Campos, G., Junqueira, V., Oliveira, H., Lourenco, D., Cardoso, F. (2026) Genomic selection models for functional longevity in Brazilian Angus cattle, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 23853: https://doi.org/10.31274/wcgalp.23853
©2026 Campos, et al. CC BY-NC-ND

Genomic selection models for functional longevity in Brazilian Angus cattle

G. S. Campos[1], V. S. Junqueira[2], H. R. Oliveira[3], D. Lourenco[4], and F. F. Cardoso[5*]

[1]Interbull Centre - Swedish University of Agriculture Science, Uppsala, Sweden; [2]Bayer Crop Science, Uberlandia, Brazil; [3]Purdue University, West Lafayette, USA; [4]University of Georgia, Athens, USA; [5]Embrapa Southern Livestock, Bage, Brazil; [*]Corresponding author:fernando.cardoso@embrapa.br

Introduction

Cow longevity is a key economic trait that significantly influences both the profitability and sustainability of beef cattle production systems. By increasing cow longevity, farms can reduce heifer rearing costs and maintain a higher proportion of productive, mature animals in the herd, thereby improving overall profitability (Garcia *et al*., 2017). In addition, longevity serves as an important indirect indicator of animal health, resilience, and welfare. However, despite its value, genetic selection for longevity remains challenging because the trait is only observable at the end of an animal's productive life.

A variety of statistical models have been employed to genetically evaluate longevity in cattle. Researchers have explored approaches such as proportional hazard, multiple-trait, and random regression models (RRM). Among these approaches, RRM has demonstrated superior predictive ability and therefore is considered the most effective method for evaluating longevity-related genetic traits in cattle (Jamrozik *et al*., 2008). In parallel, several definitions of longevity have been proposed. Measures that account for an animal's productivity over its lifetime, such as functional longevity (FL; Oliveira *et al*., 2020), are generally preferred, as production strongly influences culling decisions.

Genomic selection is often recommended to improve traits measured late in life or sex-limited, such as FL. By incorporating genomic data into genetic evaluations, prediction accuracy is enhanced, accelerating genetic gains. In this context, the single-step genomic BLUP method (ssGBLUP) offers the advantage of integrating genomic, pedigree, and phenotypic information into a single analysis to calculate genomic estimated breeding values (GEBV) for all animals in the pedigree, resulting in more accurate and less biased GEBV. Therefore, this study aimed to (1) estimate genetic parameters for FL in Brazilian Angus cattle, and (2) assess the performance of ssGBLUP using RRM under two different approaches, comparing them with traditional BLUP for predicting FL.

Materials & Methods

Phenotypic, pedigree, and genotypic data. Angus pedigree and calving records were obtained from the Promebo® Brazilian breeding program. Data quality control (QC) was first performed to discard animals with missing information on date of birth, herd, and identification. Since culling reasons were not provided in this study, cows with no recorded calving event for two consecutive years were considered to have left the herd. Additionally, calving records were retained only for cows with an age at first calving between 20 and 48 months. After data cleaning, 48,976 records were retained to create the longevity trait. FL was described by Oliveira *et al*. (2020) as a cow's capacity to stay in the herd and produce calves over time after her first calving. Therefore, FL was evaluated as a longitudinal trait from 2 to 10 years of age, using the cumulative number of calves produced by each cow at each age as

the phenotypic observation. The pedigree consisted of 180,330 animals, of which 20,446 were genotyped using 12 SNP chips with marker densities ranging from 50k to 150k. After imputation and genotype QC, 71,221 SNPs and 19,357 animals remained for the analyses.

Statistical analysis. The following single-trait RRMs model was in the analysis: **y** = **Xb** + **Hq** + **Za** + **Wp** + **e**, where **y** is the vector of phenotypes; **b** is the vector that includes a cross-classificatory effect for embryo transfer (yes or no), and the Legendre polynomial coefficients for year-season of birth. The **q** is the vector of random Legendre coefficients for herd-year-season effects, **a** is the vector of random Legendre coefficients for additive animal effect, **p** is the vector of random Legendre coefficients for permanent environmental effects, and **e** is the vector of random residuals. It was assumed that the residual variance was homogeneous. The **X**, **H**, **Z,** and **W** are the incidence matrices for **b**, **q**, **a**, and **p**, respectively. The variance components were estimated via Gibbs sampling using a pedigree-based model, and breeding values were either pedigree- or genomic-based, estimated using an iteration-on-data algorithm. Second-, third-, and fourth-order Legendre orthogonal polynomials were tested in the regressions. The genomic breeding values were estimated using the single-step genomic BLUP (ssGBLUP) considering (1) the full genomic matrix (GEBV_FULL), and (2) using the algorithm for proven and young (GEBV_APY) as defined in Misztal *et al*. (2014). The number of core animals in GEBV_APY was determined by the number of eigenvalues that explain 98% of the variance in the genomic relationship matrix (Pocrnic *et al*., 2016). The pedigree-based analyses were used to assess gains in prediction accuracy when adding genomic information. The best-fitting RRM model was selected using the Deviance Information Criterion (DIC).

Validation. Validation was carried out using the linear regression approach proposed by Legarra and Reverter (2018). Predictions were validated for young cows (born after 2018) without phenotypes in the reduced dataset (r) but with complete phenotypes in the whole dataset (w). The prediction accuracy was calculated as follows: $\widehat{acc} = \sqrt{\frac{cov[(G)EBV_w,(G)EBV_r]}{(1-\bar{F})\sigma_a^2}}$, where cov denotes the sample covariance, $\bar{F}$ is the average inbreeding coefficient of the validation cows, and σ_a^2 is the additive variance estimated from the whole dataset. The slope of the linear regression of breeding values from the whole dataset on those from the reduced dataset was used to assess the degree of prediction inflation. Model performance was also evaluated by calculating bias as the mean difference between breeding values estimated from the whole and reduced datasets. Bias estimates close to zero indicate unbiased predictions.

Results & Discussion

In this population, 9068 of the largest eigenvalues are required to explain 98% of the genetic variation. Therefore, a total of 9068 individuals was used as the core group to calculate the inverse of the genomic matrix using the APY. Based on DIC values, the RRM using fourth-order Legendre orthogonal polynomials outperformed lower-order models and was therefore selected as the optimal model in this study. These results align with findings from Oliveira *et al*. (2020) on American Angus cattle and Aliloo *et al*. (2025) in Australian Angus, who also reported that the fourth-order RRM was best when evaluating different longevity definitions, culling reasons, and statistical models. Figure 1A displays the estimated additive, herd–year–season, and permanent environmental variances across different ages. Additive variance remained consistently lower than the other variance components across all ages, increasing gradually to a peak of 0.28 at approximately 10 years of age. Permanent environmental variance accounted for the largest proportion of total variation at most ages, increasing markedly from age four through age ten, reflecting the cumulative impact of individual cow-

specific factors such as health history, reproductive conditions across the productive lifespan. The herd–year–season variance was smallest at age three, when contemporary group effects and early management decisions exerted a stronger influence on retention, and greatest at age ten, when history became a more important determinant of longevity. Consequently, heritabilities for FL were generally low, ranging from 0.02 to 0.08, with peak values observed between ages six and nine when additive genetic variance reached its maximum while total phenotypic variance stabilized. The low heritability estimates reflect the complex, multi-factorial nature of longevity, which is influenced by numerous environmental factors, management practices, and involuntary culling decisions throughout a cow's productive life. The correlation matrix revealed an age-dependent pattern, with the highest correlations (greater than 0.75) between adjacent age groups, particularly from ages five through nine. However, ages two through four exhibited notably weaker genetic correlations with all other ages, with estimates ranging from 0.07 to 0.58. This suggests that FL expressed during early reproductive life (ages 2-4) should be regarded as genetically distinct from FL at later ages. The magnitude of correlations generally increased with cow age, with the strongest inter-age relationships observed among mature cows (ages 6-9), where correlations remained above 0.80 even for non-adjacent ages. This pattern indicates that distinct genetic mechanisms may influence FL across life stages: younger cows are likely subject to selection pressures related to reproductive maturity, early adaptation, and first-calf management, whereas longevity in mature cows reflects the cumulative effects of fitness, sustained productivity, and overall robustness.

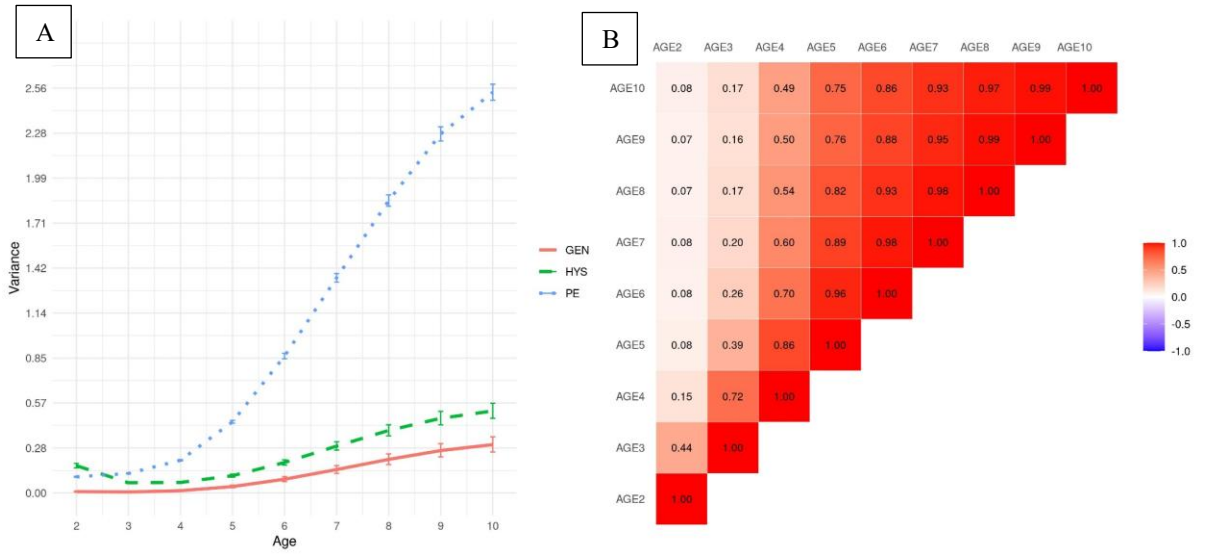

Figure 1. (A) Additive (GEN), permanent environmental (PE), and herd–year–season (HYS) variances of FL across ages in Angus cattle. (B) Heatmap of genetic correlations across ages of FL in Angus cattle.

The breeding values for accuracy were estimated at age 6 because correlations after this age surpassed 0.86, showing the trait's consistency over time and indicating that heritability was maximized at this age. Table 3 presents the estimated prediction accuracy, bias, and inflation of the validation cows. The LR statistics for both genomic methods were comparable, showing prediction accuracies of 0.41, regression coefficients of 0.89, and biases of -0.01. Genomic models achieved 17% higher prediction accuracies than pedigree-based models for young cows.

Table 1. Linear regression statistics for the validation of young cows.

Method[1]	Prediction accuracy	Inflation	Bias
BLUP	0.35	0.77	-0.02
ssGBLUP_FULL	0.41	0.81	-0.01
ssGBLUP_APY	0.41	0.81	-0.01

[1]BLUP: pedigree BLUP; ssGBLUP_FULL: single-step genomic BLUP using the full genomic matrix; ssGBLUP_APY: single-step genomic BLUP using the algorithm for proven and young.

The expected value of bias is zero under a correctly specified model. The observed bias was close to this value, which is important because estimates of genetic progress are meaningful only when predictions are unbiased (Macedo *et al*., 2020). Results indicate that predictions are inflated with the potential presence of extreme predictions. Ideally, the slope should be close to 1, with acceptable deviations within ±15%. Overall, incorporating genomic information into RRM improved prediction accuracy and reduced both bias and dispersion compared to BLUP. Future research should explore RRM with alternative effects and models (e.g., heterogeneous residual variance) and evaluate predictions for different validation groups, including young bulls, before implementing FL in genetic evaluation systems. The validation results of GEBV_FULL and GEBV_APY indicate the two approaches yielded nearly identical performance demonstrating that the computationally efficient APY algorithm can approximate the full genomic relationship matrix without compromising prediction quality.

Conclusions

The RRM using fourth-order Legendre polynomials was identified as the best fit for the genomic evaluation of FL in Brazilian Angus cattle. Although FL exhibits low heritability, genomic selection can substantially improve prediction accuracy. Enhancing FL through genomic improvement has the potential to boost productivity, profitability, and sustainability of the Brazilian beef industry.

Acknowledgements

The author acknowledge data sharing and financial support of Brazilian Breeders Association Herdbook Collares (TCA Grant 21900.25/0002-4).

References

Aliloo, H., van der Werf, J. H., and Clark, S. A. (2025). Journal of Animal Science, 103: skaf035. doi.org/10.1093/jas/skaf035

Jamrozik, J., Fatehi, J., and Schaeffer, L. R. (2008). Journal of Animal Breeding and Genetics, 125(2): 75-83. doi.org/10.1111/j.1439-0388.2007.00712.x

Legarra, A., and Reverter, A. (2018). Genetics Selection Evolution, 50(1): 53. doi.org/10.1186/s12711-018-0426-6

Macedo, F. L., Christensen, O. F., Astruc, J. M., *et al*. (2020). Genetics Selection Evolution, 52(1): 47. doi.org/10.1186/s12711-020-00567-1

Misztal, I., Legarra, A., and Aguilar, I. (2014). Journal of dairy science, 97(6): 3943-3952. doi.org/10.3168/jds.2013-7752

Garcia, D. A., Rosa, G. J. M., Valente, B. D., *et al*. (2017). Animal, 11(12): 2113-2119. doi.org/10.1017/S1751731117001136

Oliveira, H. R., Brito, L. F., Miller, S. P., *et al*. (2020). Animals, 10(12): 2410. doi.org/10.3390/ani10122410

Pocrnic, I., Lourenco, D. A., Masuda, Y., *et al*. (2016). Genetics, 203(1): 573-581. doi.org/10.1534/genetics.116.187013