

Options for evaluating multiple breeds in a single-step GBLUP for US dairy population

Alberto Cesarani, D. Lourenco, M. Bermann, PM. VanRaden, E.L. Nicolazzi, A. Legarra, S. Tsuruta, I. Misztal



**UNIVERSITY OF
GEORGIA**

Introduction

- In 2007, a multibreed genetic evaluation was implemented for US dairy cattle, and it was extended to a multistep genomic evaluation in 2009 (VanRaden et al., 2007; 2009)
- Currently, US dairy cattle are evaluated with a two-step method: a multibreed BLUP model followed by SNP effects estimation and DGV computation within each breed

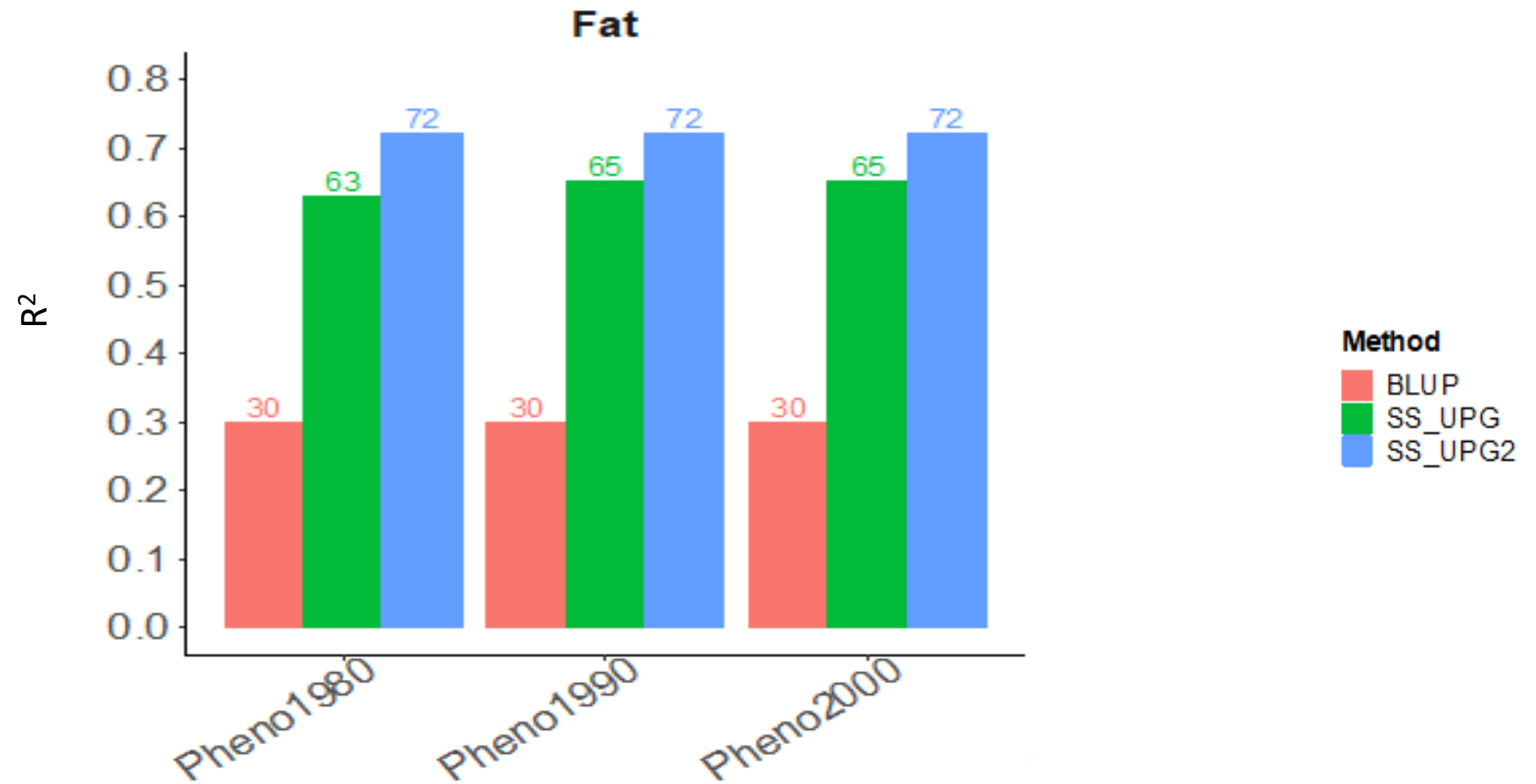
Introduction

- Single-step genomic BLUP (ssGBLUP) has been adopted instead of the multistep method for genetic evaluations in several farm animal species over the years (e.g., Macedo et al., 2020; Pimentel et al., 2021)
- Genomic models have been tested also in a multibreed context (e.g., Khansefid et al., 2020)

Which model for ssGBLUP?

- Past limitations: data size and missing parents
- Theory behind UPG is well established for the regular BLUP, but not for ssGBLUP
- Studies on removing data and different strategies to fit UPG in ssGBLUP

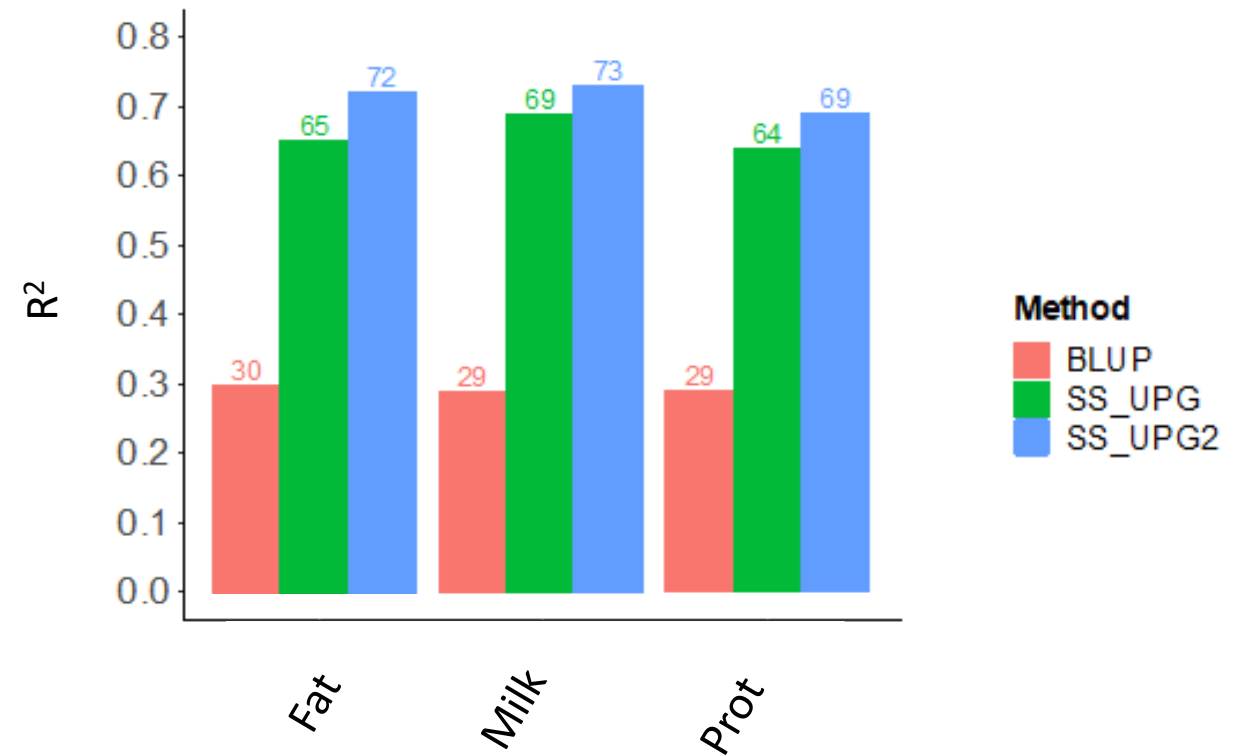
Which model for ssGBLUP?



Which model for ssGBLUP?

$$\mathbf{H}_{UPG}^* = \mathbf{A}^* + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathbf{G}_{APY}^{-1} - \mathbf{A}_{22}^{-1} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{H}_{UPG2}^* = \mathbf{A}^* + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathbf{G}_{APY}^{-1} - \mathbf{A}_{22}^{-1} & -(-\mathbf{A}_{22}^{-1})\mathbf{Q}_2 \\ 0 & -\mathbf{Q}_2'(-\mathbf{A}_{22}^{-1}) & \mathbf{Q}_2'(-\mathbf{A}_{22}^{-1})\mathbf{Q}_2 \end{bmatrix}$$



Multibreed genomic relationships

- Constructing a multibreed \mathbf{G} is challenging
- Combining breeds may improve predictions for smaller breeds
- Breeds with many more genotypes may overwhelm predictions
 - Biased predictions for smaller breeds

Aim

The **aims** of this work were:

1. develop ssGBLUP multibreed genomic predictions for US dairy cattle
2. compare single and multi breed evaluations



Data

Purebred Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), and Jersey (JE)

Data

Purebred Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), and Jersey (JE)

Breed	Phenotypes		Animals	
	N	Cows	Genotypes	Total
Ayrshire	116k	47k	9k	94k

Data

Purebred Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), and Jersey (JE)

Breed	Phenotypes		Animals	
	N	Cows	Genotypes	Total
Ayrshire	116k	47k	9k	94k
Guernsey	129k	58k	5k	100k

Data

Purebred Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), and Jersey (JE)

Breed	Phenotypes		Animals	
	N	Cows	Genotypes	Total
Ayrshire	116k	47k	9k	94k
Guernsey	129k	58k	5k	100k
Brown Swiss	328k	138k	47k	292k

Data

Purebred Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), and Jersey (JE)

Breed	Phenotypes		Animals	
	N	Cows	Genotypes	Total
Ayrshire	116k	47k	9k	94k
Guernsey	129k	58k	5k	100k
Brown Swiss	328k	138k	47k	292k
Jersey	4.1M	1.7M	427k	2.5M

Data

Purebred Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), and Jersey (JE)

Breed	Phenotypes		Animals	
	N	Cows	Genotypes	Total
Ayrshire	116k	47k	9k	94k
Guernsey	129k	58k	5k	100k
Brown Swiss	328k	138k	47k	292k
Jersey	4.1M	1.7M	427k	2.5M
Holstein	40.3M	17.5M	3.4M	26.6M

Data

Purebred Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), and Jersey (JE)

Breed	Phenotypes		Animals	
	N	Cows	Genotypes	Total
Ayrshire	116k	47k	9k	94k
Guernsey	129k	58k	5k	100k
Brown Swiss	328k	138k	47k	292k
Jersey	4.1M	1.7M	427k	2.5M
Holstein	40.3M	17.5M	3.4M	26.6M

Data

Purebred Ayrshire (AY), Brown Swiss (BS), Guernsey (GU), Holstein (HO), and Jersey (JE)

Breed	Phenotypes		Animals	
	N	Cows	Genotypes	Total
Ayrshire	116k	47k	9k	94k
Guernsey	129k	58k	5k	100k
Brown Swiss	328k	138k	47k	292k
Jersey	4.1M	1.7M	427k	2.5M
Holstein	40.3M	17.5M	3.4M	26.6M
ALL	45M	19.4M	3.9M	29.5M

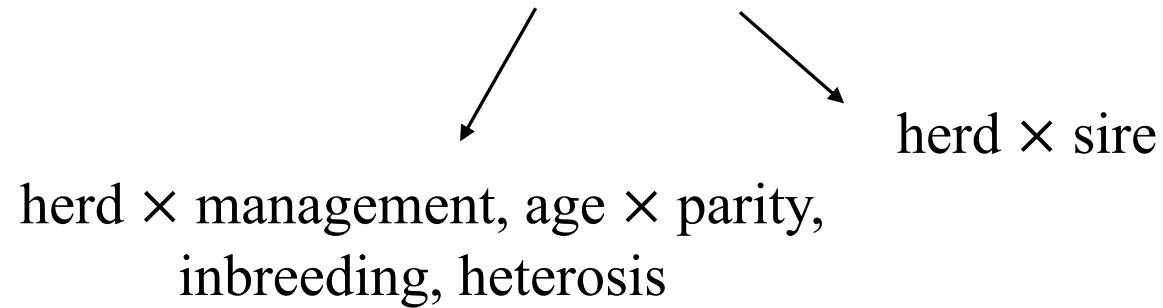
Materials and Methods I

- Milk (MY), fat (FY), and protein (PY) 305-d yields recorded from January 2000 to June 2020
- Heritability of 0.20 for the three traits
- Projected records for the final lactation and for lactations not yet completed by June 2020
- According to the real lactation length, different weights are applied to project the records to 305-d: the shorter the lactation, the lower the weight

Materials and Methods I

- Three-trait repeatability model:

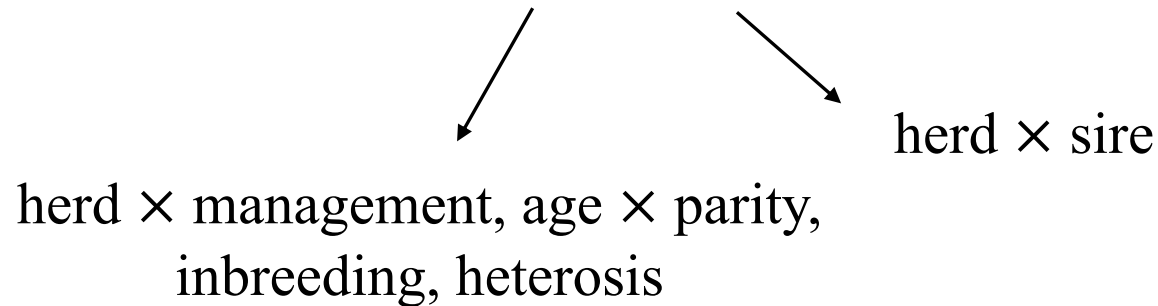
$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_h \mathbf{h} + \mathbf{Z}_a \mathbf{Q}_a \mathbf{g}_a + \mathbf{Z}_a \mathbf{a} + \mathbf{Z}_p \mathbf{p} + \mathbf{e}$$



Materials and Methods I

- Three-trait repeatability model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_h\mathbf{h} + \mathbf{Z}_a\mathbf{Q}_a\mathbf{g}_a + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_p\mathbf{p} + \mathbf{e}$$

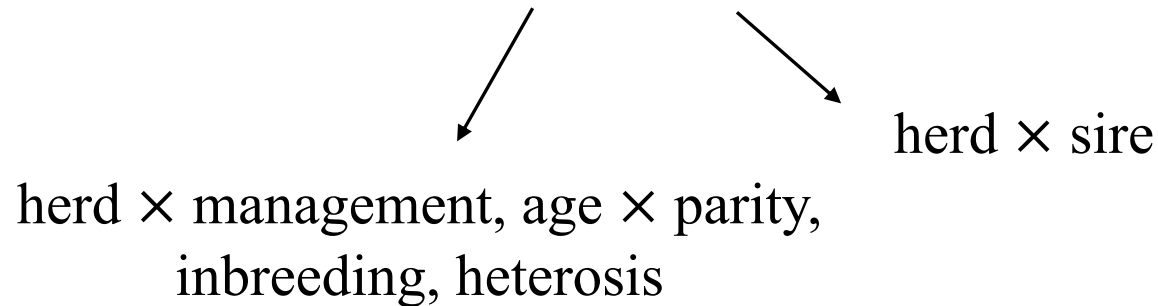


- Breed-specific fixed effects
- **Complete data**: phenotypes of cows born up to 2018
- **Reduced data**: phenotypes of cows born up to 2014

Materials and Methods I

- Three-trait repeatability model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_h\mathbf{h} + \mathbf{Z}_a\mathbf{Q}_a\mathbf{g}_a + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_p\mathbf{p} + \mathbf{e}$$



- Breed-specific fixed effects
- **Complete data**: phenotypes of cows born up to 2018
- **Reduced data**: phenotypes of cows born up to 2014
- UPG (8 groups per breed) in \mathbf{A}^{-1} and \mathbf{A}_{22}^{-1} : difference by breed, YOB, and sex

Materials and Methods III

Validation animals:

- **cows** = born from 2015 to 2018, genotyped, with no phenotypes in the reduced dataset
- **bulls** = genotyped bulls with no daughters in the reduced and at least 10 (AY, BS, GU) or 50 (HO, JE) daughters in the complete dataset

Validation method:

- **cows** → predictivity based on adjusted phenotypes (correlation)
- **bulls** → Interbull validation method based on DYD

Materials and Methods VI

- **SINGLE**
 - each breed separately
 - APY method with 15k random core animals for JE and HO
- **ALL**
 - five breeds together with breed-specific effects
 - 15k core animals at random
- **ALL45K**
 - as ALL but with 45k core animals: 5k AY + 5k BS + 5k GU + 15k HO + 15k JE

Results

Similar trends for milk, fat, and **protein**

Results: predictivity for cows

Predictivity for cows - Protein

ALL: 15k core

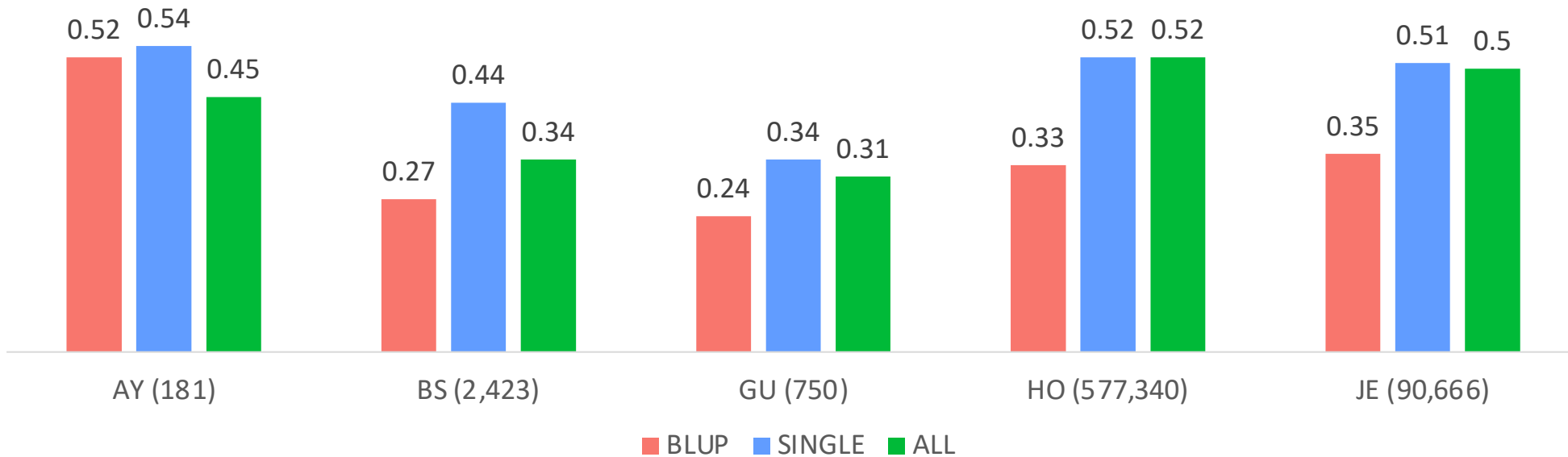
AY = 32

BS = 182

GU = 17

HO = 13k

JE = 1.7k



Results: predictivity for cows

Predictivity for cows - Protein

ALL 45k core

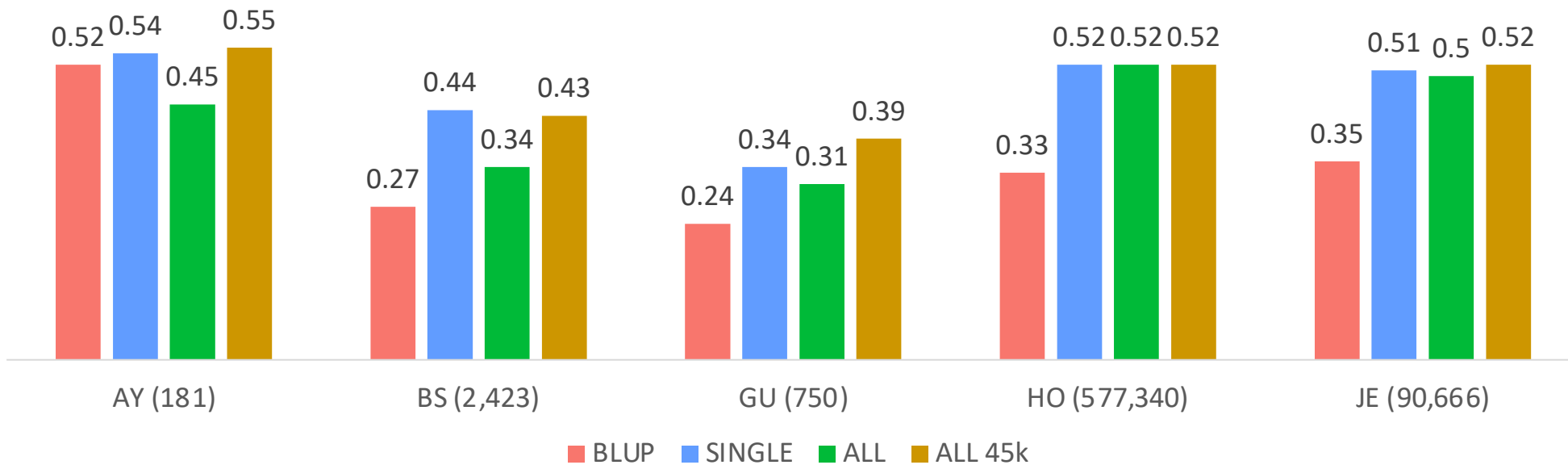
AY = 5k

BS = 5k

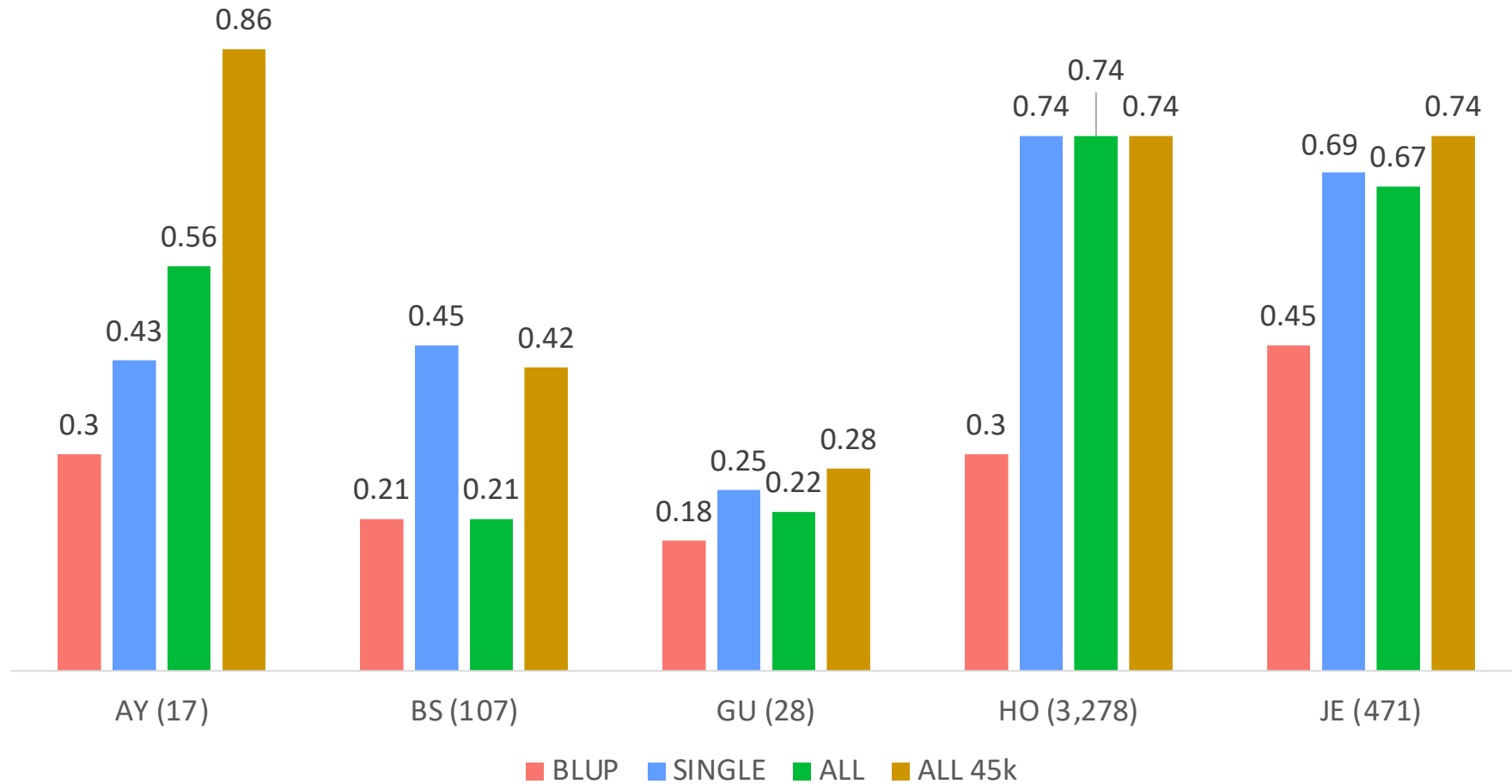
GU = 5k

HO = 15k

JE = 15k



Results: R2 for bulls



Comparison of (G)EBV

Correlation	BLUP			ssGBLUP		
AY	1.00	1.00	1.00	0.95	0.96	0.96
BS	1.00	1.00	1.00	0.95	0.96	0.97
GU	1.00	1.00	1.00	0.97	0.92	0.98
HO	1.00	1.00	1.00	1.00	1.00	1.00
JE	1.00	1.00	1.00	1.00	1.00	1.00
Slope³						
AY	1.00	1.00	1.00	1.02	0.98	0.99
BS	1.00	1.00	1.00	0.93	0.96	0.93
GU	1.00	1.00	1.00	0.95	0.95	0.95
HO	1.00	1.00	1.00	1.00	1.00	1.00
JE	1.00	1.00	1.00	0.99	0.98	0.98

³ The slope of the regressions of ALL45k on SINGLE

Discussion

- It is important to have a good representation of all breeds in the core (Mantysaari et al., 2017; Vandenplas et al., 2018).
- Modeling the breed effects in a proper way guarantees more accurate predictions in multibreed evaluations
- Breeds with fewer data slightly benefit from a multibreed estimation of SNP effects (Olson et al., 2012)
- Good match and similar genetic trends with official evaluations (Rodrigo Mota's presentation)

Conclusions

- Single-step large-scale multibreed evaluations are feasible
 - at least similar reliability and predictivity as single-breed
 - UPG accounting for breed + breed-specific effects
- Holstein evaluations were more stable because of the greatest number of genotypes

Conclusions

- Single-step large-scale multibreed evaluations are feasible
 - at least similar reliability and predictivity as single-breed
 - UPG accounting for breed + breed-specific effects
- Holstein evaluations were more stable because of the greatest number of genotypes
- A new algorithm reduced the pre-processing time from ~5 days to ~26 hours
- Considering only “useful” genotypes reduced computing time for:
 - the pre-processing from ~26h to ~8h
 - the GEBV estimation from ~60h to ~17h

Thanks!



ANIMAL BREEDING AND GENETICS GROUP

UNIVERSITY OF GEORGIA