Teodoro, M., Gubiani, G., Londoño-Gil, M., Maiorano, A., Mota, L., Cánovas, A., Baldi, F. (2026) Multi-breed genomic prediction for feed efficiency using Nellore functional variants in Guzerat beef cattle, *Proceedings 13th World Congress on Genetics Applied to Livestock Production*: 24147 https://doi.org/10.31274/wcgalp.24147
©2026 *Teodoro, et al.* CC BY-NC-ND

Multi-breed genomic prediction for feed efficiency using Nellore functional variants in Guzerat beef cattle

M. Teodoro[1,2*], G. Gubiani[3], M. Londoño-Gil[4], A.M. Maiorano[5], L.F.M. Mota[2], Á. Cánovas[1], and F. Baldi[2,3]

[1]University of Guelph, Guelph, Canada; [2]São Paulo State University, Jaboticabal, Brazil; [3]University of São Paulo, Pirassununga, Brazil; [4]Topigs Norsvin, 's-Hertogenbosch, The Netherlands; [5]Federal University of Uberlandia, Uberlandia, Brazil; [*]Corresponding author: miller.teodoro@unesp.br

Introduction

Improving feed efficiency traits in modern tropical beef cattle production is essential to meet the world's growing demand for animal protein sustainably. Efficient animals reduce production costs while mitigating environmental impacts and meeting consumer expectations for responsible production practices (Ojo *et al.,* 2014). The Nellore beef cattle breed is well established in breeding programs in Brazil and supported by extensive phenotypic and genomic records. In contrast, the Guzerat, a smaller beef breed, have limited data, constrained genomic selection, and reduced competitiveness (Schmid *et al*., 2022). Multi-breed (MB) models expand reference populations to improve prediction accuracy for smaller breeds (Londoño-Gil *et al*., 2024). However, allele-frequency differences and base-population mismatch can make the pedigree and genomic relationship matrices incompatible in single-step genomic best linear unbiased prediction (ssGBLUP). The metafounder (MF) approach addresses this issue by modelling breed-specific base populations through pseudo-ancestors, enabling coherent MB evaluations within the ssGBLUP (Legarra *et al.,* 2015). Using the extensive Nellore database and its genetic proximity to other indicine cattle breeds provides an opportunity to identify functional variants (FV) that may improve selection accuracy in the smaller breeds. Therefore, the aims of this study were to evaluate the MB genomic prediction model and the impact of integrating functional information for dry matter intake (DMI), residual feed intake (RFI), and residual gain (RG) in Guzerat beef cattle using Nellore breed as the reference population.

Materials & Methods

Phenotypic, pedigree and genotypic data. Data were provided by the National Association of Breeders and Researchers (ANCP, Brazil). No Animal Care and Use Committee approval was required for this study, as the dataset was obtained from an existing database. The Nellore dataset included 30,810 animals phenotyped for DMI, RFI, and RG between 2010 and 2023 and 53,967 animals in the pedigree file. From those, 18,567 animals with phenotypes were genotyped with 65K single-nucleotide polymorphisms (SNPs), imputed to 777K SNPs, and subsequently to whole-genome sequence (WGS) data using FImpute v3 (Sargolzaei *et al*., 2014), with 243 sequenced Nellore sires (aligned with ARS-UCD1.3) as the reference. Quality control (QC) excluded markers with a minor allele frequency < 0.05, call rate < 0.95, no extreme Hardy–Weinberg equilibrium (p-value < 1×10^{-6}), and linkage disequilibrium > 0.95. A total of 1,547,715 SNPs remaining after QC. Further details regarding WGS alignment and QC were provided in Fernandes Júnior *et al.* (2021). The dataset for Guzerat breed comprised 1,501 genotyped animals with a 65K SNPs panel and 6,924 animals with pedigree information, including 1,322 animals with DMI, RFI, and RG phenotypes measured over the same period.

Variants identification. Variance components and genomic estimated breeding values (GEBV) were estimated using ssGBLUP, followed by weighted single-step genome-wide association studies (WssGWAS; Wang *et al*., 2012). The statistical model included contemporary group and dam's age at calving as fixed effects, animal age (linear and quadratic) as covariates, and direct additive genetic and residual as random effects. The WGS data was used to perform WssGWAS analyses associated with feed efficiency traits in the Nellore population using the BLUPF90 programs (Misztal *et al.,* 2002). A total of 21 genomic windows with SNPs associated with FE, each 1 megabase pair in length and explaining more than 0.5 percent of the genetic variance, were annotated using GALLO (Fonseca *et al.,* 2020). Additionally, FV from RNA-Sequencing (RNA-Seq) analysis significantly associated with feed efficiency group (high RFI (n=9) and low RFI (n=9)) in Nellore cattle were obtained from a previous study performed by Alexandre *et al.* (2015). The CLC Genomics Workbench software (Aarhus, Denmark) was used to perform RNA-Seq analysis and the FV identification (including SNPs and insertions/deletions (INDELs) type of variants). Track tools were used to identify fixed and uniquely present SNPs and INDELs variants within each feed efficiency group of the Nellore population. Once these variants were identified, the Ensembl Variant Effect Predictor (VEP) software was used to classify the variants according to their functional consequences on transcripts. Only variants classified as high or moderate functional impact were considered as candidate FV for subsequent analysis.

Genomic prediction. Genomic prediction was assessed using the 65K SNPs panel to obtain GEBV and SNP effects in Guzerat, using ssGLBUP with the animal model described above. Two population scenarios were evaluated to compute GEBV: (i) a single-breed (SB) analysis including only Guzerat animals, and (ii) a MB analysis including Nellore and Guzerat fitted under a metafounder (MF) approach, with one MF defined for each breed. The MF relationship coefficients were 0.31 (Nellore), 0.37 (Guzerat), and 0.29 (between breeds). For each population scenario, three methods were applied: (1) standard ssGBLUP; (2) WssGBLUP, in which SNP weights were derived from SNP effects obtained in a linear weighting step; and (3) WssGBLUP-FV, which further upweighted SNPs located near FVs using the maximum λ values from the linear weighting procedure, similar to approaches that incorporate prior biological information to differentially weight SNPs in genomic prediction (Zhang *et al.*, 2014; Teodoro *et al.*, 2025). This resulted in six models: sSB and sMB (standard ssGBLUP), wSB and wMB (linear WssGBLUP), and fvSB and fvMB (WssGBLUP-FV). The WssGWAS and candidate functional variants used as biological priors were obtained from independent Nellore analyses and then used to predict phenotypes in Guzerat validation population.

Validation. The linear regression (LR) method (Legarra and Reverter, 2018) was used in the validation by computing GEBVs from a whole dataset (all animals with phenotypes) and a partial dataset in which phenotypes of the validation animals were removed. During validation, SNP effects and weights used in the weighted models were estimated from the partial dataset only, after removing phenotypes of the validation animals. The validation set consisted of 119 young genotyped Guzerat animals born in 2022–2023. Prediction accuracy ($\widehat{acc}$) was calculated as follows:

$$\widehat{acc} = \sqrt{\frac{cov(GEBV_w, GEBV_p)}{(1-\bar{F})\sigma_u^2}} \quad (1)$$

where $cov(GEBV_w, GEBV_p)$ is the covariance between whole and partial GEBV, $\bar{F}$ is the mean of the inbreeding coefficient of the validation animals, and σ_u^2 is the additive genetic variance of the whole dataset. Dispersion was assessed by the regression slope of $GEBV_w$ on $GEBV_p$ and bias was calculated as the mean difference between $GEBV_w$ and $GEBV_p$ standardized by the standard deviation of $GEBV_w$.

Results and Discussion

Variance components estimate differed between SB and MB scenarios across traits (Table 1), as expected when the number of phenotypic records and the population change across models. The heritability estimates aligned with those reported in the literature, where DMI shows a moderate-high magnitude, RFI a moderate magnitude, and RG a low magnitude (Mota *et al.,* 2024; Teodoro *et al*, 2026). The heritability standard errors were lower in the MB model for all traits, indicating that estimation can be more precise when the Nellore dataset is included.

Table 1. Estimates of additive genetic variance (σ_a^2) and heritability (h^2) for dry matter intake (DMI), residual feed intake (RFI), and residual gain (RG) in Guzerat beef cattle breed.

Parameter	DMI		RFI		RG	
	SB^1	MB^1	SB	MB	SB	MB
σ_a^2	0.270	0.454	0.120	0.122	0.007	0.006
h^2 (SE[2])	0.28 (0.08)	0.42 (0.02)	0.20 (0.07)	0.23 (0.03)	0.16 (0.05)	0.14 (0.02)

[1] SB - single breed model including only Guzerat animals and MB - multi breed model analysis including Nellore and Guzerat fitted under a metafounder approach. SE, standard error.

Regarding the identified variants, 372 SNPs were obtained from WssGWAS and 5,090 (4,805 SNPs and 285 INDELs) from RNA-Seq, being included as FV in the genomic prediction. The prediction accuracies of GEBV using ssGBLUP or WssGBLUP were presented in Table 2.

Table 2. Genomic prediction for dry matter intake (DMI), residual feed intake (RFI), and residual gain (RG) in Guzerat beef cattle breed.

Methods[1]	DMI			RFI			RG		
	acc[2]	*bias*	b_1[2]	*acc*	*bias*	b_1	*acc*	*bias*	b_1
sSB	0.29	-0.24	0.91	0.30	-0.18	1.18	0.32	-0.14	0.82
wSB	0.43	-0.30	0.75	0.43	-0.25	1.02	0.48	-0.25	0.74
fvSB	0.42	-0.29	0.73	0.43	-0.25	1.03	0.47	-0.24	0.75
sMB	0.42	-0.17	0.88	0.42	0.08	1.04	0.34	-0.17	0.91
wMB	0.56	-0.18	0.80	0.58	0.16	0.91	0.54	-0.31	0.86
fvMB	0.55	-0.19	0.77	0.57	0.16	0.87	0.56	-0.24	0.89

[1] sSB and sMB, standard ssGBLUP for single-breed and multi-breed scenarios, respectively; wSB and wMB, linear weighted ssGBLUP; fvSB and fvMB, weighted ssGBLUP using candidate functional variants.
[2] acc, accuracy; b_1, regression coefficient of GEBV from the whole dataset on GEBV from the partial dataset.

In general, accuracy improved when moving from ssGBLUP to WssGBLUP and from SB to MB across all traits, which is consistent with Londoño-Gil *et al.* (2024). Weighting SNPs and enlarging the reference population with Nellore dataset improved the capture of genetic variance and linkage disequilibrium patterns for the smaller breed. The observed accuracy varied for DMI (0.29 to 0.56), RFI (0.30 to 0.58), and RG (0.32 to 0.56) among models. These results aligned those found by Mota *et al.* (2024), that reported accuracies around 0.60 for RFI and DMI in the Nellore population. The lower accuracy in sSB was expected due to limited phenotypic and genotypic records. The improvement observed in the MB scenarios supports the premise that combining genetically related breeds such as Nellore and Guzerat can increases the effective reference population and improves the estimation of marker effects, depending on population size and genetic divergence between breeds. In that sense, MB benefited from increased sample size and broader allelic diversity, enhancing genomic prediction for complex traits (Londoño-Gil *et al.,* 2024). Regarding bias, DMI showed a slight reduction in every MB model, while RG exhibited a slight increase, and RFI shifted from

underestimation in SB to overestimation in MB. Despite trait-specific differences, an overall negative bias was observed, indicating a mild underestimation of GEBV. The dispersion approached values closer to 1 when comparing SB and MB models, indicating improved calibration of genomic predictions under the MB scenario. Incorporating FV did not alter predictions in the SB and MB models, suggesting either distinct selection signatures between the Nellore and Guzerat beef breeds or that the used functional signals are inconsistent across populations. However, fvMB for RG slightly increased prediction accuracy (from 0.54 to 0.56), while decreasing bias (-0.31 to -0.24) and increasing dispersion (0.86 to 0.89). Given the low heritability for RG (0.14), this result supports the idea that biologically informed variants may be effective for improving the genomic prediction of low-heritability traits (Zhang *et al.,* 2014).

Conclusions

Multi-breed models consistently improved accuracy and stability in genomic predictions compared with single-breed approaches. While functional variants had a limited impact across all models, they provided a slight improvement in genomic prediction for residual gain, suggesting that biologically informed variants may be most beneficial when aligned with the underlying genetic architecture of low-heritability traits. Overall, the results of this study suggest that MB strategies may improve genomic prediction in related beef populations.

Acknowledgements

The author acknowledges the São Paulo Research Foundation (FAPESP, Brazil) for financial support (grant numbers #2016/19514-2, #2017/10630-2, #2018/20026-8, #2023/13805-9, and #2025/10754-0).

References

Alexandre P.A., Kogelman L.J.A., Santana M.H.A., Passarelli D., Pulz L.H., et al. (2015) BMC Genomics 16:1073. https://doi.org/10.1186/s12864-015-2292-8

Fernandes Júnior G.A., Carvalheiro R., Oliveira H.N., Sargolzaei M., Costilla R., et al. (2021) Genetics Selection Evolution 53(1):1-10. https://doi.org/10.1186/s12711-021-00622-5

Fonseca P.A.S., Suárez-Veja A., Marras G., and Cánovas A. (2020) Gigascience 9(12):1-9. https://doi.org/10.1093/gigascience/giaa149

Legarra A., and Reverter A. (2018) Genet. Select. Evol. 50:53. https://doi.org/10.1186/s12711-018-0426-6

Legarra A., Christensen O.F., Vitezica Z.G., Aguilar I., and Misztal I. (2015) Genetics 200(2):566-468. https://doi.org/10.1534/genetics.115.177014

Londoño-Gil M., López-Correa R., Aguilar I., Magnabosco C.U., Hidalgo J., et al. (2024) Journal of Animal Breeding and Genetics 142:43-56. https://doi.org/10.1111/jbg.12882

Misztal I., Tsuruta S., Strabel T., Auvray B., Bruet T., et al. (2002) Proc. 7th WCGALP, Montpellier, France.

Mota L.F.M., Arikawa L.M., Santos S.W.B., Fernandes Júnior G.A., Alves A.A.C., et al. (2024) Scientific Reports 14:6404. https://doi.org/10.1038/s41598-024-57234-4

Ojo A.O., Mulim H.A, Campos G.S., Junqueira V.S., Lemenager R.P., et al. (2024) Animals 14(24):3633. https://doi.org/10.3390/ani14243633

Sargolzaei M., Chesnais J.P., and Schenkel F.S. (2014) BMC Genomics 15(478):1-12. https://doi.org/10.1186/1471-2164-15-478

Schmid M., Stock J., Bennewitx J., and Wellmann R. (2022) Frontiers in Genetics 13:840815. https://doi.org/10.3389/fgene.2022.840815

Teodoro M., Gubiani G., Pereira L., Gubiani A.S., Pereira A.S.C., et al. (2026) Journal of Animal Breeding and Genetics,0:1-9. https://doi.org/10.1111/jbg.70050

Teodoro M., Maiorano A.M., Campos G.S., Albuquerque L.G., and Oliveira H.N. (2025) Journal of Animal Breeding and Genetics, *142*(2), 184-199. https://doi.org/10.1111/jbg.12895

Wang H., Misztal I., Aguilar I., Legarra A., and Muir W.M. (2012). Genetical Research, 94, 73-83. https://doi.org/10.1017/S0016672312000274

Zhang Z., Ober U., Erbe M., Zhang H., Gao N., et al. (2014) PloS one 9(3) e93017. https://doi.org/10.1371/journal.pone.0093017