

# Multi-breed genomic evaluation for dairy cattle in the US using single-step GBLUP

**Alberto Cesarani, D. Lourenco, Y. Masuda, S. Tsuruta, A. Legarra, E.**

**Nicolazzi, P. VanRaden, I. Misztal**



**UNIVERSITY OF  
GEORGIA**

# Background

1. Genomic evaluations for the US dairy industry follow the multistep method
  - Requires adjustments to reduce bias (female genotypes)
  - Relies on EBV from BLUP that are biased because of genomic selection
  - May not account for pre-selection bias (Patry and Ducroq, 2011; Masuda et al., 2018)

# Background

1. Genomic evaluations for the US dairy industry follow the multistep method
  - Requires adjustments to reduce bias (female genotypes)
  - Relies on EBV from BLUP that are biased because of genomic selection
  - May not account for pre-selection bias (Patry and Ducroq, 2011; Masuda et al., 2018)
2. Official multibreed genomic evaluations for dairy cattle in the US
  - Multibreed BLUP evaluation
  - Single-breed estimation of SNP effects and DGV

# Single-step genomic BLUP

1. Single-step allows the straight computation of GEBV in multibreed populations

# Single-step genomic BLUP

1. Single-step allows the straight computation of GEBV in multibreed populations
2. ssGBLUP is being used as the main tool for genomic selection
  - Beef cattle, pigs, chicken, and dairy in some countries
  - Under constant development for over 12 years

# Single-step genomic BLUP

1. Single-step allows the straight computation of GEBV in multibreed populations
2. ssGBLUP is being used as the main tool for genomic selection
  - Beef cattle, pigs, chicken, and dairy in some countries
  - Under constant development for over 12 years
3. Feasibility for multibreed large-scale dairy evaluations should be assessed

# Aims

To develop large-scale ssGBLUP multibreed genomic predictions for US dairy cattle

To compare single and multi breed models

# 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)

<b>Breed</b>	<b>Phenotypes</b>		<b>Animals</b>	
	<b>N</b>	<b>Cows</b>	<b>Genotypes</b>	<b>Total</b>
Ayrshire	116k	47k	9k	94k

# Data

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

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

# Data

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

<b>Breed</b>	<b>Phenotypes</b>		<b>Animals</b>	
	<b>N</b>	<b>Cows</b>	<b>Genotypes</b>	<b>Total</b>
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)

<b>Breed</b>	<b>Phenotypes</b>		<b>Animals</b>	
	<b>N</b>	<b>Cows</b>	<b>Genotypes</b>	<b>Total</b>
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)

<b>Breed</b>	<b>Phenotypes</b>		<b>Animals</b>	
	<b>N</b>	<b>Cows</b>	<b>Genotypes</b>	<b>Total</b>
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
<b>ALL-breed</b>	<b>45M</b>	<b>19.4M</b>	<b>3.9M</b>	<b>29.5M</b>

# Materials and Methods I

- Milk (MY), fat (FY), and protein (PY) yields recorded from January 2000 to June 2020

# Materials and Methods I

- Milk (MY), fat (FY), and protein (PY) yields recorded from January 2000 to June 2020

- Three-trait repeatability model

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



herd × sire

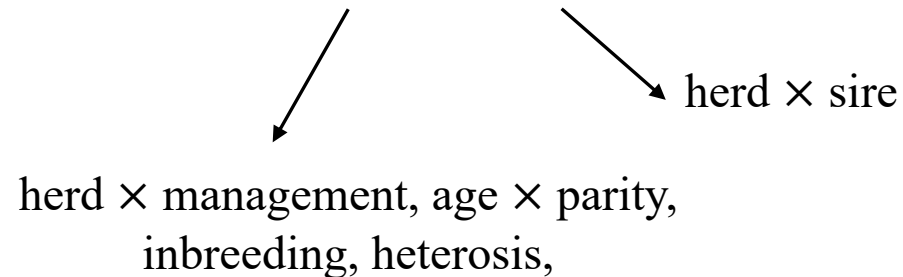
herd × management, age × parity,  
inbreeding, heterosis,



# Materials and Methods I

- Milk (MY), fat (FY), and protein (PY) yields recorded from January 2000 to June 2020

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

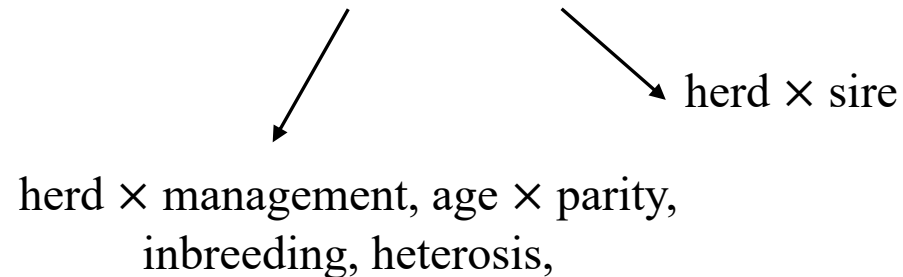


- **Complete data:** phenotypes of cows born up to 2018
- **Reduced data:** phenotypes of cows born up to 2014

# Materials and Methods I

- Milk (MY), fat (FY), and protein (PY) yields recorded from January 2000 to June 2020

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



- **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, YOY, and sex

# Materials and Methods II

Validation animals:

- **cows** = genotyped females born in 2015-2018 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
- **bulls** → Interbull validation method based on DYD

# Materials and Methods III

- **SINGLE**

- each breed separately
- APY method with 15k random core animals for JE and HO

- **ALL\_15k**

- five breeds together
- breed effect and breed-specific effects
- 15k random core animals: 32 AY + 182 BS + 17 GU + 13k HO + 1.7k JE

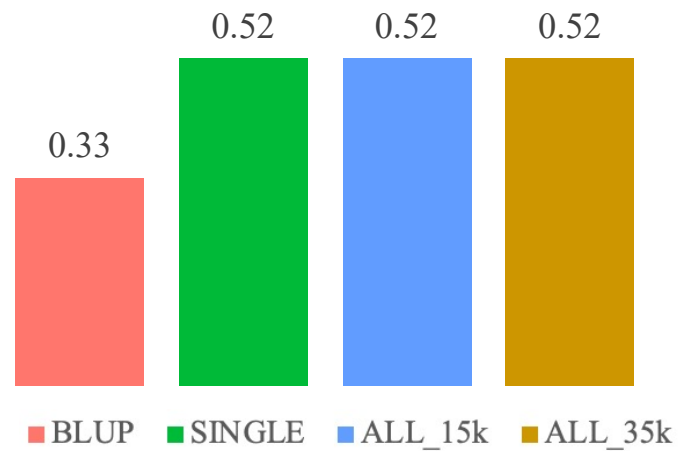
- **ALL\_35k**

- as ALL but using 35k core animals: 5k AY + 5k BS + 5k GU + 10k HO + 10k JE

# Predictive abilities for cows

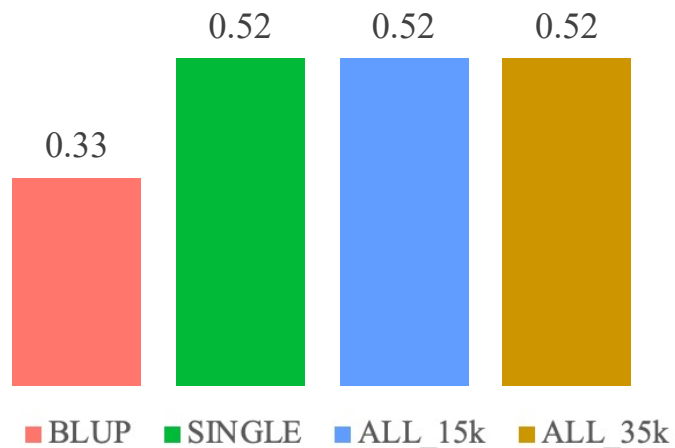
# Predictive abilities for cows

Holstein (n=577,340)

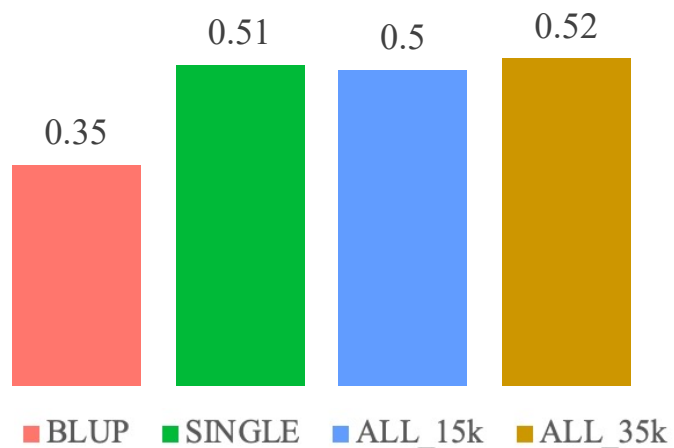


# Predictive abilities for cows

Holstein (n=577,340)

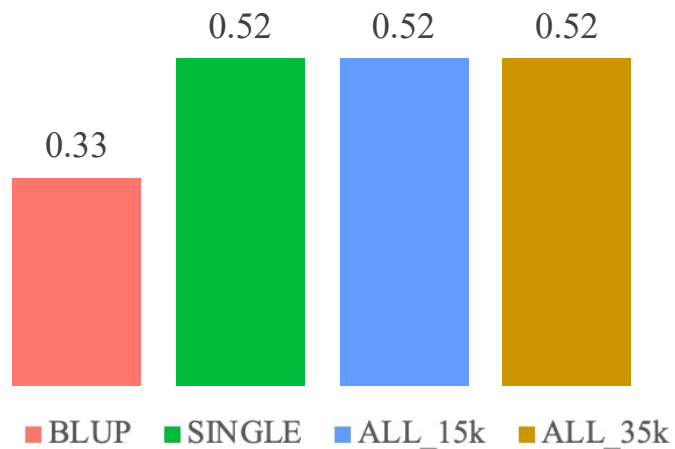


Jersey (n=90,666 )

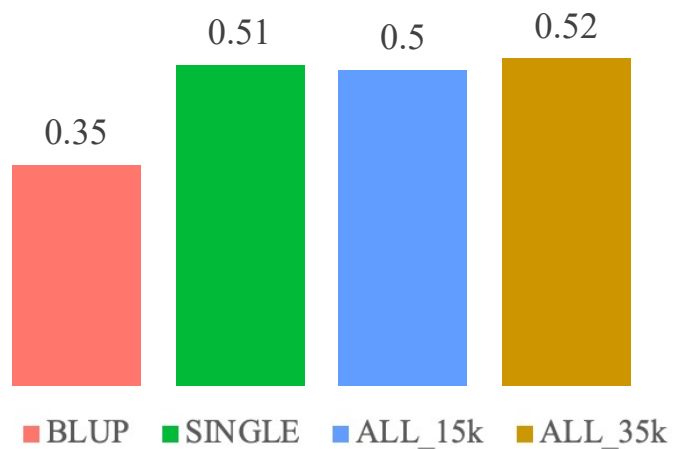


# Predictive abilities for cows

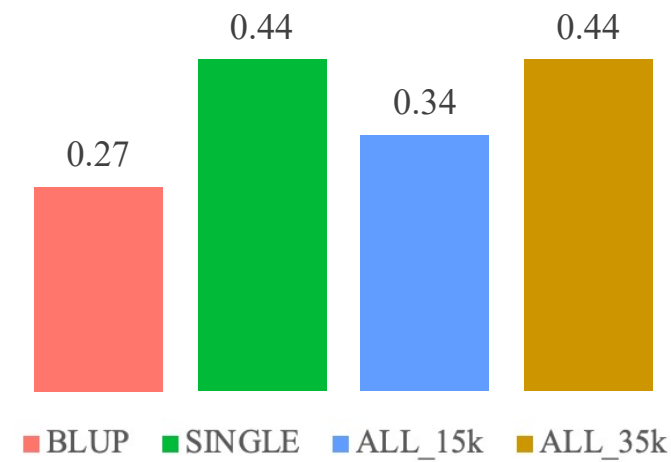
Holstein (n=577,340)



Jersey (n=90,666)



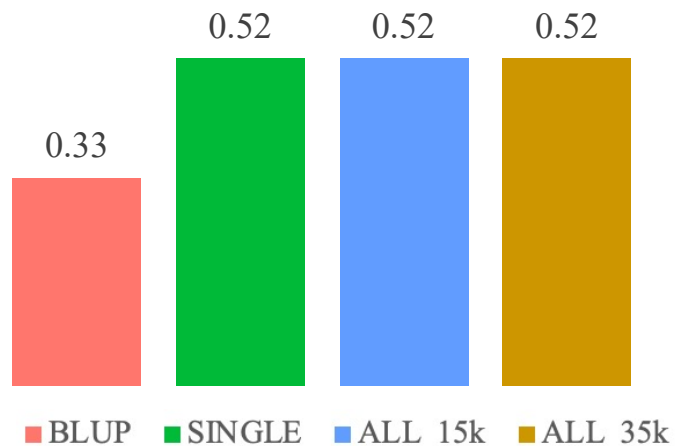
Brown Swiss (n=2,423)



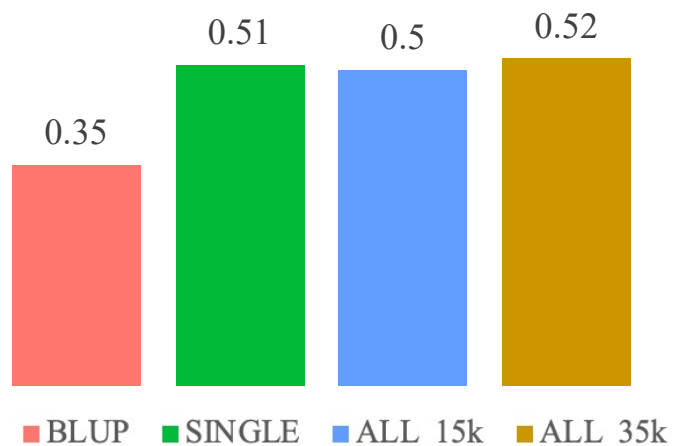


# Predictive abilities for cows

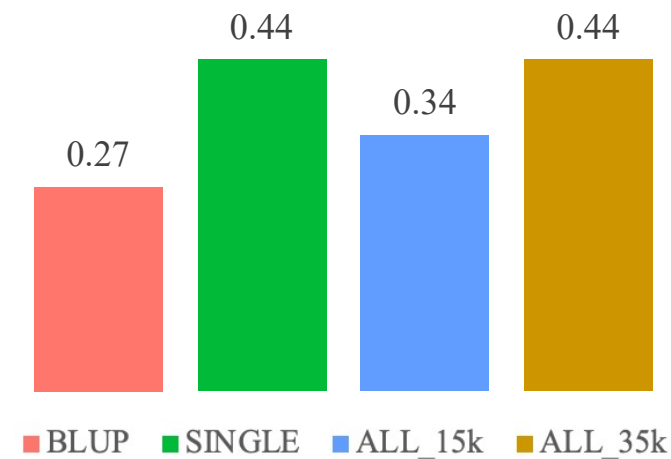
**Holstein (n=577,340)**



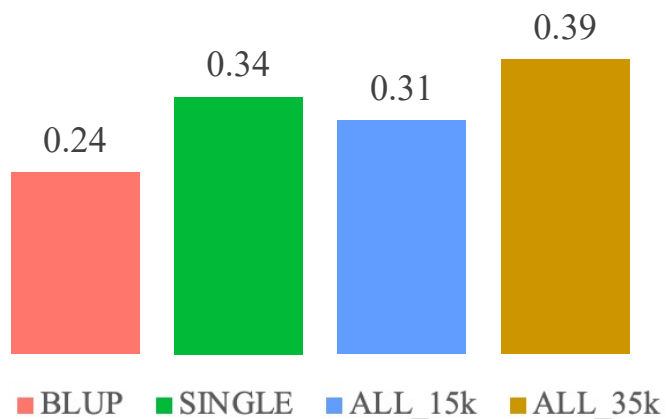
**Jersey (n=90,666 )**



**Brown Swiss (n=2,423)**

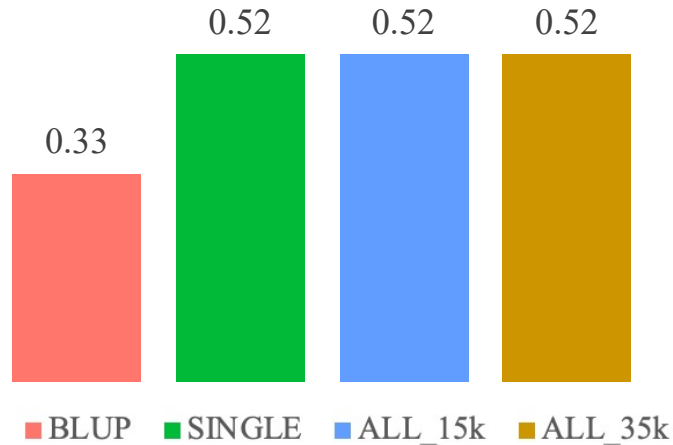


**Guernsey (n=750)**

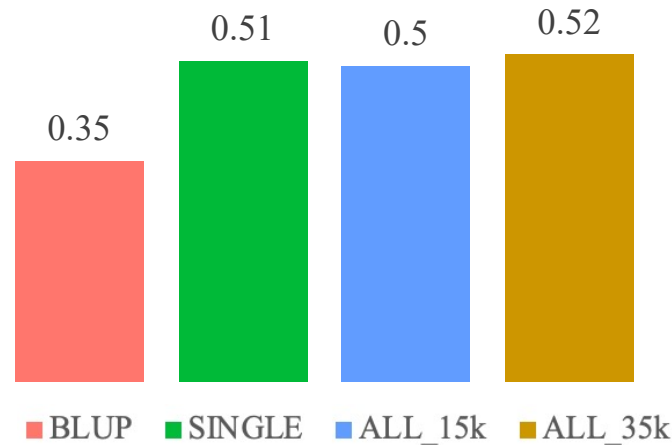


# Predictive abilities for cows

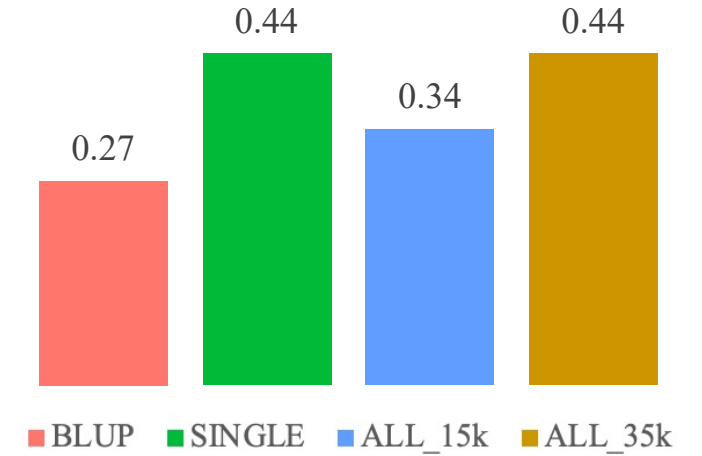
Holstein (n=577,340)



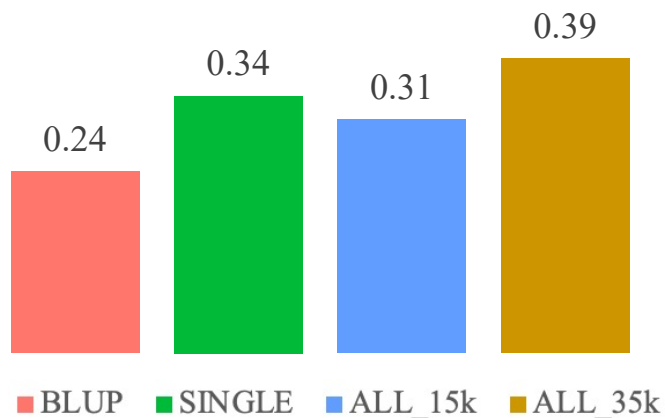
Jersey (n=90,666)



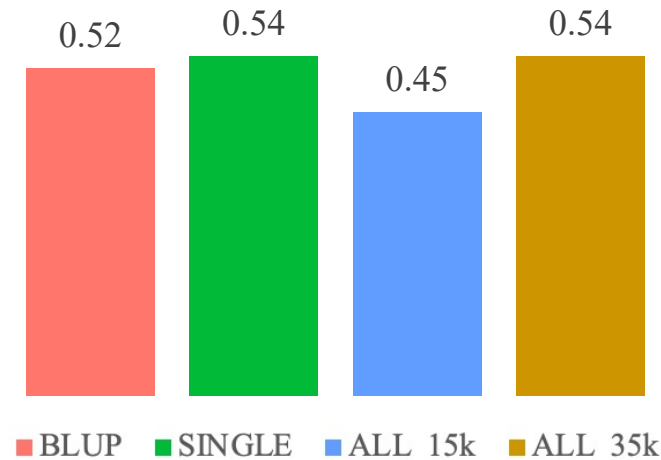
Brown Swiss (n=2,423)



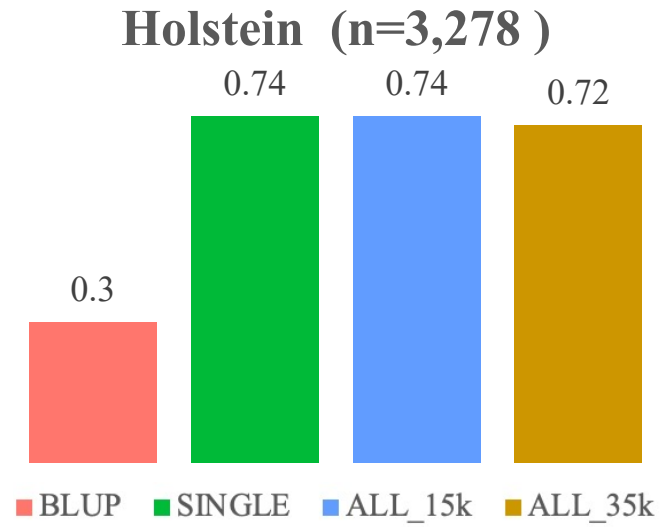
Guernsey (n=750)



Ayrshire (n=181)

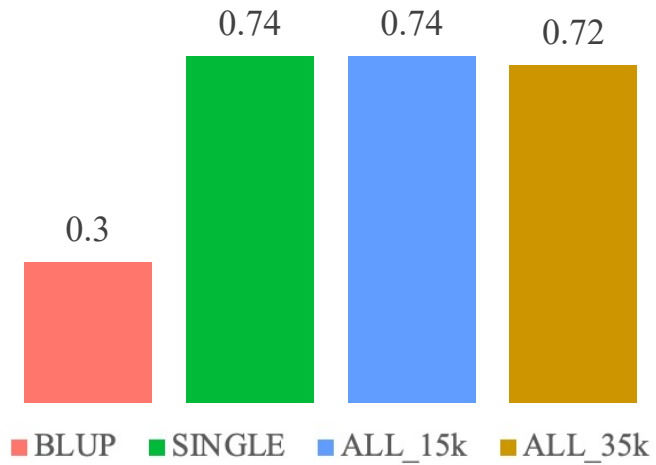


# Reliabilities for bulls

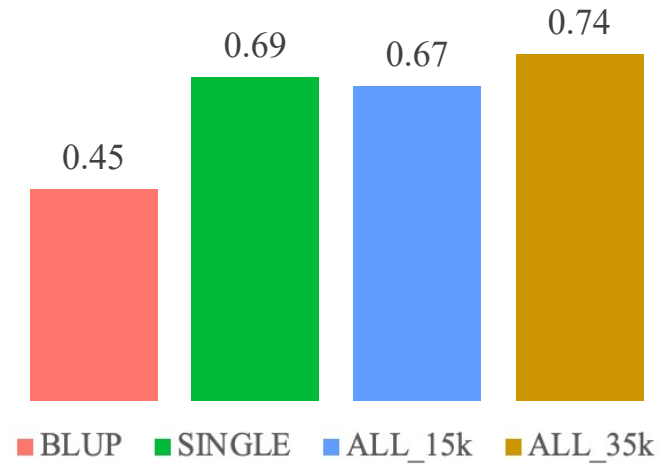


# Reliabilities for bulls

Holstein (n=3,278)

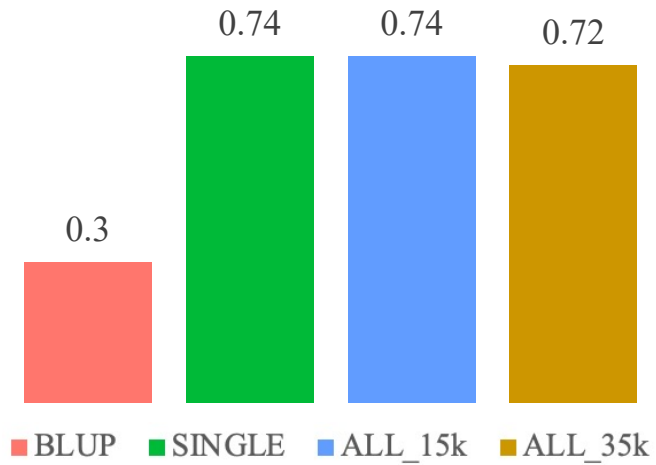


Jersey (n=471)

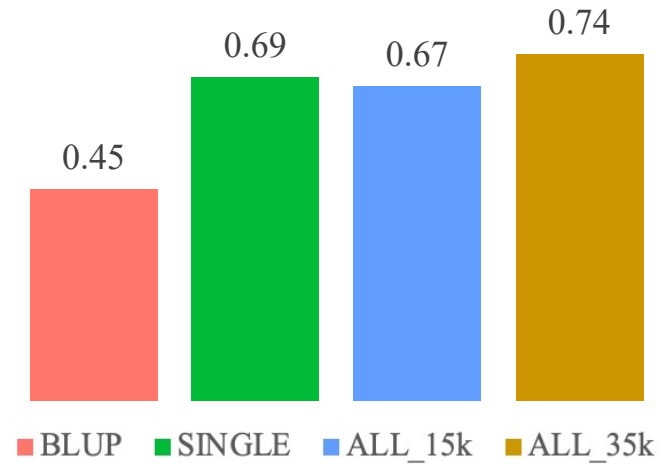


# Reliabilities for bulls

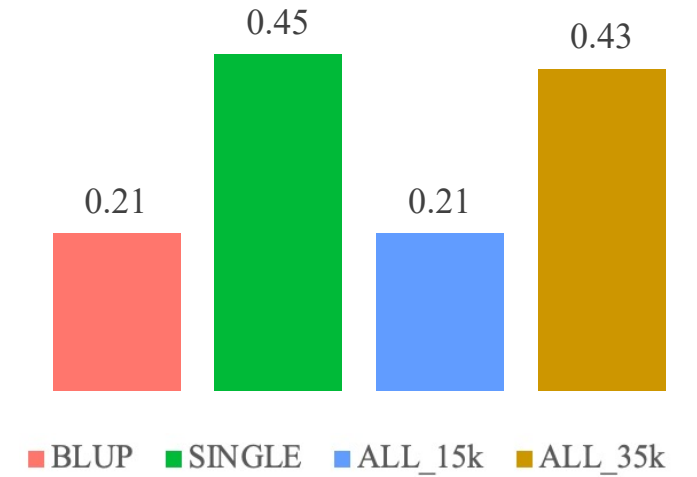
Holstein (n=3,278)



Jersey (n=471)

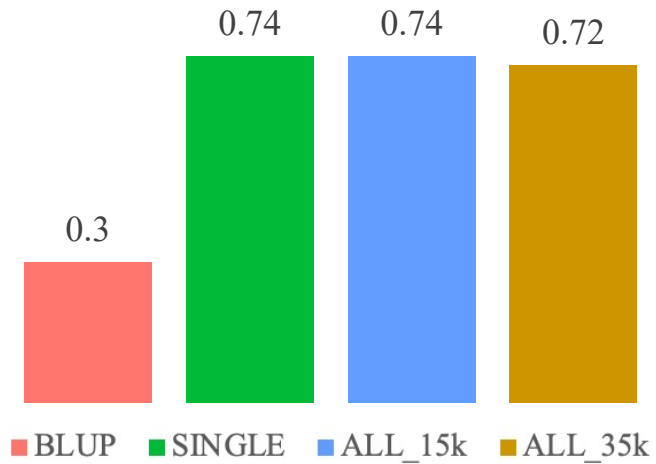


Brown Swiss (n=107)

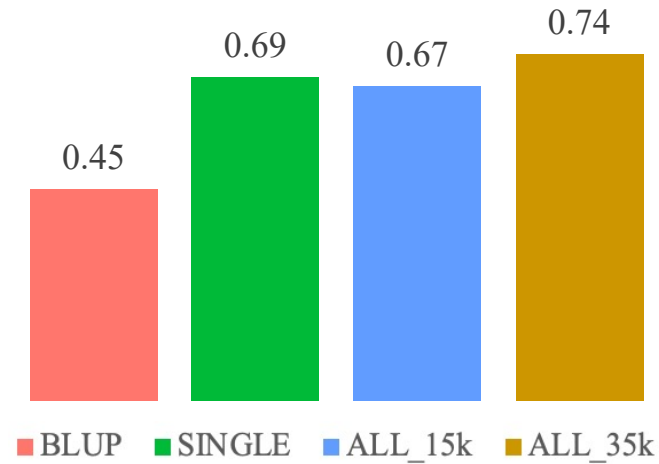


# Reliabilities for bulls

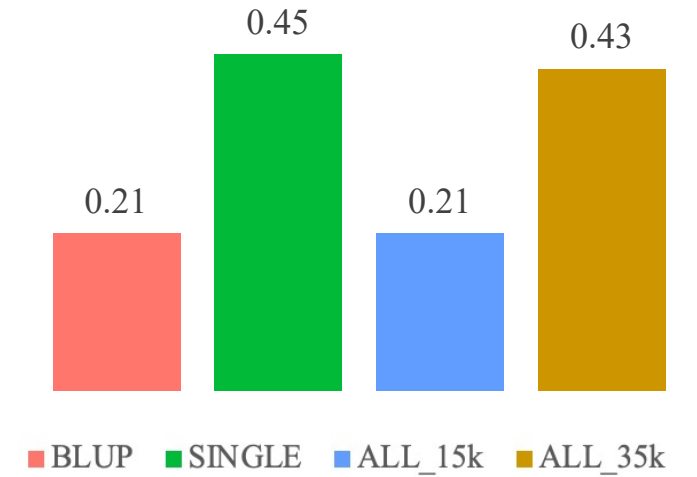
Holstein (n=3,278)



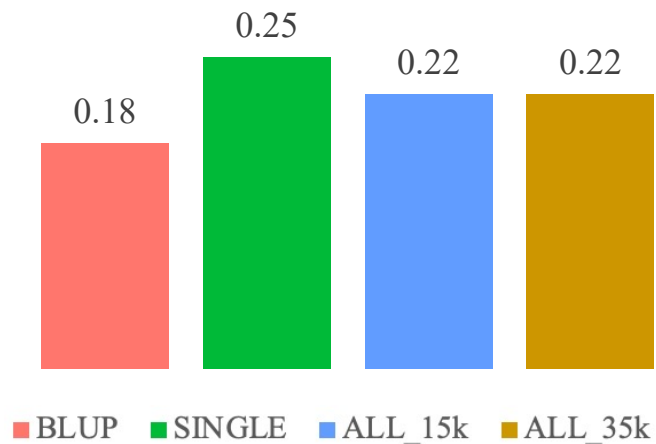
Jersey (n=471)



Brown Swiss (n=107)

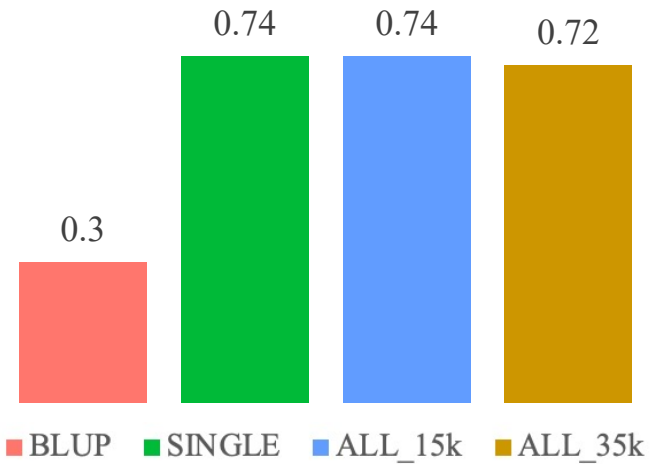


Guersney (n=28)

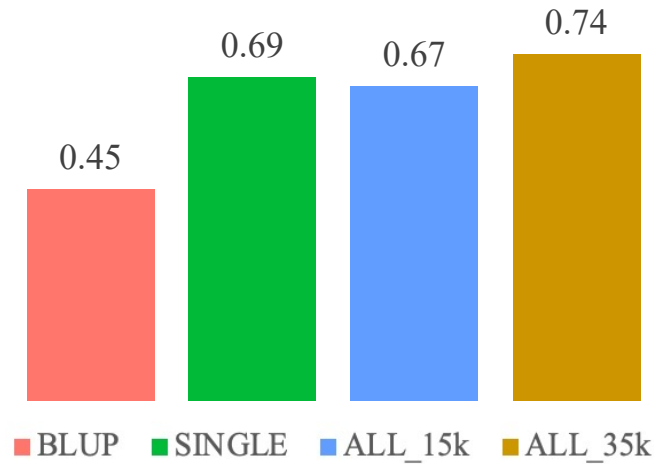


# Reliabilities for bulls

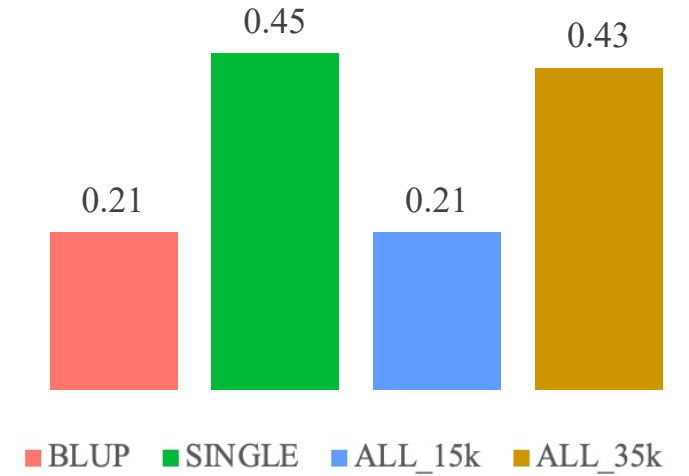
Holstein (n=3,278)



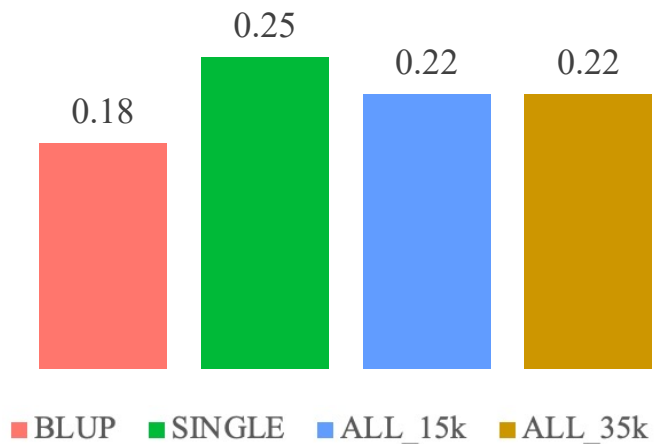
Jersey (n=471)



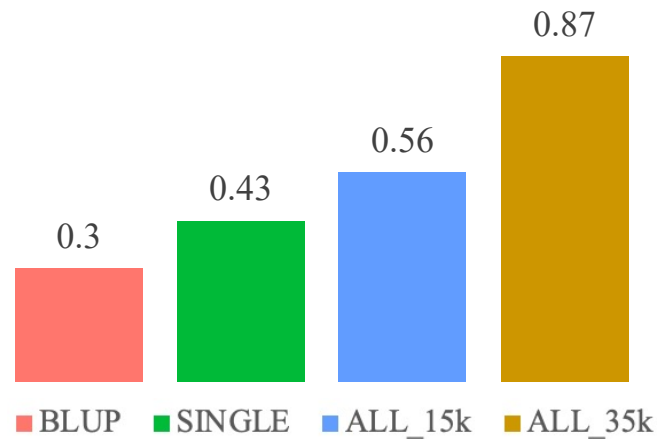
Brown Swiss (n=107)



Guersney (n=28)



Ayrshire (n=17)



# Computing cost

	Genotypes	Rounds		Time	
		BLUP	ssGBLUP	BLUP	ssGBLUP
GU	5k	345	757	< 1 min	< 1 min
AY	9k	504	863	< 1 min	< 1 min
BS	47k	364	867	1 min	~ 6 min



# Computing cost

	Genotypes	Rounds		Time	
		BLUP	ssGBLUP	BLUP	ssGBLUP
GU	5k	345	757	< 1 min	< 1 min
AY	9k	504	863	< 1 min	< 1 min
BS	47k	364	867	1 min	~ 6 min
JE	427k	586	432	~ 20 min	~ 40 min

# Computing cost

	Genotypes	Rounds		Time	
		BLUP	ssGBLUP	BLUP	ssGBLUP
GU	5k	345	757	< 1 min	< 1 min
AY	9k	504	863	< 1 min	< 1 min
BS	47k	364	867	1 min	~ 6 min
JE	427k	586	432	~ 20 min	~ 40 min
HO	3.4M	457	473	2.7 h	7.4 h

# Computing cost

	Genotypes	Rounds		Time	
		BLUP	ssGBLUP	BLUP	ssGBLUP
GU	5k	345	757	< 1 min	< 1 min
AY	9k	504	863	< 1 min	< 1 min
BS	47k	364	867	1 min	~ 6 min
JE	427k	586	432	~ 20 min	~ 40 min
HO	3.4M	457	473	2.7 h	7.4 h
ALL_15k	3.9M	643	1,142	4.8 h	~ 20 h
ALL_35K			1,600		~ 50 h

# Discussion

- Allele effects and frequencies dominated by Holstein data
  - Scaling is not proper for other breeds
  - Compromises compatibility between  $\mathbf{G}$  and  $\mathbf{A}_{22}$  → metafounders?

# Discussion

- Allele effects and frequencies dominated by Holstein data
  - Scaling is not proper for other breeds
  - Compromises compatibility between  $\mathbf{G}$  and  $\mathbf{A}_{22}$  → metafounders?
- 80k SNP possibly not covering all breeds: no benefits using more SNPs or sequence (literature)

# Discussion

- Allele effects and frequencies dominated by Holstein data
  - Scaling is not proper for other breeds
  - Compromises compatibility between  $\mathbf{G}$  and  $\mathbf{A}_{22}$  → metafounders?
- 80k SNP possibly not covering all breeds: no benefits using more SNPs or sequence (literature)
- Under ssGBLUP, predictability (reliability) was on average 4% (2%) lower in ALL\_15k compared to SINGLE model, whereas it was 1% (10%) higher in ALL\_35k compared to SINGLE

# 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
- reasonable computing time

# 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
- reasonable computing time

Holstein evaluations were more stable because of the greatest number of genotypes

- fine-tuning is needed to avoid a reduction in reliability for smaller breeds
- the number of core animals should represent the dimensionality of each breed



# Thanks!



## ANIMAL BREEDING AND GENETICS GROUP

UNIVERSITY OF GEORGIA