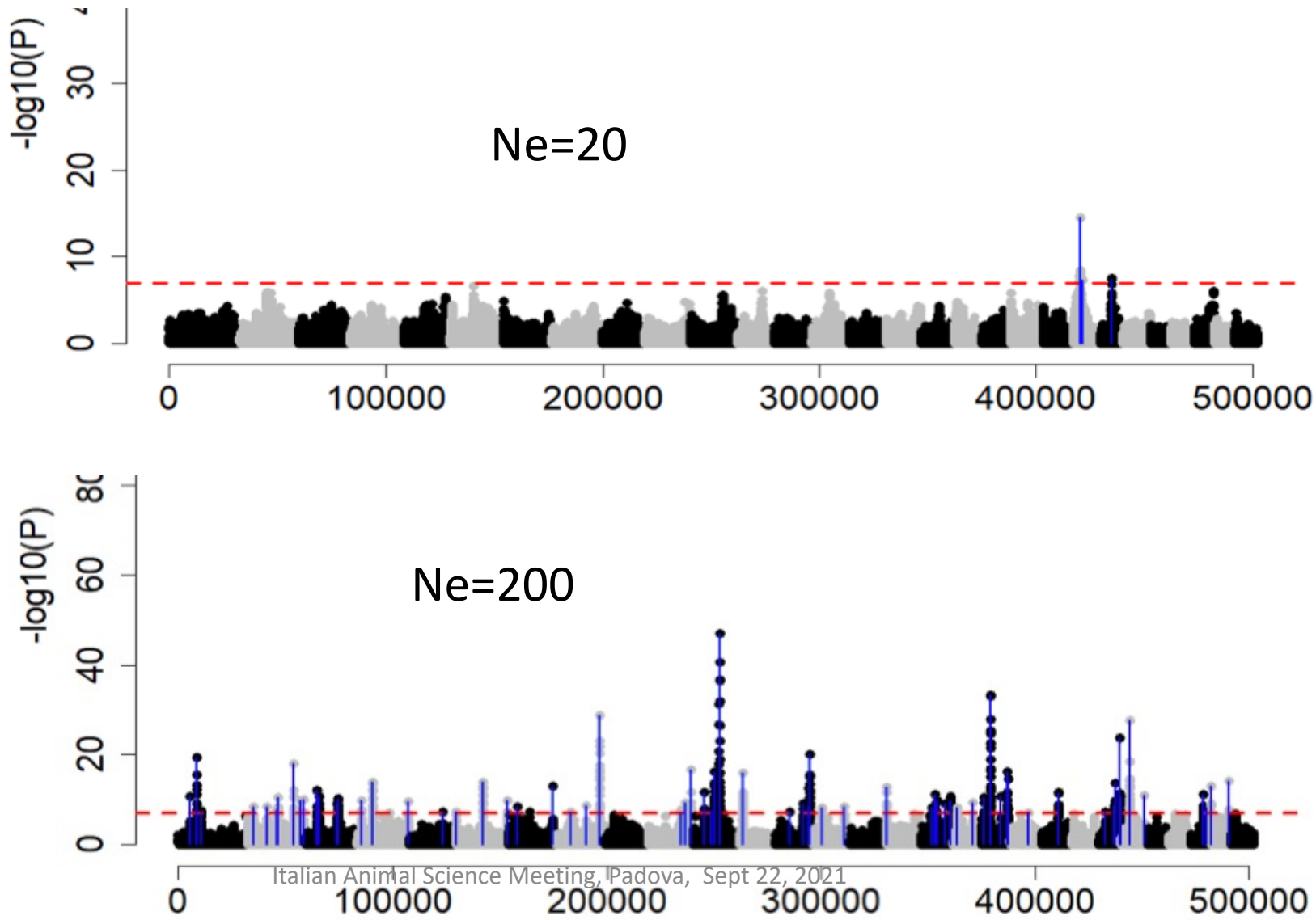
Combined

Why ssGBLUP accounts for QTN?



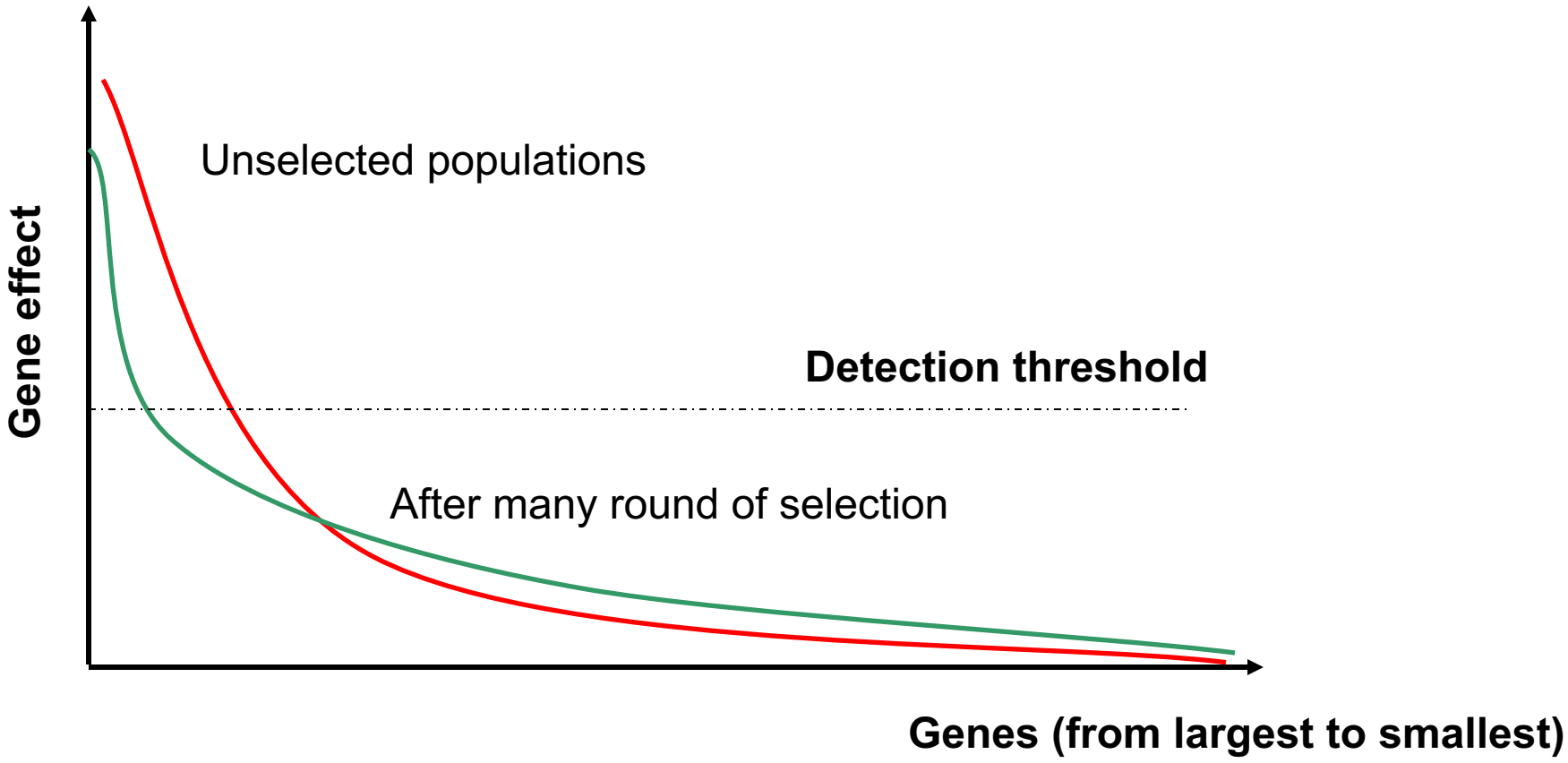
SNPs cover QTN LD curve

Effective population size affects GWAS

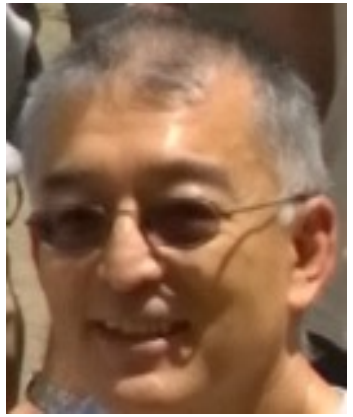


Sungbong et al., 2021

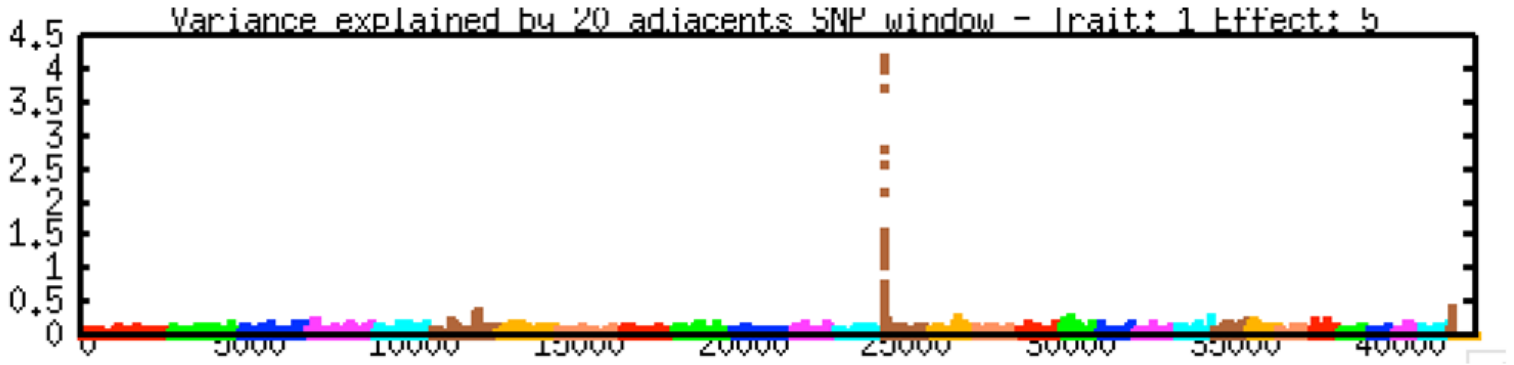
Distribution of QTL effects



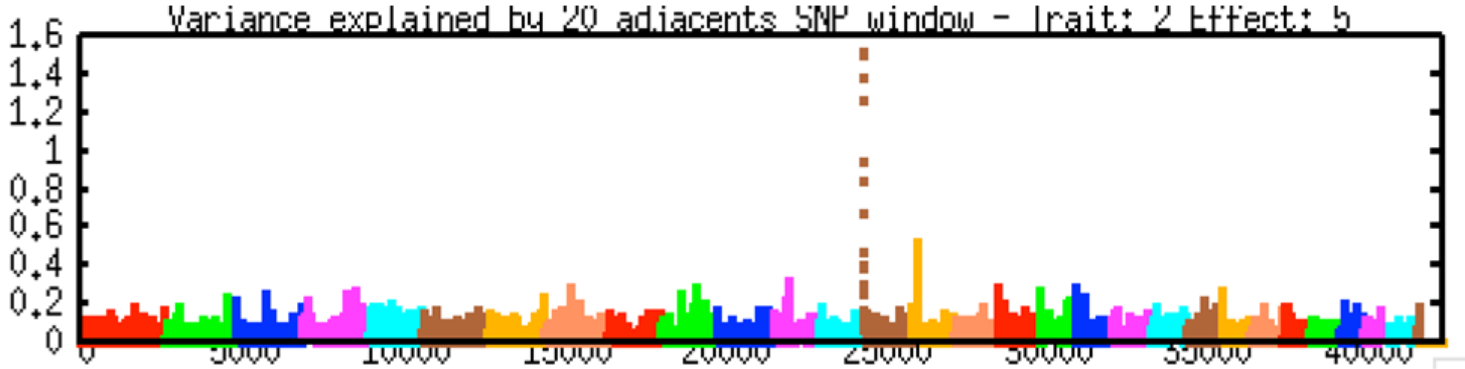
GWAS using 35k Holstein bulls



Milk – first parity



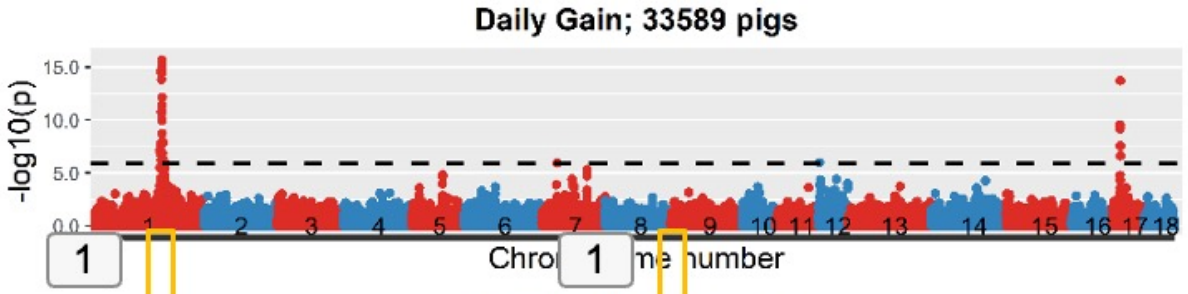
Mortality – first parity

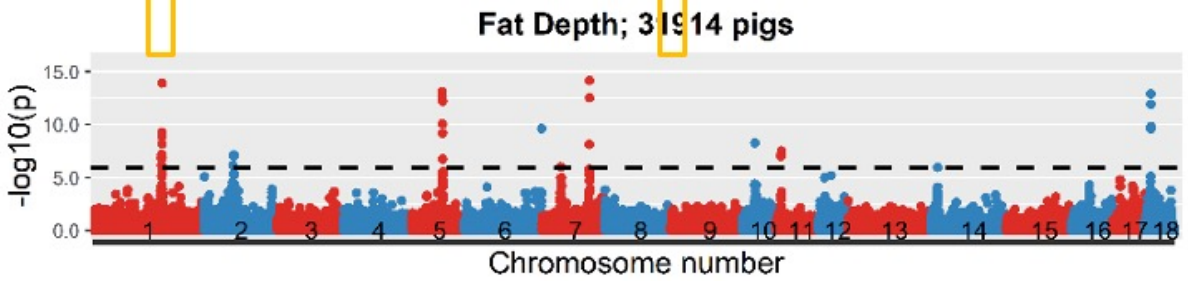


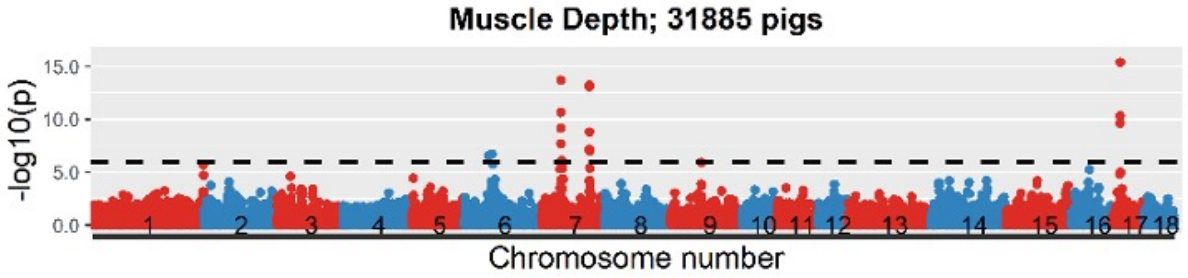
(Tokuhisa et al, 2014; Tsuruta et al., 2014)

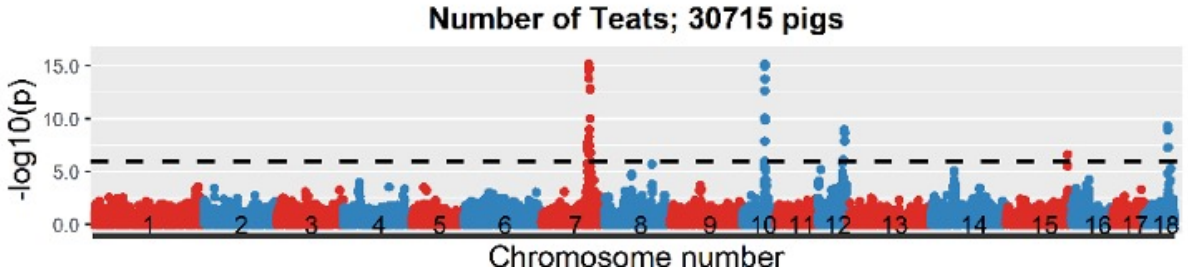
GWAS for various traits and index in pigs

Bijma, EAAP 23

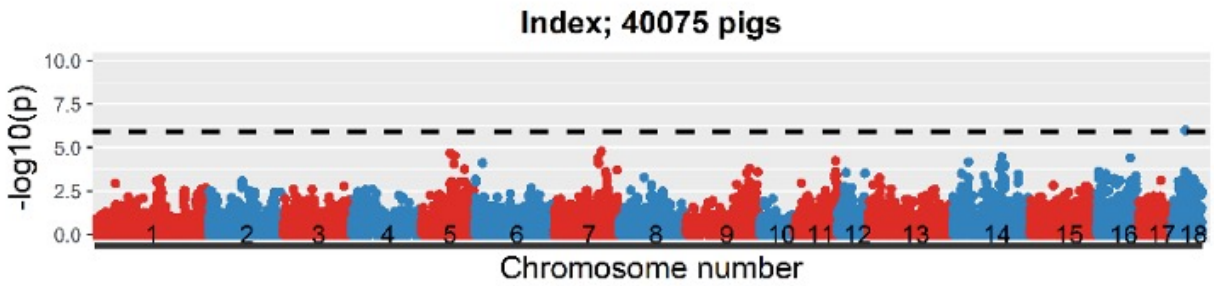








Index



- Different peaks in different lines
- Antagonistic pleiotropy

Conclusions for GWAS

- QTN profile wide with small effective population size
- Large signals in GWAS due to QTN, relationships and noise (incl. Imputation)
  - If no LD curve, probably false signal
- Large QTL show pleiotropy – QTL not visible in index
- ssGBLUP accounts for QTL with large data

Possibly Negative Impact of Genomic Selection

Ignacy Misztal and Daniela Lourenco

University of Georgia

Negative effects of genomic selection

- Informal industry reports:
  - Deteriorating sow survival and pig mortality in pigs
  - Deteriorating feet & legs in beef
  - Short teats and increased calf mortality in dairy
  - Increased sensitivity to heat stress in dairy
  - Deteriorating disease resistance across species

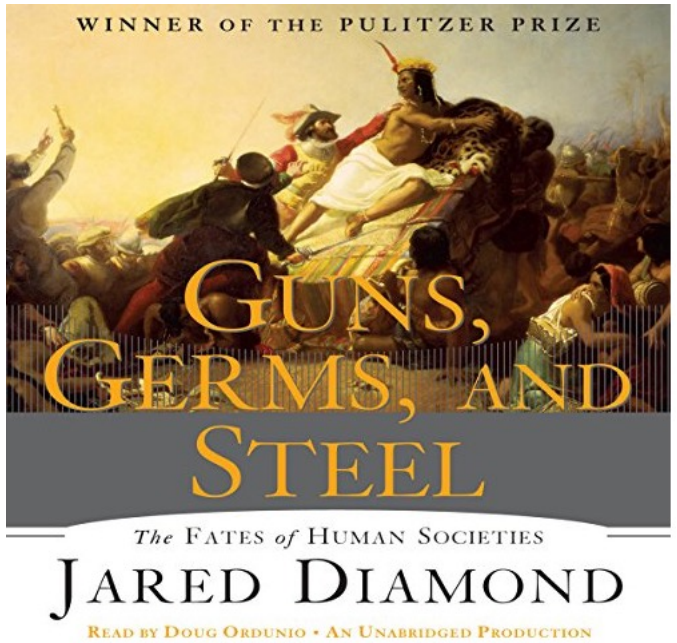
Genetic selection as optimization

- Selection for one trait or an index
- Gains on selected traits
- Losses on correlated antagonistic traits

- Losses compensated by improved environment/management

History of selection strategies

- Domestication
- Unformal
- Large-scale single-trait for production traits
- Multi-trait with fitness traits
- Genomic



Livestock Production Science
Volume 93, Issue 1, 1 April 2005, Pages 3-14

Go to Livestock Production Science on ScienceDirect

Genetics of adaptation and domestication in livestock ☆

Sandrine Mignon-Grasteau [a], Alain Boissy [b], Jacques Bouix [c], Jean-Michel Faure [a], Andrew D. Fisher [d], Geoffrey N. Hinch [e], Per Jensen [f], Pierre Le Neindre [b], Pierre Mormède [g], Patrick Prunet [h], Marc Vandeputte [i], Catherine Beaumont [a]

Domestication

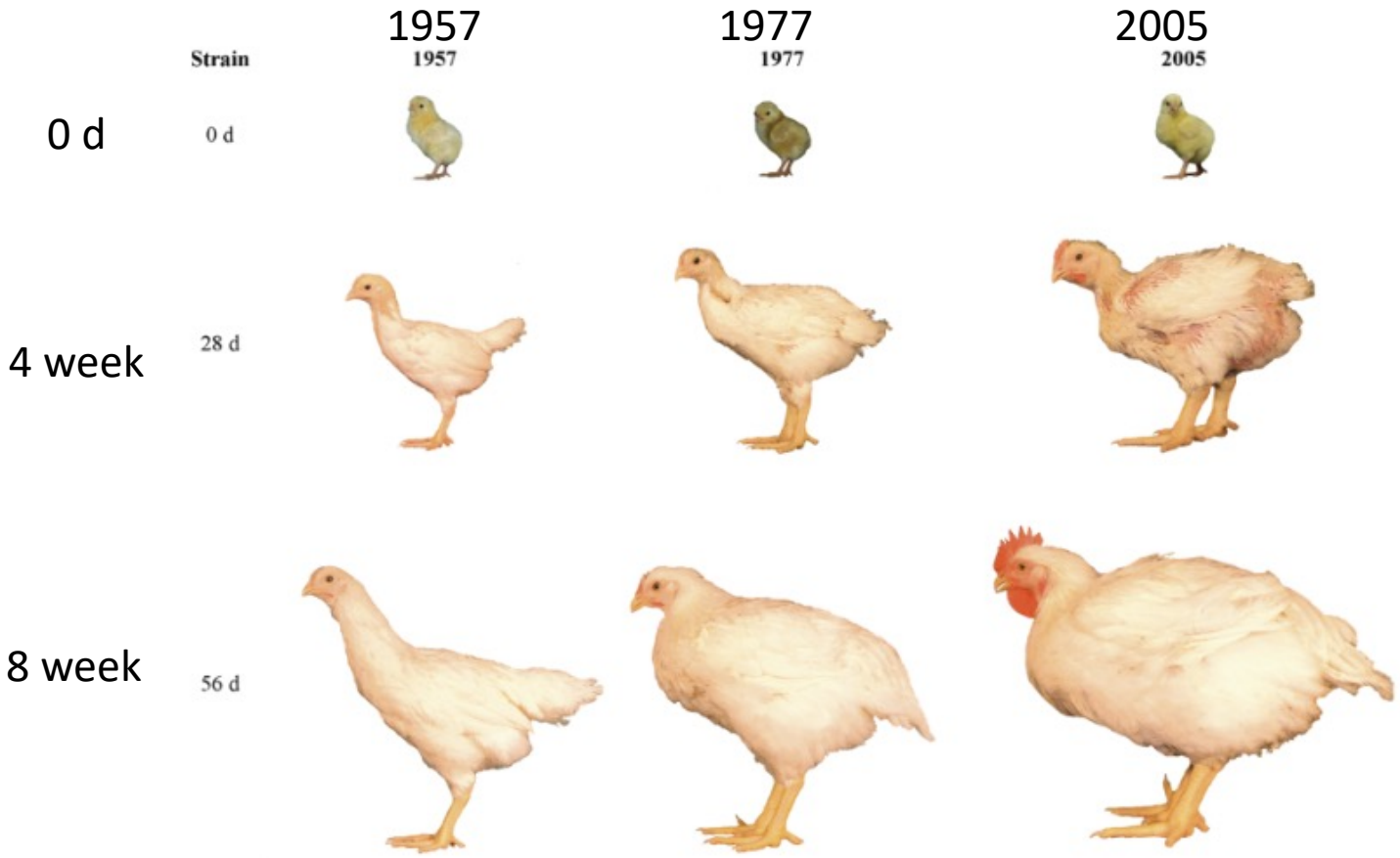
Winners

Growth
Milk
Mating procedures

Losers

Food finding
Seasonal reproduction
Predator avoidance
Brain size
…

Example of effects of mostly single-trait selection



Zuidhof et al. (2014) http://dx.doi.org/10.3382/ps.2014-04291

Side effects of intensive selection for growth in broiler chicken

- Unlimited appetite / obesity ➔ artificial lightning
- Poor survival of males ➔ male supplementation
- Increased susceptibility to diseases ➔ antibiotics
- Low hatchability ➔ alternate heating/cooling of incubators
- …

All companies – similar problems at same time

Initially problems kept confidential

Eitan and Soller, 2014

Undesirable side effects of selection for high production efficiency in farm animals: a review

W.M. Rauw[a,*], E. Kanis[b], E.N. Noordhuizen-Stassen[c], F.J. Grommers[c]

[a]*Department of Animal Science, Agricultural University of Norway, P.O. Box 5025, 1432 Ås, Norway*
[b]*Animal Breeding and Genetics Group, Wageningen Institute of Animal Science, Wageningen Agricultural University, P.O. Box 338, 6700 AH Wageningen, The Netherlands*
[c]*Department of Herd Health and Reproduction and Interdepartmental Section Veterinary Medicine and Society, University of Utrecht, P.O. Box 80151, 3508 TD Utrecht, The Netherlands*

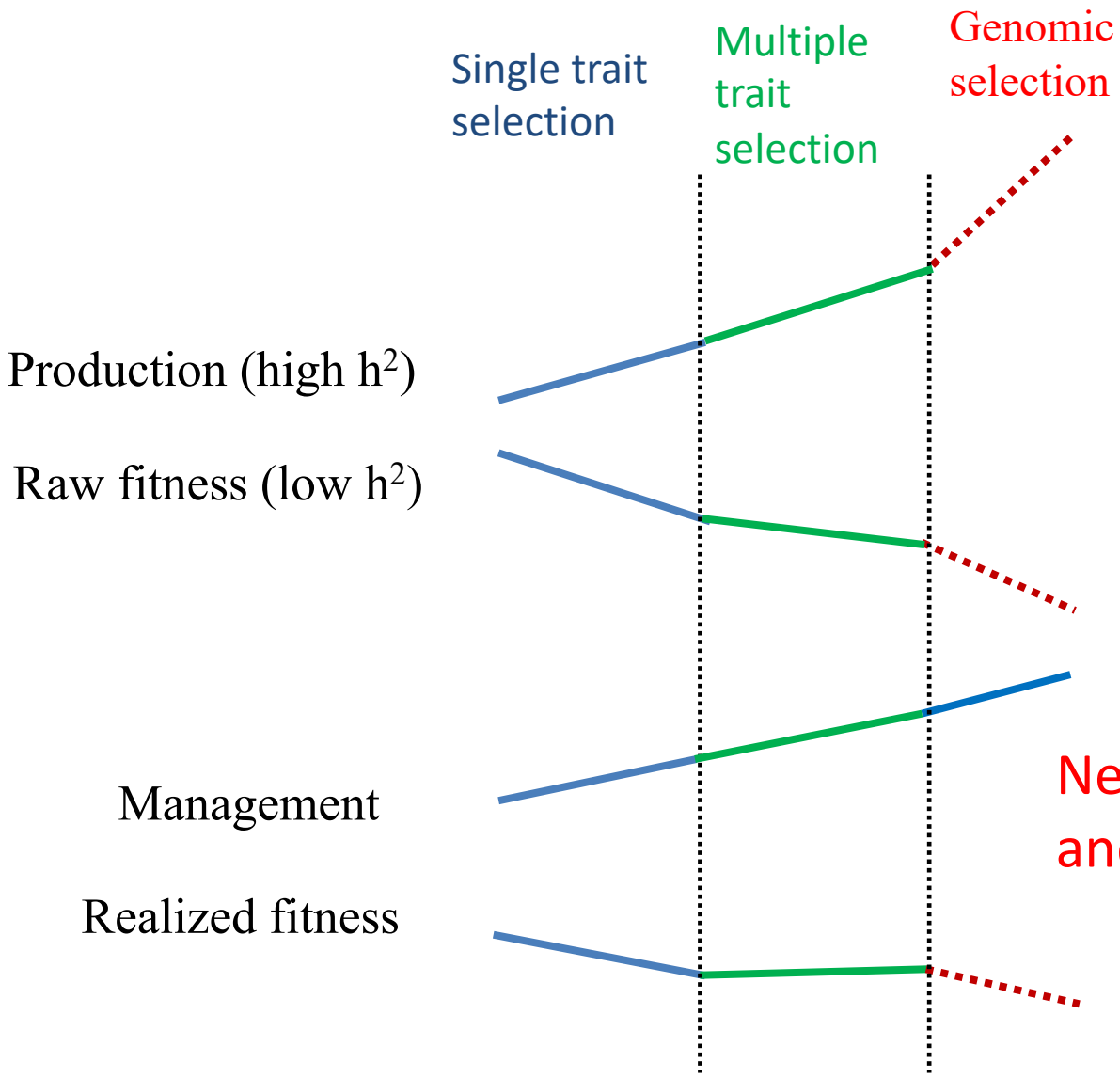
Received 4 July 1997; accepted 29 April 1998

…over 100 references on undesirable(cor)related effects of selection … in broilers, pigs and dairy cattle….

Future application … DNA-techniques .. ….more dramatic consequences….

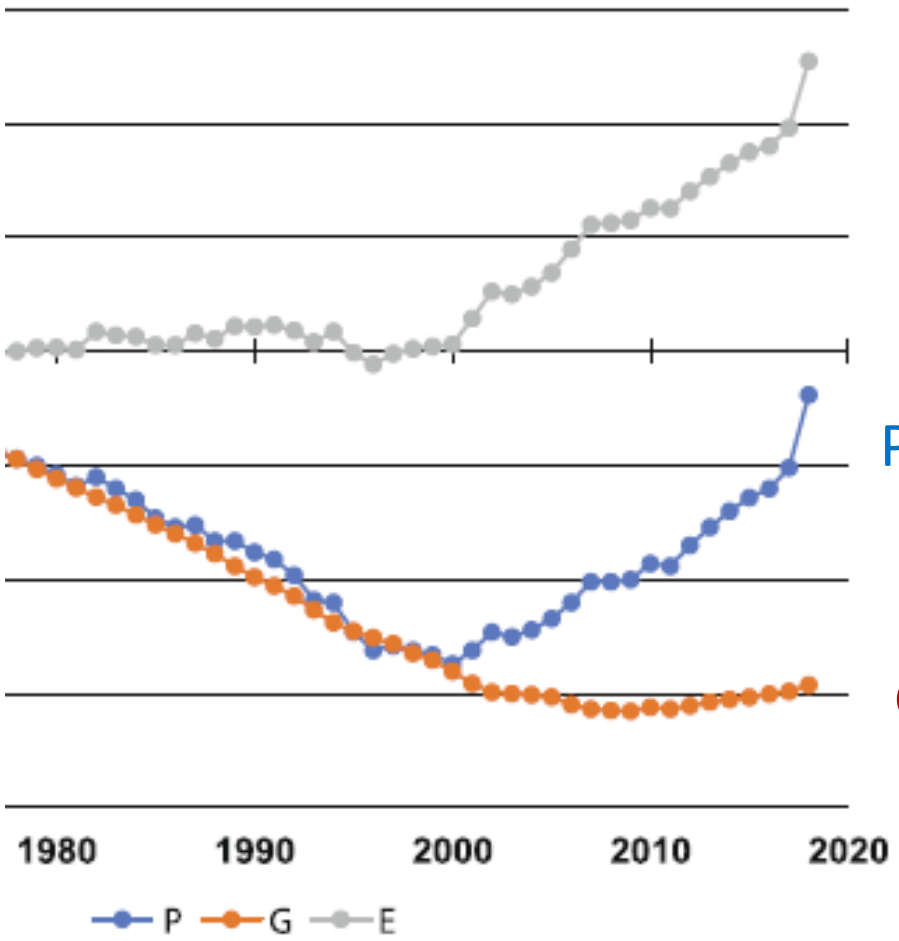
Selection for more than production traits alone may prevent such.

Hypothetical trend changes in 3 stages of genetic selection



Negative changes accelerate and management cannot catch up!

Trends for daughter pregnancy rate



Environmental - management

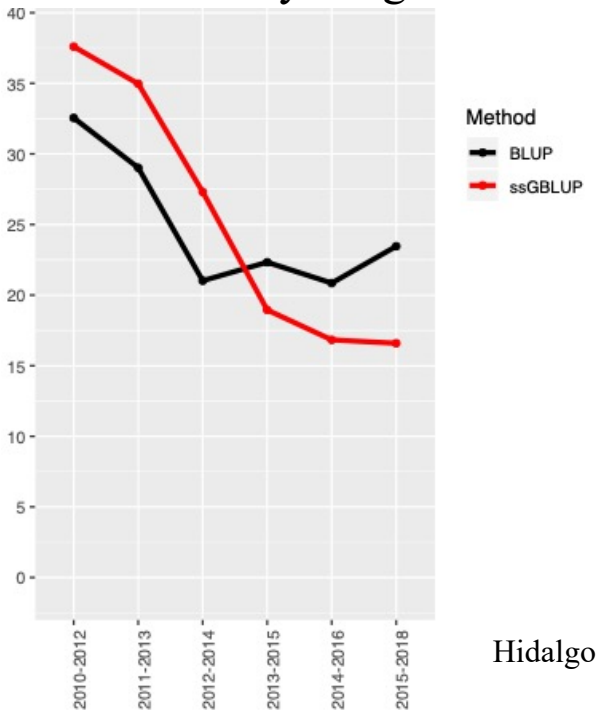
Phenotypic

Genetic

Brito et al., 2021

Changes in (co)variances in pigs due to genomic selection

Heritability for growth



Genetic correlation with reproduction

Hidalgo et al. (2019)

Heritability halved, antagonistic correlations -0.3 ➔ -0.5

Why changes in genetic parameters?

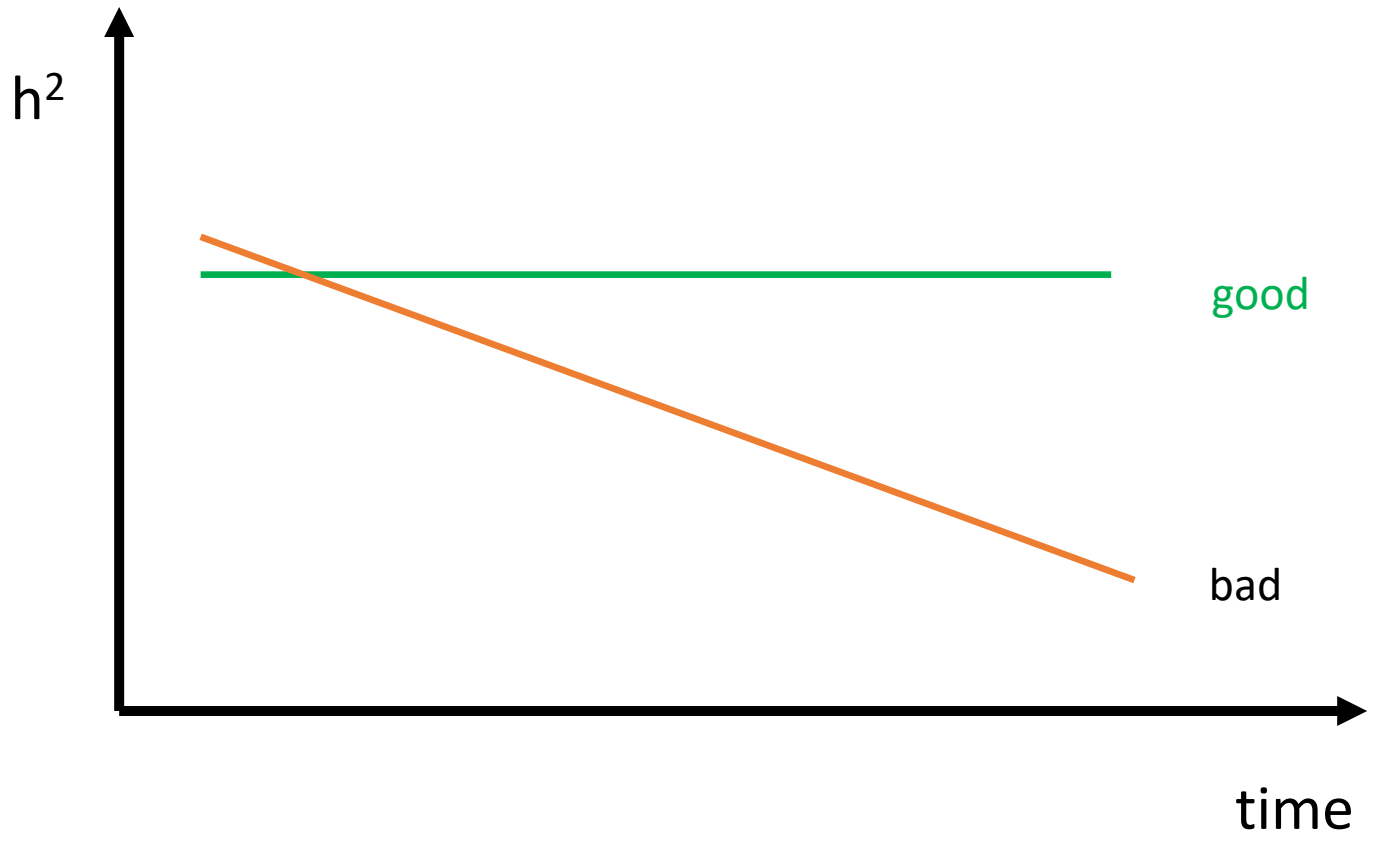
- Bulmer effect
- Changing resource allocation

- Changes in gene frequencies
- Changes in trait definitions
- G x E
- Recessives
- …

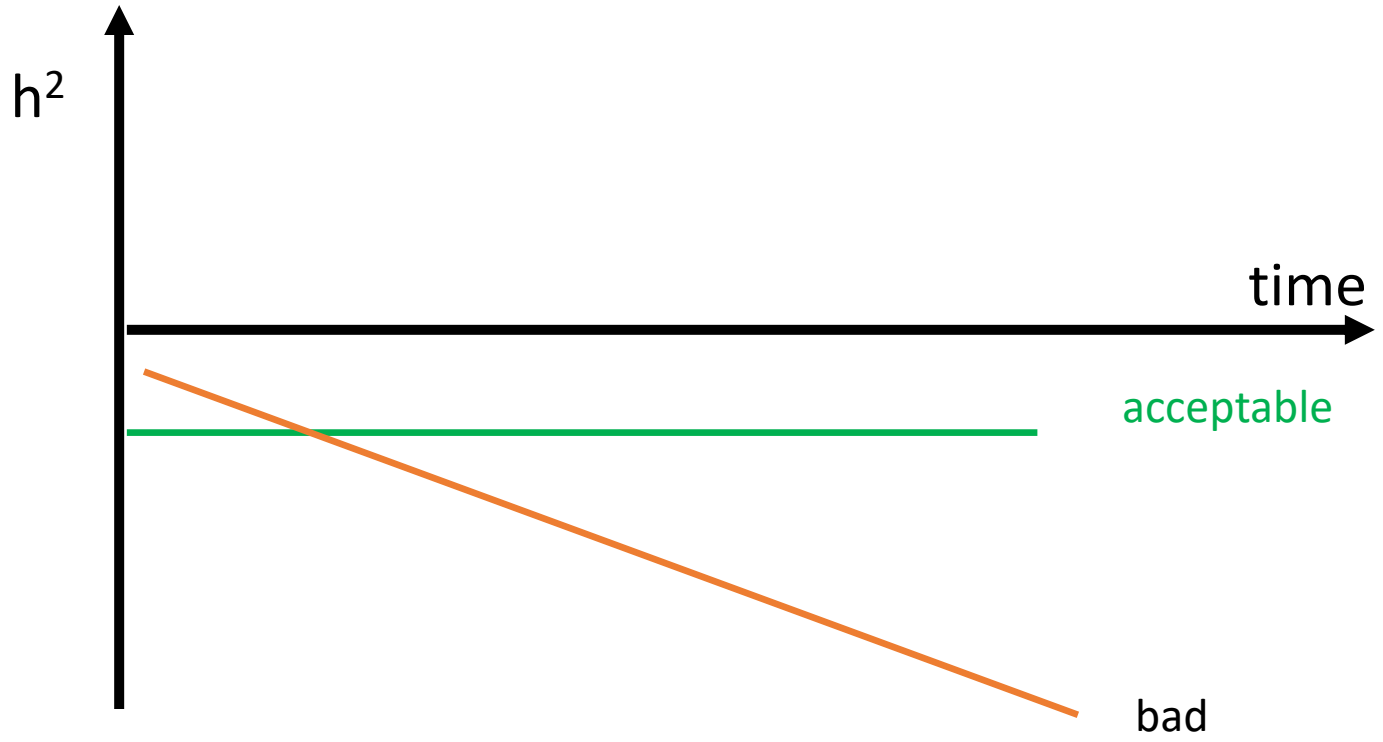
How to circumvent negative effects?

- Start or expand recording for problematic traits
- Update selection index
  - Needs estimates for last generation
- Focus on traits where the parameters are changing rapidly
  - Needs estimates generation by generation

Possible changes in heritability



Possible changes in genetic correlations



Using theoretical and realized accuracies to estimate changes in heritabilities and genetic correlations

Ignacy Misztal

University of Georgia

Realized and theoretical accuracies

Realized accuracy $acc = corr\,(y - Xb, \hat{u})/h$

Legarra et al. (2008)

y-Xb - adjusted phenotype
$\hat{u}$ - breeding value obtained without that phenotype
h^2 - heritability

Theoretical accuracy $acc = \sqrt{\frac{Nh^2}{Nh^2 + M_e}}$

Daetwyler et al. (2008)

N – number of genotyped animals with phenotypes
M_e – number of independent chromosome segments

Me ≈ 5k (chickens, pigs), 10k (beef), 15k (Holsteins)
Pocrnic et al. (2017)

Heritability by predictivity

$$\widehat{h^2} = \frac{c^2 + \sqrt{c^4 + 4c^2 M_e/N}}{2}, c = corr(y - Xb, \hat{u})$$

c - predictivity
Me – number of independent chromosome segments (about 10k in beef)
N – number of reference animals with phenotypes and genotypes

$$SE(\widehat{h^2}) \approx \frac{1}{\sqrt{N_{val}}}\left[c + \frac{2c^2 + \frac{4M_e}{N}}{\sqrt{c^2 + \frac{4M_e}{N}}}\right] \widehat{h^2}) \approx \frac{3c}{\sqrt{N_{val}}}$$

N_{val} – number of animals in validation

Heritability for milk in Holsteins

J. Dairy Sci. 104:5843–5853
https://doi.org/10.3168/jds.2020-19789
© 2021, The Authors. Published by Elsevier Inc. and Fass Inc. on behalf of the American Dairy Science Association®.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Genomic predictions for yield traits in US Holsteins with unknown parent groups

A. Cesarani,[1]* Y. Masuda,[1] S. Tsuruta,[1] E. L. Nicolazzi,[2] P. M. VanRaden,[3] D. Lourenco,[1] and I. Misztal[1]
[1]Department of Animal and Dairy Science, University of Georgia, Athens 30602
[2]Council on Dairy Cattle Breeding, Bowie, MD 20716
[3]Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

# animals with phenotypes and genotypes	580k
# animals with validation	381k
Assumed # chromosome segments M_e	15k
Predictivity	0.55
Initial h^2	0.35
Calculated h^2	0.33

How to estimate genetic correlations?

Predictivity for trait i $corr(y_i - Xb_i, \hat{u}_i) = acc_i\ h_i$

What is predictivity from trait i to trait j? $corr(y_i - Xb_i, \hat{u}_j) = ?$

.....

.....

$$corr(y_i - Xb_i, \hat{u}_j) = acc_j\ corr_{ij}\ h_i$$

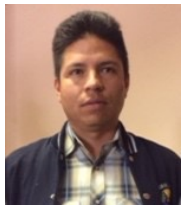
$$corr_{ij} = \frac{corr(y_i - Xb_i, \hat{u}_j)}{h_i\ acc_j}$$

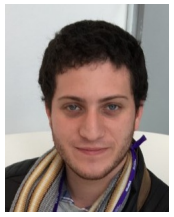
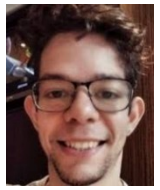
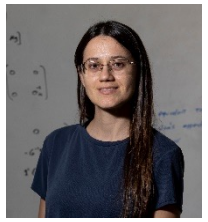
$$SD(corr_{ij}) \approx \frac{1}{h_i\ acc_j \sqrt{N_{val}}}$$

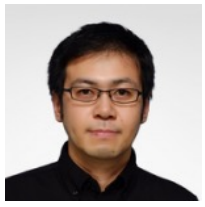
Conclusions

- Response to QTL wide for pigs & chickens– several Mb
  - Probably false QTL if no LD trail
  - ssGBLUP accounts for QTL with large data
- ”Good” large QTLs probably fixed, remaining show pleiotropy
- Potential negative effects of genomic selection on fitness traits
  - faster correlated responses
  - Potentially increased antagonism
- Need new methods to estimate genetic parameters – use of predictivity promising

UGA AB&G team





USDA NIFA

PIC

ANGUS THE BUSINESS BREED
AMERICAN ANGUS ASSOCIATION
3201 Frederick Avenue • St. Joseph, MO 64506 • (816) 383-5100 • Fax (816) 233-9703 • E-mail: angus@angus.org

AGI

USDA United States Department of Agriculture
Agricultural Research Service
Warmwater Aquaculture Research Unit

zoetis

Smithfield

USDA Agricultural Research Service
U.S. DEPARTMENT OF AGRICULTURE
Cool and Cold Water Aquaculture Research

Cobb

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

DNA GENETICS
Greater gains. Improved performance.

ABCZ

Rural Development Administration
National Institute of Animal Science

WAGYU
AMERICAN WAGYU ASSOCIATION

BAYER Crop Science

U.S. Registered Holsteins
FOR MAXIMUM PROFIT
HOLSTEIN ASSOCIATION USA, INC.

THE MASCHHOFFS
Progressive Farming. Family Style.

ABS

SEMEX
Genetics for Life

PEAK

PFHBiPM

PLEMDAT

Topigs Norsvin

methodgenetics

ANCP