



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

Alberto Cesarani¹, Daniela Lourenco¹, Yutaka Masuda¹, Shogo Tsuruta¹, Andres Legarra², Ezequiel Nicolazzi³, Paul VanRaden⁴, Ignacy Misztal¹

¹ *Department of Animal and Dairy Science, University of Georgia, Athens, GA, USA*

² *INRA, UMR1388 GenPhySE, Castanet-Tolosan, France*

³ *Council on Dairy Cattle Breeding, Bowie, MD, USA*

⁴ *Animal Genomics and Improvement Laboratory, Agricultural Research Service, USDA, Beltsville, MD, USA*



**UNIVERSITY OF
GEORGIA**

Background

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

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

Aim

The aim of this study was to develop ssGBLUP multibreed genomic predictions for US dairy cattle

This involved the use of unknown parent groups (UPG) to model the difference in genetic bases caused by breed, year of birth, and sex

Data

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

Breed	Phenotypes		Animals	
	N	Cows	Genotypes	Total
Ayrshire	116,674	47,174	9,202	94,500
Brown Swiss	328,811	138,418	47,309	292,923
Guernsey	129,422	58,554	5,032	100,643
Holstein	40,298,113	17,484,436	3,407,476	26,586,363
Jersey	4,134,973	1,704,641	427,286	2,467,946
MultiBreed	45,007,993	19,433,223	3,896,305	29,542,375

Materials and Methods I

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

A 3-trait repeatability model was applied to:

- a **complete** dataset with phenotypes of cows born from 1992 to 2018
- a **reduced** dataset with phenotypes of cows born from 1992 to 2014

Effects in the model => herd × management, age × parity, inbreeding, heterosis, herd × sire,
animal, PE, residual

Materials and Methods II

Validation for cows was based on correlations between (G)EBV and adjusted phenotypes, whereas for bulls, the latter was replaced by daughter yield deviation

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

Materials and Methods III

Evaluations were done for:

- **SINGLE** = each breed separately
- **3-BREEDS** = AY-BS-GU (with breed effect in the model and other effects breed-specific)
- **ALL** = the five breeds together (with breed effect and other effects breed-specific)

Predictive abilities for cows

Breed	n	Trait	ssGBLUP			
			BLUP	SINGLE	ALL	3-BREEDS
AY	181	milk	0.45	0.47	0.44	0.47
		fat	0.51	0.54	0.39	0.53
		prot	0.52	0.54	0.45	0.54
BS	2,423	milk	0.23	0.39	0.29	0.39
		fat	0.27	0.42	0.33	0.42
		prot	0.27	0.44	0.34	0.44
GU	750	milk	0.24	0.35	0.37	0.33
		fat	0.27	0.33	0.28	0.32
		prot	0.24	0.34	0.31	0.32
HO	577,340	milk	0.31	0.55	0.55	—
		fat	0.34	0.55	0.54	—
		prot	0.33	0.52	0.52	—
JE	90,666	milk	0.32	0.50	0.50	—
		fat	0.29	0.46	0.46	—
		prot	0.35	0.51	0.50	—

Reliabilities for bulls

Breed	n	Trait	BLUP	ssGBLUP		
				SINGLE	ALL	3-BREEDS
AY	17	milk	0.20	0.27	0.36	0.43
		fat	0.20	0.31	0.58	0.44
		prot	0.30	0.43	0.56	0.57
BS	107	milk	0.15	0.40	0.12	0.40
		fat	0.14	0.30	0.15	0.34
		prot	0.21	0.45	0.21	0.44
GU	28	milk	0.18	0.29	0.33	0.25
		fat	0.20	0.31	0.27	0.29
		prot	0.18	0.25	0.22	0.24
HO	3,278	milk	0.28	0.77	0.77	—
		fat	0.31	0.77	0.76	—
		prot	0.30	0.74	0.74	—
JE	471	milk	0.43	0.71	0.69	—
		fat	0.36	0.64	0.63	—
		prot	0.45	0.69	0.67	—

Computing cost

Breed	Genotypes	Rounds		Sec / round		Time	
		BLUP	ssGBLUP	BLUP	ssGBLUP	BLUP	ssGBLUP
AY	9,202	504	863	0.08	0.08	< 1 min	~ 1 min
BS	47,309	364	867	0.18	0.45	1 min	~ 6 min
GU	5,032	345	757	0.07	0.07	< 1 min	< 1 min
HO	3.4 M	457	473	21.25	56.31	2.7 hours	7.4 hours
JE	427,286	586	432	2.00	5.58	~ 20 min	~ 40 min
AY-BS-GU	61,543	592	1534	0.31	0.81	~ 3 min	~ 20 min
ALL	3.9 M	643	1142	27.01	64.84	4.8 hours	~ 20 hours

Lower reliability with 5 breeds

Under ssGBLUP, predictability (reliability) for AY, BS, and GU was on average 21% (9%) lower in the five-breed compared to single-breed model

- How many SNP do we need to cover all breeds?
 - 80k SNP possibly not covering all breeds
 - Steyn et al. (2019) breed-specific SNP
 - No drop in accuracy when limited information per breed
- 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?

Conclusions

No changes were observed for HO in the five-breed model because of the greatest number of animals

Combining AY-BS-GU into one evaluation resulted in predictions similar to the ones from single-breed

Single-step large-scale multibreed evaluations are computationally feasible but fine-tuning is needed to avoid a reduction in reliability when numerically dominant breeds are combined

Thanks!



ANIMAL BREEDING AND GENETICS GROUP

UNIVERSITY OF GEORGIA