One *Moo-ve* Closer: Single-Step Genomic Predictions for Crossbred Holstein and Jersey Cattle Using Metafounders

I. Ampofo[1,2], G. Vargas[1], D. Gonzalez-Peña[1], T. L. Passafaro[1], Y. L. Bernal Rubio[1], L.M. P. Sanglard[1], N. Vukasinovic[1], and B.O. Fragomeni[2,3]

[1] Zoetis Genetics, Kalamazoo, Michigan, United States

[2] Department of Animal Science, University of Connecticut, Storrs, Connecticut

[3] Institute for System Genomics, University of Connecticut, Storrs, Connecticut

Corresponding author: issabelle.ampofo@uconn.edu

Abstract

The study examined the impact of incorporating metafounders (MF) in single-step genomic BLUP (ssGBLUP) models for the genetic evaluation of Holstein (HO) and Jersey (JE) cattle with their crossbreds (CROSS). The dataset included 23,736,975 records on 8,560,986 cows. Genotypic data on 181,379 JE, 1,905,292 HO, and 53,799 CROSS animals was used for the evaluation. The genetic evaluation included five production traits, namely milk yield (MY), protein yield (PY), fat yield (FY), somatic cell score (SCS), and daughter pregnancy rate (DPR), which were analyzed using a five-trait repeatability model using ssGBLUP with or without MF. Three different MF scenarios were tested: 4MF (based on breed), 24MF (based on the combination of breed, sex, and year of birth), and 32MF (similar to 24MF but with CROSS as a separate genetic group). The three MF scenarios were compared to a conventional ssGBLUP model that did not include metafounders (NO_MF). Forward-in-time validation was carried out to evaluate predictability, inflation, and stability. For purebred Holstein and Jersey cows, the truncated dataset included phenotypes through December 2018, whereas for crossbreds the cutoff was December 2015; the complete dataset extended through December 2022. Validation targeted genotyped cows lacking records in their respective truncated dataset but with at least one record in the complete dataset, yielding 96, 295 Holsteins 26, 436 Jerseys, and 5,099 crossbreds for analysis. Results showed that including MF affected prediction metrics differently depending on the trait, breed, and MF configuration. While certain MF classifications (e.g., 4MF) reduce bias and improved predictability in crossbreds for some traits, others showed minimal effects, particularly in purebred Holsteins. For low heritability traits (SCS, DPR), MF scenarios provided better predictive ability in CROSS animals. In contrast, for high heritability traits (MY, PY, FY), stability tended to decrease in MF models, suggesting possible overfitting due to added model complexity. Overall, MF offers a promising strategy to address pedigree gaps in multibreed evaluations, but its application should be carefully tailored to trait architecture and population composition to avoid overfitting and ensure accurate genetic predictions.

Key words: base population, genomic evaluation, metafounder, single-step genomic BLUP

Introduction

Traditionally, genomic evaluations for dairy cattle have been conducted on a single-breed basis, often excluding crossbred animals. However, the growing proportion of crossbreds in U.S. herds underscores the importance of including them in evaluations to improve management decisions. From 1990 to 2018, the proportion of crossbred cows in the U.S. Dairy Herd Improvement program rose from 0.1% to 5.3% (Guinan et al., 2019). Recognizing this trend, the Council on Dairy Cattle Breeding (CDCB) extended genomic evaluations to

crossbred animals in 2019 (Wiggans et al., 2019; CDCB, 2020).

Several methods have been proposed for joint evaluations of purebred and crossbred animals (Wei & van der Werf, 1994; Christensen et al., 2014; Steyn et al., 2021; VanRaden et al., 2020). A straightforward approach combines all genotypes in a single relationship matrix (Lourenco et al., 2016). The single-step genomic BLUP (ssGBLUP) approach integrates pedigree (**A**) and genomic (**G**) matrices to estimate genomic breeding values (GEBV) (Aguilar et al., 2010; Christensen & Lund, 2010). However, ssGBLUP requires uniform scaling between **A** and **G** and a consistent base population (Christensen, 2012). Incomplete pedigrees and population stratification complicate these assumptions.

To address these issues, Thompson (1979) and Quaas (1988) introduced unknown parent groups (UPG) to account for missing pedigree information. More recently, Legarra et al. (2015) proposed metafounders (MF) to model relationships among base populations, improving compatibility between **A** and **G**. MF consider allele frequencies of 0.5 across markers and estimate relationships among pseudo-ancestors using a gamma matrix (**Γ**). Studies have shown that MF can improve prediction accuracy in multibreed populations (Garcia-Baccino et al., 2017; Xiang et al., 2016).

Despite these advances, limited work has assessed MF performance in combined Holstein-Jersey ssGBLUP models, particularly regarding crossbred evaluations. This study aims to evaluate different MF classifications and their effects on accuracy, bias, and stability in genomic predictions of purebreds and crossbreds.

Materials and Methods

Official data files from Zoetis Inc. were used for this study. Phenotypic and pedigree data were sourced from U.S. dairy producers via backups from herd management systems (DairyComp 305, PC Dart, and DHI Plus). Quality control excluded lactations with data collection ratings (DCR) <0.70 or implausible yields, and pedigree was traced back 20 generations where possible. Pedigree completeness varied: 57.2% of animals had known parents, 10.3% had missing sires, 8.6% had missing dams, and 23.8% had both parents unknown.

DNA was extracted and genotyped on Illumina BeadArray platforms (3K–80K SNPs). Low-density genotypes (<40K SNPs) were imputed to 45,245 markers using FImpute (Sargolzaei et al., 2011), achieving 97% concordance.

The genetic evaluation included five production traits: milk yield (MY), protein yield (PY), fat yield (FY), somatic cell score (SCS), and daughter pregnancy rate (DPR). Official records comprised 23.7 million observations on 8.56 million cows, with genotypes available for 1.91 million Holsteins (HO), 181,379 Jerseys (JE), and 53,799 crossbred (CROSS) animals. Table 1 summarises the total number of records and number of studied animals across traits defined by breed. Heritabilities (±SE) for the five traits were 0.35 (0.005) for MY, 0.29 (0.008) for FY, 0.31 (0.014) for PY, 0.13 (0.008) for SCS, and 0.07 (0.003) for DPR.

Table 1: number of records and cows with phenotypes and genotypes

Group	Phenotypes	Genotyped animals (ssGBLUP only)	
	N	Cows	
Holstein	20,166,782	7,298,374	1,905,292
Jersey	2,868,461	996,353	181,379
*CROSS	701,732	266,259	50,938
Total	23,736,975	8,560,986	2,137,609

*CROSS = Crossbred of Holstein x Jersey, N = Number of records, Cows = Number of cows with records

Genomic breed composition was determined using a supervised admixture model (Zoetis proprietary pipeline). Purebred HO and JE were defined as ≥80% ancestry; CROSS animals had combined HO and JE ancestry ≥80%. Three validation sets were created: 96,295 HO, 26,436 JE, and 5,099 CROSS cows. Reduced datasets included records until Dec 2018 (HO, JE) or Dec 2015 (CROSS); complete datasets extended to Dec 2022.

Models included five-trait repeatability with random animal, permanent environment, and

herd × sire interaction effects, and fixed effects included contemporary groups, heterosis and inbreeding. The genomic relationship matrix (G) and pedigree matrix (A) were combined in a single-step GBLUP (ssGBLUP) using the Algorithm for Proven and Young (APY) (Legarra et al., 2009; Aguilar et al., 2010) with a random core size of 30,000: 22,156 females and 1,931 males for HO, 5,643 females and 181 males for Jersey, and 678 females for crossbred. Models were solved based on iteration on data with the preconditioned conjugate gradient (PCG) in algorithm BLUP90IOD2OMP1 (Tsuruta et al., 2001).

Forward-in-time validation assessed (1) predictability as the correlation between adjusted phenotypes and GEBVs. Adjusted phenotypes were obtained using PREDICTf90 v1.3 (Misztal et al., 2014); (2) inflation as the regression slope of phenotypes on GEBVs (ideal slope = 1); and (3) stability as the correlation between GEBVs estimated from reduced and complete datasets. Standard errors for predictabilities and stabilities were computed following Bermann et al. (2024). All regression and correlation analyses were performed in R software (R Development Core Team, 2024). The different models with and without MFs are detailed next.

SSgblup Analyses

All computations with ssGBLUP were done using the full pedigree with 27 million animals and the genomic relationship matrix for 2,137,609 animals. The ssGBLUP allows the creation of a joint relationship matrix for genotyped and non-genotyped animals by replacing the inverse of the pedigree relationship matrix, $\mathbf{A}^{-1}$, with the inverse of the **H** matrix that combines the pedigree (**A)** and the genomic relationship matrix **G** (Legarra et al., 2009; Aguilar et al., 2010):

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

where $\mathbf{A}^{-1}$ is an inverse of the pedigree relationship matrix; $\mathbf{G}^{-1}$ is an inverse of the genomic relationship matrix (VanRaden, 2008); and $\mathbf{A}_{22}^{-1}$ is an inverse of the pedigree relationship matrix for genotyped animals only.

Single-step GBLUP with metafounders

The $\mathbf{H}^{-1}$ matrix considers relationships among MF ($\boldsymbol{\Gamma}$) in the MF approach. Hence, it is replaced with the $(\mathbf{H}^{\boldsymbol{\Gamma}})^{-1}$ matrix, as described by Legarra et al. (2015) and Christensen et al. (2014). In this way, the $\mathbf{H}^{-1}$ matrix is modified to become:

$$\left(\mathbf{H}^{\boldsymbol{\Gamma}}\right)^{-1} = \left(\mathbf{A}^{\boldsymbol{\Gamma}}\right)^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{G}_{05}^{-1} - \left(\mathbf{A}_{22}^{\boldsymbol{\Gamma}}\right)^{-1} \end{bmatrix}$$

Where $\boldsymbol{G_{05}} = \frac{(\boldsymbol{M}-\boldsymbol{P})(\boldsymbol{M}-\boldsymbol{P})^{T}}{\boldsymbol{k}}$, where **M** is the matrix of samples with SNPs encoded as 0, 1, 2 (i.e., the number of reference alleles), **P** is the matrix where each column is filled with the value 1 (i.e., assuming allele frequencies of 0.5 for all loci). The denominator $\boldsymbol{k} = 0.5\boldsymbol{s}$, where $\boldsymbol{s}$ is the total number of SNPs. This corresponds to the genomic relationship matrix proposed by VanRaden (2008) with all allele frequencies assumed to be 0.5. $\mathbf{A}^{\boldsymbol{\Gamma}}$ is pedigree relationship matrix formed with a $\boldsymbol{\Gamma}$ matrix, and $\mathbf{A}_{22}^{\boldsymbol{\Gamma}}$ is the submatrix of $\mathbf{A}^{\boldsymbol{\Gamma}}$ for the genotyped animals, and $\boldsymbol{\Gamma}$ is a variance covariance matrix of the MF estimated by $\boldsymbol{\Gamma}$ = 8*Cov*(**P**), as proposed by García-Baccino et al. (2017), where **P** is an $\boldsymbol{m}$ by $\boldsymbol{r}$ matrix of allele frequencies and **r** is the number of MF. Note that this **P** differs from the allele frequency matrix used earlier for individual SNPs in the genomic relationship matrix. Under ssGBLUP without MF, the genomic matrix **G** was constructed using the allele frequencies observed in the genotyped data. Conversely, ssGBLUP that included MF used a fixed allele frequency of 0.5 for all loci. VanRaden (2008) proposed using allele frequencies from base animals, representing an unselected population, to create the genomic matrix. Using an allele frequency of 0.5 in ssGBLUP with MF represents a relationship across individuals in the base pedigree population(s) relative to an unobserved base population with all allele frequencies equal 0.5 (Legarra et al., 2024). The only modification of the **A** matrix to include MF is the assumption that the MF have a self-relationship denoted as $\boldsymbol{\Gamma}$. The $\boldsymbol{\Gamma}$ matrix, which models the means within and across founders, was estimated using

observed genotypes and pedigree under a generalized least square (**GLS**) approach (Garcia-Baccino et al., 2017) using the gammaf90 software package (Aguilar & Misztal, 2008).

Metafounder classification

This study examined four scenarios to assess the impact of different strategies to build MF for a given data set and pedigree setting:

1) ssGBLUP without MF (NO_MF):

A ssGBLUP that did not include MF nor any UPG was implemented so that all unknown parents in the pedigree are assumed to be unrelated and from a single population, hence having unknown breeding values.

2) ssGBLUP with MF defined by breed (Γ_4):

In this approach, four MF were defined based on the breed of origin, with one MF assigned to HO, one for JE, another for CROSS, and a fourth assigned to the rest of the base animals, assuming their breed of origin was unknown. This approach treated CROSS as a distinct genetic group ("breed") alongside HO, JE, and Unknown. Thus, in the end, the variance-covariance matrix among MF was a 4x4 matrix between the means across SNP and breeds.

3) ssGBLUP with MF defined by breed, sex, and birth year (Γ_{24}):

In this approach, 24 MF were defined based on breed (HO, JE, Unknown), sex, and year of birth (≤2000, 2001–2005, 2006–2010, ≥2011). Here, the CROSS group was modelled within the covariance between HO and JE.

4) ssGBLUP with MF defined by breed, sex, year of birth and crossbreds as a breed (Γ_{32}):

This approach expanded upon $\mathbf{\Gamma_{24}}$ by explicitly treating crossbred animals (CROSS) as a distinct genetic group alongside HO, JE, and Unknown. As a result, metafounders were defined for each combination of breed (HO, JE, CROSS, Unknown), sex, and year of birth, resulting in 32 total metafounders. This distinction allowed animals with mixed ancestry and no known parents to be grouped more consistently, rather than approximating their breed origin via pedigree tracing. In this case, crossbred animals with no parent information were directly assigned to the CROSS metafounder group.

Following Legarra et al (2015), genetic variance parameters obtained from the model with unrelated founders were used to estimate corresponding parameters for the models with MF by scaling it to become;

$$\sigma^2_{related} \approx \frac{\sigma^2_{unrelated}}{1 + \frac{diag(\Gamma)}{2} - \bar{\Gamma}}$$

where the denominator is the scaling factor *k*; $\sigma^2_{unrelated}$ is the variance among unrelated founders. The variance of the breeding values can then be calculated as $\mathrm{var}(u) = \mathbf{H}^{\Gamma}.\sigma^2_{related}$ where $\mathbf{H}^{\Gamma}$ is again the combined relationship matrix described in Legarra et al. (2015).

Comparisons

The four ssGBLUP scenarios were evaluated, where three used different MF classifications and one used a conventional ssGBLUP model without the inclusion of any MF

To confirm these assumptions, we investigated the mean differences in the diagonal and off-diagonals of $\mathbf{A}_{22}$, $\mathbf{G}$, $\mathbf{A}^{\Gamma}{}_{22}$, and $\mathbf{G}^{\Gamma}$ matrices (defined by MF groups) by correlations and mean differences between these matrices.

Finally, the four sets of ssGBLUP predictions were compared using the validation metrics described above for each studied trait.

Results & Discussion

Elements of matrices

Table 2 shows the summary statistics for the different matrices used in the ssGBLUP computation using APY with a random core size of 30,000. Values of the diagonal and off-diagonal elements of $\mathbf{A}_{22}$ and $\mathbf{G}$ increased in all augmentations of $\mathbf{A}$ and $\mathbf{G}$ that considered $\mathbf{\Gamma}$. The mean, minimum, and maximum values of the diagonal and off-diagonal elements of $\mathbf{A}^{\Gamma 24}{}_{22}$, $\mathbf{A}^{\Gamma 32}{}_{22}$, $\mathbf{G}^{\Gamma 24}$, and $\mathbf{G}^{\Gamma 32}$ were similar. This similarity implied that the assignment of an MF

to the crossbred base population in $\mathbf{\Gamma}_{32}$ resulted in the little to no effect on the relationship among individuals when compared with modeling the crossbred base population within the covariance between the MF of HO and JE augmented in $\mathbf{\Gamma}_{24}$.

Incorporating MF in $\mathbf{A}_{22}$ increased the correlation between the pedigree and genomic relationship matrices. Correlation between the diagonal elements of A_{22} and $\mathbf{G}$, $\mathbf{A}^{\Gamma 4}{}_{22}$ and $\mathbf{G}^{\Gamma 4}$ $\mathbf{A}^{\Gamma 24}{}_{22}$ and $\mathbf{G}^{\Gamma 24}$, $\mathbf{A}^{\Gamma 32}{}_{22}$ and $\mathbf{G}^{\Gamma 32}$ were 0.18, 0.64, 0.28, and 0.29, respectively. In the same way, the correlation between the off-diagonal elements of A_{22} and $\mathbf{G}$, $\mathbf{A}^{\Gamma 4}{}_{22}$ and $\mathbf{G}^{\Gamma 4}$ $\mathbf{A}^{\Gamma 24}{}_{22}$ and $G^{\Gamma 24}$, $\mathbf{A}^{\Gamma 32}{}_{22}$ and $\mathbf{G}^{\Gamma 32}$ were 0.39, 0.66, 0.46, and 0.47 respectively. In all scenarios, using the $\mathbf{\Gamma}_4$ resulted in higher-than-average diagonals and off-diagonals in the elements of $\mathbf{A}$ and $\mathbf{G}$. These results were expected as including MF has been shown to improve the similarity between the pedigree and genomic relationship matrices compared to the traditional ssGBLUP model (Legarra et al., 2015).

Furthermore, the off-diagonal elements in a pedigree relationship matrix containing MF are expected to be higher than those of a pedigree without MF (Junqueira et al., 2020; Kudinov et al., 2020), as shown in table 2.

Table 2: mean, minimum, and maximum element values of $\mathbf{A}_{22}$, $\mathbf{A}^{\Gamma 4}{}_{22}$, $\mathbf{A}^{\Gamma 24}{}_{22}$, $\mathbf{A}^{\Gamma 32}{}_{22}$, $\mathbf{G}$, $\mathbf{G}^{\Gamma 4}$,$\mathbf{G}^{\Gamma 24}$,$\mathbf{G}^{\Gamma 32}$ from diagonal and off-diagonal[1].

Element	Matrix	Mean	Minimum	Maximum
Diagonal	*$\mathbf{A}_{22}$	1.004	1.000	1.286
	$\mathbf{G}$	1.004	0.779	1.453
	$\mathbf{A}^{\Gamma 4}{}_{22}$	1.324	1.266	1.551
	$\mathbf{G}^{\Gamma 4}$	1.320	1.121	1.568
	$\mathbf{A}^{\Gamma 24}{}_{22}$	1.306	1.008	1.504
	$\mathbf{G}^{\Gamma 24}$	1.319	1.120	1.568
	$\mathbf{A}^{\Gamma 32}{}_{22}$	1.306	1.163	1.504
	$\mathbf{G}^{\Gamma 32}$	1.319	1.120	1.568
Off-diagonal	$\mathbf{A}_{22}$	0.016	0.000	0.666
	$\mathbf{G}$	0.016	-0.216	1.015
	$\mathbf{A}^{\Gamma 4}{}_{22}$	0.613	0.532	1.154
	$\mathbf{G}^{\Gamma 4}$	0.632	0.397	1.386
	$\mathbf{A}^{\Gamma 24}{}_{22}$	0.604	0.385	1.073
	$\mathbf{G}^{\Gamma 24}$	0.631	0.395	1.380
	$\mathbf{A}^{\Gamma 32}{}_{22}$	0.604	0.413	1.073
	$\mathbf{G}^{\Gamma 32}$	0.631	0.396	1.380

*$\mathbf{A}_{22}$ is the pedigree relationship matrix of the genotyped animals; $\mathbf{G}^{\Gamma 4}$, $\mathbf{G}^{\Gamma 24}$, and $\mathbf{G}^{\Gamma 32}$ are the genomic relationship matrices with allele frequencies equal to 0.5 augmented by the Γ_4, Γ_{24}, and Γ_{32}, respectively; $\mathbf{G}$ is the genomic relationship matrix obtained using the VanRaden (2008) method 1; $\mathbf{A}^{\Gamma 4}{}_{22}$, $\mathbf{A}^{\Gamma 24}{}_{22}$, $\mathbf{A}^{\Gamma 32}{}_{22}$, are the pedigree relationship matrices of genotyped animals augmented by $\mathbf{\Gamma}_4$, $\mathbf{\Gamma}_{24}$, $\mathbf{\Gamma}_{32}$ respectively.

Inflation

The slope (b_1) of the regression of adjusted phenotypes on GEBV from reduced datasets measures the dispersion of predictions. A slope close to one indicates no inflation or deflation in GEBV (Mäntysaari et al., 2010). According to Interbull guidelines, b_1 should range from 0.90 to 1.10 for large populations, or be within statistical significance of 1.0 for smaller populations. Table 3 summarizes slopes across traits (DPR, FY, MY, PY, SCS), methods (NO_MF, 4MF, 24MF, 32MF), and groups (CROSS, HO, JE).

In CROSS, NO_MF exhibited severe overdispersion, with slopes well below one for MY (0.52 ± 0.08), PY (0.42 ± 0.09), and FY (0.37 ± 0.09). Introducing 4MF improved

dispersion (e.g., MY: 0.63 ± 0.06; PY: 0.51 ± 0.07). However, finer partitions (24MF, 32MF) did not consistently improve slopes and, for MY, slopes declined to 0.45 ± 0.04 (24MF) and 0.54 ± 0.05 (32MF), suggesting potential reintroduction of bias. For low-heritability traits (DPR, SCS), slopes remained far from one and highly variable across scenarios.

In HO and JE, slopes were closer to one across models. HO slopes ranged narrowly (0.68–0.77). For JE, 4MF slightly improved MY slope (0.67 ± 0.03 [NO_MF] → 0.87 ± 0.02 [4MF]), with minimal differences between 4MF, 24MF, and 32MF. These results suggest that coarser MF groupings can reduce overdispersion in CROSS, but finer granularity does not guarantee further improvement and may exacerbate bias.

Overall, slopes were significantly different from 1.0 (*P < 0.05), indicating general inflation in predictions. However, less biased results for evaluations with MF were observed as shown in other studies (e.g., Garcia-Baccino et al., 2017). A potential factor is variance scaling in MF base populations. While Legarra et al. (2015) described theoretical scaling, its practical implementation has been inconsistent (Macedo et al., 2020; Meyer, 2021). Himmelbauer et al. (2024) reported that scaled variances tend to slightly overestimate GEBV. In this study, scaling factors (k) for base animals were 1.002 (4MF), 1.011 (24MF), and 1.015 (32MF), suggesting variance scaling did not contribute to inflation. Breed-specific contributions to the base population, as noted by Kudinov et al. (2022), may explain slope differences across groups.

Suboptimal reference populations and limited crossbred genotypes that did not represent this group in the APY core likely contributed to the overdispersions observed in our study, as shown in Khansefid et al. (2020) and van den Berg et al. (2020).

Table 3: Regression coefficients (b1) and SE of cow-adjusted phenotypes on genomic estimated breeding value from different single-step genomic BLUP (ssGBLUP) scenarios for validation cows.

[1]Scenario	Group [2]	Trait				
		MY	PY	FY	SCS	DPR
NO_MF	HO	0.77	0.55	0.62	0.65	0.20
	JE	0.67	0.6	0.55	0.45	0.19
	CROSS	0.52	0.42	0.37	0.10	0.36
4MF	HO	0.77	0.55	0.70	0.79	0.27
	JE	0.87	0.76	0.69	0.67	0.22
	CROSS	0.63	0.51	0.49	0.29	0.35
24MF	HO	0.68	0.48	0.61	0.78	0.24
	JE	0.89	0.77	0.69	0.65	0.22
	CROSS	0.45	0.37	0.39	0.29	0.28
32MF	HO	0.68	0.49	0.67	0.78	0.24
	JE	0.87	0.76	0.69	0.67	0.22
	CROSS	0.54	0.43	0.44	0.28	0.30

Scenario[1]: NO_MF model (single-step genomic BLUP without metafounders); 4MF (single-step genomic BLUP with four metafounders); 24MF (single-step genomic BLUP with 24 metafounders); 32MF (single-step genomic BLUP with 32 metafounders). Group[2] = HO; Holstein (n = 96,295 animal); JE; Jersey (n = 26,436 animals); CROSS (n = 5,099). MF = Metafounder; [2]SE: HO ≤0.02 for all traits and scenarios; JE: ≤ 0.06 for all traits and scenarios; CROSS: ≤ 0.18 for all traits and scenarios; MY = milk yield; FY = fat yield; PY = protein yield; SCS = somatic cell score; DPR = daughter pregnancy rate

Predictabilities

Table 4 summarizes predictabilities for MY, PY, FY, SCS, and DPR across models (NO_MF, 4MF, 24MF, 32MF) and groups (CROSS, HO, JE). For MY, HO and JE cows showed moderate, stable predictabilities across all models (HO: 0.41–0.44; JE: 0.40–0.50). In contrast, CROSS animals demonstrated notable gains with MF inclusion, increasing from 0.33 under NO_MF to 0.44 with 4MF, and further to

0.48–0.49 under 24MF and 32MF. Incremental gains beyond 4MF were modest, suggesting diminishing returns with finer metafounder definitions.

For PY and FY, similar trends were observed. JE cows exhibited higher baseline predictabilities (PY: 0.31 NO_MF → 0.37–0.38 MF), while HO showed smaller changes. CROSS animals had the largest improvements, especially under 4MF (e.g., FY: 0.19 NO_MF → 0.26 4MF). Gains under 24MF and 32MF were limited.

For SCS and DPR, purebred predictabilities remained high and stable across all models, while CROSS animals showed improvements from low baselines (e.g., SCS: 0.04 NO_MF → 0.12 4MF → 0.14 24MF/32MF).

These results highlight that MF effects on predictability are trait- and breed-dependent. Coarser MF groupings (4MF) improve CROSS predictions, while finer partitions do not guarantee further accuracy and may introduce unnecessary complexity.

Our findings differ from Cesarani et al. (2023), who reported higher CROSS predictabilities than purebreds using UPG in ssGBLUP. They attributed this to genetic divergence between HO and JE and dense genotype panels (imputed 79K SNPs). In contrast, our study used a 45K SNP panel and a random APY core including only 678 crossbreds (<3%). These factors likely reduced CROSS prediction accuracy despite the inclusion of MF.

A more balanced APY core design, like the breed-stratified approach of Tabet et al. (2025), could better capture genetic variation in small groups like CROSS while maintaining computational efficiency. Combining variance-based core selection with breed stratification may offer a promising strategy for future multi-breed evaluations.

Table 4: Predictive ability (Pearson correlation between genomic estimated breeding values and adjusted phenotype) for the validation cows.

Scenario[1]	Group[2]	Trait				
		MY	PY	FY	SCS	DPR
NO_MF	HO	0.41	0.30	0.33	0.21	0.07
	JE	0.40	0.31	0.26	0.17	0.05
	CROSS	0.33	0.24	0.19	0.04	0.08
4MF	HO	0.44	0.31	0.37	0.25	0.08
	JE	0.50	0.38	0.34	0.22	0.06
	CROSS	0.44	0.33	0.26	0.12	0.10
24MF	HO	0.41	0.28	0.35	0.25	0.08
	JE	0.50	0.37	0.34	0.21	0.06
	CROSS	0.48	0.39	0.29	0.14	0.09
32MF	HO	0.41	0.28	0.36	0.25	0.08
	JE	0.50	0.38	0.34	0.22	0.06
	CROSS	0.49	0.39	0.30	0.14	0.09

Scenario[1]: NO_MF model (single-step genomic BLUP without metafounders); 4MF (single-step genomic BLUP with four metafounders); 24MF (single-step genomic BLUP with 24 metafounders); 32MF (single-step genomic BLUP with 32 metafounders). Group[2] = HO; Holstein (n = 96,295 animal); JE; Jersey (n = 26,436 animals); CROSS; HOxJE animals (n = 5,099). MF = Metafounder; [2]SE: HO ≤0.003 for all traits and scenarios; JE: ≤ 0.005 for all traits and scenarios; CROSS: ≤ 0.013 for all traits and scenarios; MY = milk yield; FY = fat yield; PY = protein yield; SCS = somatic cell score; DPR = daughter pregnancy rate

Stabilities

In HO, stability was high under NO_MF (≥0.87 for all traits) as shown in Table 5, reflecting strong agreement between reduced and complete datasets. Including 4MF slightly reduced stability for production traits such as PY (0.87 → 0.77) and MY (0.87 → 0.80), while traits with low heritability (SCS, DPR) remained highly stable (≥0.93). Increasing MF resolution to 24MF and 32MF had negligible additional effects, with correlations for MY and

PY ranging from 0.76 to 0.80 and SCS/DPR remaining ≥0.93. These findings suggest that, for HO, finer MF groupings increased model complexity without enhancing stability and may have even slightly destabilized predictions for certain traits.

In JE, stability was similarly high across all traits in NO_MF (e.g., MY and FY = 0.93) and remained largely unchanged with MF inclusion. Minor improvements in MY stability (0.93 → 0.94 under 4MF) were observed, but finer MF resolutions (24MF, 32MF) did not yield further gains, indicating limited impact of MF on stability in this breed.

In contrast, CROSS animals showed lower stability under NO_MF (e.g., MY = 0.59, PY = 0.52, FY = 0.69) compared to purebreds. MF inclusion modestly improved stability (e.g., MY: 0.59 → 0.61 under 4MF), with larger gains observed under 24MF (MY: 0.73) and 32MF (MY: 0.74). Similar trends were noted for other traits, suggesting that finer MF groupings may better account for heterogeneity in crossbred populations.

These results highlight potential trade-offs. In purebreds, finer MF schemes increased model complexity without clear benefits and may have introduced overparameterization relative to the data. In CROSS, finer MF improved stability but did not consistently translate to higher predictive ability or slopes closer to one. This decoupling suggests that stability alone cannot fully evaluate model performance and must be interpreted alongside other validation metrics and trait architecture.

As Legarra and Reverter (2018) emphasized, high stability does not necessarily reflect improved accuracy. For traits like MY and PY in purebreds, high stability may partly reflect that most genetic variance was captured by earlier data, limiting the impact of new phenotypes. Conversely, in traits with lower heritability (e.g., DPR, SCS), MF inclusion improved stability, indicating that such traits may benefit more from additional information introduced by metafounders.

Stability should therefore be interpreted cautiously. While desirable for routine evaluations, it primarily measures agreement between evaluations and does not indicate which evaluation is more accurate. For traits with low h^2, high stability may reflect unresponsiveness to new data, which could limit genetic progress.

Table 5: stability (correlation between genomic estimated breeding values estimated in the complete and reduced datasets) for validation cows.

Scenario[1]	Group[2]	Trait				
		MY	PY	FY	SCS	DPR
NO_MF	HO	0.87	0.87	0.89	0.95	0.91
	JE	0.93	0.92	0.93	0.92	0.89
	CROSS	0.59	0.52	0.69	0.79	0.88
4MF	HO	0.80	0.77	0.83	0.94	0.93
	JE	0.94	0.92	0.92	0.92	0.92
	CROSS	0.61	0.50	0.54	0.76	0.88
24MF	HO	0.76	0.72	0.78	0.93	0.93
	JE	0.92	0.9	0.91	0.92	0.94
	CROSS	0.73	0.65	0.62	0.82	0.91
32MF	HO	0.80	0.73	0.80	0.93	0.93
	JE	0.92	0.90	0.91	0.92	0.93
	CROSS	0.74	0.65	0.64	0.83	0.91

Scenario[1]: NO_MF model (single-step genomic BLUP without metafounders); 4MF (single-step genomic BLUP with four metafounders); 24MF (single-step genomic BLUP with 24 metafounders); 32MF (single-step genomic BLUP with 32 metafounders). Group[2] = HO (n = 96,295 animal); JE (n = 26,436 animals); CROSS (n = 5,099). MF = Metafounder; [2]SE: HO ≤0.001 for all traits and scenarios; JE: ≤ 0.001 for all traits and scenarios; CROSS: ≤ 0.011 for all traits and scenarios; MY = milk yield; FY = fat yield; PY = protein yield; SCS = somatic cell score; DPR = daughter pregnancy rate

Conclusions

This study demonstrated that incorporating metafounders (MF) into genomic evaluation models for Holstein and Jersey cattle, as well as their crossbreds, can result in differences in prediction metrics, with the effects varying by

trait, breed, and metafounder configuration. While certain MF classifications (eg, 4MF) reduced bias and improved regression slopes in crossbreds for some traits, others had minimal effects, especially for purebred Holstein. However, the added model complexity slightly reduced stability for traits with higher heritability, such as milk yield and protein yield. Overall, while MF provides a promising approach to address pedigree missingness in multibreed evaluations, its application should be tailored to the trait heritability and population composition to avoid potential overfitting and ensure accurate genetic predictions.

Acknowledgments

The authors acknowledge Zoetis Genetics for providing the data for this study. The research described here was a part of the corresponding author's internship program sponsored through the research agreement between the University of Connecticut (Storrs) and Zoetis. The authors have not stated any conflicts of interest.

References

Aguilar, I., I. Misztal, D. Johnson, A. Legarra, S. Tsuruta, and T. Lawlor. 2010. Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93:743–752. https://doi.org/10.3168/jds.2009-2730.

Aguilar, I., and I. Misztal. 2008. Technical note: recursive algorithm for inbreeding coefficients assuming nonzero inbreeding of unknown parents. *J. Dairy Sci.* 91:1669–1672. https://doi.org/10.3168/jds.2007-0575.

Bermann, M., A. Legarra, A. A. Munera, I. Misztal, and D. Lourenco. 2024. Confidence intervals for validation statistics with data truncation in genomic prediction. *Genet. Sel. Evol.* 56:18. https://doi.org/10.1186/s12711-024-00883-w.

CDCB. 2020. Description of national genetic evaluation systems. Accessed Aug. 1, 2020. https://www.uscdcb.com/wp-content/uploads/2020/04/USApro.pdf.

Cesarani, A., D. Lourenco, M. Bermann, E. L. Nicolazzi, P. M. VanRaden, and I. Misztal. 2023. Single-step genomic predictions for crossbred Holstein and Jersey cattle in the United States. *J. Dairy Sci.* https://doi.org/10.3168/jdsc.2023-0399.

Christensen, O.F., Lund, M.S. 2010. Genomic prediction when some animals are not genotyped. *Genet Sel Evol* **42**, 2. https://doi.org/10.1186/1297-9686-42-2

Christensen, O. F. 2012. Compatibility of pedigree-based and marker based relationship matrices for single-step genetic evaluation. Genet. Sel. Evol. 44:1–10. https://doi.org/10.1186/1297-9686-44-37.

Christensen, O. F., P. Madsen, B. Nielsen, and G. Su. 2014. Genomic evaluation of both purebred and crossbred performances. *Genet. Sel. Evol.* 46:23. https://doi.org/10.1186/1297-9686-46-23.

Garcia-Baccino, C. A., A. Legarra, O. F. Christensen, I. Misztal, I. Pocrnic, Z. G. Vitezica, and R. J. Cantet. 2017. Metafounders are related to Fst fixation indices and reduce bias in single-step genomic evaluations. *Genet. Sel. Evol.* 49:34. https://doi.org/10.1186/s12711-017-0309-2.

Guinan, F. L., D. Norman, and J. Dürr. 2019. Changes occurring in the breed composition of U.S. dairy herds. *Interbull Bull.* 55:11–16.

Himmelbauer, J., H. Schwarzenbacher, C. Fuerst, and B. Fuerst-Waltl. 2024. Parent groups and metafounders in single-step genomic BLUP: Insights from a simulated cattle population. *J. Dairy Sci.* https://doi.org/10.3168/jds.2024-24891.

Junqueira, V. S., P. S. Lopes, D. Lourenco, F. F. E. Silva, and F. F. Cardoso. 2020. Applying the Metafounders Approach for Genomic Evaluation in a Multibreed Beef Cattle Population. *Front. Genet.* 11:556399. https://doi:10.3389/fgene.2020.556399

Khansefid, M., M. E. Goddard, M. Haile-Mariam, K. V. Konstantinov, C. Schrooten, G. de Jong, E. G. Jewell, E. O'Connor, E. J. Pryce, H. D. Daetwyler, and I. M. MacLeod. 2020. Improving genomic prediction of crossbred and purebred dairy cattle. *Front. Genet.* 11:598580. https://doi.org/10.3389/fgene.2020.598580

Kudinov, A. A., M. Koivula, G. P. Aamand, I. Strandén, and E. A. Mäntysaari. 2022. Single-step genomic BLUP with many metafounders. Front. Genet. 13:1012205. https://doi.org/10.3389/fgene.2022.1012205

Kudinov, A. A., E. A. Mäntysaari, G. P. Aamand, P. Uimari, and I. Strandén. 2020. Metafounder approach for single-step genomic evaluations of Red Dairy cattle. J. Dairy Sci. 103:6299–6310. https://doi .org/10.3168/jds.2019-17483.

Legarra, A., M. Bermann, Q. Mei, and O. F. Christensen. 2024. Redefining and interpreting genomic relationships of metafounders. *Genet. Sel. Evol.* 56:34. https://doi.org/10.1186/s12711-024-00891-w.

Legarra, A., Aguilar, I., & Misztal, I. 2009. A relationship matrix including full pedigree and genomic information. Journal of dairy science, 92(9), 4656-4663. https://doi.org/10.3168/jds.2009-2061

Legarra, A., O. F. Christensen, Z. G. Vitezica, I. Aguilar, and I. Misztal. 2015. Ancestral relationships using metafounders: Finite ancestral populations and across population relationships. *Genetics* 200:455–468. https://doi.org/10.1534/genetics.115.177014.

Legarra, A., Reverter, A. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genet Sel Evol* **50**, 53 https://doi.org/10.1186/s12711-018-0426-6

Lourenco, D. A., S. Tsuruta, B. Fragomeni, C. Chen, W. Herring, and I. Misztal. 2016. Crossbreed evaluations in single-step genomic best linear unbiased predictor using adjusted realized relationship matrices. J. Anim. Sci. 94:909–919. https://doi.org/10.2527/jas.2015-9748.

Macedo, F. L., Christensen, O. F., Astruc, J. M., Aguilar, I., Masuda, Y., & Legarra, A. 2020. Bias and accuracy of dairy sheep evaluations using BLUP and SSGBLUP with metafounders and unknown parent groups. Genetics Selection Evolution, 52(1), 47.

Mäntysaari, E. A., Liu, Z., & VanRaden, P. 2010. Interbull validation test for genomic evaluations. *Interbull bulletin*, (41), 17-17.

Misztal, I., Tsuruta, S., Lourenco, D. A. L., Masuda, Y., Aguilar, I., Legarra, A., & Vitezica, Z. 2014. Manual for BLUPF90 family programs. University of Georgia. http://nce.ads.uga.edu/wiki/doku.php?id=documentation

Meyer, K. 2021. Impact of missing pedigrees in single-step genomic evaluation. Animal Production Science, 61(11), 1182–1191. https://doi.org/10.1071/AN21045

Quaas, R. L. 1988. Additive genetic model with groups and relationships. *J. Dairy Sci*. 71:1338–1345. https://doi.org/10.3168/jds.S0022-0302(88)79691-5

R Development Core Team. 2024. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2011. FImpute: An efficient imputation algorithm for dairy cattle populations. *J. Dairy Sci.* 94(E-Suppl. 1):421. (Abstr.)

Steyn, Y., D. Gonzalez-Pena, Y. L. Bernal Rubio, N. Vukasinovic, S. K. DeNise, D. A. L. Lourenco, and I. Misztal. 2021. Indirect genomic predictions for milk yield in crossbred Holstein-Jersey dairy cattle. *J. Dairy Sci.* 104:5728–5737. https://doi.org/10.3168/jds.2020-19451.

Tabet, J. M., Lourenco, D., Bussiman, F., Bermann, M., Misztal, I., VanRaden, P. M., Vitezica, Z. G., & Legarra, A. 2025. All-

breed single-step genomic best linear unbiased predictor evaluations for fertility traits in US dairy cattle. *J. Dairy Sci*, *108*(1),694–706. https://doi.org/10.3168/jds.2024-25281.

Thompson, R. 1979. The estimation of heritability with unbalanced data: II. Data available on more than two generations. *Biometrics* 33:497–504. https://doi.org/10.2307/2529364

Tsuruta, S., I. Misztal, and I. Stranden. 2001. Use of the preconditioned conjugate gradient algorithm as a generic solver for mixed-model equations in animal breeding applications. *J. Anim. Sci.* 79:1166–1172. https://doi.org/10.2527/2001.7951166x

van den Berg, I., I. M. MacLeod, C. M. Reich, E. J. Breen, and J. E. Pryce. 2020. Optimizing genomic prediction for Australian Red dairy cattle. *J. Dairy Sci.* 103:6276–6298. https://doi.org/10.3168/jds.2019-17914

VanRaden, P. M. 2008. Efficient Methods to Compute Genomic Predictions. *J. Dairy Sci.*91:4414–4423. https://doi.org/10.3168/jds.2007-0980

VanRaden, P. M., M. E. Tooker, T. C. S. Chud, H. D. Norman, J. H. Megonigal Jr., I. W. Haagen, and G. R. Wiggans. 2020. Genomic predictions for crossbred dairy cattle. J. Dairy Sci. 103:1620–1631. https://doi.org/10.3168/jds.2019-16634

Wei, M., and J. H. J. van der Werf. 1994. Maximizing genetic response in crossbreds using both purebred and crossbred information. *Anim. Sci.* 59:401–413. https://doi.org/10.1017/S0003356100007923

Wiggans, G. R., P. M. VanRaden, E. L. Nicolazzi, M. E. Tooker, J. H. Megonigal Jr., and L. M. Walton. 2019. Extending genomic evaluation to crossbred dairy cattle – US implementation. *Interbull Bull*. 55:46–49.

Xiang, T., B. Nielsen, G. Su, A. Legarra, and O. F. Christensen. 2016. Application of single-step genomic evaluation for crossbred performance in pig. J. Anim. Sci. 94:936–948. https://doi.org/10.2527/jas.2015-9930